

На правах рукописи



Костромин Роман Олегович

**Модели, алгоритмы и инструментальные средства поддержки
мультиагентного управления потоками вычислительных заданий**

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

Автореферат диссертации на соискание ученой степени
кандидата технических наук

Иркутск – 2020

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

- Научный руководитель:** кандидат технических наук, доцент,
Феоктистов Александр Геннадьевич,
ИДСТУ СО РАН, заведующий
лабораторией 5.1 параллельных и
распределенных вычислительных систем
- Официальные оппоненты:** доктор физико-математических наук,
доцент,
Белеванцев Андрей Андреевич,
Институт системного программирования
им. В.П. Иванникова РАН, ведущий
научный сотрудник
- кандидат технических наук, доцент,
Ковтуненко Алексей Сергеевич,
Уфимский государственный авиационный
технический университет, доцент
- Ведущая организация:** **Вычислительный центр ДВО РАН –
обособленное подразделение
Хабаровского Федерального
исследовательского центра ДВО РАН
(г. Хабаровск)**

Защита состоится «04» февраля 2021 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 «15» декабря 2020 г.

Отзывы и замечания по автореферату в двух экземплярах, заверенные печатью, просьба высылать по вышеуказанному адресу на имя ученого секретаря диссертационного совета.

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



Т.В. Груздева

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы диссертации. В настоящее время как в России, так и за рубежом, ведутся активные исследования, связанные с созданием и применением высокопроизводительных вычислительных систем различного назначения на основе парадигм Grid и облачных вычислений. Результаты этих исследований представлены в работах С.М. Абрамова, А.И. Аветисяна, А.П. Афанасьева, И.В. Бычкова, Вл.В. Воеводина, Б.М. Глинского, В.П. Иванникова, И.А. Каляева, И.И. Левина, А.И. Легалова, Г.И. Радченко, Л.Б. Соколинского, В.В. Топоркова, Г.А. Опарина, А.Н. Черных, Д. Андерсена, Р. Байя, К. Кессельмана, Э. Танненбаума, Я. Фостера, а также многих других российских и зарубежных ученых. В связи с высокой интенсивностью потоков вычислительных заданий в таких системах необходимо эффективное и гибкое управление ими на метауровне.

Эффективность управления состоит в обеспечении высокого качества обслуживания очередей заданий, повышении надежности выполнения заданий, минимизации времени решения задач, поддержки равномерной балансировки загрузки ресурсов и достижения других заданных показателей. Гибкость управления заключается в рациональном предоставлении ресурсов, необходимых для выполнения заданий, в динамически изменяющихся системах.

В процессе управления потоками заданий требуется детальный учет и согласование критериев пользователей, определяемых спецификой решаемых задач, и предпочтений владельцев ресурсов, вытекающих из их характеристик. Как правило, распределенная вычислительная система (РВС), интегрирующая модели Grid и облачных вычислений, обладает рядом свойств, существенно усложняющих унификацию процесса управления заданиями. К свойствам такого рода, например, относятся архитектурно-функциональная разнородность, неполнота описания и динамичность ресурсов, широта спектра задач, решаемых с помощью этих ресурсов, наличие различных категорий пользователей, преследующих свои цели и задачи эксплуатации вычислительной среды.

Известные модели, методы, алгоритмы и программные средства централизованного управления потоками заданий не решают перечисленные выше проблемы полностью. Как правило, это обусловлено следующими причинами:

- ограниченным контролем централизованной системы над ее распределенными ресурсами, не позволяющим в полной мере оценить состояние этих ресурсов и воздействовать на него;

- отсутствием необходимой информации о разнородных ресурсах (степени их надежности, показателях производительности, вычислительной истории выполнения конкретных заданий, дисциплинах обслуживания очередей, административных политиках и других важных характеристиках) или существенным увеличением накладных расходов при получении таких сведений, которое приводит к снижению эффективности функционирования системы управления;

– недостаточной надежностью самой системы управления, так как при отказе центрального узла вся вычислительная среда становится неработоспособной;

– резким снижением производительности системы при значительном увеличении числа заданий потока.

В связи с этим возникает необходимость разработки новой эффективной и гибкой системы управления потоками заданий. Для решения данной проблемы целесообразно использование принципов организации распределенного группового управления. Качество управления, осуществляемого отдельными компонентами, обеспечивается наличием у них более полных локальных знаний об управляемых ими ресурсах по сравнению с централизованной системой.

Групповое управление потоками заданий в процессе решения сложной прикладной задачи в РВС осуществляется посредством передачи сигналов (сообщений) в коммуникационной среде между компонентами распределенной управляющей системы. Учитывая динамическую природу РВС, целесообразно применять адаптивное предоставление ресурсов в процессе управления потоками заданий. В этом случае перспективным подходом является использование мультиагентных технологий. В рамках такого подхода отдельные ресурсы представляются специализированными приложениями (агентами), образующими в совокупности мультиагентную систему (МАС) управления.

Анализ результатов исследований в области самоорганизации вычислительных систем (см., например, работы Д.В. Винса, Т.А. Гавриловой, В.И. Гальперова, В.И. Городецкого, И.А. Каляева, А.С. Ковтуненко, Л.В. Массель, Д.А. Поспелова, В.Б. Тарасова, В.Ф. Хорошевского, N. Jennings, S.J. Russell, P. Norvig, G. Di Marzo Serugendo, M. Wooldridge, F. Zambonelli) показывает, что эффективное управление потоками заданий с помощью МАС достигается за счет применения алгоритмов работы агентов, адаптирующихся в процессе их выполнения к текущим условиям и состоянию функционирования среды в соответствии с заданной агентам целью, определенными критериями качества решения задачи и использования ресурсов, а также знаниями об особенностях предметных областей решаемых задач.

Построение проблемно-ориентированной самоорганизующейся МАС порождает ряд проблем, связанных с автоматизацией разработки агентов и агентных платформ, реализации алгоритмов функционирования агентов, а также накопления и применения предметных знаний агентами. Существующие в настоящее время инструменты для создания МАС не позволяют решить вышеперечисленные проблемы в полной мере.

