

## Отзыв на автореферат диссертации Климентова Алексея Анатольевича

“Методы обработки сверхбольших объемов данных в распределенной гетерогенной компьютерной среде для приложений в ядерной физике и физике высоких энергий”, представленной на соискание ученой степени доктора физико-математических наук по специальности 05.13.11 — математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Установки класса мегасайнс во многом определяют направление научных исследований в области физики высоких энергий и в ядерной физике. В первую очередь это большой адронный коллайдер (БАК) в европейском центре ЦЕРН (Женева, Швейцария), Европейский рентгеновский лазер на свободных электронах XFEL (DESY, Гамбург, Германия), релятивистский тяжелоионный коллайдер (RHIC, США), в скором будущем коллайдер НИКА (ОИЯИ, Дубна, Россия). Российские научные группы играют ведущую роль в научной программе и в создании (усовершенствовании) детекторов на этих установках. За годы работы большого адронного коллайдера получены сотни петабайт экспериментальных данных. Объемы данных экспериментов на БАК будут расти и скоро достигнут эксабайтной отметки. Информация должна быть доступна для всех членов международных научных сообществ (а это тысячи ученых), информационный поиск ведется по многочисленным хранилищам данных и мета-информации, одновременно идет обмен и обсуждение результатов между университетами и институтами, входящими в международные сотрудничества, для проверки результатов исследований, научных подходов и гипотез. Все это накладывает новые требования к системам обработки, анализа и моделирования данных, к компьютерной модели (модели обработки данных) современных физических экспериментов, и во многом влияет на развитие информационных технологий в других информационно-научных областях науки.

Для управления и обработки больших массивов данных необходимы многоуровневые интеллектуальные системы, системы управления потоками данных и заданий, системы контроля и мониторинга, а также системы хранения информации.

Диссертация А.А. Климентова посвящена рассмотрению и решению вопросов разработки компьютерной модели и архитектуры распределенных вычислительных систем для обработки сверхбольших объемов данных, интеграции высокопроизводительных вычислительных систем и высокоскоростных вычислительных сетей для обработки физических данных. В диссертации рассмотрены основополагающие принципы и модели таких комплексов применительно к области физики высоких энергий и ядерной физики.

Разработанная автором и представленная в диссертации масштабируемая система для обработки, анализа и моделирования данных современного физического эксперимента позволяет объединить в единую вычислительную среду разнородные вычислительные мощности, такие как суперкомпьютеры, ресурсы облачных вычислений, грид и университетские кластеры. Такая возможность была впервые реализована группой А.А.Климентова на основе представленных в диссертации методов при сопровождении экспериментов на БАК. В результате, тысячи пользователей получили возможность обрабатывать и анализировать эксабайтные объёмы данных, при этом любые инфраструктурные изменения вычислительных ресурсов, на которых выполняются задачи анализа и обработки данных, «скрыты» от пользователя, и решение любых системных задач не влияет на прикладные инструменты обработки данных эксперимента. Автором диссертации предложены, разработаны и внедрены комплекс методик и методов, на основе которых создана система для управления потоками заданий. Созданная система обеспечивает высокую эффективность обработки данных физических экспериментов в масштабе более 2 миллионов задач в день в 200 вычислительных центрах для 1000 пользователей.

Научные результаты представленной диссертационной работы вносят значительный вклад в возможности анализа и обработки сверхбольших объёмов данных в будущие десятилетия на этапе работы БАК и коллайдера NICA с высокой светимостью. В частности, в 2026-2038 годах ожидаемое увеличение светимости БАК приведет к повышению объёмов данных на порядок. Новая компьютерная модель, предложенная А.А. Климентовым, способна стать решением этой сложной для экспериментов в области физики частиц задачи. Разработанная система, использующая гетерогенные вычислительные мощности, с применением динамического управления ресурсами, на сегодняшний день не имеет аналогов в мире и является не просто желаемым улучшением, а необходимым требованием современных научных исследований. Вычислительные средства такого масштаба и облачные платформы могут стать реальной альтернативой очень большим центрам обработки данных, принадлежавшим и управляемым научным сообществом.

Многие результаты работы являются пионерскими (в частности использование суперкомпьютеров для приложений ФВЭ и ЯФ и их интеграция с системой грид) и используются в ведущих экспериментах, таких как ATLAS, ALICE, COMPASS. В диссертации показано применение разработанной системы для анализа ДНК мамонта на суперкомпьютере НИЦ КИ и в проекте по биоинформатике BlueBrain. Также

разработанная система и компьютерная модель могут быть использованы для новых комплексов, таких как FAIR, NICA, EIC, и в других проектах класса мегасайенс - XFEL, LSST.

По теме диссертации опубликовано более 150 печатных работ, по основным результатам более 50 работ. Научная новизна и практическая значимость, полученные в результате более чем 20-летней научной деятельности автора, не подлежат сомнению. Климентов А.А. является ведущим специалистом в области разработки распределенных систем обработки данных, разработке и созданию пакетов программ в области физики высоких энергий, методов анализа данных. Диссертационная работа отвечает требованиям ВАК РФ и является законченным научным трудом по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и сетей. Считаю, что автор заслуживает присуждения ученой степени доктора физико-математических наук по указанной специальности.

Зам. Директора ИПМ им. М.В. Келдыша РАН  
(125047, Москва, Миусская пл., д.4)  
д.ф.-м.н., чл.-корр. РАН, профессор  
(тел. +7-499-220-78-23, e-mail: lira@imamod.ru)



Якобовский Михаил Владимирович

27.02.2018 г.