

Мультиагентная модель распределения ресурсов высокопроизводительной вычислительной среды в процессе обработки потоков заданий

А.Г.Феоктистов¹, Р.О. Костромин¹

¹Институт динамики систем и теории управления им. В.М. Матросова СО РАН, Лермонтова 134, Иркутск, Россия, 664033

Аннотация. В настоящее время задача эффективного распределения вычислительной мощности ресурсов высокопроизводительной среды в процессе решения задач конечных пользователей с целью улучшения показателей использования ресурсов (время, стоимость, надежность) по-прежнему остается актуальной. Она характеризуется высокой вычислительной сложностью. Поэтому на практике управление ресурсами зачастую осуществляется с использованием эвристических подходов, основанных на знаниях. Одним из таких подходов является применение мультиагентных технологий. В связи с этим в докладе представлена мультиагентная модель распределения ресурсов высокопроизводительной вычислительной среды при обработке потоков заданий, которые генерируются распределенными пакетами прикладных программ. Среда интегрирует ресурсы высокопроизводительных кластеров суперкомпьютерного центра коллективного доступа. В статье проведен сравнительный анализ эффективности управления вычислениями с помощью традиционных метапланировщиков (GridWay и Condor DAGMan) и предложенной мультиагентной системой. Экспериментальные результаты показывают преимущества управления мультиагентной системой по ускорению и эффективности вычислений, загрузке процессоров и ее балансировке.

1. Введение

В настоящее время проведение научных экспериментов, связанных с математическим моделированием сложных систем, как правило, обоснованно требует применения высокопроизводительных вычислений. Проблема обеспечения массового проведения подобных экспериментов актуализирует исследования, направленные на повышение степени открытости и доступности высокопроизводительных вычислительных систем с целью обеспечения удобного и гибкого доступа для специалистов различных категорий (математиков, разработчиков прикладного и системного программного обеспечения, администраторов и конечных пользователей информационно-вычислительных ресурсов и научных приложений) к системной программно-аппаратной инфраструктуре, данным и приложениям.

В рамках решения данной проблемы разрабатываются вычислительные системы на основе технологий облачных и грид-вычислений [1-3], создаются суперкомпьютерные центры коллективного пользования (ЦКП). Зачастую ресурсы таких систем и центров интегрируются в единую вычислительную среду (ВС). В ВС, как правило, возникает необходимость согласования критериев эффективности использования ресурсов и выполнения научных приложений на уровне ВС. Эти критерии зачастую являются противоречивыми. Данная многокритериальная задача является NP-трудной. Поэтому на практике, как правило,

используются различные эвристические подходы к ее решению, позволяющие находить некоторое субоптимальное распределение ресурсов в процессе обработки потоков заданий приложений.

Одним из таких подходов является применение мультиагентной технологии, в рамках которой агент представляет собой аппаратную или программную сущность, способную действовать в интересах достижения целей, поставленных перед ним субъектами виртуальной организации [4]. В этом случае, агент выступает в качестве посредника между человеком (владельцем ресурсов, их пользователем, разработчиком приложений, администратором среды или другим субъектом), другими программами и ВС. Таким образом, образуется мультиагентная система (МАС) – виртуальное сообщество агентов, представляющих различных субъектов, которые взаимодействуют в рамках ВС.

В статье рассматривается новая мультиагентная модель распределения ресурсов предметно-ориентированной ВС, организованной на базе высокопроизводительных ресурсов ЦКП «Иркутский суперкомпьютерный центр СО РАН» (ИСКЦ) [5], в процессе обработки потоков заданий научных приложений специального класса – распределенных пакетов прикладных программ (РППП).

