

УТВЕРЖДАЮ

Генеральный директор АО «ИТМиВТ
им. С.А. Лебедева РАН»,

доктор физико-математических наук,
профессор



А.В. Князев

ноября 2025 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию Баранова Антона Викторовича

«Методы и средства управления вычислительными ресурсами в суперкомпьютерных системах коллективного пользования», представленной на соискание ученой степени доктора технических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Актуальность темы исследования

Создание инфраструктуры и условий для проведения научных исследований и разработок, внедрения наукоемких технологий, отвечающих современным принципам организации научной, научно-технической и инновационной деятельности является одной из актуальных задач научно-технологического развития России, вынесенной на государственный уровень. Неотъемлемой частью научно-исследовательской инфраструктуры страны являются суперкомпьютерные центры коллективного пользования (СКЦ).

Высокопроизводительные вычислительные системы (суперкомпьютеры), эксплуатируемые в СКЦ, представляют собой сложные аппаратно-программные комплексы малой серийности и высокой стоимости с иерархической организацией вычислений. На каждом уровне иерархии необходимо эффективно распределять вычислительную работу по составляющим этот уровень вычислительным ресурсам. При этом система управления суперкомпьютерными ресурсами должна принимать во внимание разнообразие решаемых пользователями вычислительных задач и различие пользовательских требований к объему и составу необходимых для решения этих задач вычислительных ресурсов.

Повышение эффективности использования вычислительных ресурсов суперкомпьютерных систем на каждом уровне иерархии в условиях обработки потоков пользовательских запросов с различными требованиями является актуальной научно-технической проблемой, решению которой посвящена диссертационная работа А.В. Баранова.

Структура и содержание диссертации

Во введении обоснована актуальность работы. Представлены основные этапы обработки информации в суперкомпьютерной системе коллективного пользования в виде технологического тракта обработки потоков пользовательских заданий как запросов на предоставление для проведения расчетов подсистемы вычислительных узлов суперкомпьютера. В технологическом тракте выделены наиболее важные этапы, требующие для обеспечения эффективного использования вычислительных ресурсов новых научно обоснованных технических и технологических решений. К этим этапам соискатель отнес планирование заданий, распределение вычислительных процессов задания по узлам суперкомпьютера и организацию параллельных вычислений с распараллеливанием по данным в рамках иерархически организованной подсистемы узлов суперкомпьютера. Предлагаемые для каждого из перечисленных этапов решения представлены как единый комплекс в виде системы управления заданиями (СУЗ). Отмечена необходимость сохранения и развития научно-технических компетенций в области построения отечественных систем управления заданиями. Сформулированы цель и задачи исследования, представлены основные положения диссертационной работы, выносимые на защиту, раскрыта их научная новизна. Кроме этого, во введении приведены сведения о примененных методах исследования, теоретической и практической значимости результатов диссертации и их апробации.

В первой главе рассмотрена типовая архитектура суперкомпьютерной системы коллективного пользования как совокупность автономных вычислительных узлов (ВУ), объединенных коммуникационной средой. Введено понятие «задание» как пользовательский запрос на доступ к суперкомпьютерным ресурсам, включающий прикладную программу с входными данными и требования к ресурсам. Рассмотрены жизненный цикл задания и основные задачи управления вычислительными ресурсами суперкомпьютерной системы, решаемые системами управления заданиями.

В главе сформулированы основные требования к СУЗ. Среди требований выделена универсальность, как отсутствие ограничений на выполняемые пользователями прикладные программы и возможность функционирования СУЗ на произвольной вычислительной системе, соответствующей типовой архитектуре суперкомпьютера. С универсальностью связаны требования надежности и императивности управления, как технической невозможности совершения пользователем неразрешенных в системе действий, и модульности архитектуры СУЗ, обеспечивающей переносимость СУЗ между различными суперкомпьютерами типовой архитектуры.

Основное внимание в первой главе сфокусировано на пятиуровневой иерархической модели управления вычислительными ресурсами

суперкомпьютерной системы коллективного пользования и предложенной соискателем на основе этой модели архитектуры Системы управления прохождением параллельных заданий (СУППЗ). Показано соответствие предложенной архитектуры ранее выдвинутым требованиям к системам управления заданиями, а также развитие архитектуры СУППЗ, начиная от варианта, разработанного для массивно-параллельных суперкомпьютеров серии МВС-1000 (1999 г.).

