

На правах рукописи

Феоктистов Александр Геннадьевич

**Организация предметно-ориентированных распределенных
вычислений в гетерогенной среде на основе мультиагентного
управления заданиями**

Специальность 05.13.11 – Математическое и программное
обеспечение вычислительных машин, комплексов и
компьютерных сетей

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
доктора технических наук

Иркутск – 2021

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук (ИДСТУ СО РАН)

- Научный консультант:** **Бычков Игорь Вячеславович** – академик РАН, доктор технических наук, профессор, Институт динамики систем и теории управления имени В.М. Матросова СО РАН, директор
- Официальные оппоненты:** **Каляев Игорь Анатольевич** – академик РАН, доктор технических наук, профессор, Научно-исследовательский институт многопроцессорных вычислительных систем имени академика А.В. Каляева федерального государственного автономного образовательного учреждения высшего образования «Южный федеральный университет», главный научный сотрудник
Легалов Александр Иванович – доктор технических наук, профессор, Национальный исследовательский институт «Высшая школа экономики», профессор
Топорков Виктор Васильевич – доктор технических наук, профессор, Национальный исследовательский университет «МЭИ», заведующий кафедрой вычислительных технологий
- Ведущая организация:** **Вычислительный центр ДВО РАН – обособленное подразделение Хабаровского Федерального исследовательского центра ДВО РАН (г. Хабаровск)**

Защита состоится 7 апреля 2022 г. в 14:00 часов на заседании диссертационного совета Д 003.021.01 при Федеральном государственном бюджетном учреждении науки Институте динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук (ИДСТУ СО РАН) по адресу: 664033, г. Иркутск, ул. Лермонтова, 134.

С диссертацией можно ознакомиться в библиотеке и на официальном сайте www.idstu.irk.ru ИДСТУ СО РАН.

Автореферат разослан 25 января 2022 г.

Ученый секретарь
диссертационного совета,
к.ф.-м.н.

Т.В. Груздева

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В настоящее время решение крупномасштабных фундаментальных и прикладных задач, обеспечивающих прорывные результаты и конкурентное преимущество в научной, производственной, экономической и других сферах человеческой деятельности, обоснованно требует применения высокопроизводительных вычислений. Анализ современных тенденций развития параллельных и распределенных вычислительных систем, базирующийся на теоретических и практических результатах исследований ведущих российских и зарубежных ученых, в том числе С.М. Абрамова, А.И. Аветисяна, А.П. Афанасьева, В.Б. Бетелина, И.В. Бычкова, А.В. Бухановского, Вл.В. Воеводина, В.П. Гергеля, Б.М. Глинского, В.П. Иванникова, В.А. Ильина, В.П. Ильина, И.А. Каляева, В.Н. Коваленко, Д.А. Корягина, В.В. Коренькова, В.Д. Корнеева, И.И. Левина, А.И. Легалова, Ю.Е. Малашенко, В.Э. Малышкина, Г.А. Опарина, Г.И. Радченко, С.И. Смагина, Л.Б. Соколинского, А.А. Сорокина, В.В. Стегайлова, О.В. Сухорослова, В.В. Топоркова, В.Г. Хорошевского, А.Н. Черных, Б.Н. Четверушкина, М.В. Якобовского, D. Abramson, D. Andersen, P. Bouvry, K. Bubendorfer, R. Buyya, H. Casanova, T.L. Casavant, R.S. Chang, E. Deelman, J. Dongarra, I. Foster, K. Kesselman, Y.C. Lee, P. Sloot, D. Talia, T. Tannenbaum, H. Topcuoglu, A. Zomaya и других известных специалистов, позволяет сделать ряд важных выводов.

Тенденция развития современных гетерогенных распределенных вычислительных сред (ГРВС), таких как грид-системы или облачные инфраструктуры, заключается в стремительном росте их масштаба и суммарной производительности, которые зачастую достигаются путем увеличения числа их элементов. В связи с этим одним из ключевых свойств современных научных приложений, таких как распределенные пакеты прикладных программ (РППП) и системы, основанные на применении научного рабочего процесса (англ., workflow), является их масштабируемость (обеспечение роста ускорения вычислений в процессе решения задачи при увеличении числа узлов и поддержка близкой к единице эффективности использования ресурсов – отношения ускорения к числу узлов).

Как правило, разработка приложений и генерация заданий по решению задач осуществляется с помощью специализированных инструментальных средств. В то же время выполнение заданий и выделение им ресурсов в ГРВС производится традиционными системами управления распределенными вычислениями, установленными в узлах среды. В их числе PBS, LSF, HTCCondor, SGE и SLURM. В качестве связующего программного обеспечения (ПО) зачастую используются различные метапланировщики (GridWay, Condor DAGMan, PanDA, Nimrod/G, gLite и др.), обеспечивающие взаимодействие между вычислительными подсистемами инструментальных комплексов и вышеупомянутыми системами управления в узлах среды. Успешным примером реализации подобного подхода является грид-инфраструктура Национальной нанотехнологической сети (ГридННС).

Для повышения гибкости и управляемости гетерогенной среды выполнения распределенных вычислений зачастую прибегают к виртуализации. В

виртуализированных средах используются различные гипервизоры, такие как Xen, KVM и VMware ESXi, средства управления контейнерами LXC и Docker, а также программные комплексы, подобные платформам OpenStack и Open Nebula.

Конечные пользователи приложений, запускающие задания по решению своих задач в ГРВС, являются, как правило, специалистами в разных предметных областях. Они имеют свои субъективные требования к выполнению приложений и эксплуатации общих ресурсов среды. В роли пользовательских критериев качества решения задачи в диссертации рассматриваются общее время и стоимость решения задач, а также ускорение вычислений. Наличие большого числа пользователей ГРВС и разнообразие спектра решаемых ими задач обуславливают необходимость предоставления им разных уровней обслуживания в зависимости от заданных критериев качества решения задачи, учитывая в то же время предпочтения владельцев ресурсов среды. В рамках диссертации к таким предпочтениям относятся эффективность их использования, средняя загрузка, а также степень ее балансировки, оцениваемая среднеквадратическим отклонением. Эти показатели, а также ускорение вычислений рассчитываются традиционными методами¹.

Обеспечение требуемых уровней обслуживания в ГРВС невозможно без комплексного решения ряда проблем: гибкого администрирования процесса обслуживания заданий; учета специфики предметных областей задач пользователей; устранения различного рода неопределенностей, возникающих при планировании вычислений и распределении ресурсов из-за полного или частичного отсутствия информации о характеристиках и свойствах вычислительных процессов; обеспечения масштабируемости вычислений в разнородной среде и др. Решение этих проблем с помощью широко используемых сегодня вышеперечисленных систем управления неосуществимо в полной мере, так как они изначально создавались для других целей. Их информационно-управляющие модели, как правило, не совместимы с моделями, описывающими предметные области приложений.

Поэтому активно развиваются подходы к решению проблем обеспечения удовлетворительного уровня обслуживания в гетерогенной вычислительной среде, базирующиеся на интеллектуализации методов и средств управления с помощью мультиагентных технологий. Большой вклад в развитие теории и практики интеллектуальных систем управления внесли С.Н. Васильев, Т.А. Гаврилова, В.И. Городецкий, А.И. Дивеев, И.А. Каляев, Л.В. Массель, Д.А. Поспелов, К.А. Пупков, П.О. Скобелев, В.Б. Тарасов, В.Ф. Хорошевский, В.П. Хранилов, И.А. Шеремет, N. Jennings, A. Kusiak, M. De Weerdт, G. Weiss, M. Wooldridge и другие известные специалисты в этой области.

При создании интеллектуальных систем управления распределенными вычислениями особое внимание привлекает использование экономических механизмов управления (см. работы В.В. Топоркова, В. Abramson и R. Buyya), обеспечивающих повышение качества выполнения вычислительных процессов

¹ Воеводин, В.В. Параллельные вычисления / В.В. Воеводин, Вл.В. Воеводин. — СПб.: БХВ–Петербург, 2002. — 608 с.

(сокращение времени и стоимости решения задач, повышение эффективности использования ресурсов, балансировки их нагрузки и др.).

Успешные практические решения в области мультиагентного управления распределенными вычислениями в гетерогенных средах известны (см., например, работы В.И. Городецкого, И.А. Каляева, Э.В. Мельника, А.В. Тимофеева, J. Cao, E.H. Durfee, S. Jarvis и A. Singh), однако многие мультиагентные системы создаются с ориентацией на специальные программно-аппаратные инфраструктуры и не могут в полной мере поддерживать требуемое функционирование в средах с иными характеристиками.

Все вышесказанное позволяет сделать вывод об актуальности разработки новых моделей, алгоритмов, инструментальных средств и технологии предметно-ориентированных распределенных вычислений и управления ими в гетерогенной среде.

Цель работы состоит в разработке технологии предметно-ориентированных распределенных вычислений, позволяющей согласовать критерии качества решения задачи и предпочтения владельцев ресурсов и улучшить показатели этих критериев и предпочтений по сравнению с известными метапланировщиками, такими как GridWay и Condor DAGMan, за счет использования мультиагентных технологий управления заданиями и привлечения дополнительных знаний в процессе планирования вычислений и распределения ресурсов.

Основные задачи диссертационного исследования, решаемые для достижения поставленной цели:

- анализ современных подходов к организации распределенных предметно-ориентированных приложений, а также известных методов и средств управления заданиями таких приложений в ГРВС;
- разработка модели предметно-ориентированной ГРВС;
- разработка моделей для определения показателей качества выполнения заданий в такой среде;
- разработка системы классификации заданий;
- разработка алгоритмов мультиагентного управления заданиями РППП;
- разработка подхода к созданию приложений для решения крупномасштабных задач в ГРВС;
- разработка технологии предметно-ориентированных распределенных вычислений и мультиагентного управления в такой среде.

Объектом исследования являются методы и средства организации вычислений и распределения ресурсов в ГРВС.

Предметом исследования выступают модели, алгоритмы и инструментальные средства разработки РППП, а также мультиагентного управления их заданиями в ГРВС.

Методы исследования. При решении поставленных задач использовались методы и средства концептуального и имитационного моделирования, организации распределенных вычислений, реализации мультиагентных технологий, а также экономические механизмы регулирования спроса и предложения вычислительных ресурсов.

Научную новизну диссертации представляют следующие результаты исследования, выносимые на защиту и расширяющие существующий базис теории и практики распределенных вычислений:

- 1) разработана агрегированная модель ГРВС, которая в сравнении с подобными моделями обеспечивает взаимосвязанное представление алгоритмических знаний предметных областей решаемых задач, а также знаний о программно-аппаратной инфраструктуре среды и административных политиках использования ее ресурсов;
- 2) предложены модели и алгоритмы определения показателей качества решения задач в ГРВС, базирующиеся на применении предложенной агрегированной модели среды и набора специализированных методов прогнозирования времени выполнения заданий;
- 3) создана система классификации заданий, позволяющая в отличие от известных систем подобного назначения привлечь дополнительные экспертные знания администраторов узлов ГРВС для детализации спецификаций вычислительных процессов решения пользовательских задач относительно особенностей ресурсов среды с целью снижения неопределенности в распределении заданий по узлам;
- 4) предложен мультиагентный алгоритм планирования вычислений и распределения ресурсов, характерными особенностями которого являются применение экономических механизмов регулирования спроса и предложения этих ресурсов, а также возможность его адаптации к различным моделям сочетания критериев качества решения задач и предпочтений владельцев ресурсов на основе методов дискретного многокритериального выбора;
- 5) разработан пакетный подход к организации предметно-ориентированных вычислений, базирующийся на построении РППП, среда функционирования которых в отличие от других подобных сред может включать ресурсы НРС-кластеров, грид-систем, облачных инфраструктур и других программно-аппаратных компонентов, а также обеспечивать их интегрированное использование для решения крупномасштабных задач;
- 6) создана технология предметно-ориентированных распределенных вычислений в ГРВС, интегрирующая вышеупомянутые модели, алгоритмы, систему классификации заданий и пакетный подход, специализированные инструментальные средства создания и применения РППП, а также программно-аппаратные ресурсы среды в рамках единой технологической цепочки решения крупномасштабных задач.

Практическая значимость. Применение вышеперечисленных результатов диссертационных исследований позволяет обеспечить управление заданиями в предметно-ориентированной ГРВС на основе согласования заданных критериев качества решения задач и предпочтений владельцев ресурсов.

Основные результаты диссертационных исследований были использованы в рамках следующих научно-технических работ:

- гранта № 075-15-2020-787 Министерства науки и высшего образования РФ на выполнение крупного научного проекта по приоритетным направлениям научно-технологического развития (проект «Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории»);
- проектов РФФИ № 01-07-90220-в «Разработка инструментальной среды для создания и поддержки функционирования распределенных пакетов знаний в сети Интернет», № 10-07-00146-а «Средства создания и поддержки проблемно-ориентированных распределенных систем, основанных на знаниях», № 15-29-07955-офи_м «Разработка методов, алгоритмов и инструментальных средств планирования выполнения масштабируемых приложений в разнородной кластерной Grid», № 16-07-00931-а «Методология и инструментальные средства разработки и применения проблемно-ориентированных мультиагентных систем управления масштабируемыми вычислениями в разнородной распределенной вычислительной среде» и № 19-07-00097-а «Фундаментальные проблемы непрерывной интеграции функционального наполнения распределенных пакетов прикладных программ на основе инженерии знаний»;
- регионального проекта РФФИ и Правительства Иркутской области № 20-47-380002 р_а «Математическое и информационное моделирование инфраструктурных объектов Байкальской природной территории»;
- проектов фундаментальных исследований Президиума РАН № 21.6 «Разработка фундаментальных основ создания научной распределенной информационно-вычислительной среды на основе технологий GRID» (2004 г.), № 13.3 «Концептуальные основы и программные средства разработки проблемно-ориентированных распределенных вычислительных сред» (2009-2011 гг.) и № 14.1 «Модели, методы и инструментальные средства для испытания и оценки надежности функционирования проблемно-ориентированных распределенных вычислительных сред» (2012-2014 гг.), проект «Разработка новых подходов к созданию и исследованию моделей сложных информационно-вычислительных и динамических систем с приложениями» программы 1.33П (2016 и 2017 гг.), проект «Фундаментальные проблемы решения сложных практических задач с помощью суперкомпьютеров» программы № 27 (2018 г.), проект «Методы, алгоритмы и инструментальные средства децентрализованного группового решения задач в вычислительных и управляющих системах» программы № 30 (2018 г.);
- НИР № 111-09-103 «Создание и внедрение в учебный процесс вычислительного кластера международного факультета» ФГБОУ ВО «ИГУ» (2008-2010 гг.) и № 111-15-701 «Разработка и внедрение комплексов автоматизированного анализа и систематизации научных данных» (2017-2019 гг.);
- госзаданий и базовых тем исследований ИДСТУ СО РАН.

Под руководством автора разработаны следующие программные комплексы и вычислительные среды, функционирующие на базе интегрированных ресурсов Центра коллективного пользования (ЦКП) «Иркутский суперкомпьютерный центр СО РАН» (ИСКЦ), ИДСТУ СО РАН, Иркутского научного центра (ИНЦ) СО РАН и Международного института экономики и лингвистики (МИЭЛ) ФГБОУ ВО «Иркутский государственный университет» (ИГУ):

- инструментальный комплекс DISCENT для организации распределенных вычислительных сред;
- инструментальные комплексы DISCOMP и Orlando Tools, применяющиеся в качестве сред параллельного программирования ЦКП ИСКЦ;
- инструментальный комплекс SIRIUS III для автоматизации имитационного моделирования систем массового обслуживания;
- экспериментальная Грид ИДСТУ СО РАН;
- учебная информационно-вычислительная среда МИЭЛ ИГУ;
- предметно-ориентированная среда для исследования живучести энергетических инфраструктур.

