
ПРОДВИЖЕНИЕ ТЕХНОЛОГИЙ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ

УДК 004

А. В. БУХАНОВСКИЙ, С. В. МАРЬИН, К. В. КНЯЗЬКОВ, А. А. СИДНЕВ, С. Н. ЖАБИН,
А. П. БАГЛИЙ, Р. Б. ШТЕЙНБЕРГ, А. В. ШАМАКИНА, В. В. ВОЕВОДИН,
Е. Н. ГОЛОВЧЕНКО, Р. Т. ФАЛАЛЕЕВ, А. В. ДУХАНОВ, А. А. ТАРАСОВ,
Л. В. ШАМАРДИН, А. И. МОИСЕЕНКО

РЕЗУЛЬТАТЫ РЕАЛИЗАЦИИ ПРОЕКТА „МОБИЛЬНОСТЬ МОЛОДЫХ УЧЕНЫХ“ В 2010 ГОДУ: РАЗВИТИЕ ФУНКЦИОНАЛЬНЫХ ЭЛЕМЕНТОВ ТЕХНОЛОГИИ iPSE И РАСШИРЕНИЕ СОСТАВА ПРИКЛАДНЫХ СЕРВИСОВ

Представлены результаты коллективных исследований в области применения и развития технологий проблемно-ориентированных распределенных инструментальных сред нового поколения для компьютерного моделирования сложных систем. Исследования выполнены в рамках мероприятия „Мобильность молодых ученых“ ФЦП „Кадры инновационной России на 2009—2013 гг.“ в 2010 г.

Ключевые слова: распределенные вычисления, компьютерное моделирование, сложная система.

Введение. Современный этап эволюции представлений о программных комплексах компьютерного моделирования и обработки данных научных экспериментов тесно связан с продвижением парадигмы „электронной науки“ (eScience). Понятие eScience ассоциируется с проблемой организации разнородными группами специалистов совместных научных исследований, требующих консолидации вычислительных и программных ресурсов для решения сложных междисциплинарных задач на основе технологий распределенных вычислений и систем. Инструментарий компьютерного моделирования, ориентированного на персональные компьютеры и вычислительные кластеры монопольного применения, в настоящее время претерпевает существенные изменения, связанные, в первую очередь, с глобализацией сферы его использования, ограниченной охватом сети Интернет. Это касается особенностей архитектуры программных комплексов, специфики их развертывания на вычислительных платформах, характеристик информационного обеспечения, поддержки жизненного цикла и продвижения на рынке [1].

Принципиальным аспектом развития и эффективного использования прикладных программных комплексов в рамках парадигмы eScience является введение единой технологической *проблемно-ориентированной* платформы, адаптируемой под специфику различных *предметных* областей. Характерным примером такой платформы является уникальная для Российской Федерации разработка — технология HPC-NASIS, заложенная в основу высокопроизводительного программного комплекса нового поколения для квантово-механических расчетов и моделирования наноразмерных структур и комплексов [2].

В настоящей статье рассматриваются результаты исследований, выполненных на основе технологии HPC-NASIS в рамках мероприятия „Мобильность молодых ученых“, проект „Развитие технологии проблемно-ориентированных распределенных инструментальных сред нового поколения для компьютерного моделирования сложных систем на основе высокопроизводительного программного комплекса HPC-NASIS“ в 2010 г. в рамках ФЦП „Кадры инновационной России на 2009—2013 гг.“.

Технология iPSE как основа мультидисциплинарных распределенных исследований. В отличие от большинства пакетов программ для моделирования сложных систем, комплекс HPC-NASIS основан на перспективном подходе к организации высокопроизводительного программного обеспечения сложных систем — технологии iPSE. Технология iPSE (Intelligent Problem Solving Environment) [3] строится на развитии традиционных проблемно-ориентированных оболочек и определяет принципы построения композитных приложений для eScience как интеллектуальной оболочки управления параллельными вычислительными процессами в распределенной иерархической среде, включающей в себя системы различной архитектуры. Такой подход позволяет расширить возможности традиционных проблемно-ориентированных оболочек за счет обеспечения эффективного параллельного исполнения композитных приложений, в силу того что использует для управления параллельными вычислениями симбиотические знания об особенностях предметной области и специфике процесса.