Цель работы заключается в разработке новых инструментальных средств, обеспечивающих снижение трудозатрат при построении МАС по сравнению с существующими инструментариями, а также моделей и алгоритмов работы агентов создаваемых систем, позволяющих улучшить показатели (качество обслуживания очереди заданий, время и надежность их выполнения, балансировку загрузки ресурсов) управления потоками вычислительных заданий в разнородной РВС по сравнению с известными метапланировщиками GridWay и

Condor Directed Acyclic Graph Manager (DAGMan).

Основные задачи для достижения поставленной цели:

- проведение сравнительного анализа известных методов и средств организации МАС, а также известных методов и средств мультиагентного управления потоками вычислительных заданий;
- разработка модели поведения агентов;
- разработка алгоритмов функционирования агентов;
- разработка инструментального комплекса построения МАС для управления потоками вычислительных заданий;
- создание МАС управления потоками заданий в экспериментальной РВС;
- оценка надежности и эффективности управления потоками заданий с помощью разработанной МАС, а также трудоемкости ее построения.

Объектом исследования является процесс управления потоками вычислительных заданий в РВС.

Предметом исследования выступают мультиагентные модели, алгоритмы и система управления потоками заданий в разнородной РВС, а также инструментальные средства построения МАС.

Методы исследования. При решении поставленных задач использовались методы концептуального, имитационного, конкретизирующего и автоматного программирования, организации распределенных вычислений и мультиагентных технологий, а также машинного обучения.

Научная новизна диссертации состоит в интеграции уникальной совокупности методов и средств концептуального, имитационного, конкретизирующего и автоматного программирования, классификации заданий и параметрической настройки алгоритмов работы агентов в качестве основы их машинного обучения, организации распределенных вычислений и управления ими в процессе создания и применения оригинальных мультиагентных моделей, алгоритмов и системы управления потоками заданий в разнородной РВС, а также инструментальных средств их разработки.

На защиту выносятся следующие основные результаты:

1) ролевая модель поведения агентов, базирующаяся, в отличие от известных, на использовании конечных управляющих автоматов с динамическим планированием их состояний-действий в разнородной РВС на концептуальной модели среды и применении механизма порождения дочерних автоматов для реализации специфических ролей в возникающих виртуальных сообществах;

2) система машинного обучения агентов, которая основывается на применении новой гибридной модели представления знаний, обеспечивающей интегрированное использование концептуального и имитационного моделирования, классификации заданий и параметрической настройки алгоритмов работы агентов в качестве методов обучения в сочетании с процессами самостоятельного извлечения и передачи знаний агентами;

3) мультиагентный алгоритм перераспределения ресурсов разнородной РВС в случае отказа ее программно-аппаратных средств, реализующий адаптивное

мультиценарное решение данной проблемы на основе конкретизирующего программирования и тем самым существенно повышающий отказоустойчивость процесса выполнения заданий под управлением МАС по сравнению с метапланировщиками GridWay и Condor DAGMan;

4) инструментальный комплекс построения МАС, обеспечивающий по сравнению с известными инструментариями сокращение трудозатрат на реализацию разработанных моделей, алгоритмов и системы в целом путем автоматизации основных этапов разработки, настройки, конфигурации и применения агентов.

Достоверность и обоснованность полученных в работе результатов подтверждается корректным применением классических методов исследования, а также анализом адекватности разработанных моделей и алгоритмов на основе полунатурного моделирования.

Соответствие диссертации паспорту специальности. Тема и основные результаты диссертации соответствуют следующим областям исследований, включенным в паспорт 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.

Практическая значимость. Применение вышеперечисленных результатов диссертационного исследования обеспечивает существенное улучшение качества обслуживания очередей заданий, минимизацию времени решения задач, повышение надежности их выполнения и сбалансированную загрузку ресурсов по сравнению с известными метапланировщиками GridWay и CondorDAGMan.

Исследование, разработка и применение рассматриваемых в диссертации программных средств выполнялись в рамках следующих научно-технических работ: проектов РФФИ № 14-08 3162-мол_а «Методические подходы и комплекс программ для оптимизации режимов работы крупных ТЭЦ», № 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»; проекта «Фундаментальные проблемы решения сложных практических задач с помощью суперкомпьютеров» программы фундаментальных исследований президиума РАН № 27; проекта «Методы, алгоритмы и инструментальные средства децентрализованного группового решения задач в вычислительных и управляющих системах» программы фундаментальных исследований президиума РАН № 30; базовых тем исследований ИДСТУ СО РАН № IV.38.1.2 «Разработка проблемно-ориентированных технологий, систем и сервисов поддержки научных

исследований на основе интеллектуальных методов и алгоритмов организации параллельных и распределенных вычислений» и № IV.38.1.1 «Технологии разработки проблемно-ориентированных самоорганизующихся мультиагентных систем группового управления: методы, инструментальные средства, приложения».

Практическое использование разработанных моделей, алгоритмов, методов и инструментальных средств в процессе решения научных и прикладных задач подтверждено справкой № 11-054 от 06.05.2019 Международного института экономики и лингвистики Иркутского государственного университета об использовании результата интеллектуальной деятельности для автоматизации проведения научных исследований.

Апробация результатов работы. Основные результаты работы докладывались автором на следующих научных мероприятиях: 13th International Symposium «Intelligent Systems» (INTELS-2018, Санкт-Петербург, Россия, 2018 г.), International Symposium on Cloud Computing and Services for High Performance Computing Systems (HPCS-2018, Орлеан, Франция, 2018 г.), 41th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO-2018, Опатия, Хорватия, 2018 г.), 4th International Conference on Information Technology and Nanotechnology (ITNT-2018, Самара, Россия, 2018 г.), XLVI Международная конференция «Информационные технологии в науке, образовании и управлении» (IT+S&E`17, Гурзуф, Россия, 2017 г.), IV Всероссийская мультikonференция по проблемам управления (МКПУ-2017, Дивноморское, Россия, 2017 г.), XXII и XXIII Байкальские Всероссийские конференции с международным участием «Информационные и математические технологии в науке и управлении» (ИМТ, Иркутск – Байкал, Россия, 2017–2018 гг.), Multidisciplinary Youth Academic Research Conference on Science Present and Future: Research Landscape in the 21st Century (Иркутск, Россия, 2017 г.), XIII Всероссийская конференция молодых ученых «Моделирование, оптимизация и информационные технологии» (Иркутск, Россия, 2017 г.), конференции «Ляпуновские чтения» (Иркутск, Россия, 2016–2019 гг.), IV и V Всероссийские научно-технические конференции «Суперкомпьютерные технологии» (СКТ, Дивноморское, Россия, 2016, 2018 гг.), XVI и XVIII Всероссийские конференции молодых ученых по математическому моделированию и информационным технологиям (Красноярск, Россия, 2015 г.; Иркутск, Россия, 2017 г.), XV Молодежная научно-практическая конференция «Российская цивилизация: история, проблемы, перспективы» (Иркутск, Россия, 2015 г.), а также семинарах ИДСТУ СО РАН.