Результаты диссертации применены в ИСЭМ СО РАН, МИЭЛ ИГУ, Обществе с ограниченной ответственностью (ООО) «Терминал комплекс» и ООО «Катанна». Получены акты о внедрении данных результатов и справки об использовании программ для ЭВМ, разработанных в рамках диссертационного исследования.

Достоверность и обоснованность полученных в диссертации результатов подтверждается корректным применением классических методов исследования, анализом адекватности разработанных моделей и алгоритмов на основе имитационного и полунатурного моделирования, а также в процессе решения практических задач.

Соответствие диссертации паспорту специальности. Тема и основные результаты диссертации соответствуют следующим областям исследований паспорта специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»:

- модели, методы, алгоритмы, языки и программные инструменты для организации взаимодействия программ и программных систем;
- модели и методы создания программ и программных систем для параллельной и распределенной обработки данных, языки и инструментальные средства параллельного программирования.

Апробация. Основные результаты диссертационного исследования докладывались автором на следующих научных мероприятиях: 4th ACM International Conference on Future Networks and Distributed Systems (Санкт-Петербург, 2020 г.); 6th Latin America High Performance Computing Conference (Турриальба, Коста-Рика, 2019 г.); 1st and 2nd International Workshops on Information, Computation, and Control Systems for Distributed Environments (Иркутск, 2019 и 2020 гг.); V Международной конференции «Информационные и нанотехнологии» (Самара, 2019 г.); 5th Latin American Conference on High

Performance Computing (Букараманга, Колумбия, 2018 г.); Международных конференций «Суперкомпьютерные дни в России» (Москва, 2018 и 2019 гг.); научной конференции «Фундаментальные проблемы организации распределенных облачных вычислений при решении крупномасштабных научных задач» (Дивноморское, 2018 г.); IV и V Всероссийских научно-технических конференциях «Суперкомпьютерные технологии» (Дивноморское, 2016 и 2018 гг.); Международной конференции «Информационные технологии в науке, образовании и управлении» (Гурзуф, 2017 г.); научно-технической конференции «Управление в распределенных и сетевых системах» в рамках X и XI Всероссийских Мультиконференций по проблемам управления (Дивноморское, 2017 и 2019 гг.); XVIII Всероссийской конференции молодых ученых по математическому моделированию и информационным технологиям (Иркутск, 2017 г.); Международной конференции «Вычислительная и прикладная математика» в рамках «Марчуковских чтений» (Новосибирск, 2017 г.); XI Международной научной конференции «Параллельные вычислительные технологии» (Казань, 2017 г., Калининград, 2019 г.); 41th International Convention on «Information and communication technology, electronics and microelectronics» (Риека, Хорватия, 2016 г.), 12th International Symposium on Intelligent Systems (Москва, 2016 г.); III Российско-монгольской конференции молодых ученых по математическому моделированию, вычислительно-информационным технологиям и управлению (Иркутск, 2015 г.) и XV Байкальской всероссийской конференции «Информационные и математические технологии в науке и управлении» (Иркутск, 2010 г.), а также семинарах ИДСТУ СО РАН.

Публикации. Результаты научных исследований автора отражены в 76 научных работах. Основные публикации представлены в российских журналах [1-22], рекомендованных ВАК для опубликования научных результатов диссертации, а также проиндексированы в международных базах цитирования Web of Science и Scopus [23-44]. Получено 19 свидетельств о государственной регистрации программ для ЭВМ.

Личный вклад автора. Все выносимые на защиту научные положения получены соискателем лично. Из совместных работ в диссертацию включены только те результаты, которые принадлежат непосредственно автору.

Структура работы. Диссертация состоит из введения, шести глав, заключения, библиографии из 402 наименований, списка принятых сокращений и 11 приложений. Общий объем основного текста работы – 262 страницы, включая 28 таблиц и 80 рисунков.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность диссертационного исследования, сформулированы его цель и основные задачи, показана научная и практическая значимость работы. Кратко изложено содержание диссертации.

В первой главе определены ключевые особенности предметно-ориентированной ГРВС, организуемой на базе ресурсов ЦКП. В зависимости от масштабности решаемых задач в вычислительную инфраструктуру ГРВС могут

быть включены персональные компьютеры (ПК), серверы, кластеры, грид-системы и облачные платформы (рисунок 1). В качестве основных компонентов среды выступают кластеры, которые организуются на базе как выделенных, так и невыделенных узлов.

В рамках ЦКП решают задачи как локальные пользователи отдельных кластеров среды, так и глобальные – те, кому нужны ее интегрированные ресурсы. В обоих случаях они разделяют общие ресурсы среды при решении своих задач. Для решения задачи пользователю необходимо сформировать задание системе, представляющее собой спецификацию вычислительного процесса, содержащую информацию о требуемых ресурсах, исполняемых прикладных программах, входных/выходных данных, а также другие необходимые сведения. Он также определяет критерии качества решения своей задачи: время, стоимость, надежность, безопасность и другие характеристики. На кластерах установлены системы управления прохождением заданий (СУПЗ). Они осуществляют обработку заданий, поступающих в их общие очереди и затем оттуда направляемых на выполнение в назначенные ресурсы среды.

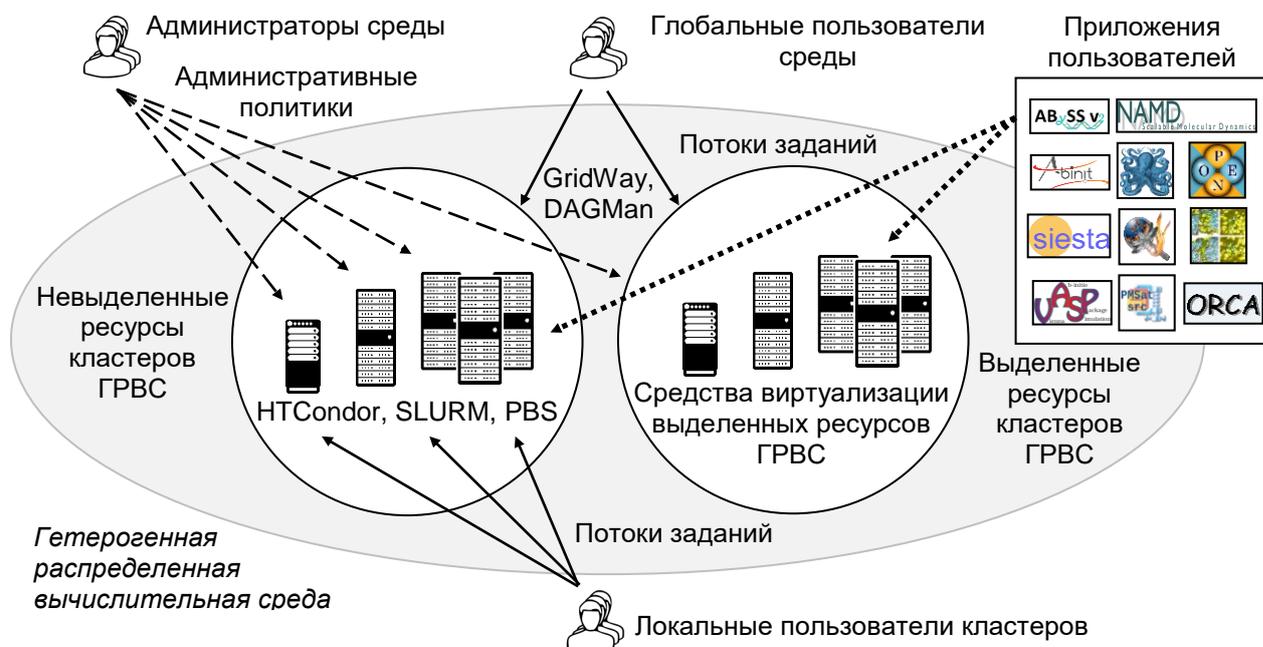


Рисунок 1 – Схема взаимодействия основных субъектов и объектов ГРВС

Интеграция кластерных ресурсов среды может базироваться на следующих технологиях: использование СУПЗ, например, HTCondor, SLURM или PBS Torque, поддерживающих проведение вычислений одновременно на нескольких кластерах; распределение потока заданий с помощью метапланировщиков GridWay или Condor DAGMan; виртуализация выделенной части ресурсов ГРВС с помощью платформы OpenStack.

В диссертации рассматриваются задачи, для решения которых специалисту в предметной области не хватает ресурсов доступных ему компьютеров. Для решения таких задач в ГРВС возникает необходимость представления в модели

предметной области сценариев отображения методов их решения на архитектуру вычислительной среды. При этом одной из важных проблем организации распределенных вычислений является концептуальное моделирование предметной области. Эта проблема характерна как для областей исследований, связанных с инженерией знаний, создания баз данных и информационных систем, так и для предметно-ориентированных технологий программирования. В связи с этим выполнен анализ подходов к построению моделей управления заданиями в распределенных средах, включая методы и средства описания вычислительных моделей в пакетах прикладных программ (см., например, работы М.М. Горбунова-Посадова, В.П. Ильина, Г.А. Опарина, Э.Х. Тыугу) и системах управления рабочими процессами, таких как, например, Kepler, Pegasus или Taverna.

Исходя из специфики предметно-ориентированной ГРВС, предложена агрегированная модель такой среды, в структуру которой включено три компонента знаний: алгоритмические знания предметных областей решаемых задач; знания о программно-аппаратной инфраструктуре среды; экспертные знания администраторов среды. Компоненты знаний и категории субъектов, формирующих эти знания, представлены на рисунке 2. Алгоритмические знания в свою очередь включают прикладные и системные знания.



Рисунок 2 – Структура компонентов знания агрегированной модели

Структура алгоритмических знаний предложена Г.А. Опариным для пакетов прикладных программ, создаваемых в рамках САТУРН-технологии, и в дальнейшем адаптирована автором диссертации для распределенных вычислительных сред. В диссертации схемные знания расширены системными параметрами и операциями для описания процессов, реализуемых системными модулями. Такое расширение обеспечивает принципиально новые возможности адаптации схем решения задач к текущему состоянию ГРВС в процессе их обработки путем мониторинга ресурсов среды, перепланирования вычислений, динамической декомпозиции решаемых задач, формирования новых заданий, прогнозирования их времени выполнения и переназначения ресурсов с помощью системных модулей.

Вторая глава посвящена формализации используемой в диссертации вычислительной модели как фрагмента агрегированной модели ГРВС. Данная модель имеет следующую структуру:

$$M = \langle Z, F, M, PR, S, J, R, A, Q, PLCS, MH, O \rangle,$$

где $Z, F, M, S, J, R, A, Q, O$ – это соответственно множества параметров, операций, программных модулей, схем решения задач, заданий, ресурсов, агентов, ограничений на выполнение заданий и использование ресурсов, накладываемых множеством $PLCS$ административных политик, а также отношений между перечисленными объектами. Структура данных MH отражает вычислительную историю работы модулей из M .

Операции из F определяют отношения между параметрами из множества Z . Каждая операция $f_i \in F$ реализуется модулем $m_j \in M$, где $i \in \overline{1, n_f}, j \in \overline{1, n_m}, n_f$ – число операций, n_m – число модулей. Один модуль может реализовать несколько операций. С каждой операцией f_i связано два подмножества параметров $Z_i^{in}, Z_i^{out} \subset Z$. Подмножество Z_i^{in} включает параметры, значения которых необходимо задать, чтобы получить значения параметров из подмножества Z_i^{out} . Параметры подмножеств Z_i^{in} и Z_i^{out} отражают назначение и семантику формальных параметров модуля m_j , реализующего операцию f_i . В множествах Z, F и M выделены подмножества системных объектов $Z^s \subset Z, F^s \subset F$ и $M^s \subset M$. Такие объекты создаются в вычислительной модели администраторами ГРВС. Модули из M функционируют на ресурсах из R , которые представлены в ГРВС агентами из A . Продукции из PR определяют условия применения операций. Схема $s_k \in S$ выполнения операций из F является аналогом строгой параллельной формы графа алгоритма. Обработка схемы решения задачи осуществляется в асинхронном режиме по готовности данных для ее операций или в режиме fork/join. Схемы решения задач строятся на основе формализованных постановок задач.

Модель M позволяет сформулировать многокритериальную постановку задачи управления распределенными вычислениями: «вычислить $Y = \{y_1, y_2, \dots, y_{n_{out}}\}$ по $X = \{x_1, x_2, \dots, x_{n_{in}}\}$, выполнив $F = \{f_1, f_2, \dots, f_{n_f}\}$ с критериями качества решения задачи $C^u = \{c_1^u, c_2^u, \dots, c_{n_{cu}}^u\}$, указанными пользователем ГРВС, на ресурсах $R = \{r_1, r_2, \dots, r_{n_r}\}$ с критериями качества их

использования $C^o = \{c_1^o, c_2^o, \dots, c_{n_{co}}^o\}$, заданными владельцами этих ресурсов, и учетом $Q = \{q, q_2, \dots, q_{n_q}\}$. Наличие неопределенностей различного рода в условиях задачи (в множествах Y, X, F, C^u, C^o, R и Q) приводит к многообразию вариантов ее постановки.

В общем случае $c_1^u, c_2^u, \dots, c_{n_{cu}}^u$ и $c_1^o, c_2^o, \dots, c_{n_{co}}^o$ представляют собой субъективные критерии качества решения задач пользователей и предпочтения владельцев ресурсов соответственно. Они могут накладывать противоречивые ограничения на схему решения задачи. Предложены модели их определения на основе сочетания экспериментальных данных и предсказательного моделирования.

Модель прогнозирования времени выполнения модулей схемы решения задачи. Модуль в процессе его тестирования запускается на эталонном узле в программной среде для профилирования программ. В процессе работы модуля измеряются его временные затраты на взаимодействие с различными компонентами узла с учетом его вычислительных характеристик. Путем сравнения характеристик эталонного и целевого узлов прогнозируется время выполнения модуля на целевом узле с некоторой погрешностью.

В рамках данной модели формируются наборы $CR = \{c_{r,1}, c_{r,2}, \dots, c_{r,l}\}$ и $CT = \{c_{t,1}, c_{t,2}, \dots, c_{t,l}\}$ характеристик эталонного и целевого узлов и выявляются их значения. Между $c_{r,i}$ и $c_{t,i}$ устанавливается взаимно однозначное соответствие, $i = \overline{1, l}$. Выбирается множество элементов $h_1(x), h_2(x), \dots, h_n(x)$, отражающих вычислительную нагрузку узла при выполнении модуля (число целочисленных операций, число операций с плавающей точкой, число обращений к оперативной памяти и кэш-памяти разных уровней, число промахов таких обращений, число сессий чтения с диска и записи на диск и другие показатели) в зависимости от объема x данных, обрабатываемых модулем. Реализуются функции $t_{r,i}(c_{r,1}, c_{r,2}, \dots, c_{r,l}, h_1(x), h_2(x), \dots, h_m(x))$, рассчитывающие время функционирования l -го компонента узла при обработке вычислительной нагрузки, $l = \overline{1, k}$. Время выполнения модуля в эталонном и целевом узлах оценивается выражениями $\hat{T}_r(x) = \sum_{i=1}^n t_{r,i}(c_{r,1}, c_{r,2}, \dots, c_{r,l}, h_1(x), h_2(x), \dots, h_m(x))$ и $\hat{T}_t(x) = \sum_{i=1}^n t_{t,i}(c_{t,1}, c_{t,2}, \dots, c_{t,l}, h_1(x), h_2(x), \dots, h_m(x))$.