— В рамках iPSE формализуются не только методы и вычислительные алгоритмы, но и экспертные знания об изучении данного явления средствами компьютерного моделирования. Другими словами, iPSE реализует функции интеллектуальной системы поддержки принятия решений исследователя, что важно для практического внедрения такого комплекса.

— iPSE предоставляет единый интерфейс взаимодействия для предметно-ориентированных программных модулей и компонентов, которые могут разрабатываться различными коллективами и иметь разные условия распространения и использования.

— iPSE ориентирована на поддержку высокопроизводительных вычислений, причем не только для суперкомпьютерных систем с традиционной (кластерной) архитектурой, но и для неоднородных систем, например, гиперкластеров (суперкомпьютеров, объединенных высокоскоростным каналом). При этом управление эффективностью выполнения сценария является прерогативой iPSE, что позволяет избежать конфликтных ситуаций при разделении ресурсов между различными вычислительными модулями и разными пользователями.

Оболочка iPSE может быть отнесена к классу интеллектуальных систем, ориентированных на учет и формализацию неопределенности (информационная сложность), обладающих распределенной структурой (структурная сложность) и многоцелевым характером преобразования информации (функциональная сложность).

Концепция создания систем класса iPSE базируется на принципах, определяющих архитектуру системы и уровни ее управления. Наряду с традиционными математическими методами построения расписаний здесь широко применяются достижения искусственного интеллекта, определяющие принципы функционирования систем, основанных на знаниях. Важная роль принадлежит принципу открытости, позволяющему обеспечить наиболее сложные уровни иерархической структуры системы — самоорганизацию и самообучение, благодаря чему возможна интерпретация сложных процессов и явлений при анализе альтернатив и принятии решений, а также при моделировании действий системы в процессе решения задачи и „обучении“ на своем опыте.

Фактически программный комплекс на основе iPSE представляет собой интеллектуальную интегрирующую оболочку, которая позволяет встраивать, комбинировать и эффективно запускать предметно-ориентированные модули, решающие задачи, разработанные разными группами специалистов и функционирующие (как сервисы) на различных вычислительных системах в рамках

модели метакомпьютинга. Это обеспечивает открытость архитектуры комплекса и коммунальный характер его использования как ASP-приложения (Application Service Provider).

Развитие функциональных компонентов технологии iPSE. Программный комплекс HPC-NASIS представляет собой открытую информационную систему, динамично развиваемую усилиями разработчиков и виртуального сообщества пользователей [4]. Отдельные сервисы могут исполняться как на отдельных вычислительных узлах (серверах приложений или даже суперкомпьютерах), так и в распределенных вычислительных средах. Архитектура комплекса позволяет включать в его состав сервисы на основе систем следующих архитектур: SMP, MPP (кластеры), GPGPU (и кластеры на их основе), FPGA, а также СВЕА. Такие возможности в целом определяют пути дальнейшего развития комплекса и разработки его новых функциональных элементов в рамках концепции iPSE. В частности, элементы могут быть ориентированы на:

— обеспечение механизмов, позволяющих сторонним разработчикам удобно, эффективно и безопасно (как с технологической, так и с правовой точки зрения) интегрировать предметно-ориентированные модули в комплекс HPC-NASIS. Это включает выбор и адаптацию автоматизированного тестирования и отладки программных интерфейсов, средств для оптимизации структур хранилищ данных, а также средств виртуализации;

— развитие дополнительных системных инструментов в рамках технологии HPC-NASIS, повышающих надежность, реактивность, защищенность и эргономичность программных комплексов на ее основе.

В рамках исследований по проекту „Развитие технологии проблемно-ориентированных распределенных инструментальных сред нового поколения для компьютерного моделирования сложных систем на основе высокопроизводительного программного комплекса HPC-NASIS“ был спроектирован и реализован ряд экспериментальных образцов дополнительных программных компонентов HPC-NASIS, а также на основе результатов исследований показана их применимость и полезность.

— На основе анализа и развития программного компонента управления вычислительными платформами предложены и обоснованы принципы усовершенствования технологии мониторинга в составе HPC-NASIS, позволяющие расширить число контролируемых характеристик доступных платформ исполнения, реализовать более гибкий механизм предоставления актуальных данных о состоянии платформ, расширить возможности технической поддержки и частично автоматизировать ее, а также увеличить число доступных для интеграции стандартных средств мониторинга и управления платформами.