Публикации. Результаты научных исследований автора отражены в 34 научных работах. В их числе 6 публикаций [1–6] в российских журналах, рекомендованных ВАК для опубликования научных результатов диссертации, 8 публикаций [7–14], проиндексированных в международных базах цитирования Web of Science и Scopus. Получены 2 свидетельства о государственной регистрации программ для ЭВМ [15, 16].

Личный вклад автора. Все выносимые на защиту научные положения

получены соискателем лично. В основных научных работах по теме диссертации, опубликованных в соавторстве, лично соискателем представлены: в [4, 17, 18] – сравнительный анализ методов и средств организации МАС; в [2, 10, 11–14, 16] – ролевая модель поведения агентов и алгоритмы их функционирования; в [1, 3, 5, 7–9, 15] – инструментальные средства создания агентов и результаты вычислительных экспериментов в рамках полунатурного моделирования мультиагентной системы.

Структура работы. Диссертация состоит из введения, четырех глав, заключения, списка используемых сокращений, глоссария, библиографии из 125 наименований и семи приложений. Общий объем работы – 173 страницы, из которых 117 страниц основного текста, включающего 55 рисунков и 13 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В главе 1 рассматриваются общие вопросы организации и применения МАС для управления потоками вычислительных заданий в РВС, исследуются известные МАС, проводится сравнительный анализ методов и средств построения таких систем, обосновывается необходимость разработки нового инструментального комплекса создания МАС, формулируются системные и функциональные требования к разрабатываемому комплексу.

Глава 2 посвящена новым моделям и алгоритмам мультиагентного управления потоками вычислительных заданий, разрабатываемым в рамках диссертационного исследования.

В п. 2.1 представлена иерархическая структура МАС для управления потоками вычислительных заданий в РВС, на каждом уровне которой могут функционировать агенты с различными ролями и функциями. Агенты одного уровня могут объединяться в виртуальные сообщества, кооперироваться и конкурировать в рамках этих сообществ. Координация действий агентов осуществляется с помощью общих правил группового поведения. Агенты работают в соответствии с заданными ролями по своим правилам поведения в виртуальном сообществе. Данные о текущем состоянии РВС передаются в ее базу знаний системой метамониторинга. МАС ориентирована на управление потоками заданий распределенных пакетов прикладных программ (РППП), разрабатываемых с помощью специализированных инструментальных комплексов.

В п. 2.2 дано описание специализированной концептуальной модели, которая, в отличие от известных вычислительных моделей, обеспечивает взаимосвязанное представление проблемно-ориентированного, программно-аппаратного, имитационного и управляющего слоев знаний о РВС. Тем самым обеспечивается проблемная ориентация МАС. Предложенная модель является частным случаем семантической сети. Она обеспечивает спецификацию знаний о программных модулях для решения задач в предметных областях и работы с объектами РВС, схемных знаний о модульной структуре модели и алгоритмов, продукционных знаний для поддержки принятия решений по выбору

оптимальных алгоритмов в зависимости от состояния среды, а также знаний о программно-аппаратной инфраструктуре системы и административных политиках в ее узлах.

Пусть Z , F и M – это множества параметров, операций и программных модулей модели. Множество модулей является элементом вычислительных знаний. Параметры, операции и их взаимосвязи отражают схемные знания. Операции из F определяют вычислительные действия на множестве параметров Z . Каждой операции $f_i \in F$ соответствует модуль $m_j \in M$, где $i \in \overline{1, n_f}$, $j \in \overline{1, n_m}$, n_f – число операций, n_m – число модулей. Спецификация модуля включает язык программирования, тип и семантику входных, выходных и транзитных параметров, способы передачи параметров, модуль представления, требуемый компилятор и другие сведения. С каждой операцией f_i связаны два множества параметров $Z_i^{in}, Z_i^{out} \subset Z$. Множество Z_i^{in} определяет параметры, значения которых необходимо задать, чтобы получить значения параметров, представленных множеством Z_i^{out} . Множества Z_i^{in} и Z_i^{out} являются множествами соответственно входных и выходных параметров модуля m_j , реализующего операцию f_i . Постановки задач могут формулироваться в полной или сокращенной (процедурной или непроцедурной) форме. По сформулированной постановке задачи строится схема ее решения (абстрактная программа) на основе методов статического, динамического или статико-динамического планирования вычислений. В множествах параметров и операций введены подмножества соответственно системных параметров, отражающих характеристики объектов РВС, и операций, представляющих алгоритмы планирования вычислений, мониторинга и распределения ресурсов, моделирования поведения среды и других действий.

В п. 2.3 предложена новая ролевая модель поведения агентов. Она представлена структурой $M^{agent} = \langle sm^p, \{sm_{i,j}^c : i \in \overline{1, n_{vc}}, j \in \overline{1, n_{rol}}\}, MES \rangle$, где sm^p – родительский автомат, $sm_{i,j}^c$ – дочерние автоматы, n_{vc} – число виртуальных сообществ, в которых состоит агент, n_{rol} – число ролей, которые может играть агент, MES – множество сообщений агента. Основной функцией родительского автомата sm^p является создание дочернего автомата $sm_{i,j}^c$ при каждом включении агента в новое виртуальное сообщество, где i и j – это соответственно номера виртуального сообщества и роли агента.