2. Предметно-ориентированная вычислительная среда

Под предметно-ориентированной ВС понимается совокупность программно-аппаратных средств, позволяющих ее конечным пользователям решать определенный класс задач для некоторой предметной области. Такая ВС должна обеспечивать ее конечным пользователям гибкие языковые возможности и программные средства построения модели предметной области, формулировки задач и их решения с использованием необходимых ресурсов. Каждый такой ресурс ВС, например, вычислительный кластер, работает под управлением системы управления прохождением заданий (СУПЗ) и обслуживает дополнительный поток заданий локальных пользователей данного ресурса. Менеджеры ресурсов (метапланировщики), функционирующие на уровне ВС обслуживают поток заданий от глобальных пользователей среды.

Задание представляет собой спецификацию процесса решения задачи, содержащую информацию о требуемых ресурсах, исполняемых прикладных программах (модулях), входных и выходных структурах данных, а также другие необходимые сведения. В общем случае задание может состоять из набора взаимосвязанных подзаданий. Множество заданий приложения образуют поток заданий, который в дальнейшем может сливаться с потоками заданий других приложений, образуя новые потоки. В статье в качестве приложений рассматриваются РППП, разрабатываемые с помощью инструментального комплекса Orlando Tools [6]. Под РППП понимается комплекс взаимосвязанных прикладных программ и средств системного обеспечения (программных и языковых), предназначенный для автоматизации решения определенного класса задач в ВС.

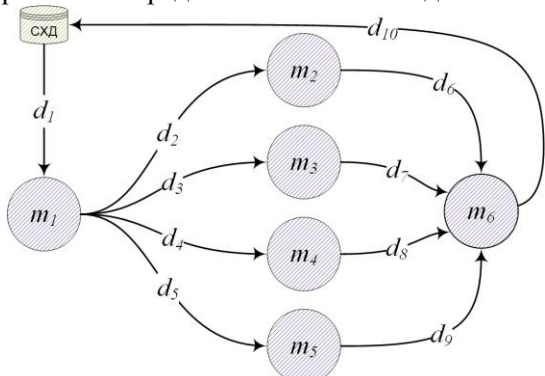


Рисунок 1. Схема решения задачи.

Имя характеристики	Комментарии	Тип	Область допустимых значений
Число ядер на узле	Число вычислительных элементов (ядер) на каждом узле	integer	[1; 36]
Число узлов	Число вычислительных устройств	integer	[1; 20]
Имя приложения	Имя программного приложения, выполняемого в задании	string	any value
Число виртуальных машин	Число виртуальных машин, которые будут запущены на одном узле	integer	[1; 36]
Тип гостевой ОС	Тип операционной системы для виртуальных машин	string	1-256
Промежуточное ПО для виртуальных машин	Используемый менеджер ресурсов	string	1-256

Рисунок 2. Характеристики заданий системы классификации.

Схема решения задачи в РППП коррелирует с понятием рабочего процесса (workflow) и представляется в виде ациклического направленного графа – direct acyclic graph (DAG), который может быть представлен в ярусно-параллельной форме. На рисунке 1 показан пример схемы решения задачи, где $m_1 - m_6$ – это модули схемы, а $d_1 - d_9$ – данные, передаваемые между модулями. Задания и их подзадания классифицируются в соответствии с указанными в них требованиями по выполнению модулей. Примеры характеристик классов заданий системы классификации, используемой в ЦКП ИСКЦ, приведены на рисунке 2. Аспекты ее работы детально рассмотрены в [7].

В рамках данного исследования, ориентированного на проведение сравнительного анализа эффективности обработки потоков заданий различными метапланировщиками, потоки заданий специфицируются в формате Standard Workload Format (SWF) [8]. Он позволяет достаточно точно отражать характеристики исполняемых модулей и требования к ресурсам в заданиях. Описание потока заданий в формате SWF разделено на две основные части: описание моделируемой ВС и спецификация потока заданий, выполняемого в ней. SWF-формат включает набор predetermined характеристик, каждая из которых имеет свой уникальный порядковый номер. Характеристики, используемые в первой части SWF-формата, могут иметь как числовые, так и текстовые значения. Характеристики, используемые во второй части SWF-формата, могут иметь только числовые значения.