Во второй главе представлены технические и технологические решения, предложенные автором для управления вычислительными ресурсами суперкомпьютерных систем коллективного пользования на каждом уровне иерархической модели. Показана реализация предложенных решений в составе СУППЗ.

Представлен сравнительный анализ качественных характеристик СУППЗ и ведущих мировых систем управления заданиями SLURM, PBS, OpenLava, Sun Grid Engine и IBM Spectrum LSF. Результаты анализа показывают, что СУППЗ обладает подавляющим большинством качественных характеристик ведущих систем управления заданиями, и это позволяет говорить о соответствии СУППЗ мировому уровню развития таких систем. Кроме этого, в главе представлены результаты сравнения СУППЗ и ведущей системы SLURM по ряду количественных показателей: загрузке вычислительных ресурсов, среднему времени ожидания задания в очереди, среднему коэффициенту замедления заданий, среднему значению времени нахождения задания в очереди относительно заказанного времени выполнения и относительной доли своевременно обработанных заданий. Результаты сравнения свидетельствуют о примерном паритете по количественным показателям между СУППЗ и SLURM.

В завершении главы рассмотрены итоговые результаты эксплуатации СУППЗ на множестве суперкомпьютерных систем Межведомственного суперкомпьютерного центра РАН (МСЦ РАН, ныне – отделения МСЦ Курчатовского высокопроизводительного вычислительного комплекса НИЦ «Курчатовский институт»), Института прикладной математики им. М.В. Келдыша РАН (ИПМ им. М.В. Келдыша РАН) и других научных и образовательных организаций. Рассмотрены состав и характеристики суперкомпьютерных систем под управлением СУППЗ, отмечено разнообразие архитектур этих систем, построенных в том числе на базе универсальных процессоров, ВУ с графическими ускорителями, ВУ с ускорителями линейки Intel Xeon Phi обоих поколений. Общее число исследователей, воспользовавшихся услугами СУППЗ, составляет более 1240 человек из 135 научных и образовательных организаций, а число реализованных ими исследовательских проектов превышает 530. Эти цифры свидетельствуют о высокой практической значимости результатов диссертации и востребованности этих результатов у широкого круга пользователей. Соискатель справедливо делает вывод о возможности

формировании на основе СУППЗ цифровой экосистемы научного суперкомпьютерного центра коллективного пользования.

В третьей главе рассмотрены предложенные соискателем методы планирования заданий в суперкомпьютерных системах коллективного пользования.

Метод планирования заданий с фиксированными параметрами основан на разделении входного потока на классы ординарных (пакетных), отладочных и фоновых заданий. Одна из идей метода состоит в выделении плавающего резерва ВУ суперкомпьютера для заданий небольшого размера, запускаемых, как правило, с целью отладки. В отсутствие отладочных заданий выделенные под резерв узлы могут быть выделены заданиям других классов, что положительно сказывается на загрузке вычислительных ресурсов. Второй оригинальной идеей является разделение обслуживания длительных по времени заданий, названных в диссертации фоновыми, на отдельные временные отрезки – кванты. В промежутках между квантами фоновое задание может быть снято с выполнения и возвращено в очередь.

Реализованный в составе планировщика СУППЗ метод планирования применяется на суперкомпьютерах, установленных в МСЦ РАН и ИПМ им. М.В. Келдыша РАН. В главе приведен анализ статистики работы этих суперкомпьютеров, результаты которого позволяют утверждать, что средний коэффициент замедления отладочных и фоновых заданий в подавляющем большинстве случаев кратно ниже, чем общий коэффициент замедления. Предложенные механизмы фоновых заданий и плавающего резерва ВУ для отладочных заданий положительно сказываются на загрузке вычислительных ресурсов. Кроме этого, доля своевременно обработанных заданий выше для класса отладочных заданий.