Модель родительского автомата представлена структурой $sm^p = \langle STS^p, sts_o^p, X_{inputs}^p, X_{outputs}^p, g^p, h^p, ACT^p \rangle$, где STS^p – конечное множество управляющих состояний родительского автомата, $sts_o^p \in STS^p$ – начальное состояние, $X_{inputs}^p \subset Z$ – конечное множество входных воздействий, порождаемых РВС, $X_{outputs}^p \subset Z$ – конечное множество выходных воздействий,

$g^p : STS^p \times X_{inputs}^p \times ACT^p \rightarrow X_{outputs}^p$ – функция выходов, $h^p : STS^p \times X_{inputs}^p \rightarrow STS^p$ – функция переходов, $g^p, h^p \in F$, $ACT^p \subset F$ – конечное множество действий родительского автомата. В множестве STS^p выделяется конечное состояние sts_{end}^p , в котором агент завершает свою работу. В множестве $X_{inputs}^p \cup X_{outputs}^p$ выделяется набор глобальных переменных GV родительского автомата, доступных дочерним автоматам. В общем случае $X_{inputs}^p \cap X_{outputs}^p \neq \emptyset$. При создании агента все схемы выполнения действий являются абстрактными программами, которые генерируются на языке программирования Java.

Модель дочернего автомата представлена структурой $sm_{i,j}^c = \langle STS_{i,j}^c, sts_{i,j,0}^c, X_{i,j,inputs}^c, X_{i,j,outputs}^c, g_{i,j}^c, h_{i,j}^c, ACT_{i,j}^c \rangle$, где $sm_{i,j}^c$ – конечное множество управляющих состояний дочернего автомата, $sts_{i,j,0}^c \in sts_{i,j}^c$ – начальное состояние, $X_{i,j,inputs}^c \subset Z$ – конечное множество входных воздействий, порождаемых РВС, $X_{i,j,outputs}^c \subset Z$ – конечное множество выходных воздействий, $g_{i,j}^c : STS_{i,j}^c \times X_{i,j,inputs}^c \times ACT_{i,j}^c \rightarrow X_{i,j,outputs}^c$ – функция выходов, $h_{i,j}^c : STS_{i,j}^c \times X_{i,j,inputs}^c \rightarrow STS_{i,j}^c$ – функция переходов, $g_{i,j}^c, h_{i,j}^c \in F$, $ACT_{i,j}^c \subset F$ – конечное множество действий дочернего автомата. В множестве $STS_{i,j}^c$ выделяется конечное состояние $sts_{i,j,end}^c$, в котором агент завершает свою работу.

В общем случае $X_{i,j,inputs}^c \cap X_{i,j,outputs}^c \neq \emptyset$. Дочерние автоматы разных агентов, входящих в одно виртуальное сообщество, взаимодействуют путем обмена сообщениями, передаваемыми через родительские автоматы. Автоматы, являющиеся потомками одного и того же родительского автомата, обмениваются информацией об использовании общих ресурсов агента через глобальные переменные родительского автомата и агентную базу знаний.

Рассмотренные выше родительский и дочерний автоматы являются частными случаями конечного управляющего автомата. Их принципиальным отличием является то, что в качестве состояний автомата используются состояния-действия. В рамках одного состояния агент может выполнять одно или несколько действий. В каждом состоянии родительский (дочерний) автомат изменяет значения параметров функций переходов h^p ($h_{i,j}^c$) с помощью функции g^p ($g_{i,j}^c$), а затем осуществляет переход в новое состояние.

При вступлении агента в i -е виртуальное сообщество для соответствующего дочернего автомата создается система логического времени SLT_i , определяемая структурой $SLT_i = \langle T, T_m, g_t, g_m, g_r \rangle$, где T – область значений логического времени, T_m – область значений временных маркеров датировки сообщений, $T_m \subseteq T$, g_t , g_m и g_r – функции соответственно датировки событий автомата, маркировки сообщений и сравнения значений логического времени $\forall i \in \overline{1, n_{vc}}$.

Система логического времени использует векторные часы, в которых число компонент вектора времени равно числу агентов виртуального сообщества, и обеспечивает отношение частичного порядка на множестве событий виртуального сообщества с учетом их обусловленности. Применение функции датировки сообщений позволяет осуществлять их обработку в установленной логической последовательности, а не в произвольном порядке поступления их в общий пул.

В п. 2.4 рассматривается расширение концептуальной модели РВС. С целью поддержки построения и применения ролевой модели поведения агентов, концептуальная модель дополнена новыми объектами: агентами, их ролями, виртуальными сообществами и отношениями между ними, а также состояниями, функциями и графами переходов автоматов. В качестве состояния агента используется состояние-действие – последовательность системных операций, выполняемых над полем системных параметров модели. Фрагмент такой модели, описывающей объекты, необходимые для построения графов переходов агентов, представлен на рисунке 1. Здесь G – множество графов переходов, A – множество агентов, VC – множество виртуальных сообществ агентов, R – множество ролей агентов и STS – множество состояний автоматов. На рисунке 1 отношения между объектами обозначены $o_1 - o_8$. Отношение o_4 отображает взаимосвязь состояний с операциями, реализующими функции переходов. Отношение o_5 представляет взаимосвязь состояний с остальными системными операциями. База знаний агента создается на основе фрагмента модели РВС.

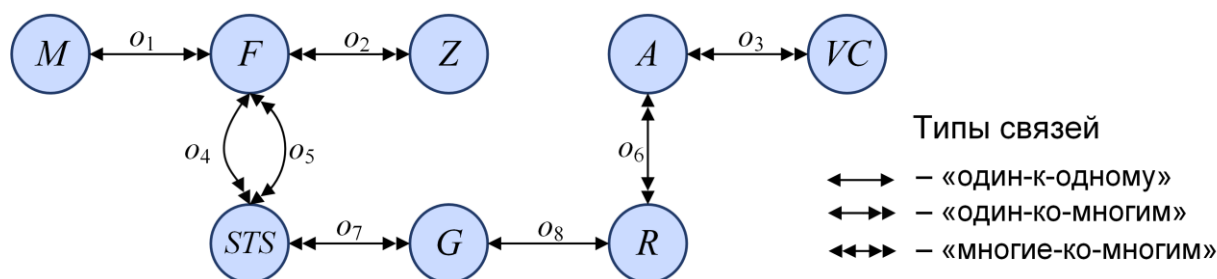


Рисунок 1 – Фрагмент описания расширенной концептуальной модели

В п. 2.5 подробно описаны алгоритмы функционирования агентов, базирующиеся на разработанной ранее ролевой модели поведения агентов (см. п. 2.3). Для каждого алгоритма приведены соответствующие им графы переходов, функции переходов и состояния-действия агентов. Также приводятся ограничения корректности функций, определяющих условия переходов, и соответствующие им таблицы истинности.