Синтетический поток заданий в SWF-формате сгенерирован на основе вычислительной истории, накопленной в процессе решения ряда научных и практических задач. В их числе задачи исследования направлений развития топливно-энергетических комплексов России [9] и Вьетнама [10] с позиций обеспечения энергетической безопасности, а также оптимизации процессов складской логистики [6].

Под синтетическим потоком заданий понимается искусственно-сгенерированный поток, где в качестве приложений в заданиях используются программы-имитаторы, выполняющие в ВС реальную загрузку ресурсов и обмен заданными объемами данных. Данный поток был поочередно выполнен под управлением трех систем: метапланировщиков GridWay [11] и Condor DAGMan [12], а также рассматриваемой в статье MAC. Характеристики заданий синтетического потока (время поступления, число ядер, размер оперативной и дисковой памяти, время выполнения и другие сведения) соответствуют характеристикам заданий реального потока.

Задания пользователей ВС поступают через ее общий шлюз (выделенный узел), на котором также установлены управляющие компоненты GridWay, Condor DAGMan и MAC. Перечисленные метапланировщики осуществляют управление потоками заданий на уровне ВС, распределяя их между ресурсами среды, которые работают под управлением локальных СУПЗ. Функционирование метапланировщика определяется его конфигурационными параметрами, определяющими число заданий для обработки за одну итерацию планирования, периодичность процессов планирования, поиска свободных ресурсов, обновления сведений о ресурсах, обновления сведений о статусах заданий и других операций.

3. Сравнительный анализ работы метапланировщиков

Распределение потока заданий в MAC выполняется на основе тендера вычислительных работ [7]. Тендер проводится в несколько раундов при поступлении задания в ВС. В каждом раунде распределяются модули очередного уровня ярусно-параллельной схемы решения задачи. Эффективность работы тендера базируется на реализации следующих решений:

- предварительной классификации заданий, обеспечивающей формирование виртуального сообщества агентов (участников тендера), каждый из которых удостоверен в наличии и заинтересован в выполнении соответствующих его ресурсам подзаданий вычислительной работы;
- разбиения схемы решения задачи, представленной ярусно-параллельной формой, на подсхемы, каждая из которых включает один ярус схемы, и проведения тендера для каждого яруса в отдельности, что позволяет агентам учитывать в процессе торгов для

подзаданий некоторого яруса результаты распределения ресурсов для выполнения подзаданий предыдущих ярусов;

- использования параллельных операций в вычислительной модели Orlando Tools, обеспечивающих доленое распределение вычислительной нагрузки.

В таблице 1 представлены характеристики процесса обработки заданий сравниваемыми метапланировщиками GridWay, Condor DAGMan и MAC. Все метапланировщики выполняют подзадания в асинхронном режиме по готовности данных. Однако MAC в отличие от GridWay и Condor DAGMan сразу ставит подзадания в локальные очереди, поддерживает непосредственную пересылку данных между агентами, распределяет подзадания по ресурсам с учетом зависимости по данным в DAG и прогнозирования времени выполнения подзаданий.

Преимущества распределения подзаданий одного задания по ограниченным ресурсам ВС под управлением MAC показано на примере схемы, показанной ранее на рисунке 1. На рисунке 3 схематично представлен процесс обработки заданий в MAC, где каждый прямоугольник представляет модуль подзадания, поступающий из общей очереди ВС в локальную очередь СУПЗ. Модули передаются на выполнение на ресурсы по мере их освобождения. Фиолетовой штриховкой обозначены модули задания для выполнения вышеупомянутой схемы решения задачи. Красной штриховкой обозначены прочие задания потока. На рисунке 3 ресурсы (кластеры) условно представлены прямоугольниками K1, K2, K3 и K4. Соответствующие процессы обработки заданий в GridWay и Condor DAGMan отражены на рисунках 4 и 5. На разных этапах обработки задания на рисунках 3-5 приводятся время t_0 начала обработки задания, время t_q нахождения подзаданий в общей очереди, время t_d передачи данных и время t_k завершения выполнения задания. Результаты выполнения задания, приведенные в таблице 2, демонстрируют очевидное преимущество MAC.