Погрешность α оценки $\hat{T}_r(x)$ находится с учетом реального времени $T_r(x)$ работы модуля: $\alpha = T_r(x) - \hat{T}_r(x)$. Итоговая оценка $\hat{T}'_t(x)$ времени выполнения модуля в целевом узле определяется выражением $\hat{T}'_t(x) = \hat{T}_t(x) + \alpha \frac{\hat{T}_r(x)}{\hat{T}_t(x)}$. Оценка времени функционирования программы для переменного объема данных $x \in (x_0, x_k)$ осуществляется с помощью функции $\hat{T}^*_t(x) = f^*_t(x_0, x_k, x, \hat{T}'_t(x))$, где f^*_t – это некоторая интерполяционная функция, которая, как правило, индивидуально подбирается для каждого класса программ. Оценки времени выполнения модуля $\hat{T}'_t(x)$ и $\hat{T}^*_t(x)$ являются достаточно грубыми. Тем не менее такие прогнозные оценки позволяют, как правило, существенно улучшить пользовательские оценки, а также такие оценки, скорректированные на основе вычислительной истории.

Дополнительно предложено использовать результаты тестирования модулей в узлах среды для получения оценок времени их выполнения. Экспериментально показано, что в случае проведения тестирования полученные на его основе оценки могут улучшить соответствующие прогнозные оценки.

Модель оценки времени выполнения схемы решения задачи. Данная модель включает набор подмоделей для оценки времени обработки схемы решения задачи в двух режимах: асинхронном режиме запуска ее операций по готовности данных и режиме fork/join. Ниже приведены обобщенные формулы расчета времени выполнения схемы решения задачи на основе подмоделей, позволяющие учитывать оценки дополнительных временных затрат на виртуализацию вычислений и проведение рестартов модулей в случае отказов программно-аппаратных средств. Оценка дополнительных временных затрат в разных узлах ГРВС производится экспериментальным путем.

Пусть матрица \mathbf{P} размерности $k \times k$ представляет сведения о предшествовании k операций схемы решения задачи. Элемент $p_{ij} = 1$ ($p_{ij} = 0$) означает, что завершение операции f_j в схеме предшествует (не предшествует) выполнению операции f_i и $Z_i^{in} \cap Z_j^{out} \neq \emptyset$. Матрица $\hat{\mathbf{D}}$ размерности $k \times k$ отражает оценки объемов передаваемых данных между операциями. Элемент $\hat{d}_{ij} \geq 0$ показывает объем данных, передаваемых операцией f_i операции f_j . Матрица \mathbf{W} размерности $k \times k$ предоставляет информацию о пропускной способности интерконнекта между узлами, в которых запускаются модули, реализующие операции схемы. Элемент $w_{ij} \geq 0$ демонстрирует пропускную способность интерконнекта между узлами, в которых работают модули, реализующие операции f_i и f_j . Тогда оценка \hat{T}_s времени выполнения схемы в асинхронном режиме по готовности данных определяется следующим образом:

$$\hat{T}_s = \max_{i=1, \dots, k} \hat{t}_i, \quad \hat{t}_i = \hat{q}_i + \hat{t}_i + \hat{t}_i^{add} + \max_{\forall j \in \{1, \dots, k\}: p_{ij}=1, i \neq j} \left(\hat{t}_j + \frac{\hat{d}_{ji}}{\omega_{ji}(t)w_{ji}} \right),$$

где \hat{t}_i (\hat{t}_j) – оценка времени, прошедшего с начала обработки схемы до завершения операции f_i (f_j), \hat{q}_i – оценка времени нахождения в очереди модуля, реализующего операцию f_i , \hat{t}_i – оценка времени выполнения этого модуля (прогнозная, указанная пользователем, скорректированная с учетом вычислительной истории или полученная на основе тестирования), \hat{t}_i^{add} – дополнительные временные затраты, k – число операций схемы, $0 < \omega_{ji}(t) \leq 1$ – коэффициент снижения пропускной способности интерконнекта в момент времени t между узлами, в которых функционируют модули, реализующие операции f_i и f_j .

Пусть матрица \mathbf{S} размерности $m \times k$ представляет собой ярусно-параллельную форму, описывающую обработку схемы решения задачи в режиме fork/join, m – число ярусов схемы. Элемент $s_{li} = 1$ означает, что операция f_i должна быть выполнена на l -м ярусе. Переход к операциям $(l+1)$ -го яруса возможен при условии завершения всех операций на l -м ярусе. Тогда оценка \hat{T}'_s времени выполнения схемы в режиме fork/join находится следующим образом:

$$\hat{T}'_s = \sum_{l=1}^m \hat{t}_l, \quad \hat{t}_l = \max_{\forall i \in \overline{1, k}: s_{li}=1} \left(\hat{q}_i + \hat{t}_i + \hat{t}_i^{add} + \sum_{\forall j \in \overline{1, k}: p_{ij}=1, i \neq j} \frac{\hat{a}_{ji}}{\omega_{ji}(t)w_{ji}} \right).$$

Модель оценки стоимости выполнения схемы решения задачи. Данная оценка \hat{C}_s определяется по формуле $\hat{C}_s = \sum_{i=1}^k \pi_i \hat{t}_i + \pi_i^*$, где \hat{t}_i – это оценка времени выполнения модуля, реализующего операцию f_i , $\pi_i \geq 0$ – себестоимость единицы процессорного времени узла, на котором он работает, $\pi_i^* \geq 0$ – торговая наценка, закладываемая владельцем данного узла.

Модель надежности схемы решения задачи. Данная модель базируется на использовании методов логико-вероятностного анализа, применяемых к сложным техническим системам в научных работах И.А. Рябины. Пусть $s \in S$ – схема решения задачи, $p^*(t)$ – требуемый показатель надежности схемы в момент времени t , n_s – число сегментов узлов ГРВС, участвующих в выполнении схемы s , а x – набор булевых переменных (параметров) x_{ijk} , отражающих события завершения ($x_{ijk} = 1$) или незавершения ($x_{ijk} = 0$) операции $f_i \in F$ в k -м узле j -го сегмента, $i = \overline{1, n_f}$, $j = \overline{1, n_s}$. Индекс k переменной x_{ijk} указывает соответственно основной ($k = 1$) и резервные ($k \geq 2$) узлы j -го сегмента, выделенные для операции f_i . В каждом j -м сегменте можно назначить n_j резервных узлов для схемы s . Предполагается, что все узлы одного и того же сегмента являются однородными. Возможность запуска модулей в основных или резервных узлах обуславливает различные сценарии выполнения схемы s .

Введем следующие обозначения: $y_i(x)$ – булева функция, определяющая условия применения операции f_i в процессе вычислений по схеме s ; $p_{ijk}(t)$ – вероятность выполнения операции f_i на k -м узле j -го сегмента в момент времени t (стационарный коэффициент готовности, характеризующий k -й узел в произвольный момент времени t по статистическим результатам его работы). Логическая схема надежности схемы s описывается формулами

$$y_1(x) \equiv 1, \quad y_i(x) = \begin{cases} h_i(x), & \text{если } i = 2, \\ x_{ij_1} h_i(x), & \text{если } i > 2, \end{cases} \quad h_i(x) = \bigwedge_{\forall k: w_{ik}=1} y_k(x), \quad (1)$$

где $i = \overline{1, n_f}$, $j_i = \overline{1, n_s}$, $k = \overline{1, n_f}$. Первоначально в этой схеме основной узел j -го сегмента выделяется для операции f_i . Булева функция $y_2(x)$, задающая условия выполнения целевого модуля, вычисляет показатель надежности схемы s . После проведения всех подстановок в (1) функция $y_2(x)$ принимает вид

$$y_2(x) = \bigwedge_{i=3}^{n_f} x_{ij_1}, \quad j_i = \overline{1, n_s}. \quad (2)$$

Для получения показателя надежности схемы s производится переход с помощью соответствующих правил преобразования² от функции $y_2(x)$ к вероятностной функции $P(t) = \prod_{i=3}^{n_f} p_{ij_1}(t)$.

Функция $P(t)$ вычисляет вероятность осуществления единственного имеющегося сценария выполнения схемы s . Если $P(t) < p^*(t)$, то требуется преобразование функции $y_2(x)$ путем улучшения показателей надежности

² Поспелов, А.А. Логические методы анализа и синтеза схем / А.А. Поспелов. — М.—Ленинград: Энергия, 1964. — 320 с.

элементов ее структуры. Функция $y_2(x)$ из (2) для схемы s без резервирования узлов преобразуется в функцию $y'_2(x)$ в дизъюнктивной нормальной форме без отрицаний, которая является монотонной и соответствует схеме с резервированием узлов. Свойство монотонности обеспечивает отсутствие в структуре функции $y'_2(x)$ элементов, для которых улучшение показателей их надежности ухудшает показатели надежности плана s в целом. Функция $y'_2(x)$ имеет следующий вид:

$$y'_2(x) = \bigvee_{l=1}^n K_l, \quad K_l = \bigwedge_{i=3}^{n_f} x_{ijie}, \quad n = \prod_{i=3}^{n_f} \sum_{j=1}^{n_s} n_{ij}, \quad (3)$$

где $e = n_{ij_i}$, n_{ij_i} – число узлов, выделенных в j -м сегменте для операции f_i .

Каждая элементарная конъюнкция K_l в (3) представляет собой один из сценариев выполнения схемы s . Конъюнкции функции $y'_2(x)$ пронумерованы от 1 до n в соответствии с их рангом по возрастанию. С целью обеспечения несовместности этих сценариев производится ортогонализация $y'_2(x)$ с помощью специального алгоритма, предложенного А.А. Поспеловым². В результате применения этого алгоритма функция $\tilde{y}'_2(x)$, ортогональная функции $y'_2(x)$, принимает вид $\tilde{y}'_2(x) = K_1 \vee \bar{K}_1 K_2 \vee \dots \vee \bar{K}_1 \bar{K}_2 \dots \bar{K}_{n-1} K_n$. Затем функция $\tilde{y}'_2(x)$ упрощается путем удаления тождественно равных нулю и поглощаемых конъюнкций. На заключительном этапе реализуется переход от функции $\tilde{y}'_2(x)$ к вероятностной функции $P'(t) = \sum_{i \in I} p'_i(t)$, где $p'_i(t)$ – вероятность осуществления i -го сценария выполнения схемы s . Такой переход осуществляется в соответствии с заданными правилами символического перехода от булевой функции к вероятностной функции. Если $P'(t) < p^*(t)$, то процесс резервирования узлов может быть продолжен при наличии необходимых ресурсов.

Рассмотренный процесс резервирования узлов обеспечивает достижение показателя надежности вычислительного процесса, максимально приближенного к заданному критерию с учетом ограничений на число выделяемых резервных узлов. Данные ограничения гарантируют сходимость процесса резервирования узлов.

Механизмы многокритериального выбора. Важными характеристиками такого рода механизмов для лица, принимающего решение (специалиста предметной области), являются легкость его использования и скорость вычислений, необходимых для выработки управляющих воздействий. Простота механизмов обуславливает в свою очередь требование минимального объема входной информации, предоставляемой специалистом предметной области и необходимой для его применения.

Пусть имеется n критериев $y_1 = \phi_1(\mathbf{x}), y_2 = \phi_2(\mathbf{x}), \dots, y_n = \phi_n(\mathbf{x})$, показывающих качество вычислительного процесса, где $\mathbf{x} = (x_1, x_2, \dots, x_l)$ – это вектор входных переменных его модели, $x_1, x_2, \dots, x_l, y_1, y_2, \dots, y_n \in Z$. Для различных y_j абстрактная связь $\phi_j(\mathbf{x})$ может являться функциональным, статистическим, неоднозначным или иным отображением, $j = \overline{1, n}$. Для каждого критерия заданы его предельные значения y_j^{min} и y_j^{max} такие, что $y_j^{min} \leq y_j \leq y_j^{max}$, а также указано условие его оптимальности

$$y_j \rightarrow y_j^{min} (y_j^{max}), \quad j = \overline{1, n}. \quad (4)$$

Предельные значения и условия оптимальности представлены матрицей \mathbf{Y}^* размерности $3 \times n$, где y_{1j}^* и y_{2j}^* содержат соответственно y_j^{min} и y_j^{max} , а $y_{1j}^* = 0$ ($y_{1j}^* = 1$) определяет условие оптимальности $y_j \rightarrow y_j^{min}$ ($y_j \rightarrow y_j^{max}$).

Матрица \mathbf{V} размерности $m \times n$ содержит множество значений критериев y_1, y_2, \dots, y_n , полученных путем варьирования значений переменных x_1, x_2, \dots, x_l при моделировании вычислительного процесса. Ее i -я строка $\mathbf{V}(i,) = (v_{i1}, v_{i2}, \dots, v_{in})$ соответствует i -му варианту значений этих критериев, а ее j -й столбец $\mathbf{V}(, j) = (v_{1j}, v_{2j}, \dots, v_{mj})$ включает значения критерия $y_j(\mathbf{x})$ в разных вариантах, $i = \overline{1, m}$, $j = \overline{1, n}$.

Матрица $\widehat{\mathbf{V}}$ размерности $m \times n$ включает оценки значений критериев y_1, y_2, \dots, y_n . Ее i -я строка $\widehat{\mathbf{V}}(i,) = (\hat{v}_{i1}, \hat{v}_{i2}, \dots, \hat{v}_{in})$ соответствует $\mathbf{V}(i,)$, а ее j -й столбец $\widehat{\mathbf{V}}(, j) = (\hat{v}_{1j}, \hat{v}_{2j}, \dots, \hat{v}_{mj})$ соответствует $\mathbf{V}(, j)$, $i = \overline{1, m}$, $j = \overline{1, n}$.

Для оценки значений критерия y_j множество элементов столбца $\mathbf{V}(, j)$ разбивается на подмножества V_1, V_2, \dots, V_k , попарно непересекающиеся. Если $y_j \rightarrow y_j^{max}$, то эти подмножества упорядочиваются по возрастанию значений $v_{1j}, v_{2j}, \dots, v_{mj}$. В противном случае, когда $y_j \rightarrow y_j^{min}$, они упорядочиваются по убыванию значений $v_{1j}, v_{2j}, \dots, v_{mj}$. В соответствии с этим упорядочением каждое подмножество V_s получает свой индекс $s \in \overline{1, k}$, который используется для оценки значений $v_{ij} \in V_s$. Таким образом, $\hat{v}_{ij} = s \forall v_{ij} \in V_s$, $i = \overline{1, m}$, $j = \overline{1, n}$, $s \in \overline{1, k}$. Чем больше значение s , тем лучше качество значений $v_{ij} \in V_s$.

Оценка значений для разных критериев y_1, y_2, \dots, y_n реализуется системной операцией $f_1^s(\mathbf{V}, m, n, \mathbf{Y}^* \rightarrow \widehat{\mathbf{V}})$, которая может выполняться параллельно.

Сформулированы следующие правила многокритериального выбора вариантов значений критериев y_1, y_2, \dots, y_n , представленных строками матрицы \mathbf{V} :

$$\mathbf{V}(i,): (\nexists p: (\hat{v}_{i1} \leq \hat{v}_{p1}) \wedge (\hat{v}_{i2} \leq \hat{v}_{p2}) \wedge \dots \wedge (\hat{v}_{in} \leq \hat{v}_{pn}) \wedge (\exists q: \hat{v}_{iq} \neq \hat{v}_{pq})), \quad (5)$$

$$y_j^{min} \leq v_{ij} \leq y_j^{max}, \quad i \in \overline{1, m}, \quad j = \overline{1, n}, \quad p \in \overline{1, m}, \quad i \neq p, \quad q \in \overline{1, n};$$

$$\mathbf{V}(i,): (\nexists p: \sum_{j=1}^n \text{sign}(\hat{v}_{pj} - \hat{v}_{ij}) > 0), \quad (6)$$

$$y_j^{min} \leq v_{ij} \leq y_j^{max}, \quad i \in \overline{1, m}, \quad j = \overline{1, n}, \quad p \in \overline{1, m}, \quad i \neq p;$$

$$\mathbf{V}(i,): \forall \mathbf{V}(p,) \exists q: \quad (7)$$

$$(\hat{v}_{i1} = \hat{v}_{p1}) \wedge (\hat{v}_{i2} = \hat{v}_{p2}) \wedge \dots \wedge (\hat{v}_{iq} = \hat{v}_{pq}) \wedge (\hat{v}_{i(q+1)} > \hat{v}_{p(q+1)}),$$

$$y_j^{min} \leq v_{ij} \leq y_j^{max}, \quad i \in \overline{1, m}, \quad j = \overline{1, n}, \quad p \in \overline{1, m}, \quad i \neq p, \quad q \in \overline{1, n-1}.$$

Правило (5) определяет оптимальный выбор по Парето в случае отсутствия информации о значимости критериев. Правило (6) применяется при условии равнозначности критериев. Оно предполагает отбор вариантов значений критериев, превосходящих по числу оптимальных оценок другие варианты. Правило (7) предназначено для упорядоченного множества критериев. Использование данных правил обусловлено тем, что они обладают наименьшей сложностью с вычислительной точки зрения по сравнению с другими известными

методами решения подобной задачи, просты в реализации и требуют минимальную дополнительную информацию от эксперта.