В п. 2.6 представлен новый алгоритм перераспределения ресурсов РВС для остаточной схемы решения задачи в случае отказа ее программно-аппаратных средств, определено понятие остаточной схемы решения задачи, формализован процесс ее построения. Рассмотрена информационно-вычислительная модель системы диагностики, приведены основные виды отказов. Алгоритм включает

следующие этапы работы: обнаружение отказа; обработка отказа; выбор сценария для обеспечения отказоустойчивости выполнения схемы решения задачи (рисунок 2); построение остаточной схемы решения задачи.

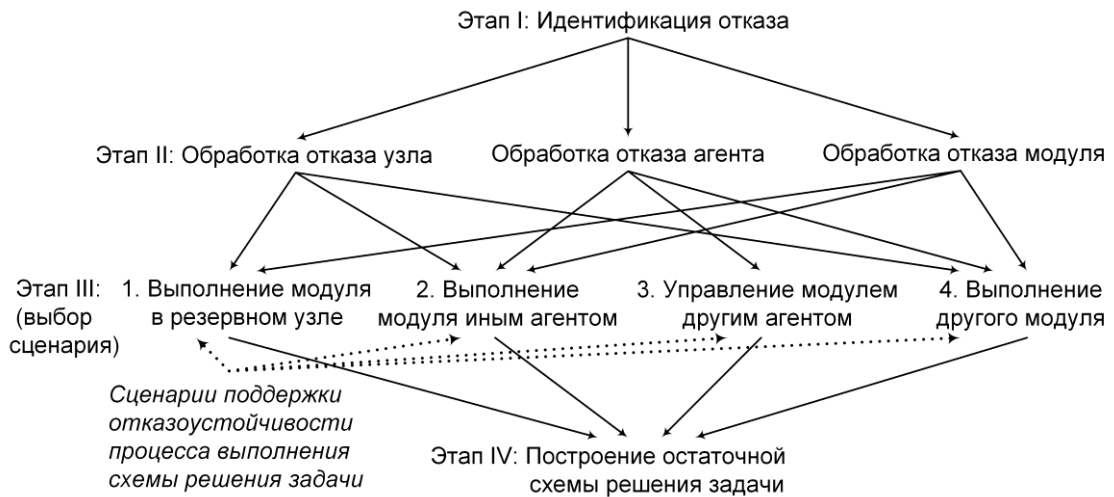


Рисунок 2 – Сценарии обработки отказов

В п. 2.7 предложена система машинного обучения агентов. Процесс обучения агентов строится на комплексном использовании методов концептуального моделирования, классификации заданий и параметрической настройки системы управления. Представлены методы, средства обучения агентов, рассматривается система классификации заданий, схема параметрической настройки, раскрыты аспекты их работы, особенности реализации и преимущества применения машинного обучения. В частности, в рамках параметрической настройки агенты способны самостоятельно корректировать отдельные управляющие параметры своих алгоритмов в пределах их допустимых значений. Классификация позволяет агентам осуществлять рациональное назначение ресурсов для выполнения наиболее подходящих им классов заданий.

В главе 3 рассмотрена агентная платформа JADE и предложена интеллектуальная надстройка над JADE, позволяющая автоматизировать процесс конструирования агентов. Освещены аспекты реализации инструментального комплекса организации MAC с применением платформы JADE, а также предложена методика применения инструментального комплекса.

В п. 3.1 обсуждаются ключевые особенности архитектуры платформы JADE, а также основные трудности, сопровождающие разработку агентов в JADE.

В п. 3.2. рассматривается реализация автоматной модели поведения агентов с помощью встроенного класса JADE FSMBehaviour.

В п. 3.3 представлен инструментальный комплекс для автоматизации разработки MAC. Одним из основных элементов данного комплекса является надстройка над JADE, которая включает библиотеку алгоритмов работы агентов,

реализующую их основные поведенческие функции: формулировку постановок задач, планирование вычислений, распределение ресурсов, обеспечение отказоустойчивости вычислительных процессов и др. Данная надстройка отвечает за генерацию агентов MAS на основе базовых библиотек JADE, специальных библиотек поведения агентов, концептуальной модели PBC и прикладных модулей решаемой задачи. Генерация кода агента осуществляется в рамках каркасного подхода к конструированию программ (рисунок 3).

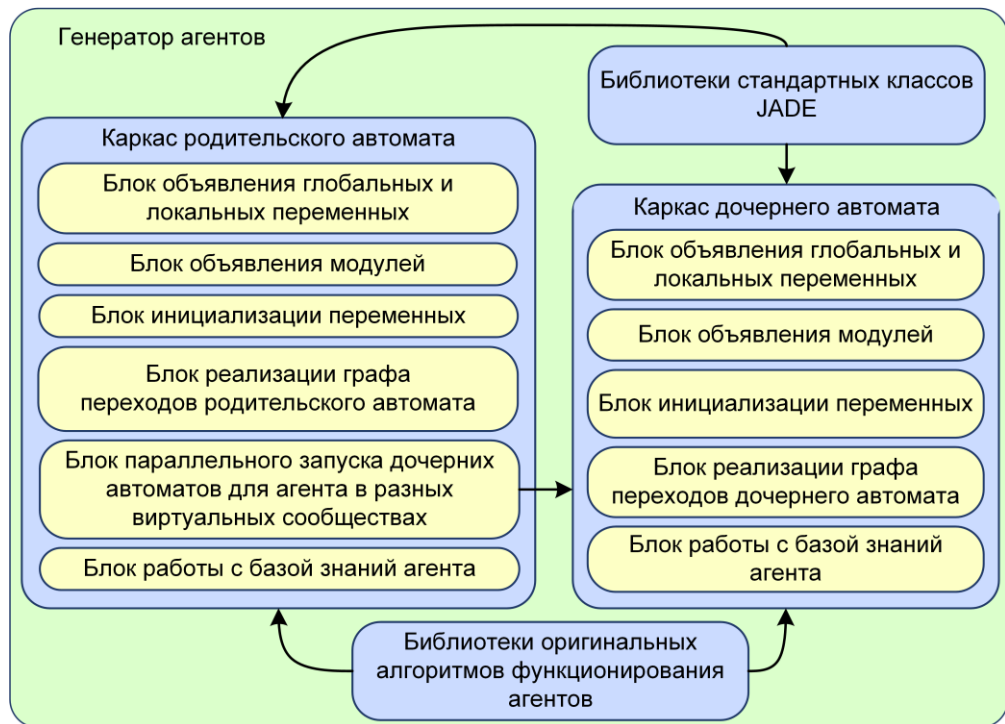


Рисунок 3 – Генерация программного кода агента

JADE предоставляет набор стандартных библиотек, обеспечивающих функционирование агентной платформы, в том числе обмен сообщениями. С помощью данных библиотек реализуются базовые функции управления жизненным циклом агентов и взаимодействия с платформой. За наполнение агента поведением отвечает разработчик MAS.