Таблица 1. Характеристики процесса обработки заданий метапланировщиками.

Параметр	GridWay	Condor DAGMan	MAC
Постановка подзаданий в локальную очередь	В порядке общей очереди по готовности данных	В порядке общей очереди по готовности данных	При поступлении задания
Передача данных между подзаданиями	Через локальное хранилище; через общую систему хранения данных	Через локальное хранилище или общую систему хранения данных	Непосредственная пересылка данных между агентами
Выполнение подзаданий в асинхронном режиме по готовности данных	Да	Да	Да
Учет зависимости по данным в DAG	Нет	Нет	Да
Прогнозирование времени выполнения подзаданий	Нет	Нет	На основе вычислительной истории или результатов профилирования модулей

Таблица 2. Показатели выполнения задания.

Параметр	t_0	t_q	t_d	t_k	Время выполнения задания
GridWay	15:00:00	00:07:29	00:01:18	15:12:49	12 мин. 49 сек.
Condor DAGMan	14:00:00	00:07:29	00:01:17	14:12:21	12 мин. 21 сек.
MAC	17:00:00	00:00:16	00:00:10	17:09:26	9 мин. 26 сек.

Обработка синтетического потока заданий под управлением MAC позволило ускорить процесс вычислений на 35.8 и 30.92% по сравнению с GridWay и Condor DAGMan

соответственно. Применение MAC позволило также повысить среднюю загрузку процессора узлов ВС на 6 и 5% относительно GridWay и Condor DAGMan.