Вторым предложенным методом является метод постпланирования адаптивных (эластичных) заданий, которые могут быть выполнены на параллельном ресурсе произвольного размера в некотором заданном диапазоне. За счет адаптивных заданий предложенный метод позволяет заполнять окна простаивающих ВУ в расписании запусков заданий. Такие окна неизбежно образуются из-за разных размеров планируемых заданий. В соответствии с методом постпланирования в обнаруженное в расписании запусков заданий основного потока окно помещается очередное адаптивное задание. Время выполнения и число узлов адаптивного задания подбираются в соответствии с размером обнаруженного окна, что за счет принципа обратного заполнения влечёт практически немедленный запуск адаптивного задания основным планировщиком.

Как показало проведенное соискателем имитационное моделирование, метод постпланирования позволяет повысить загрузку суперкомпьютера для входных потоков любой интенсивности при незначительном ухудшении показателей среднего времени ожидания заданий в очереди и среднего коэффициента замедления. Результаты исследования позволили определить

границу применимости метода постпланирования: если загрузка вычислителя менее 85%, за счет потока адаптивных заданий метод позволит максимизировать загрузку без существенного замедления заданий в очереди.

В главе также представлены решения для обработки входного потока пользовательских заданий, в котором совмещаются стандартные и нестандартные задания. Последние требуют для своего выполнения отдельной программной платформы, с отдельными операционной системой, прикладным и инструментальным программным обеспечением. Соискателем предложено два метода совмещения потоков стандартных и нестандартных заданий. Первый метод заключается в представлении системы управления заданиями в виде гипервизора виртуального многоядерного вычислительного узла с большим объемом оперативной памяти. Второй метод основан на использовании облачной платформы, как инструмента автоматического развертывания при старте нестандартного задания виртуальных машин и контейнеров на вычислительных узлах суперкомпьютера. Проведённые эксперименты показали, что применённые при реализации метода средства виртуализации не оказывают значительного влияния на производительность стандартных заданий, что позволяет успешно совмещать обработку стандартных и нестандартных заданий в одном входном потоке.

В четвертой главе рассмотрены методы и алгоритмы отображения параллельной программы на вычислительные узлы суперкомпьютера, преследующие цель распределить по узлам процессы расчетной программы из состава задания таким образом, чтобы минимизировать время выполнения этой программы. Эта задача сведена в диссертации к поиску отображения информационного графа параллельной программы на граф свободных ВУ суперкомпьютера путем минимизации значения целевой функции отображения. Решение задачи требуется обеспечить в системе управления заданиями суперкомпьютера в условиях обработки потока пользовательских заданий. Поскольку состав графа свободных ВУ динамичен, заранее неизвестно, какие конкретно узлы будут свободны при поступлении очередного задания. При этом, несмотря на экспоненциальную сложность задачи, отображение должно строиться за приемлемое время, не превышающее значения системных таймаутов.

Соискателем в главе предложены двухэтапный метод и алгоритмы решения сформулированной задачи отображения. На первом этапе метода решается задача выделения для очередного задания подсистемы ВУ из множества свободных на момент запуска задания узлов. Эта задача сведена к разрезанию графа свободных ВУ на минимально связанные между собой подграфы, для чего предложен оригинальный эвристический алгоритм, основанный на имитации отжига. На втором этапе производится поиск оптимального отображения программного графа на граф выделенных заданию ВУ. Для этого используется предложенный соискателем параллельный эвристический алгоритм, комбинирующий методы имитации

отжига и генетического отбора. Для выполнения параллельного алгоритма предлагается использовать выделенные для задания вычислительные узлы.

Глава также содержит результаты экспериментальных исследований характеристик предложенных алгоритмов, демонстрирующие рост быстродействия расчетов как для тестовых программ, так и для реальных приложений. Циклическое повторение фаз отжига и генетического отбора в параллельном алгоритме второго этапа отображения позволило достичь более высокой точности по сравнению с известными алгоритмами при времени поиска отображения, сравнимом с системными таймаутами.