Если в результате применения любого из правил (5)-(7) ни один вариант значений критериев y_1, y_2, \dots, y_n не выбран, то задача неразрешима. В этом случае нужно менять постановку задачи и проводить новый эксперимент.

В противном случае в результате выполнения каждого из правил (5)-(7) формируется матрица \mathbf{V}^* размерности $m \times r$, содержащая $r \leq m$ строк матрицы \mathbf{V} , представляющих варианты значений критериев y_1, y_2, \dots, y_n и удовлетворяющих условиям отбора. Если $r > 1$, то дополнительно необходимо избрать единственный подходящий вариант.

Пусть $y_j^{opt} = \min_{i=1, \dots, m} (\max) v_{ij}^*$, $j = \overline{1, n}$ в соответствии с условием (4)

оптимальности критерия $y_j(\mathbf{x})$. Значения $y_1^{opt}, y_2^{opt}, \dots, y_n^{opt}$ принимаются в качестве идеальных значений критериев y_1, y_2, \dots, y_n . Для каждой строки $\mathbf{V}^*(i, \cdot)$ определяется евклидова метрика относительно варианта с идеальными значениями

$$y_1^{opt}, y_2^{opt}, \dots, y_n^{opt}: d_i = \sqrt{\sum_{j=1}^n (y_j^{opt} - v_{ij}^*)^2}.$$

В третьей главе предложена система классификации заданий, позволяющая использовать экспертные знания администраторов узлов ГРВС для детализации спецификаций вычислительных процессов решения пользовательских задач относительно особенностей ресурсов среды с целью снижения неопределенности в распределении заданий по узлам.

Модель классификации. Пусть имеется конечное множество $H = \{h_1, h_2, \dots, h_k\}$ характеристик заданий. Каждая характеристика h_i описывается областью D_i ее допустимых значений, включающей символ неопределенности θ , целочисленным рангом $r_i \geq 1$, показывающим степень важности данной характеристики, и весом $w_i \geq 0$, представляющим собой численное выражение ее важности. Элементы множества H частично упорядочены по их рангу по убыванию: $r_i \geq r_{i+1} \forall i \in \overline{1, k-1}$. На множестве H построено конечное множество $C = \{c_1, c_2, \dots, c_m\}$ классов заданий. Каждый класс c_j определяется базовым (обязательным) и дополнительным (необязательным) наборами характеристик из H . Характеристики базового и дополнительного набора для классов удобно представить в виде булевых матриц A и B размерности $k \times m$, элементы которых $a_{ij} = 1$ или $b_{ij} = 1$ означают, что характеристика h_i принадлежит базовому или дополнительному набору, используется в классе c_j и для этого класса имеет конкретизированную область допустимых значений $D_{ij}^* \subseteq D_i \setminus \{\theta\}$. Если $a_{ij} \vee b_{ij} = 0$, то $D_{ij}^* \equiv \{\theta\}$. Матрицы A и B должны удовлетворять следующим условиям:

$$\begin{aligned} \bigvee_{j=1}^m \bigwedge_{i=1}^k \bar{a}_{ij} &= 0, \\ \bigvee_{i=1}^k \bigvee_{j=1}^m (a_{ij} \wedge b_{ij}) &= 0. \end{aligned}$$

Формирование множества H и создание на его основе множества C осуществляется администратором ГРВС. Каждому классу сопоставляются ресурсы среды, подходящие для выполнения заданий данного класса с точки зрения

предпочтений администратора. При поступлении нового задания в ГРВС специальная программа (классификатор заданий) автоматически проверяет, какие характеристики из H использованы в его спецификации, и извлекает для каждой из них указанную область значений. Задание представляется булевым вектором x размерности k . Между индексами элементов вектора x и индексами характеристик из H установлено взаимно однозначное соответствие. Значение i -го элемента вектора x находится следующим образом:

$$x_i = \begin{cases} 0, & \text{если } (D'_i \equiv \{\theta\}) \vee (D'_i \cap (D_i \setminus \{\theta\}) = \emptyset), \\ 1 & \text{в противном случае,} \end{cases}$$

где D'_i – это область значений характеристики h_i , требуемых в данном задании. Вектор x должен удовлетворять условию $\bigwedge_{i=1}^k \bar{x}_i = 0$.

Соответствие требуемых областей допустимых значений характеристик задания областям допустимых значений характеристик j -го класса устанавливается с помощью характеристической функции

$$\chi_j(x) = \begin{cases} 0, & \text{если } \exists i: (a_{ij} \vee b_{ij} = 1) \wedge (D'_i \cap \bar{D}_{ij}^* \neq \emptyset), \\ 1 & \text{в противном случае,} \end{cases}$$

где $i \in \overline{1, k}$, $j \in \overline{1, m}$. Для первичной классификации задания достаточно применения функции χ . После проведения первичной классификации может оставаться неопределенность относительно класса задания из-за того, что оно может быть соотнесено сразу с несколькими классами. С целью ее смягчения введены дополнительные характеристические функции, используемые для проведения вторичной классификации задания:

$$\begin{aligned} \chi_j^\rho(x, y) &= \begin{cases} 0, & \text{если } \max_{\forall l: y_l=1} \{\rho_l(x)\} - \rho_j(x) > \delta_\rho, \\ 1 & \text{в противном случае,} \end{cases} \\ \chi_j^\sigma(x, y) &= \begin{cases} 0, & \text{если } \max_{\forall l: y_l=1} \{\sigma_l(x)\} - \sigma_j(x) > \delta_\sigma, \\ 1 & \text{в противном случае,} \end{cases} \\ \chi_j^\omega(x, y) &= \begin{cases} 0, & \text{если } \max_{\forall l: y_l=1} \{\omega_l(x)\} - \omega_j(x) > \delta_\omega, \\ 1 & \text{в противном случае,} \end{cases} \\ \chi_j^\varphi(x, y, z) &= \begin{cases} 0, & \text{если } \max_{\forall l: y_l=1} \{\varphi_l(x, z)\} - \varphi_j(x, z) > \delta_\varphi, \\ 1 & \text{в противном случае,} \end{cases} \end{aligned}$$

где функции $\rho_j(x)$, $\sigma_j(x)$, $\omega_j(x)$ и $\varphi_j(x, z)$ соответственно определяют вероятность отнесения задания к классу c_j на основе числа характеристик, удовлетворяющих по всем параметрам данному классу, агрегированный числовой показатель важности характеристик для класса c_j с учетом рангов этих характеристик, сумму весов характеристик для класса c_j и вероятность отнесения задания к классу c_j с учетом его вычислительной истории z , а переменные δ_ρ , δ_σ , δ_ω и δ_φ представляют верхние границы эквивалентности значений функций ρ , σ , ω и φ , $l \in \overline{1, m}$.

Алгоритм классификации задания. На вход данному алгоритму поступает вектор x . Выходом является вектор y , содержащий информацию о классах, к которым относится задание. Алгоритм включает пять основных этапов работы:

- 1) инициализация элементов вектора y : $y_j = 0 \forall j \in \overline{1, m}$;
- 2) выполнение первичной классификации – проверки принадлежности задания каждому из m классов: $y_j = \chi_j(x) \forall j \in \overline{1, m}$;
- 3) если $y_j = 0 \forall j \in \overline{1, m}$, то завершение работы алгоритма (задание не может быть классифицировано);
- 4) иначе – проведение вторичной классификации:
 - а. $y_j = \chi_j^p(x, y) \forall j: y_j = 1$;
 - б. $y_j = \chi_j^s(x, y) \forall j: y_j = 1$;
 - в. $y_j = \chi_j^\omega(x, y) \forall j: y_j = 1$;
 - г. $y_j = \chi_j^\varphi(x, y, z) \forall j: y_j = 1$;
- 5) завершение работы алгоритма (задание классифицировано).

Иллюстративный пример демонстрирует смягчение неопределенности в процессе распределения заданий по ресурсам. Пусть по результатам первичной классификации найдены 8 классов, характеристикам которых удовлетворяет задание, поступившее в среду для выполнения указанного в задании модуля. Эти классы различаются между собой областью допустимых значений одной характеристики, определяющей допустимое время выполнения заданий. Значение данной характеристики не указано в спецификации задания. Среда включает 3 разнородных ресурса (3 кластера, узлы которых различаются по вычислительным характеристикам). Определим энтропию $E^c(x_c) = \sum_{i=1}^m p_i(x_c) E_i^c(x_c)$, отражающую степень неопределенности принадлежности задания одному из классов, где x_c – случайное событие, которое обуславливает принадлежность задания одному из m классов, $p_i(x_c)$ – вероятность такого события относительно i -го класса, $\sum_{i=1}^m p_i(x_c) = 1$, $E_i^c(x_c)$ – энтропия, соответствующая данному событию. Значение $E_i^c(x_c)$ вычисляется следующим образом:

$$E_i^c(x_c) = \begin{cases} 0, & \text{если } p_i(x_c) = 0, \\ -\log_2 p_i(x_c) & \text{в противном случае.} \end{cases}$$

Тем же способом определим энтропию $E^r(x_r) = \sum_{j=1}^n p_j(x_r) E_j^r(x_r)$, отражающую степень неопределенности принадлежности ресурса одному из классов, где x_r – случайное событие, которое обуславливает возможность назначения заданию одного из n ресурсов, $p_j(x_r)$ – вероятность такого события относительно j -го ресурса, $E_j^r(x_r)$ – энтропия, соответствующая данному событию. Значения $p_j(x_r)$ и $E_j^r(x_r)$ вычисляются следующим образом:

$$p_j(x_r) = \sum_{\forall l: y_{jl}=1} p_l(x_c) / \sum_{j=1}^n \sum_{\forall l: y_{jl}=1} p_l(x_c), \quad l \in \overline{1, m}, \quad \sum_{j=1}^n p_j(x_r) = 1,$$

$$E_j^r(x_r) = \begin{cases} 0, & \text{если } p_j(x_r) = 0, \\ -\log_2 p_j(x_r) & \text{в противном случае.} \end{cases}$$

Таблица 1 – Показатели энтропии и вероятности принадлежности к классам

№	$p_1(x_c)$	$p_2(x_c)$	$p_3(x_c)$	$p_4(x_c)$	$p_5(x_c)$	$p_6(x_c)$	$p_7(x_c)$	$p_8(x_c)$	E^c	$p_1(x_r)$	$p_2(x_r)$	$p_3(x_r)$	E^r
1	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	3.00	0.333	0.333	0.333	1.06
2	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	3.00	0.444	0.556	0	0.99
3	0.115	0.741	0.132	0.012	0	0	0	0	1.14	0.988	0.012	0	0.09
4	0	0.936	0.064	0	0	0	0	0	0.34	1.000	0	0	0

На первом этапе вероятности принадлежности задания одному из 8 классов равны между собой, так же как равны и вероятности назначения заданию одного из 3 ресурсов. В этом случае $E^c = 3.00$ и $E^r = 1.06$. На следующем этапе выбираются два ресурса, наиболее подходящие для выполнения задания с точки зрения экспертного опыта администраторов ресурсов. При этом степень неопределенности назначения ресурсов заданию понижается до $E^r = 0.99$. Вторичная классификация задания на основе анализа статистической информации о вычислительной истории конкретизирует вероятности принадлежности задания классам, полученные на третьем этапе. Значение E^c становится равным 1.14. В соответствие с этим изменяются вероятности назначения ресурсов заданию. Степень неопределенности назначения ресурсов заданию уменьшается до $E^r = 0.09$. На последнем этапе результаты вторичной классификации уточняются на основе прогнозного времени выполнения модуля или полученных ранее данных его тестирования. В результате этого $E^c = 0.34$ и $E^r = 0.09$. Таким образом, ресурс 1 назначается для выполнения задания.

В четвертой главе проведен сравнительный анализ известных мультиагентных методов и средств, используемых для управления распределенными вычислениями. Сформулирована постановка, а также разработаны модели и алгоритм мультиагентного управления в ГРВС. Проведены экспериментальные исследования, показавшие преимущества предложенного подхода к мультиагентному управлению по сравнению с традиционными средствами управления распределенными вычислениями в среде с разнородными ресурсами.

Постановка задачи. Пусть на модели M построена схема решения задачи, представленная матрицей S , описывающей схему решения задачи в виде ярусно-параллельной формы, и матрицей P , отражающей сведения о предшествовании операций схемы решения задачи. Для операций f_1, f_2, \dots, f_{n_o} , входящих в схему решения задачи, найдены оценки времени их выполнения на эталонном узле ГРВС. Значения параметров передаются между операциями в виде файлов данных, для которых получены оценки их размера.

Агентам необходимо распределить свои ресурсы для модулей, реализующих операции схемы решения задачи.

Процесс вычислений должен удовлетворять заданным критериям c_1, c_2, \dots, c_{n_c} качества решения задачи. Для каждого критерия приведены его предельные значения $c_l^{min} \geq 0$ и $c_l^{max} \geq 0$ такие, что $c_l^{min} \leq c_l \leq c_l^{max}$, а также указано условие его оптимальности:

$$c_l \rightarrow c_l^{min} (c_l^{max}), \quad (1)$$

где $l = \overline{1, n_c}$. Предельные значения и условия оптимальности критериев пользователя определяются им при формулировке постановки задачи. Предельные значения и условия оптимальности предпочтений владельцев ресурсов задаются администратором, представляющим данные ресурсы.

Задание по решению задачи включает набор взаимосвязанных подзаданий разных классов.

Модели. Предложены три модели тендера вычислительных работ, поддерживающие следующие три уровня обслуживания на основе мультиагентного управления:

- 1) модель, ориентированная на пользовательские критерии качества решения задачи,
- 2) модель, направленная на удовлетворение предпочтений владельцев ресурсов,
- 3) модель, поддерживающая согласование пользовательских критериев качества решения задачи с предпочтениями владельцев ресурсов.

Для проведения и оценки результатов тендера используется виртуальная стоимость ресурсов. В качестве такого показателя (или наряду с ним) могут выступать другие показатели – пользовательские критерии качества решения задачи, предпочтения владельцев ресурсов и другие характеристики, включая стоимость процессорного времени, передачи и хранения данных.

Каждый агент (участник торгов) делает предложения по выполнению отдельных подзаданий, класс которых соответствует ресурсам данного агента. Предложение выражается в виде ставки, элементы которой отражают критерии c_1, c_2, \dots, c_{n_c} применительно к подзаданию.

Дополнительно любой агент может сделать предложения по выполнению наборов подзаданий, классы которых соответствуют его ресурсам. Такое предложение осуществляется в виде ставки, элементы которой отражают критерии c_1, c_2, \dots, c_{n_c} применительно к каждому подзаданию, входящему в данный набор.