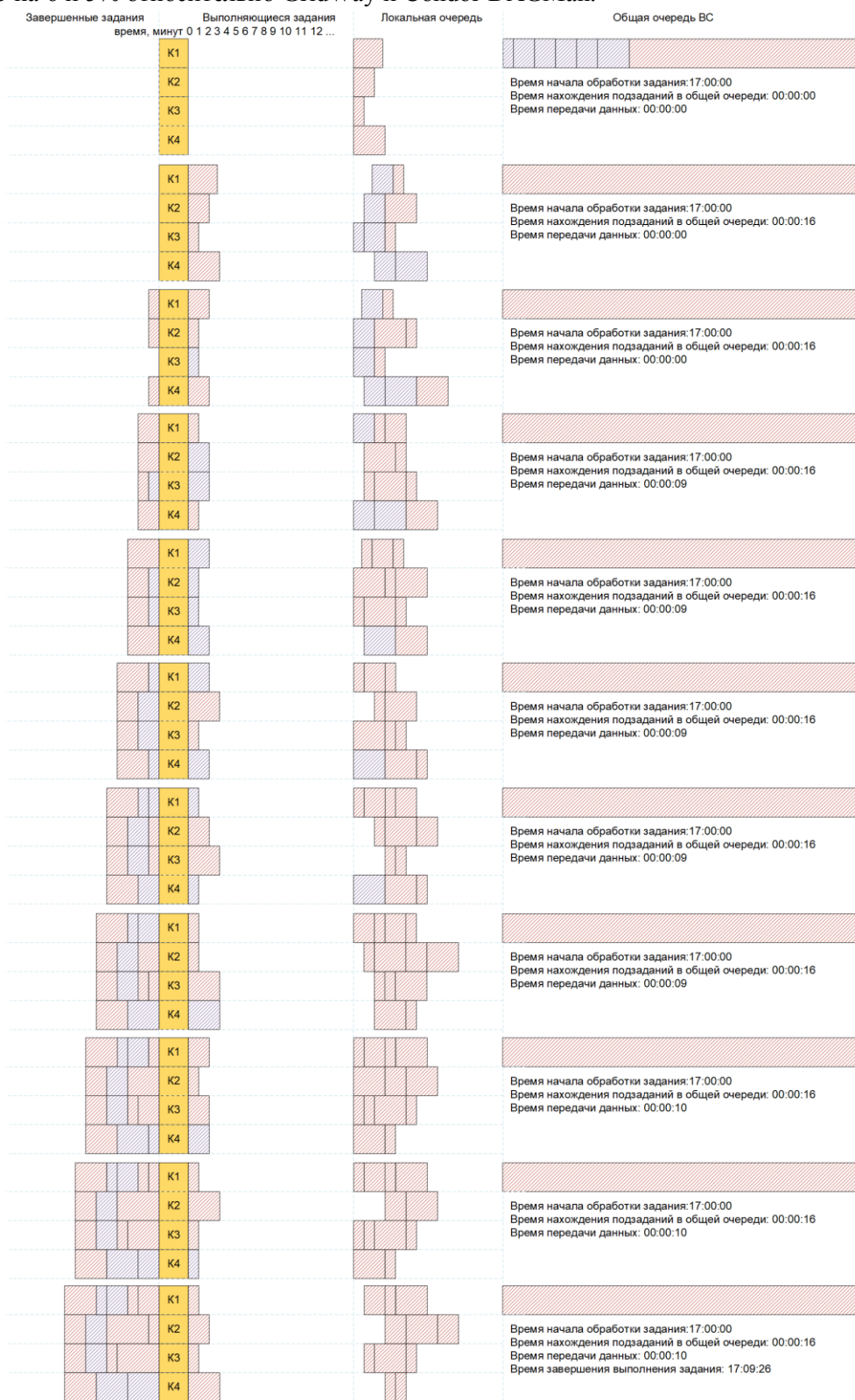


Рисунок 3. Обработка очереди заданий в MAC.

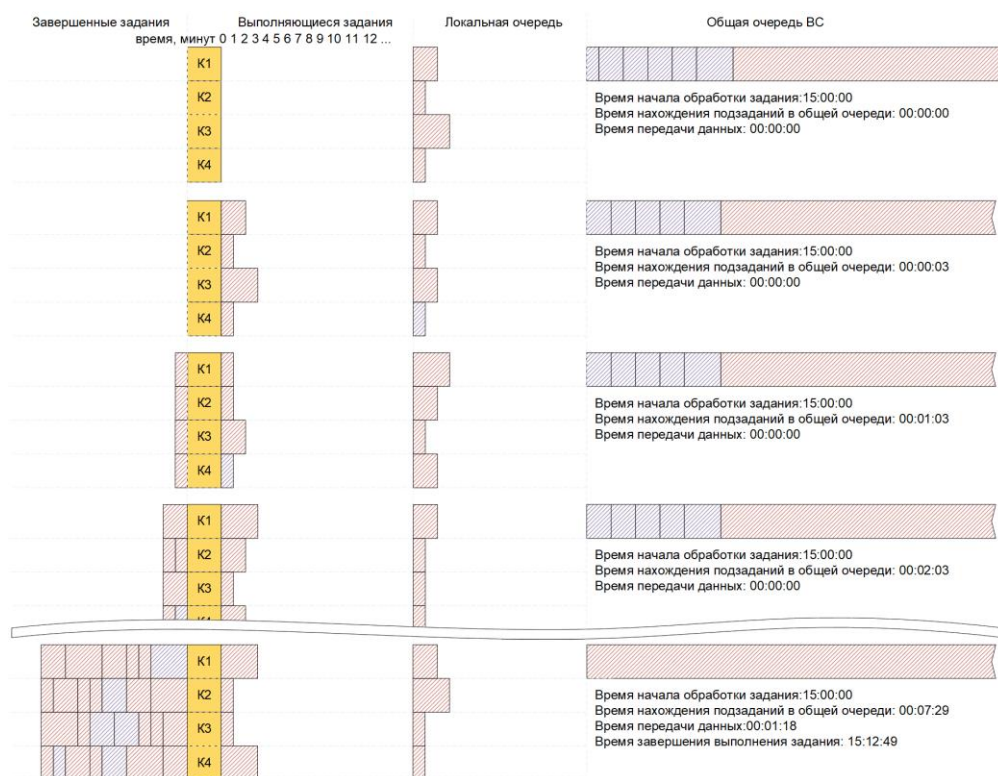


Рисунок 4. Обработка очереди заданий в GridWay.