В пятой главе исследованы вопросы организации параллельных вычислений с распараллеливанием по данным, в которых одна и та же последовательность вычислений (прикладной алгоритм) выполняется над каждым элементом множества (пула) входных данных. В главе вводится модель одной последовательной программы (ОПП), реализующей прикладной алгоритм. Порция входных данных для ОПП определяется значениями одного или нескольких параметров. Из таких порций вычислительной работы складывается пул входных данных - множество всех возможных значений параметров ОПП во всех их комбинациях. Организация параллельных вычислений заключается в осуществлении запусков экземпляров ОПП на множестве доступных вычислительных узлов и распределении порций входных данных запущенным экземплярам.

Соискателем предложен метод рекурсивного разделения данных, основанный на организации вычислительного процесса в виде иерархической системы менеджеров, каждый из которых контролирует либо группу менеджеров нижестоящего уровня иерархии, либо запускаемые на вычислительном узле экземпляры ОПП. Метод подразумевает деление на каждом уровне иерархии входного пула данных на порции заданного размера и асинхронное распределение этих порций между менеджерами нижестоящего уровня иерархии в качестве входных пулов. Пулы данных при этом представляются виде упорядоченного набора нумерованных слайсов – элементарных порций данных. Каждая порция (пул) данных при таком представлении однозначно определяется номерами первого и последнего слайсов порции. При разделении данных на каждом уровне иерархии фиксируется компактная контрольная точка в виде текущих выделенных порций и остатка входного пула данных, что позволяет обойтись без специализированной базы учета обработанных порций данных и одновременно свести к минимуму накладные расходы на распараллеливание. За счет асинхронного распределения и механизма перераспределения порций данных обеспечиваются отказоустойчивость вычислений и возможность применения для решения одной задачи вычислительных средств разной производительности и архитектуры.

Метод иерархического деления данных был реализован в составе программного комплекса (ПК) распараллеливания по данным «Пирамида». В

главе приведены результаты экспериментального сравнения ПК «Пирамида» с другими известными программными решениями. Результаты показали, что ПК «Пирамида» вносит существенно меньшие накладные расходы при распараллеливании, основанном на переборе комбинаций параметров, и уступает ПК X-SOM разработки НИВЦ МГУ им. М.В. Ломоносова только при обработке пула входных данных, представленного в виде набора строк в файле. Для устранения этого недостатка соискателем была предложена и реализована архитектура гибридного ПК XP-SOM, совмещающего достоинства программных комплексов X-SOM и «Пирамида». Приведенные в главе результаты экспериментов свидетельствуют, что ПК XP-SOM опережает по производительности другие известные из открытых источников решения.

В заключении обобщены основные результаты и выводы диссертации. Приведены сведения о внедрении результатов диссертации в практическую деятельность ряда научных организаций и в вузовский учебный процесс подготовки специалистов.

Таким образом, можно заключить, что **научно-техническая новизна диссертационной работы** состоит в следующем.

1. Разработаны новые методы планирования прохождения заданий, совмещающие обработку классов ординарных, отладочных, фоновых и адаптивных заданий, в том числе представленных в виде виртуальных машин и контейнеров.

2. Разработан новый двухэтапный метод отображения параллельной программы на вычислительные узлы суперкомпьютера: на первом этапе производится выделение вычислительных узлов для параллельной программы и тем самым сокращается размер задачи отображения, на втором этапе выделенные узлы используются для выполнения параллельного алгоритма поиска отображения.

3. Разработан новый эвристический алгоритм выделения вычислительных узлов для задания, основанный на разрезании графа свободных вычислительных узлов на два минимально связанных друг с другом подграфа при помощи имитации отжига.

4. Разработан новый параллельный алгоритм поиска отображения программного графа на граф вычислительных узлов с использованием циклической смены фаз имитации отжига и генетического отбора.

5. Разработан новый метод иерархического деления данных для организации параллельных вычислений с распараллеливанием по данным, за счет разделения входного пула данных на наборы упорядоченных элементарных вычислительных работ обеспечивающий компактное представление состояния вычислений и низкие накладные расходы на распараллеливание.

6. Разработана новая архитектура системы управления заданиями пользователей, основанная на иерархической модели управления

вычислительными ресурсами суперкомпьютерной системы коллективного пользования.