Для создания дополнительных методов стандартных классов JADE, представляющих функции (операции) агентов MAS, применяются библиотека «встроенных» алгоритмов, оформленных в виде модулей, а также подключаемые исполняемые модули пользователей (рисунок 4). Каждому агенту назначается модель поведения, связанная с имеющимися алгоритмами, а также передается полная информация о модели PBC. Сгенерированный таким образом агент в результате сборки и компиляции автоматически размещается на выбранном узле и подключается к MAS.

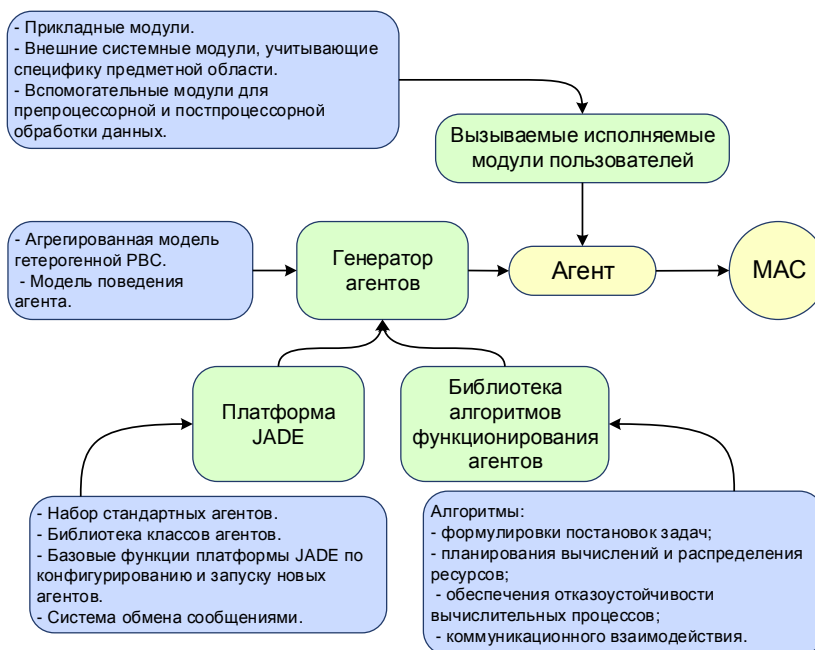


Рисунок 4 – Надстройка над JADE

В п. 3.4 предлагается методика применения разработанного инструментального комплекса, включающая следующие основные этапы: установку и запуск платформы JADE; настройку конфигурации MAS; разработку моделей поведения агентов; генерацию программного кода агентов; размещение агентов и их подключение к JADE; подключение агентов к выбранному инструментальному комплексу разработки РППП.

Глава 4 посвящена экспериментальному анализу результатов практической апробации разработанных моделей, алгоритмов и инструментальных средств мультиагентного управления потоками заданий при решении задач в экспериментальной РВС, организованной с использованием ресурсов Центра коллективного пользования (ЦКП) «Иркутский суперкомпьютерный центр СО РАН» (ИСКЦ) и других научно-образовательных организаций. Структура РВС, характеристики ее ресурсов и схема управления вычислениями в ней представлены в п. 4.1.

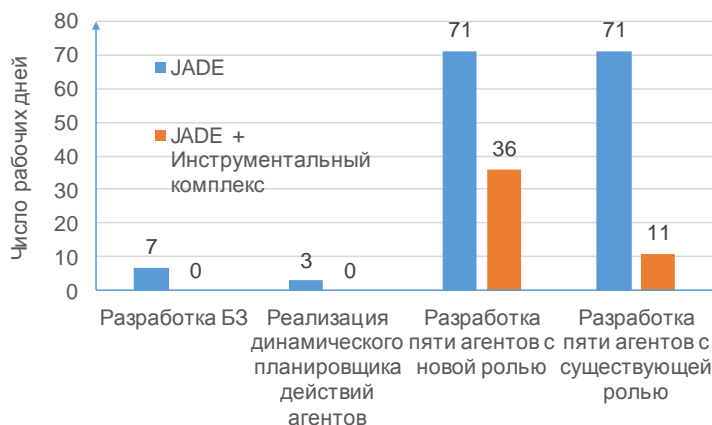


Рисунок 5 – Трудозатраты на создание MAS

В п. 4.2 осуществляется сравнительный анализ трудоемкости построения MAC на основе JADE с использованием разработанного в диссертации инструментального комплекса и без него. Показано, что трудозатраты на создание MAC в первом случае значительно ниже (рисунок 5).

В п. 4.3 производится оценка качества функционирования MAC при управлении потоками заданий. С этой целью проведен сравнительный анализ отказоустойчивости PBC при обработке потоков заданий с помощью полунатурного моделирования. Разработано приложение, генерирующее синтетический поток заданий с использованием формата Standard Workload Format (SWF) на основе вычислительной истории. Статистика вычислений собрана в процессе решения типовых практических задач исследования направлений развития топливо-энергетических комплексов России и Вьетнама с позиций обеспечения энергетической безопасности, а также оптимизации процессов складской логистики. Поток заданий был поочередно выполнен под управлением трех систем: метапланировщиков GridWay и CondorDAGMan, а также рассматриваемой в диссертации MAC с обучением агентов и без него. В процессе работы систем моделировались типовые отказы их компонентов и узлов разнородной PBC.

В рамках экспериментов суммарное число ядер изменялось от 320 до 352 единиц. Разные узлы характеризовались различной степенью их надежности. Рисунки 6 и 7 демонстрируют преимущество MAC с точки зрения отказов узлов и компонентов управляющих систем, что особенно очевидно при использовании MAC с обучением агентов.

При этом средняя загрузка процессора узлов экспериментальной PBC под управлением систем GridWay и Condor DAGMan составляет соответственно 81% и 77%. Применение MAC с обучением агентов и без него позволило повысить данный показатель до 87% и 82% соответственно.