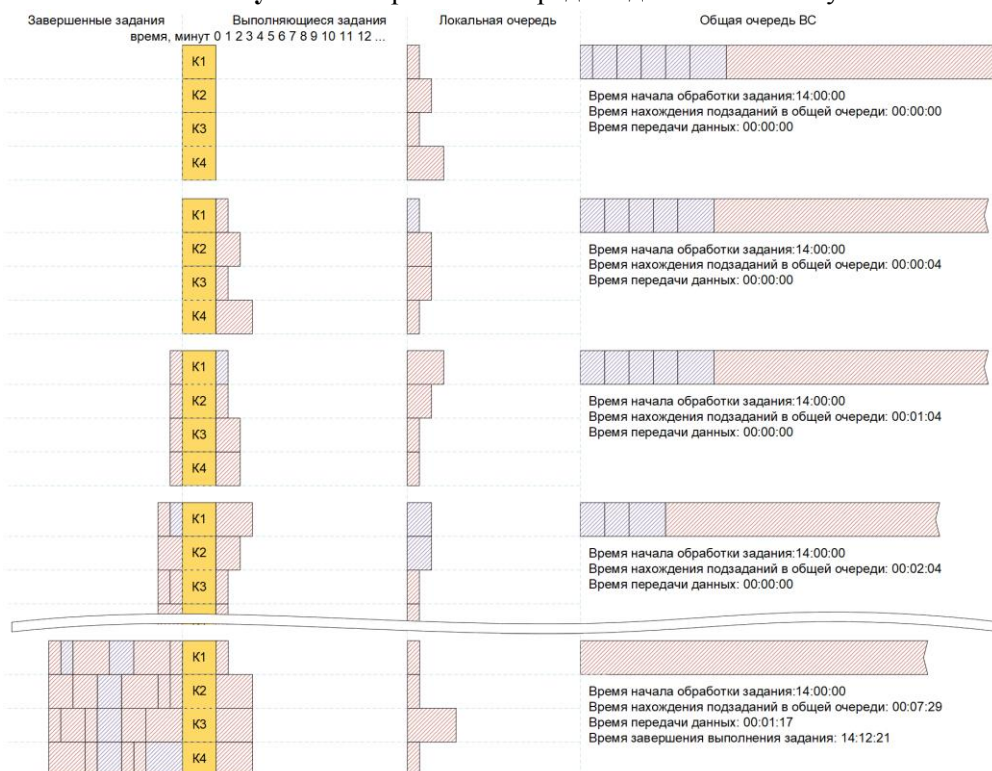


Рисунок 5. Обработка очереди заданий в Condor DAGMan.

4. Заключение

В статье представлен сравнительный анализ эффективности обработки синтетического потока заданий с помощью MAC и традиционных метапланировщиков GridWay и Condor DAGMan. Показано, что в условиях распределения ограниченных ресурсов вычислительной среды

мультиагентное управление, основанное на применении рыночных механизмов регулирования спроса и предложения ресурсов (тендера вычислительных работ), демонстрирует очевидное преимущество.

5. Благодарности

Исследование выполнено в рамках проекта IV.38.1.1 программы фундаментальных исследований СО РАН, а также поддержано РФФИ, проект № 19-07-00097-а.