Введем следующие обозначения:

- n_{job} – число подзаданий;
- $B_i^j = (b_{i1}^j, b_{i2}^j, \dots, b_{in_c}^j)$ – i -я ставка на выполнение j -го подзадания в отдельности, сделанная ξ_i -м агентом, $j \in \overline{1, n_{job}}$;
- $b_{il}^j = \varphi(\xi_i, j, k, l, \mathbf{S}, \mathbf{P}, S_{\xi_i})$ – элемент ставки B_i^j , относящийся к l -му критерию выполнения j -го подзадания, $l \in \overline{1, n_c}$;
- S_{ξ_i} – расписание обслуживания заданий ξ_i -м агентом;
- I_j – множество индексов агентов, подавших предложения по j -му подзаданию;
- $\tilde{B}_t = (\tilde{b}_{i1}^{j_1}, \tilde{b}_{i2}^{j_1}, \dots, \tilde{b}_{in_c}^{j_1}, \tilde{b}_{i1}^{j_2}, \tilde{b}_{i2}^{j_2}, \dots, \tilde{b}_{in_c}^{j_2}, \dots, \tilde{b}_{i1}^{j_{n_{\eta_t r_t}}}, \tilde{b}_{i2}^{j_{n_{\eta_t r_t}}}, \dots, \tilde{b}_{in_c}^{j_{n_{\eta_t r_t}}})$ – t -я ставка на совместное выполнение подзаданий $j_1, j_2, \dots, j_{n_{\eta_t r_t}}$, сделанная η_t -м агентом, который включил их в свой r_t -й набор ($0 \leq r_t \leq n_{max_set}$, где n_{max_set} – максимально допустимое число наборов подзаданий);

- $n_{\eta_l r_l}$ – число подзаданий в r_l -м наборе η_l -го агента;
- $\tilde{b}_{il}^j = \tilde{\varphi}(\xi_i, j, k, l, b_{il}^j, J_{\eta_l r_l}, \mathbf{S}, \mathbf{P}, S_{\eta_l})$ – элемент ставки \tilde{B}_l , относящийся к l -му критерию выполнения j -го подзадания, такой, что

$$\left((\forall \xi_i = \eta_l \tilde{b}_{il}^j \leq b_{il}^j) \wedge (\exists k \in \overline{1, n_c}: \tilde{b}_{ik}^j < b_{ik}^j), c_l \rightarrow c_l^{\min} \right) \vee$$

$$\vee \left((\forall \xi_i = \eta_l b_{il}^j \leq \tilde{b}_{il}^j) \wedge (\exists k \in \overline{1, n_c}: b_{ik}^j < \tilde{b}_{ik}^j), c_l \rightarrow c_l^{\max} \right);$$
- $J_{\eta_l r_l}$ – множество индексов подзаданий, входящих в r_l -й набор η_l -го агента;
- S_{η_l} – расписание обслуживания заданий ξ_i -м агентом;
- I_s – множество агентов, сделавших предложения по выполнению наборов подзаданий;
- $n_a \geq 1$ – число агентов виртуального сообщества, выполняющих общее задание;
- $n_c \geq 1$ – число критериев выполнения общего задания;
- $\xi_1, \xi_2, \dots, \xi_{n_j}, \eta_1, \eta_2, \dots, \eta_{m_j} \in \overline{1, n_a}$ – индексы агентов;
- $1 \leq n_j \leq n_a$ ($1 \leq m_j \leq n_a$) – число агентов виртуального сообщества, сделавших ставки по выполнению j -го подзадания в отдельности (в составе набора заданий).

Реализация функций φ и $\tilde{\varphi}$ на модели M для вычисления ставок b_{il}^j и \tilde{b}_{il}^j базируется на подходах к нахождению оценок критериев, рассмотренных во второй главе диссертации.

Тендер проводится в два этапа. На первом этапе осуществляется предварительное определение победителя по каждому подзаданию в отдельности. При наличии ставок на наборы подзаданий по возможности производится улучшение результатов, полученных на первом этапе, с учетом данных ставок.

Этап I. Предварительное выявление победителя по каждому подзаданию схемы решения задачи в отдельности.

Модель 1. В данной модели критерий c_1 – это стоимость выполнения задания. Дополнительные критерии c_2, c_3, \dots, c_{n_c} упорядочены следующим образом:

- c_2, c_3, \dots, c_i – предпочтения пользователей;
- $c_{i+1}, c_{i+2}, \dots, c_{n_c}$ – предпочтения владельцев ресурсов.

Множество B претендентов за право выполнения j -го задания в отдельности определяется с помощью лексикографического правила:

$$B = \{B_i^j: \forall e \in \overline{1, n_j} \exists k \in \overline{1, n_c - 1}: \\ (\hat{b}_{i1}^j = \hat{b}_{e1}^j) \wedge (\hat{b}_{i2}^j = \hat{b}_{e2}^j) \wedge \dots \wedge (\hat{b}_{ik}^j = \hat{b}_{ek}^j) \wedge (\hat{b}_{i(k+1)}^j < \hat{b}_{e(k+1)}^j)\},$$

$$\hat{b}_{il}^j = \begin{cases} \frac{b_{il}^j - b_l^{\min}}{b_l^{\max} - b_l^{\min}} & \text{при } c_l \rightarrow c_l^{\min}, \\ \frac{c_l^{\max} - b_{il}^j}{b_l^{\max} - b_l^{\min}} & \text{при } c_l \rightarrow c_l^{\max}, \end{cases}$$

$$b_l^{\min} = \min \{b_{1l}^j, b_{2l}^j, \dots, b_{n_j l}^j, \tilde{b}_{1l}^j, \tilde{b}_{2l}^j, \dots, \tilde{b}_{m_j l}^j\},$$

$$b_l^{max} = \max \{b_{1l}^j, b_{2l}^j, \dots, b_{n_j l}^j, \tilde{b}_{1l}^j, \tilde{b}_{2l}^j, \dots, \tilde{b}_{m_j l}^j\},$$

где \hat{b}_{il}^j – оценка удаленности значения элемента ставки b_{il}^j от идеальных показателей, нормированная в диапазоне $[0,1]$, $i \in \overline{1, n_j}$, $i \neq e$.

Индекс агента-победителя за право выполнения j -го задания в отдельности находится следующим образом:

$$k_j = \underset{\forall i \in \overline{1, n_j}: B_i^j \in B}{\operatorname{argmin}} \{t_{\xi_i}\}, \quad (1)$$

где t_{ξ_i} – это время поступления ставки ξ_i -го агента.

Стоимость π_j j -го подзадания вычисляется по формуле

$$\pi_j = b_{q1}^j, \quad (2)$$

$$q = \underset{\forall i \in \overline{1, n_j}, \xi_i \neq k_j}{\operatorname{argmin}} \{b_{i1}^j\}. \quad (3)$$

Модель 2. В данной модели критерий c_1 – это стоимость задания. Дополнительные критерии c_2, c_3, \dots, c_{n_c} упорядочиваются следующим образом:

- c_2, c_3, \dots, c_i – предпочтения владельцев ресурсов;
- $c_{i+1}, c_{i+2}, \dots, c_{n_c}$ – предпочтения пользователей.

Построение множества претендентов за право выполнения j -го задания в отдельности, получение индекса k_j агента-победителя и вычисление π_j осуществляется аналогично модели 1.

Модель 3. В данной модели критерий c_1 – это стоимость задания. Дополнительные критерии c_2, c_3, \dots, c_{n_c} не упорядочиваются.

Определение. Будем говорить, что ставка $B_{i_1}^j$ предпочтительнее ставки $B_{i_2}^j$ ($B_{i_1}^j > B_{i_2}^j$), если выполняется условие

$$\left(\forall l \in \overline{1, n_c} \hat{b}_{i_1 l}^j \leq \hat{b}_{i_2 l}^j \right) \wedge \left(\exists k \in \overline{1, n_c}: \hat{b}_{i_1 k}^j < \hat{b}_{i_2 k}^j \right).$$

В рамках третьей модели множество B претендентов на право выполнения j -го задания в отдельности строится следующим образом:

$$B = \{B_i^j \in B': \nexists k \in I_j: B_k^j > B_i^j\},$$

$$B' = \{B_i^j: \nexists k \in I_j: \hat{b}_{k1}^j < \hat{b}_{i1}^j\}, i \in I_j.$$

Индекс агента-победителя за право выполнения j -го задания и его стоимость находятся соответственно по формулам (1)-(3).

В рамках каждой из трех моделей определяется множество ставок агентов, победивших на первом этапе, и их оценок:

$$\varpi_1^j = \pi_j, \hat{\varpi}_1^j = \hat{b}_{q1}^j, \varpi_l^j = b_{il}^j, \hat{\varpi}_l^j = \hat{b}_{il}^j, \forall i: \xi_i = k_j, j = \overline{1, n_{job}}, l = \overline{2, n_c}.$$

Результаты торгов, полученные на первом этапе, представим вектором

$$\varpi = (\varpi_1^1, \varpi_2^1, \dots, \varpi_{n_c}^1, \varpi_1^2, \varpi_2^2, \dots, \varpi_{n_c}^2, \varpi_1^{n_{job}}, \varpi_2^{n_{job}}, \dots, \varpi_{n_c}^{n_{job}}).$$

Критерии выполнения задания на основе ставок агентов вычисляются следующим образом: $c_l = \phi(l, \mathbf{S}, \mathbf{P}, \varpi_l^1, \varpi_l^2, \dots, \varpi_l^{n_{job}})$, $l = \overline{1, n_c}$.

Реализация функции ϕ на модели \mathbb{M} для получения значений c_1, c_2, \dots, c_{n_c} базируется на подходах к оценке критериев, рассмотренных в разделе 2.3 диссертации. Предполагается, что функция ϕ является строго возрастающими по каждой из своих переменных $\varpi_l^1, \varpi_l^2, \dots, \varpi_l^{n_{job}}$:

$$\begin{aligned} \varpi_l^{j^*} > \varpi_l^j &\Rightarrow \phi(l, \varpi_l^1, \varpi_l^2, \dots, \varpi_l^{j-1}, \varpi_l^{j^*}, \varpi_l^{j+1}, \dots, \varpi_l^{n_{job}}) > \\ &> \phi(l, \varpi_l^1, \varpi_l^2, \dots, \varpi_l^{j-1}, \varpi_l^j, \varpi_l^{j+1}, \dots, \varpi_l^{n_{job}}) \quad \forall l \in \overline{1, n_c}. \end{aligned}$$

Если нет предложений агентов по наборам подзаданий, то проведение тендера завершается. В противном случае начинается второй этап тендера.

Этап II. Улучшение результатов, полученных на первом этапе.

Включим в множество \tilde{B}' ставок, претендующих на улучшение результатов I этапа проведения тендера, ставки $\tilde{B}_l \forall j_k \in I_{\eta_l r_l}$, которые удовлетворяют условиям

$$\begin{aligned} \hat{b}_{l1}^{j_k} &\leq \hat{\omega}_1^{j_k} \quad \forall j_k \in I_{\eta_l r_l}, \quad k = \overline{1, n_{\eta_l r_l}}, \\ (\exists j_k \in I_{\eta_l r_l}: \hat{\omega}_1^{j_k} < \hat{b}_{l1}^{j_k}) &\vee (d_{\eta_l r_l}^{\varpi} > d_{\eta_l r_l}), \\ d_{\eta_l r_l} &= \sum_{k=1}^{n_{\eta_l r_l}} \sum_{l=2}^{n_c} \hat{b}_{ul}^{j_k}, \quad d_{\eta_l r_l}^{\varpi} = \sum_{k=1}^{n_{\eta_l r_l}} \sum_{l=2}^{n_c} \hat{\omega}_l^{j_k}, \\ \hat{b}_{ul}^{j_k} &= \begin{cases} \frac{\tilde{b}_{ul}^{j_k} - b_l^{\min}}{b_l^{\max} - b_l^{\min}} & \text{при } c_l \rightarrow c_l^{\min}, \\ \frac{b_l^{\max} - \tilde{b}_{ul}^{j_k}}{b_l^{\max} - b_l^{\min}} & \text{при } c_l \rightarrow c_l^{\max}, \end{cases} \end{aligned}$$

где $\hat{b}_{ul}^{j_k}$ – оценка удаленности значения элемента ставки $\tilde{b}_{ul}^{j_k}$ от идеальных показателей, нормированная в диапазоне $[0, 1]$.

Множество \tilde{B}' частично упорядочено. Для его упорядочения используются значения показателей k_{l1} , k_{l2} и k_{l3} :

$$k_{l1} = \sum_{k=1}^{n_{\eta_l r_l}} (\hat{b}_{l1}^{j_k} - \hat{\omega}_1^{j_k}), \quad k_{l2} = d_{\eta_l r_l} - d_{\eta_l r_l}^{\varpi}, \quad k_{l3} = t_l,$$

где t_l – время поступления l -й ставки. Данные показатели перечислены в порядке уменьшения их значимости для упорядочения множества \tilde{B}' .

Пусть булев вектор $\tilde{\omega}$ размерности n_{job} отражает факт использования ставок на подзадания в составе набора для улучшения результатов торгов, полученных на первом этапе. Элемент вектора $\tilde{\omega}(j) = 1$ ($\tilde{\omega}(j) = 0$) показывает, что ставка по подзаданию j улучшена (не улучшена). При инициализации вектора $\tilde{\omega}$ задаются нулевые значения его элементов $\tilde{\omega}(j) = 0$, $j = \overline{1, n_{job}}$.

Первая ставка \tilde{B}'_1 из упорядоченного множества \tilde{B}' ставок безусловно используется для улучшения результатов торгов: $\varpi_l^{j_k} = \tilde{b}_{l1}^{j_k}$, $k = \overline{1, n_{\eta_1 r_1}}$, $l = \overline{1, n_c}$.

После этого корректируются значения вектора $\tilde{\omega}$: $\tilde{\omega}(j_k) = 1$, $k = \overline{1, n_{\eta_1 r_1}}$.

Циклическая процедура применения остальных ставок $\tilde{B}'_l \in \tilde{B}'$, $l = \overline{2, |\tilde{B}'|}$ приведена ниже:

$$\forall l: (\forall j_k \in J_{\eta_l r_l} \tilde{\omega}(j_k) = 0) \tilde{\omega}_l^{j_k} = \tilde{b}_{ul}^{j_k}, \quad k = \overline{1, n_{\eta_l r_l}}, \quad l = \overline{1, n_c}, \\ \tilde{\omega}(j_k) = 1, \quad k = \overline{1, n_{\eta_l r_l}}.$$

Алгоритм. Предложен алгоритм мультиагентного распределения ресурсов, базирующийся на рассмотренном выше тендере. Он включает шесть основных этапов его работы.

I. *Планирование вычислений.* На основе процедурной или не процедурной постановки задачи, сформулированной пользователем пакета на модели M, агентом планирования вычислений строится схема s решения задачи.

II. *Формирование виртуального сообщества агентов.* Объединение агентов в виртуальное сообщество осуществляется на основе знаний о соответствии классов заданий и ресурсов ГРВС. Эти знания заносятся в модель ГРВС администраторами среды. В виртуальное сообщество включаются агенты, представляющие ресурсы, на которых могут быть запущены модули, реализующие операции схемы s .

III. *Выбор координатора.* Избрание агента-координатора участниками виртуального сообщества производится в процессе их локальных взаимодействий на основе модифицированного древесного алгоритма с учетом коммуникационной топологии сети агентов.

IV. *Проведение тендера вычислительных работ.* Формирование агентами предложений по выполнению модулей схемы s и выявление победителей торгов в рамках тендера реализуется на одной из трех рассмотренных моделей тендера. Задание модели осуществляет администратор ГРВС.

V. *Подведение итогов торгов.* Определение агентов (победителей торгов), участвующих в выполнении модулей схемы s , производится на основе правил многокритериального выбора с учетом заданных критериев, упорядоченных в соответствии с выбранной моделью.