Проведен сравнительный анализ эффективности передачи сообщений между агентами в различных типах сетей (локальная, городская и глобальная). Число агентов для каждого типа сети составило 10, 100 и 1000. Результаты проведенных экспериментов показали высокую надежность и эффективность процесса обмена сообщениями между агентами MAC.

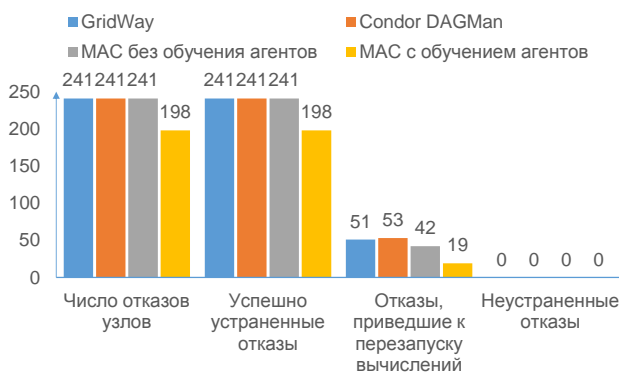


Рисунок 6 – Устранение отказов узлов

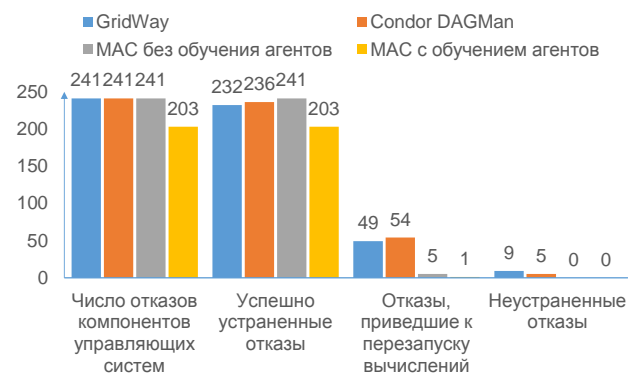


Рисунок 7 – Устранение отказов компонентов системы управления

Представлены результаты сравнения процессов функционирования разработанной MAC с системами GridWay и Condor DAGMan с точки зрения теории массового обслуживания. Разработано приложение, включающее 20 тестовых модулей. Модули моделируют работу прикладных программ рабочих процессов Montage, BLAST и Broadband. Приложение формирует синтетический поток заданий для PBS. Результаты показывают существенное улучшение следующих четырех показателей качества обслуживания очереди заданий при использовании MAC: числа заданий с нулевым временем ожидания (a), среднего времени пребывания задания в очереди (b), коэффициента полезного использования ресурсов PBS (c) и общего время выполнения заданий потока (d). Улучшение перечисленных выше четырех показателей обслуживания очередей заданий составляет от 1.01% до 21.34% по сравнению с системой Condor DAGMan и от 0.30% до 7.09% по сравнению с системой GridWay (рисунок 8).

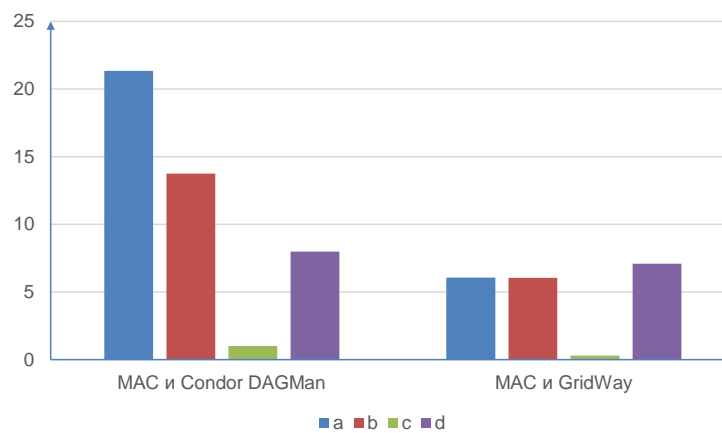


Рисунок 8 – Улучшение показателей качества обслуживания очереди заданий

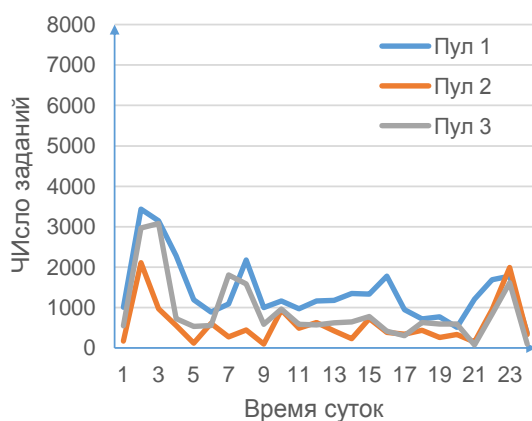


Рисунок 9 – Балансировка нагрузки с применением дополнительного уровня управления, реализуемого с помощью MAC

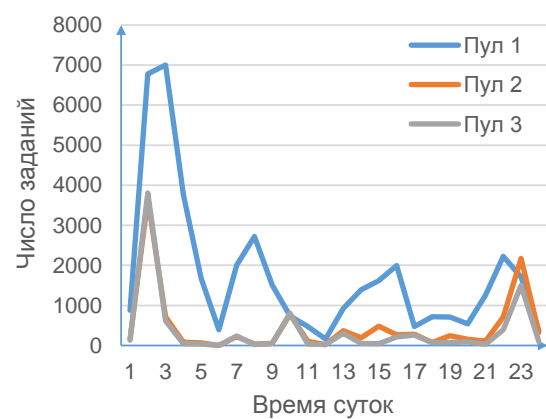


Рисунок 10 – Балансировка загрузки с помощью традиционных средств HTCondor и PBS Torque, используемых в пулах ресурсов PBS

Выполнена оценка балансировки загрузки ресурсов PBC. Показано, что применение MAC позволяет улучшить балансировку вычислительных ресурсов PBC (рисунок 9) по сравнению с балансировкой, обеспечиваемой традиционными системами HTCondor и PBS Torque (рисунок 10), используемыми в узлах PBC.