6. Литература

- [1] Krasovec, B. Enhancing the Grid with Cloud Computing / B. Krasovec, A. Filipcic // *Journal of Grid Computing*. – 2019. – Vol. 17. – P. 119-135. DOI: 10.1007/s10723-018-09472-w.
- [2] Dakkak, O. From grids to clouds: Recap on challenges and solutions / O. Dakkak, S.A. Nor, M.S. Sajat, Y. Fazea, S. Arif // *AIP Conference Proceedings*. – 2016. – P. 020-040. DOI: 10.1063/1.5055442.
- [3] El-Sayed, H. Edge of Things: The Big Picture on the Integration of Edge, IoT and the Cloud in a Distributed Computing Environment / H. El-Sayed, S. Sankar, M. Prasad, D. Puthal, A. Gupta, M. Mohanty, C.-T. Lin // *IEEE Access*. – 2018. – Vol. 6. – P. 1706-1717. DOI: 10.1109/ACCESS.2017.2780087.
- [4] Wooldridge, M. Intelligent agents: Theory and practice / M. Wooldridge, N.R. Jennings // *The knowledge engineering review*. – 1995. – Vol. 10(2). – P. 115-152.
- [5] ЦКП Иркутский суперкомпьютерный центр СО РАН [Электронный ресурс] – Режим доступа: <http://hpc.icc.ru/> (08.12.2019).
- [6] Tchernykh, 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 Information Science*. – 2019. – Vol. 979. – P. 265-279. DOI: 10.1007/978-3-030-16205-4_20.
- [7] Бычков, И.В. Обучение агентов на основе параметрической настройки их алгоритмов управления распределенными вычислениями / И.В. Бычков, А.Г. Феоктистов, И.А. Сидоров, А.В. Еделев, С.А. Горский, Р.О. Костромин // *Сборник трудов IV Международной конференции и молодежной школы “Информационные технологии и нанотехнологии” (ИТНТ) – Самара: Новая техника, 2018. – С. 2237-2247.*
- [8] Feitelson, D.G. Experience with using the parallel workloads archive / D.G. Feitelson, D. Tsafir, D. Krakov // *Journal of Parallel and Distributed Computing*. – 2014. – Vol. 74(10). – P. 2967-2982. DOI: 10.1016/j.jpdc.2014.06.013.
- [9] Еделев, А.В. 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 связи, управления и безопасности*. – 2018. – № 3. – С. 197-231.
- [10] Feoktistov, A. Multi-Agent Approach for Dynamic Elasticity of Virtual Machines Provisioning in Heterogeneous Distributed Computing Environment / A. Feoktistov, I. Sidorov, A. Tchernykh, A. Edelev, V. Zorkalzev, S. Gorsky, R. Kostromin, I. Bychkov, A. Avetisyan // *Proceedings of the International Conference on High Performance Computing and Simulation*. – IEEE, 2018. – P. 909-916. DOI: 10.1109/HPCS.2018.00145.
- [11] GridWay Metascheduler [Электронный ресурс]. – Режим доступа: <http://www.gridway.org> (08.12.2019).
- [12] Condor DAGMan [Электронный ресурс]. – Режим доступа: <https://research.cs.wisc.edu/htcondor/dagman/dagman.html> (08.12.2019).

A multi-agent model of allocating resources of a high-performance computing environment in the processing of job flows

A.G. Feoktistov¹, R.O. Kostromin¹

¹Matrosov Institute for System Dynamics and Control Theory SB RAS, Lermontov street 134, Irkutsk, Russia, 664033

Abstract. Nowadays, the problem of effective allocating resources of a high-performance computing environment to achieve their efficient use and improve criteria (time, cost, reliability, etc.) of solving end-user problems is still very relevant. It is characterized by high computational complexity. Therefore, in practice, resource allocation is often implemented using knowledge-based heuristic approaches. One such approach is applying a multi-agent technology. In this regard, we represent a multi-agent model of allocating resources of a high-performance computing environment in the processing of job flows that are generated by the distributed applied software packages. The environment integrates the resources of high-performance clusters of the public access supercomputer center. We provide a comparative analysis of effectiveness for computation management with traditional meta-schedulers (GridWay and Condor DAGMan) and our multi-agent system. Experimental results show advantages of management under the multi-agent system in speedup and efficiency of computations, processor load, and its balancing.