VI. *Процесс вычислений.* Выполнение модулей схемы s в назначенных агентами ресурсах осуществляется в асинхронном режиме по готовности данных.

Экспериментальный анализ. Оценка функционирования мультиагентной системы (МАС) при обработке потоков заданий для трех предложенных моделей тендера проведена в экспериментальной ГРВС, включающей два пула ресурсов: 8 ВМ со следующими характеристиками: 1 процессор Intel Xeon E5506 (1 core, 2.13 GHz, 4 MB L3 cache, 2 GB RAM DDR3-800, 4 FLOP/cycle); 10 ВМ со следующими характеристиками: 1 процессор AMD Opteron 6276 (8 core, 2.3 GHz, 16 MB L3 cache, 32 GB RAM DDR3-1600, 4 FLOP/cycle). В роли схем решения задач выступают типовые рабочие процессы Montage, CyberShake, Epigenomics, LIGO и SIPHT, объединенные в рамках одного РППП. Предпочтение данных рабочих процессов обусловлено тем, что они наиболее близко отражают характерные особенности схем решения практических задач, рассматриваемых ниже в главах 5 и 6. Каждый поток включал 30 рабочих процессов одного типа.

Для сравнительного анализа качества обработки потоков схем решения задач отобраны следующие критерии: стоимость и время выполнения потока, ускорение вычислений, эффективность использования ресурсов, средняя загрузка процессора и среднеквадратическое отклонение от нее, характеризующее степень балансировки загрузки ресурсов. Первые три критерия представляют требования пользователя. Остальные критерии отражают предпочтения владельцев ресурсов.

Результаты, приведенные на рисунке 3, показывают, что MAC с моделью 1 показывает преимущество по всем критериям пользователя для каждого потока заданий в сравнении с Condor DAGMan и GridWay.

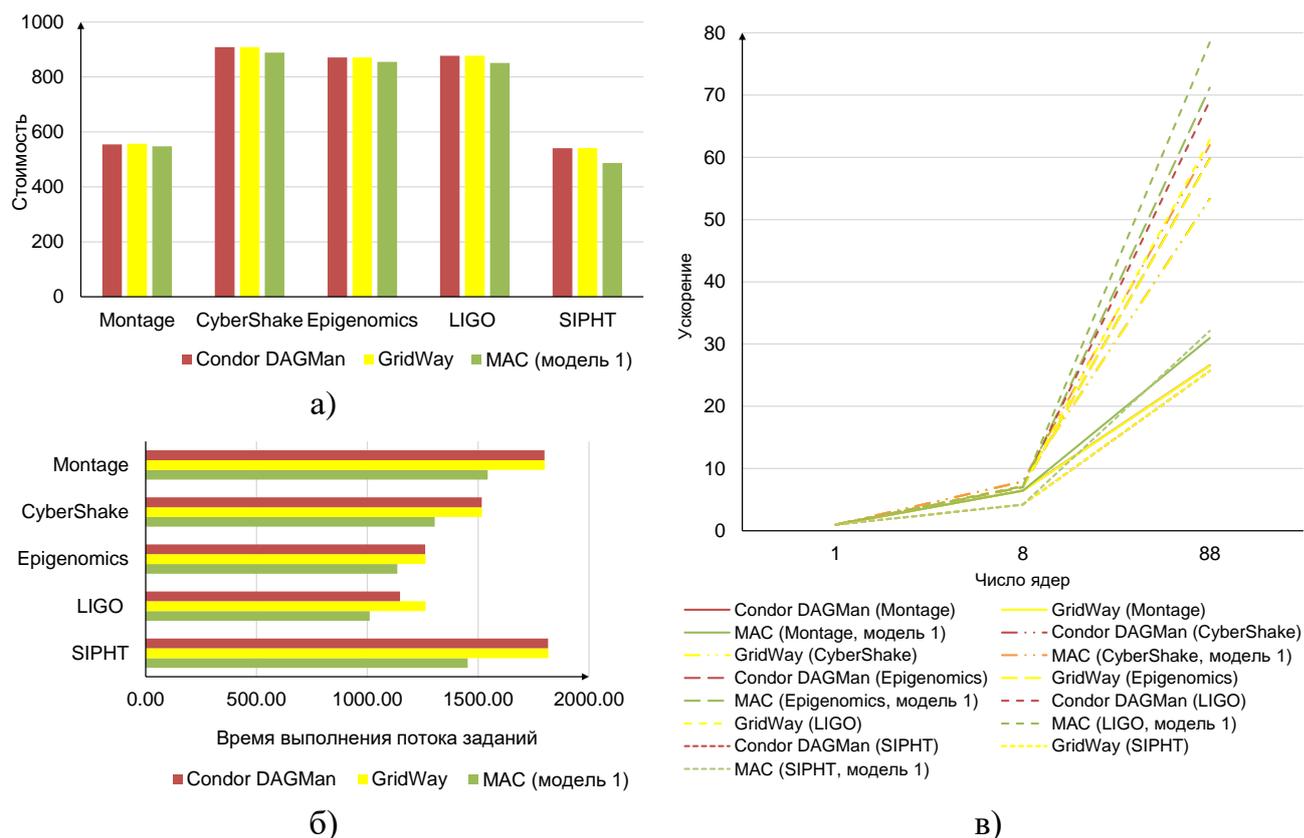


Рисунок 3 – Показатели качества решения задачи: стоимость (а) и время (б) выполнения потока заданий; ускорение вычислений (в)

Результаты, приведенные на рисунке 4, означают, что MAC с моделью 2 показывает преимущество по всем показателям, отражающим предпочтения владельцев ресурсов, для каждого потока в сравнении с Condor DAGMan и GridWay. Для модели 3 с неупорядоченными критериями качества выполнения потоков заданий MAC показывает улучшение по большему числу критериев по сравнению с Condor DAGMan и GridWay.

Таким образом, преимущества предложенного мультиагентного управления по сравнению с традиционными средствами управления распределенными вычислениями в среде с разнородными ресурсами очевидны. Каждая из трех моделей тендера позволяет достичь требуемого уровня обслуживания в зависимости от способа упорядочения критериев.

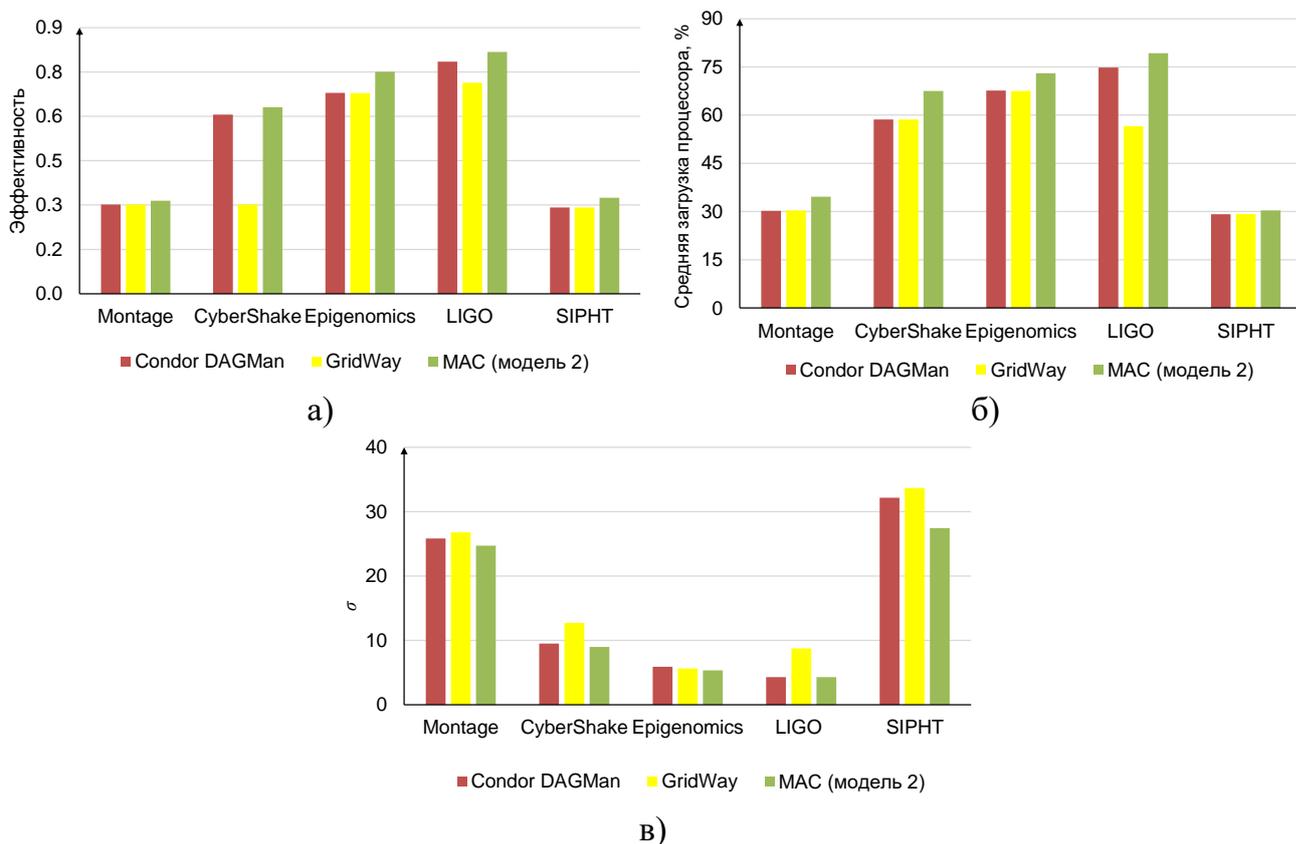


Рисунок 4 – Показатели качества использования ресурсов: эффективность (а), средняя загрузка процессора (б), среднеквадратическое отклонение σ от средней загрузки процессора (в)

В пятой главе разработан пакетный подход к организации высокопроизводительных вычислений, базирующийся на применении специализированных инструментальных комплексов DISCENT, DISCOMP и Orlando Tools в процессе создания и эксплуатации РППП, ориентированных на решение крупномасштабных задач в ГРВС. Предложена технология создания таких предметно-ориентированных вычислений. Среда функционирования создаваемых пакетов может включать НРС-кластеры, грид-системы, облачные инфраструктуры и другие программно-аппаратные компоненты, а также обеспечивать их совместное использование.

На рисунке 5 приведена схема интеграции вычислительных систем в рамках проблемно-ориентированной среды для решения трудных переборных задач на основе многовариантных расчетов. Интеграция осуществляется с помощью рассмотренных выше инструментальных комплексов организации вычислений в Грид и разработки РППП.

Системная часть является общей для всех РППП, разработанных в рамках конкретной ГРВС. Она представлена сервером ГРВС, поддерживающим веб-интерфейс для доступа пользователей к ее компонентам. Основу системной части составляют конструктор концептуальной модели, планировщики вычислений и интерпретаторы схем решения задач. Библиотеки модулей РППП включают прикладное ПО для решения задач предметной области.

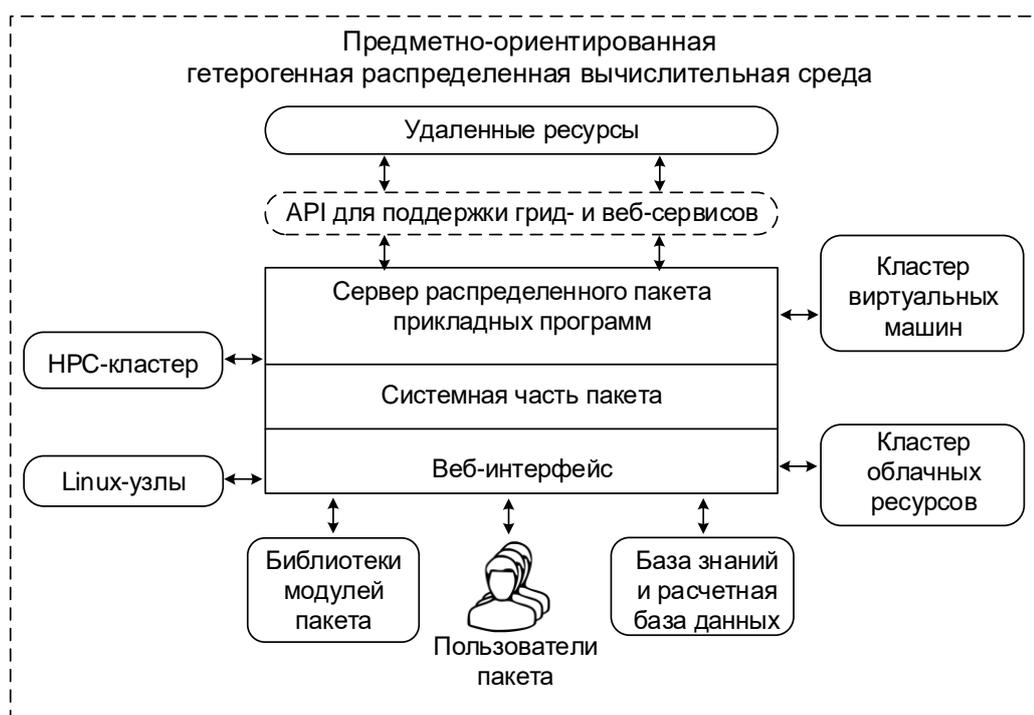


Рисунок 5 – Схема интеграции вычислительных систем

База знаний содержит следующие компоненты знаний концептуальной модели: информацию о прикладных программных модулях для решения задач предметной области, а также препроцессорной и постпроцессорной обработки данных; структурные свойства алгоритмов решения задач в терминах параметров и операций предметной области; условия выбора оптимальных алгоритмов решения задачи в зависимости от конфигурации ресурсов среды и их состояния; сведения о программно-аппаратной инфраструктуре (характеристики узлов, каналов связи, сетевых устройств и топологии сети, а также данные о сбоях ПО и аппаратных средств); данные о системах управления распределенными вычислениями, используемых в среде; параметры административных политик, применяемых в отношении ресурсов и пользователей. API для поддержки грид- и веб-сервисов обеспечивает подключение разнообразных удаленных информационно-вычислительных ресурсов.

Расчетная база данных хранит информацию, необходимую для проведения экспериментов, и результаты решения задач.

В зависимости от функциональных и системных требований, предъявляемых к РППП их разработчиками, данная технология может быть реализована как с помощью Orlando Tools, так и DISCOMP. Доступ к удаленным ресурсам Грид организуется с помощью сервис-ориентированных средств, являющихся развитием системного ПО инструментального комплекса DISCENT.

В шестой главе продемонстрировано применение результатов диссертационного исследования при разработке двух РППП для выявления критических элементов отраслевых систем энергетики и решения задач складской логистики.

Пакет для выявления критических элементов газотранспортной сети России. Критическими элементами сети являются те ее объекты, нарушение

штатного функционирования которых приводит к возникновению существенной недопоставки газа. Сложность выявления критических элементов значительно возрастает при необходимости учета одновременных отказов нескольких элементов, так как их синергетические последствия могут изменять степень критичности отказавших элементов. Рассматриваемый в задаче граф системы газоснабжения России состоит из 378 узлов. В их числе 28 источников природного газа, 64 потребителя (регионы Российской Федерации), 24 подземных хранилища газа, 266 магистральных компрессорных станций. Данный граф также содержит 486 дуг, представляющих сегменты магистральных трубопроводов. 415 дуг отобраны, как элементы с возможными отказами. Число n_k комбинаций элементов с отказами может быть оценено формулой

$$n_k = \frac{n_e!}{(n_e - n_f)! n_f!},$$

где n_f – число одновременно отказавших элементов, n_e – число дуг, выбранных в качестве элементов с возможными отказами. Число n_f варьировалось от 2 до 3.