Достоверность результатов и выводов, представленных в диссертации, подтверждается следующими факторами:

- реализованный в составе Системы управления прохождением параллельных заданий (СУППЗ) комплекс технических и технологических решений нашел успешное применение в наиболее мощных российских суперкомпьютерах, установленных в 1999-2024 гг. в МСЦ РАН, ИПМ им. М.В. Келдыша РАН и ряде других научных и образовательных организаций;

- статистические данные о работе суперкомпьютеров под управлением СУППЗ, а также результаты экспериментов и имитационного моделирования свидетельствуют об эффективности предложенных соискателем решений;

- основные результаты работы неоднократно докладывались и обсуждались на международных и всероссийских научных конференциях;

- по теме диссертации опубликовано 45 печатных работ статей, из них 27 работ опубликовано в рецензируемых изданиях, рекомендованных ВАК.

Рекомендации по использованию результатов работы

Результаты исследований, представленные в диссертации, могут заинтересовать ученых, занимающихся исследованиями и разработками систем управления заданиями суперкомпьютеров, а также вопросами организации параллельных вычислений с распараллеливанием по данным, создания и развития суперкомпьютерных центров коллективного пользования.

Соответствие диссертации и автореферата требованиям «Положения о порядке присуждения ученых степеней». Диссертация А.В. Баранова обладает внутренним единством. Автореферат диссертации адекватно отражает ее содержание. Диссертация и автореферат удовлетворяют требованиям «Положения о присуждении ученых степеней».

По тексту диссертации есть следующие замечания.

1. В работе недостаточно уделено внимания исследованиям развития суперкомпьютерных центров коллективного пользования как цифровых экосистем. Соискатель в выводах второй главы справедливо утверждает, что основу такой цифровой экосистемы могут составлять системы управления заданиями, но при этом:

- не определяет круг пользователей и состав сервисов позволяющим им удовлетворять свои потребности в рамках реализации единого бесшовного бизнес-процесса;
- не приводит никаких предложений по протоколам защиты персональных данных;
- нет предложений по накоплению и выработке знаний;

- не приводит возможных путей достижения некоторых других характеристик эко-систем IT-предприятий.

2. При описании специализированной СУППЗ «Пирамида» использование в качестве множества обрабатываемых данных декартова произведения является искусственным ограничением и преодолевалось ранее различными авторами путем реализации специальных процедур генерации исходных данных.

3. Среди количественных характеристик СУППЗ, особенно специализированных, хотелось бы видеть характеристики эффективности использования технического ресурса, реальных затрат «Инвестора».

Заключение

Отмеченные замечания не меняют общего положительного впечатления о диссертационной работе. В ней изложены востребованные практикой решения, позволяющие эффективно использовать вычислительные ресурсы суперкомпьютерных систем в режиме коллективного пользования.

Диссертация А.В. Баранова обладает внутренним единством, представляет завершенное исследование на актуальную тему и является научно-квалификационной работой, в которой на основании выполненных автором исследований предложены новые научно обоснованные архитектурные, технические и технологические решения, внедрение которых вносит значительный вклад в развитие научных суперкомпьютерных центров коллективного пользования, как неотъемлемой части исследовательской инфраструктуры страны. Автореферат диссертации адекватно отражает ее содержание.

Диссертационная работа А.В. Баранова удовлетворяет требованиям, предъявляемым к диссертациям по специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей», а ее автор заслуживает присуждения ему ученой степени доктора технических наук по этой специальности.

Баранов А.В не был и не является сотрудником АО «ИТМиВТ»

Отзыв составлен доктором физико-математических наук, профессором А.Ф Ронжиным.

Отзыв обсужден и утвержден на заседании объединенного научно-технического семинара АО «ИТМиВТ» 05.11.2025 протокол № 5.

Главный научный сотрудник группы анализа перспективных направлений исследований и центра инженерного анализа и специальных разработок, доктор физико-математических наук, профессор

Ронжин Александр Фёдорович

07 ноября 2025 г.

Сведения об организации:

Акционерное общество «Институт точной механики и вычислительной техники имени С.А. Лебедева Российской академии наук» (АО «ИТМиВТ»)
119991, Российская Федерация, Москва, Ленинский проспект, д. 51;
info@ipmse.ru; +7 (495) 649-12-70