Разработан специализированный алгоритм определения длительности окон в расписании системы PBS Torque. Применение алгоритма продемонстрировано на примере решения задачи развития энергетического сектора Вьетнама для оптимизации энергетической безопасности на определенный период. Использование окон позволило ускорить процесс решения задачи на 16%.

Детальное представление процесса мультиагентного управления потоками заданий в PBC и его преимущества по сравнению с метапланировщиками GridWay и Condor DAGMan приведены в соответствующем приложении.

В **заключении** сформулированы основные выводы и результаты диссертационной работы, а также приведены возможные направления их развития.

Приложения содержат скан-копии акта об использовании и свидетельств о государственной регистрации программ для ЭВМ, результаты сравнительного анализа инструментариев для создания MAC, таблицы истинности для функций, определяющих три или более переходов из одного состояния, примеры развертывания агентной платформы JADE и подключения к ней агентов, детальное описание SWF-формата, а также пример обработки конкретного задания под управлением MAC и метапланировщиков GridWay и Condor DAGMan.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в журналах из перечня ВАК

1. Костромин, Р.О. Мультиагентный алгоритм построения остаточной схемы решения задачи в распределенных пакетах прикладных программ / А.Г. Феоктистов, Р.О. Костромин, И.А. Сидоров, С.А. Горский // Известия ЮФУ. Технические науки. — 2018. — № 8. — С. 59–69.

2. Костромин, Р.О. Управление заданиями в гетерогенной распределенной вычислительной среде на основе знаний / А.Г. Феоктистов, Р.О. Костромин, Ю.А. Дядькин // Вестник компьютерных и информационных технологий. — 2018. — № 2. — С. 10–17.

3. Kostromin, R. Virtualization of Heterogeneous HPC-clusters Based on OpenStack Platform / A. Feoktistov, I. Sidorov, R. Kostromin, V. Sergeev, V. Bogdanova // Вестник Южно-Уральского гос. ун-та. Сер. Вычислительная математика и информатика. — 2017. — Т. 6, № 2. — С. 37–48.

4. Костромин, Р.О. Разработка и применение предметно-ориентированных мультиагентных систем управления распределенными вычислениями / А.Г. Феоктистов, Р.О. Костромин // Известия ЮФУ. Технические науки. — 2016. — № 11. — С. 65–75.

5. Костромин, Р.О. Мультиагентный алгоритм перераспределения вычислительных ресурсов для остаточной схемы решения задачи в Grid / А.Г. Феоктистов, Р.О. Костромин // Современные наукоемкие технологии. — 2016. — № 9, ч. 2. — С. 244–248.

6. Костромин, Р.О. Модели, методы и средства управления вычислениями в интегрированной кластерной системе / Р.О. Костромин // Фундаментальные исследования. — 2015. — № 6, ч. 1. — С. 35–38.

Публикации в изданиях, индексируемых в Web of Science и Scopus

7. Kostromin, R. 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. — V. 150. — P. 171–178.

8. Kostromin, R. 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 Information Science. — 2019. — V. 979. — P. 265–279.

9. Kostromin, R. Agent Behavior Model for Distributed Computing Management in the Environment with Virtualized Resources / A. Feoktistov, R. Kostromin, A. Chernykh // Proceedings of the 41th International Convention on information and communication technology, electronics and microelectronics (MIPRO-2018). — Rieyka: IEEE, 2018. — P. 1153–1158.

10. Kostromin, R. 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 // Proceedings of the 41th International Convention on information and communication technology, electronics and microelectronics (MIPRO-2018). — Riejka: IEEE, 2018. — P. 256–261.

11. Kostromin, R. Machine Learning in a Multi-Agent System for Distributed Computing Management / I. Bychkov, A. Feoktistov, R. Kostromin, I. Sidorov, A. Edelev, S. Gorsky // Data Science Information Technology and Nanotechnology 2018: CEUR-WS Proc. — 2018. — V. 2212. — P. 89–97.

12. Kostromin, R. Multi-Agent Approach for Dynamic Elasticity of Virtual Machines Provisioning in Heterogeneous Distributed Computing Environment / A. Feoktistov, I. Sidorov, A. Tchernykh, V. Zorkalzev, S. Gorsky, R. Kostromin, I. Bychkov, A. Avetisyan // Proceedings of the International Conference on High Performance Computing and Simulation (HPCS-2018). — IEEE, 2018. — P. 909–916.

13. Kostromin, R. 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.

14. Kostromin, R. Knowledge Elicitation in Multi-Agent System for Distributed Computing Management / A. Feoktistov, A. Tchernykh, S. Gorsky, R. Kostromin // Proceedings of the 40th International Convention on information and communication technology, electronics and microelectronics (MIPRO-2017). — Riejka: IEEE, 2017. — P. 1350–1355.

Свидетельства о государственной регистрации программ для ЭВМ

15. Программа мониторинга очередей заданий в гетерогенной распределенной вычислительной среде: Свидетельство о государственной регистрации программы для ЭВМ от 10.04.2018 № 2018616092 / А.Г. Феоктистов, Р.О. Костромин. — М.: Федеральная служба по интеллектуальной собственности (РОСПАТЕНТ), 2018.

16. Библиотека алгоритмов для эффективного извлечения и применения проблемно-ориентированных знаний агентами: Свидетельство о государственной регистрации программы для ЭВМ от 11.12.2017 № 2017663706 / А.Г. Феоктистов, Р.О. Костромин. — М.: Федеральная служба по интеллектуальной собственности (РОСПАТЕНТ), 2017.

Публикации в других изданиях

17. Костромин, Р.О. Обзор инструментальных средств организации мультиагентных систем / А.Г. Феоктистов, Р.О. Костромин // Фундаментальные проблемы науки: Сб. статей Междунар. науч.-практ. конф.: в 2-х ч. — Уфа: АЭТЕРНА, 2016. — Ч. 1 — С. 68–72.

18. Костромин, Р.О. Обзор мультиагентных систем управления масштабируемыми приложениями / А.Г. Феоктистов, Р.О. Костромин // Фундаментальные проблемы науки: Сб. статей Междунар. науч.-практ. конф.: в 2-х ч. — Уфа: АЭТЕРНА, 2016. — Ч. 1 — С. 72–76.

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 26.11.2020
Формат бумаги 60 x 84 1/16, объем 1 п.л. Заказ № 3. Тираж 100 экз.

Отпечатано в ИДСТУ СО РАН