ГРВС, используемая в процессе решения задачи, включает выделенные кластерные ресурсы ЦКП ИСКЦ (приватное облако под управлением платформы OpenStack) и невыделенные ресурсы под управлением PBS Torque. ВМ управляются гипервизором KVM. Эксперименты проводились в трех пулах ресурсов: 1) 10 узлов (невыделенные ресурсы) со следующими характеристиками – 2 процессора Intel Xeon 5345 EM64T (4 ядра, 2.3 ГГц, 8 Гбайт RAM); 2) 10 узлов (выделенные ресурсы) со следующими характеристиками – 2 процессора AMD Opteron 6276 (16 ядер, 2.3 ГГц, 64 Гбайт RAM); 3) 10 узлов (выделенные и невыделенные ресурсы) со следующими характеристиками – 2 процессора Intel Xeon CPU X5670 (18 ядер, 2.1 ГГц, 128 Гбайт RAM). Узлы разных пулов имеют разные типы интерконнекта (1 GigE, QDR Infiniband) и жестких дисков (HDD, SSD). Все выделенные ресурсы виртуализированы.

Рисунки 6 а и 6 б показывают число использованных ядер и слотов в процессе решения задач при $n_f = 3$, $k = 400$ и $k = 760$. Распределение потока заданий между пулами с учетом производительности их узлов обеспечило сокращение времени решения задач более чем на 15%.

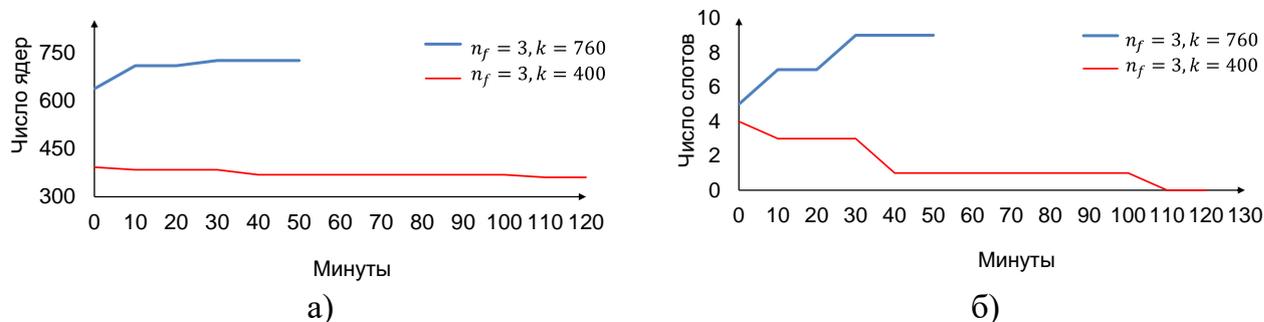


Рисунок 6 – Число использованных ядер (а) и слотов (б)

Рисунок 7 демонстрирует высокую среднюю загрузку процессора. Она немного меньше в пулах 1 и 3 из-за накладных расходов на виртуализацию. В то

же время накладные расходы на виртуализацию в узлах всех пулов не превышают 5%. Рисунок 8 показывает улучшение (сокращение) времени решения задачи при $n_f = 3$ как для $k = 400$, так и для $k = 760$.

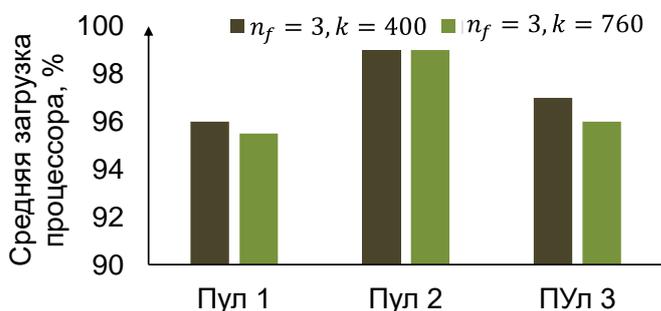


Рисунок 7 – Средняя загрузка процессора

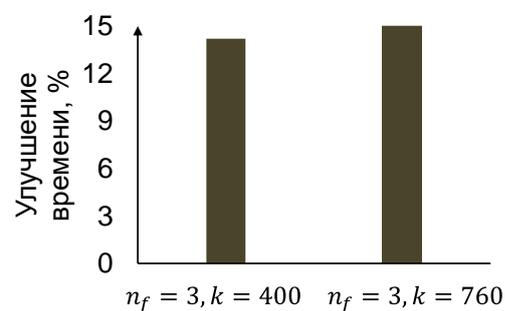


Рисунок 8 – Улучшение времени решения задачи

Решение рассмотренной задачи позволило экспертам из ИСЭМ СО РАН составить список критически важных объектов газотранспортной сети России и сформулировать некоторые мероприятия по повышению живучести системы газоснабжения России. В частности, предложены рекомендации по технологическому доустройству отдельных сегментов магистральных трубопроводов с целью обеспечения режима их работы, позволяющего повысить длительность кратковременного увеличения их пропускной способности в случае необходимости компенсации последствий крупных возмущений на других участках сети.

Пакет для решения задач складской логистики. Разработан РППП для решения задач складской логистики на промышленном рефрижераторном складе предприятия ООО «Иркутский хладокомбинат», который является вторым по величине единовременного хранения товаров в России на территории от Урала до Дальнего Востока. Пакет создан на основе разработанной автором диссертации технологии предметно-ориентированных распределенных вычислений и мультиагентного управления в гетерогенной среде. Все поставленные задачи складской логистики успешно решены.

Рисунок 9 демонстрирует основные результаты решения трех типовых задач (совершенствования процессов разгрузки и отгрузки товаров, реструктуризации процесса обслуживания клиентов и реорганизации складских объектов).

В процессе эксперимента в ГРВС использовались четыре компонента: ПК (Intel Core i3-4160, 2 ядра с гипертрейдингом, 3.6 GHz, 4 GB RAM) предприятия ООО «Иркутский хладокомбинат»; ПК (Intel Core i7-4770, 4 ядра с гипертрейдингом, 3.4 GHz, 4 GB RAM) предприятия ООО «Иркутский хладокомбинат»; кластер ПК (16 узлов с одним процессором Intel Core i3-4000M, 2 ядра с гипертрейдингом, 2.4 GHz, 2 GB RAM) вычислительной среды МИЭЛ ИГУ; сегмент НРС-кластера «Академик В.М. Матросов» ЦКП ИСКЦ (15 узлов с двумя процессорами AMD Opteron 6276, 16 ядер, 2.3 GHz, 64 GB of RAM).

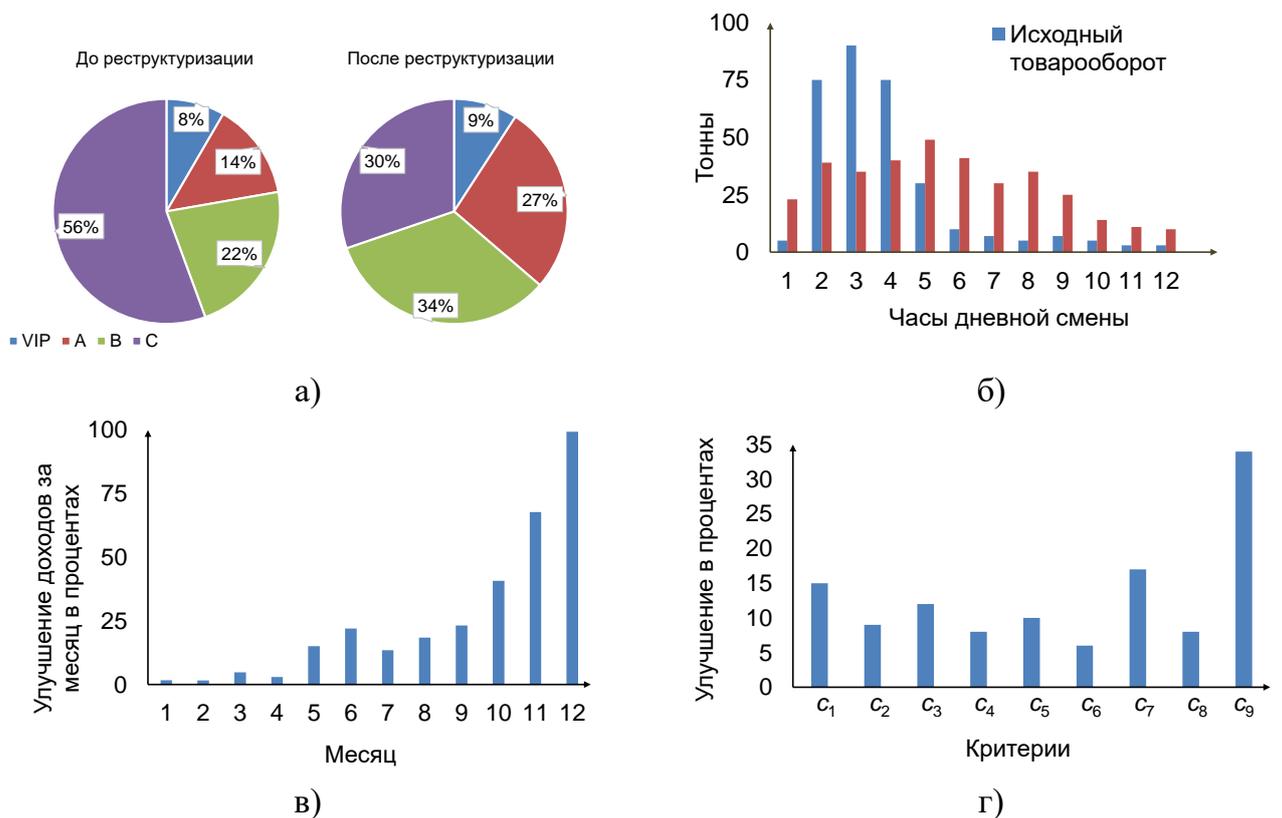


Рисунок 9 – Результаты решения задач: улучшение структуры клиентов (а), увеличение и балансировка суточного товарооборота (б), увеличение дохода от сдачи в аренду дополнительных объектов коммерческой недвижимости (в), улучшение критериев функционирования склада $c_1 - c_9$: коэффициентов полезного использования кладовщиков, грузчиков, электропогрузчиков, рабочих бригад и лифтов, а также среднего времени выполнения логистических операций, дохода от этих операций, наполнения склада и дохода от складских объектов (г).

Рисунки 10 а и 10 б демонстрируют рост ускорения вычислений и стабильность эффективности использования ресурсов ГРВС в процессе решения задач по мере увеличения числа используемых ядер. Эти результаты показывают высокую масштабируемость вычислений при выполнении заданий РППП. Задачи 1-3 упорядочены по возрастанию относительно числа обрабатываемых вариантов данных. Очевидно, что более ресурсоемкая задача 3 обеспечивает наилучшую масштабируемость.

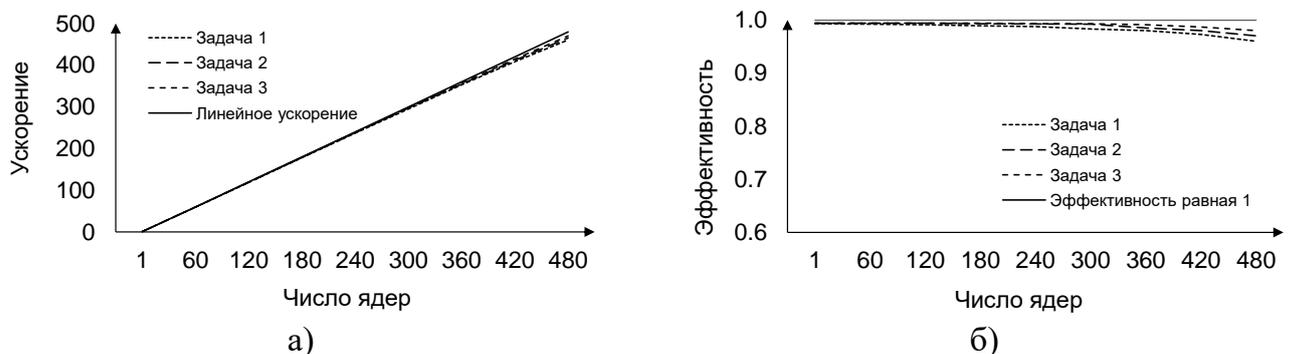


Рисунок 10 – Изменение ускорения вычислений (а) и эффективности использования ресурсов (б) по мере увеличения числа используемых ядер

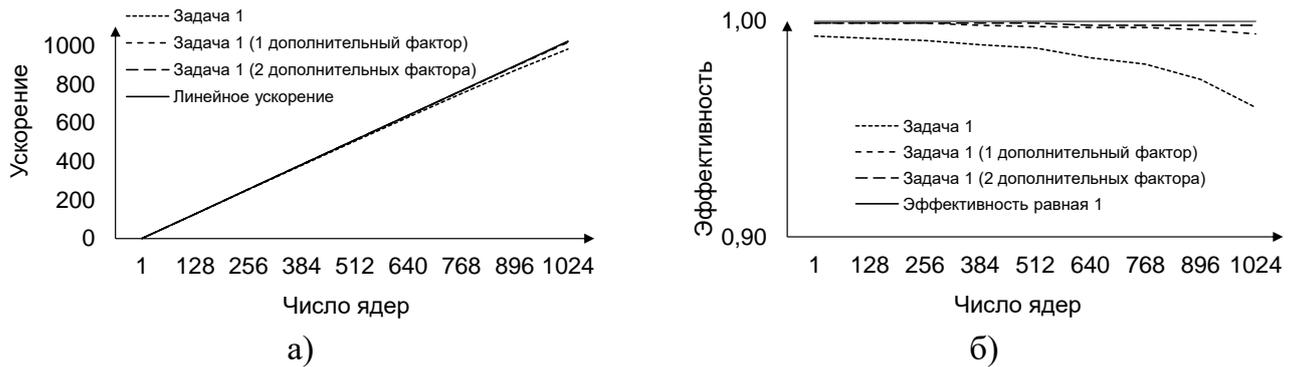


Рисунок 11 – Изменение ускорения вычислений (а) и эффективности использования ресурсов (б) в процессе решения задач при росте числа факторов

На рисунках 11 а и 11 б приведены оценки ускорения и эффективности, полученные для задачи 1 с разным числом факторов (входных переменных модели). Добавление одного нового фактора увеличивает число исходных вариантов данных в k раз, где k – это число уровней (допустимых значений) фактора. Это позволяет стабилизировать ускорение процесса вычислений и существенно повысить эффективность использования ресурсов. Для оценки масштабируемости вычислений проведено имитационное моделирование его работы в ГРВС, включающей в совокупности 30000 ядер. При этом для выполнения вычислений выделялось от 1024 до 9110 ядер разнородных узлов.

На рисунках 12 а и 12 б приведены ускорение и эффективность, полученные в результате эксперимента. Результаты моделирования подтверждают масштабируемость вычислений при росте числа ядер.

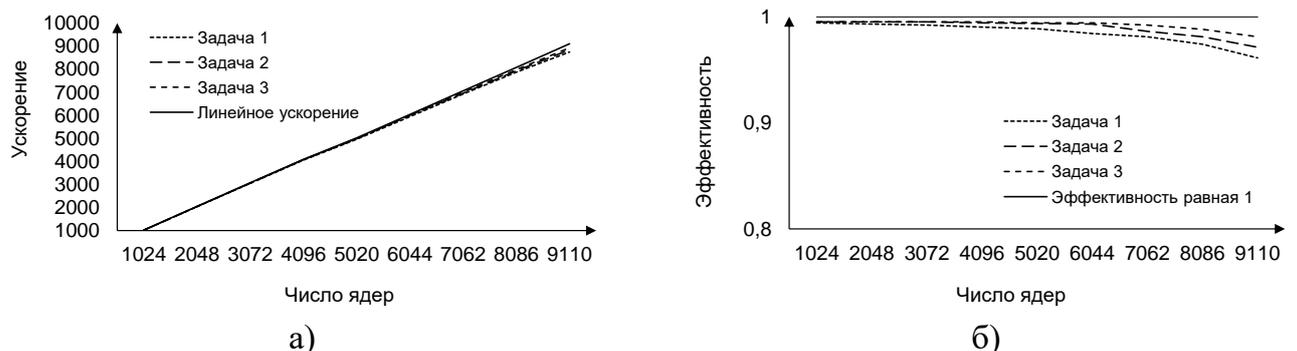


Рисунок 12 – Изменение ускорения вычислений (а) и эффективности использования ресурсов (б) при моделировании процесса решения задач с ростом числа факторов

В заключении подытожены основные результаты диссертационного исследования. **Приложения** включают копии актов о внедрении результатов диссертационного исследования, описание агрегированной модели ГРВС, экспериментальные данные, а также другие дополнительные материалы.

ЗАКЛЮЧЕНИЕ

Главным результатом диссертации является разработка моделей, алгоритмов и инструментальных средств создания и применения РППП, а также мультиагентного управления потоками их заданий в ГРВС.

Характерной особенностью предложенного подхода к организации мультиагентного управления в ГРВС является применение экономических механизмов регулирования спроса и предложения ресурсов среды.

Предложенная в диссертации технология предметно-ориентированных распределенных вычислений и мультиагентного управления ими в ГРВС допускает свое естественное развитие и обобщение применительно к другим классам вычислительных и управляющих систем.

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

В изданиях, рекомендованных ВАК:

1. Феоктистов, А.Г. Поддержка управления живучестью систем энергетики на основе комбинаторного подхода / И.В. Бычков, С.А. Горский, А.В. Еделев, Р.О. Костромин, И.А. Сидоров, А.Г. Феоктистов, Е.С. Фереферов, Р.К. Федоров // Известия РАН. Теория и системы управления. — 2021. — № 6. — С. 122–135.
2. Феоктистов, А.Г. Непрерывная интеграция функционального наполнения распределенных пакетов прикладных программ в Orlando Tools / А.Г. Феоктистов, С.А. Горский, И.А. Сидоров, Р.О. Костромин, Е.С. Фереферов, И.В. Бычков // Тр. Института системного программирования РАН. — 2019. — Т. 31, № 2. — С. 83–96.
3. Феоктистов, А.Г. Управление заданиями в гетерогенной распределенной вычислительной среде на основе знаний / А.Г. Феоктистов, Р.О. Костромин, Ю.А. Дядькин // Вестник компьютерных и информ. технологий. — 2018. — № 2. — С. 10–17.
4. Феоктистов, А.Г. Масштабируемое приложение для поиска глобальных минимумов многоэкстремальных функций / И.В. Бычков, Г.А. Опарин, А.Н. Черных, А.Г. Феоктистов, С.А. Горский, Р. Ривера-Родригес // Автометрия. — 2018. — Т. 54, № 1. — С. 98–105.
5. Феоктистов, А.Г. Распределенная вычислительная среда для анализа уязвимости критических инфраструктур в энергетике / А.В. Еделев, С.М. Сендеров, Н.М. Береснева, И.А. Сидоров, А.Г. Феоктистов // Системы управления, связи и безопасности. — 2018. — № 3. — С. 197–231.
6. Feoktistov, A. Virtualization of Heterogeneous HPC-clusters Based on OpenStack Platform / A. Feoktistov, I. Sidorov, V. Sergeev, R. Kostromin, V. Bogdanova // Вестник Южно-Уральского гос. ун-та. Сер. Вычисл. математика и информатика. — 2017. — Т. 6, № 2. — С. 37–48.

7. Феоктистов, А.Г. Логико-вероятностные аспекты алгоритма управления распределенными вычислениями / А.Г. Феоктистов // Вычисл. технологии. — 2016. — Т. 21, № 3. — С. 91–102.
8. Феоктистов, А.Г. Мультиагентное управление вычислительной системой на основе метамониторинга и имитационного моделирования / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, И.А. Сидоров, В.Г. Богданова, С.А. Горский // Автометрия. — 2016. — Т. 52, № 2. — С. 3–9.
9. Феоктистов, А.Г. Язык спецификации вычислительных моделей в масштабируемых пакетах прикладных программ / А.Г. Феоктистов, С.А. Горский // Современные наукоемкие технологии. — 2016. — № 7. — С. 84–88.
10. Феоктистов, А.Г. Мультиагентный подход к управлению сервис-ориентированными высокопроизводительными вычислениями / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, В.Г. Богданова, И.А. Сидоров, А.А. Пашинин // Вестник компьютерных и информ. технологий. — 2016. — № 9. — С. 35–41.
11. Феоктистов, А.Г. Разработка и применение проблемно-ориентированных мультиагентных систем управления распределенными вычислениями / А.Г. Феоктистов, Р.О. Костромин // Известия ЮФУ. Техн. науки. — 2016. — № 11. — С. 65–74.
12. Феоктистов, А.Г. Сервис-ориентированный мультиагентный подход к управлению распределенными вычислениями / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, В.Г. Богданова, А.А. Пашинин // Автоматика и телемеханика. — 2015. — № 11. — С. 118–131.
13. Феоктистов, А.Г. Автоматизация имитационного моделирования сложных систем в распределенной вычислительной среде / А.Г. Феоктистов, О.Ю. Башарина // Программные продукты и системы. — 2015. — № 3. — С. 75–79.
14. Феоктистов, А.Г. Мультиагентный подход к управлению распределенными вычислениями в кластерной Grid-системе / В.Г. Богданова, И.В. Бычков, А.С. Корсуков, Г.А. Опарин, А.Г. Феоктистов // Известия РАН. Теория и системы управления. — 2014. — № 5. — С. 95–105.
15. Феоктистов, А.Г. Мультиагентный алгоритм распределения вычислительных ресурсов на основе экономического механизма регулирования их спроса и предложения / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, А.Н. Кантер // Вестник компьютерных и информ. технологий. — 2014. — № 1. — С. 39–45.
16. Феоктистов, А.Г. Сервис-ориентированный подход к организации распределенных вычислений с помощью инструментального комплекса DISCENT / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, В.Г. Богданова, А.С. Корсуков // Информ. технологии и вычисл. системы. — 2014. — № 2. — С. 7–15.
17. Феоктистов, А.Г. Мультиагентные методы и инструментальные средства управления в сервис-ориентированной распределенной вычислительной среде / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, В.Г. Богданова, А.А. Пашинин //

- Тр. Института системного программирования РАН. — 2014. — Т. 26, вып. 5. — С. 65–82.
18. Феоктистов, А.Г. Испытание и оценка надежности интегрированных кластерных систем на основе их комплексного моделирования / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, А.С. Корсуков // Вестник компьютерных и информ. технологий. — 2013. — № 3. — С. 3–8.
 19. Феоктистов, А.Г. Распределение заданий в интегрированной кластерной системе на основе их классификации / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, А.С. Корсуков // Вычисл. технологии. — 2013. — Т. 18, № 2. — С. 25–32.
 20. Феоктистов, А.Г. Децентрализованное управление потоками заданий в интегрированной кластерной системе / И.В. Бычков, Г.А. Опарин, А.Г. Феоктистов, А.С. Корсуков // Вестник Новосибирского гос. ун-та. Сер. Информ. технологии. — 2011. — Т. 9, вып. 2. — С. 42–54.
 21. Феоктистов, А.Г. Инструментальный комплекс для организации гетерогенных распределенных вычислительных сред / И.В. Бычков, А.С. Корсуков, Г.А. Опарин, А.Г. Феоктистов // Информ. технологии и вычисл. системы. — 2010. — № 1. — С. 45–54.
 22. Феоктистов, А.Г. Разработка и применение распределенных пакетов прикладных программ / И.А. Сидоров, Г.А. Опарин, А.Г. Феоктистов // Программные продукты и системы. — 2010. — № 2. — С. 108–111.

В изданиях, входящих в базы цитирования Web of Science и Scopus

23. Feoktistov, A. Microservice-Based Approach to Simulating Environmentally-Friendly Equipment of Infrastructure Objects Taking into Account Meteorological Data / R. Kostromin, O. Basharina, A. Feoktistov, I. Sidorov // Atmosphere. — 2021. — Vol. 12, № 9: 1217. — P. 1–24.
24. Feoktistov, A.G. Static-dynamic algorithm for managing asynchronous computations in distributed environments / A.G. Feoktistov, S.A. Gorsky // Proceedings of the 1st International Workshop on Advanced Information and Computation Technologies and Systems. — CEUR-WS Proceedings. — 2021. — Vol. 2858. — P.64–73.
25. Feoktistov, A. Service-Oriented Tools for Automating Digital Twin Development / R. Kostromin, A. Feoktistov, M. Voskoboinikov // Proceedings of the 4th Scientific-practical Workshop on Information Technologies: Algorithms, Models, Systems. — CEUR-WS Proceedings. — 2021. — Vol. 2984. — P. 95–100.
26. Feoktistov, A. Agent-Based DevOps of Software and Hardware Resources for Digital Twins of Infrastructural Objects / R. Kostromin, A. Feoktistov // Proceedings of the 4th International Conference on Future Networks and Distributed Systems (ICFNDS 2020). — ACM, 2020. — P. 1–6.
27. Feoktistov, A. Collaborative Development and Use of Scientific Applications in Orlando Tools: Integration, Delivery, and Deployment / A. Feoktistov, S. Gorsky, I. Sidorov, I. Bychkov, A. Tchernykh, A. Edelev // Communications in Computer and Information Science. — 2020. — Vol. 1087. — P. 18–32.

28. Feoktistov, A. Supercomputer Engineering for Supporting Decision-making on Energy Systems Resilience / I. Bychkov, A. Feoktistov, S. Gorsky, A. Edelev, I. Sidorov, R. Kostromin, E. Fereferov, R. Fedorov // Proceedings of the 14th IEEE International Conference on Application of Information and Communication Technologies. — IEEE, 2020. — P. 1–6.
29. Feoktistov, A.G. Predicting runtime of computational jobs in distributed computing environment / A.G. Feoktistov, O.Yu. Basharina // Proceedings of the 2nd International Workshop on Information, Computation, and Control Systems for Distributed Environments. — CEUR-WS Proceedings. — 2020. — Vol. 2638. — P. 109–117.
30. Feoktistov, A. Tender of computational works in heterogeneous distributed environment / A. Feoktistov // Proceedings of the 2nd International Workshop on Information, Computation, and Control Systems for Distributed Environments. — CEUR-WS Proceedings. — 2020. — Vol. 2638. — P. 99–108.
31. Feoktistov, A. Methods and tools for evaluating the reliability of information and computation processes in grid and cloud systems / A. Feoktistov, I. Sidorov, R. Kostromin, G. Oparin, O. Basharina // Proceedings of the 1st International Workshop on Information, Computation, and Control Systems for Distributed Environments. — CEUR-WS Proceedings. — 2019. — Vol. 2430. — P. 60–69.
32. Feoktistov, A. Orlando Tools: Development, Training, and Use of Scalable Applications in Heterogeneous Distributed Computing Environments / A. Tchernykh, A. Feoktistov, S. Gorsky, I. Sidorov, R. Kostromin, I. Bychkov, O. Basharina, A. Alexandrov, R. Rivera-Rodriguez // Communications in Computer and Information Science. — 2019. — Vol. 979. — P. 265–279.
33. Feoktistov, A. Multi-Agent Algorithm for Re-Allocating Grid-Resources and Improving Fault-Tolerance of Problem-Solving Processes / A. Feoktistov, R. Kostromin, I. Sidorov, S. Gorsky, G. Oparin // Procedia Computer Science. — 2019. — Vol. 150. — P. 171–178.
34. Feoktistov, A. Configurable cost-quality optimization of cloud-based VoIP / A. Tchernykh, J.M. Cortés-Mendoza, I. Bychkov, A. Feoktistov, L. Didelot, P. Bouvry, G. Radchenko, K. Borodulin // J. of Parallel and Distributed Computing. — 2018. — Vol. 133. — P. 319–336.
35. Feoktistov, A. Operating cost and quality of service optimization for multi-vehicle-type timetabling for urban bus systems / D. Pena, A. Tchernykh, S. Nesmachnow, R. Massobrio, A. Feoktistov, I. Bychkov, G. Radchenko, A. Drozdov, S. Garichev // J. of Parallel and Distributed Computing. — 2018. — Vol. 133. — P. 272–285.
36. Feoktistov, A. Development of Distributed Subject-Oriented Applications for Cloud Computing through the Integration of Conceptual and Modular Programming / A. Feoktistov, R. Kostromin, I. Sidorov, S. Gorsky // Proc. of the 41st Intern. Convention on information and communication technology, electronics and microelectronics (MIPRO-2018). — Riejka: IEEE, 2018. — P. 256–261.
37. Feoktistov, A.G. Distributed Computing Environment for Vulnerability Analysis of Energy Critical Infrastructures / A.V. Edelev, I.A. Sidorov, A.G. Feoktistov // Advances in Intelligent Systems Research. — 2018. — Vol. 158. — P. 37–42.

38. Feoktistov, A. Framework for preparing subject data in testing modules of scientific applications / E.S. Fereferov, A.G. Feoktistov, I.V. Bychkov // Proceedings of the 1st Intern. Workshop on Information, Computation, and Control Systems for Distributed Environments. — CEUR-WS Proceedings. — 2018. — Vol. 2212. — P. 70–77.
39. Feoktistov, A.G. Conceptual Model of Problem-Oriented Heterogeneous Distributed Computing Environment with Multi-Agent Management / I.V. Bychkov, G.A. Oparin, A.N. Tchernykh, A.G. Feoktistov, V.G. Bogdanova, S.A. Gorsky // Procedia Computer Science. — 2017. — Vol. 103. — P. 162–167.
40. Feoktistov, A. Simulation Modeling in Heterogeneous Distributed Computing Environments to Support Decisions Making in Warehouse Logistics / I. Bychkov, G. Oparin, A. Tchernykh, A. Feoktistov, V. Bogdanova, Yu. Dyadkin, V. Andrukhova, O. Basharina // Procedia Engineering. — 2017. — Vol. 201. — P. 524–533.
41. Feoktistov, A. Job Flow Management for Virtualized Resources of Heterogeneous Distributed Computing Environment / I. Bychkov, A. Feoktistov, I. Sidorov, R. Kostromin // Procedia Engineering. — 2017. — Vol. 201. — P. 534–542.
42. Feoktistov, A.G. The Service-Oriented Multiagent Approach to High-Performance Scientific Computing / I.V. Bychkov, G.A. Oparin, A.G. Feoktistov, V.G. Bogdanova, I.A. Sidorov // Lecture Notes in Computer Science. — 2017. — Vol. 10187. — P. 256–263.
43. Feoktistov, A. Automation of Multi-Agent Control for Complex Dynamic Systems in Heterogeneous Computational Network / G. Oparin, A. Feoktistov, V. Bogdanova, I. Sidorov // AIP Conference Proceedings. — 2017. — Vol. 1798. — P. 0201171–02011710.
44. Feoktistov, A.G. Logical-Probabilistic Analysis of Distributed Computing Reliability / A.G. Feoktistov, I.A. Sidorov // Proc. of the 39th Intern. Convention on information and communication technology, electronics and microelectronics (MIPRO-2016). — Riejska: IEEE, 2016. — P. 247–252.

Научно-организационный отдел
Федерального государственного бюджетного учреждения науки
Института динамики систем и теории управления имени В.М. Матросова СО РАН
664033, Иркутск, ул. Лермонтова, д. 134
e-mail: rio@icc.ru

Подписано к печати 22.12.2021 г.
Формат бумаги 60×84 1/16, объем 1 п.л.
Заказ № 5. Тираж 100 экз.

Отпечатано в ИДСТУ СО РАН