— Обоснована целесообразность тестирования проблемно-ориентированных модулей в HPC-NASIS по методу „белого ящика“ и „черного ящика“ (с варьированием входных данных сложной структуры), реализован набор соответствующих инструментов. Это позволит применять автоматическое тестирование программных сервисов при включении новых проблемно-ориентированных модулей в комплекс HPC-NASIS, тем самым расширяя его функциональность и круг решаемых с его помощью проблем.

— Предложены механизмы разработки и модификации высокопроизводительных программных сервисов, применяемые при организации работы с памятью ЭВМ для повышения производительности приложений. В частности, такие подходы могут быть применены при разработке новых сервисов в составе программного комплекса HPC-NASIS и модификации существующих.

— Предложены и обоснованы подходы и методы интеграции проблемно-ориентированных модулей в комплекс HPC-NASIS, которые основаны на парадигме Грид 2.0 и подходе REST для организации сервисной инфраструктуры. Это позволит расширять и модифицировать сервисную инфраструктуру HPC-NASIS с целью снижения накладных расходов на передачу данных (без использования протокола SOAP) и повышения прозрачности и простоты удаленных вызовов.

— Продемонстрирована возможность применения технологий облачных вычислений согласно нескольким сценариям (использование существующих облачных решений, предоставление SaaS или PaaS услуг). Сформулированы рекомендации для развития подсистем комплекса HPC-NASIS в соответствии с этими сценариями.

— Разработан механизм интеграции вычислительной инфраструктуры Национальной нанотехнологической сети (ННС) Грид в состав комплекса HPC-NASIS. Предложенное решение обеспечивает эффективный доступ к высокопроизводительным сервисам в рамках Грид-инфраструктуры, который отличается низкими накладными расходами, вносимыми управляющим ядром программного комплекса.

Проведенные исследования показали нетипичное распределение времени собственных накладных расходов среды ГридННС при испытании интегрирующего решения. На рис. 1 приведены ядерные оценки распределения: *a* — собственного времени исполнения пакета ORCA и *b* — общего времени исполнения композитного приложения в ГридННС (*1* — бензол, *2* — нафталин, *3* — пирен).

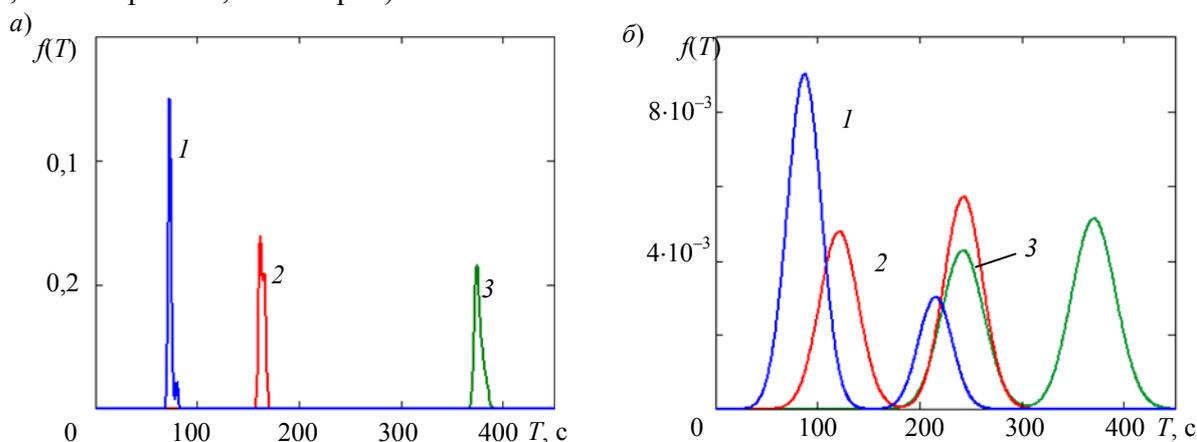


Рис. 1

На графике распределения времени исполнения в ГридННС хорошо заметны два отдельных пика, связанных с возникновением циклических процессов обработки очередей заданий в Грид. При этом накладные расходы сопоставимы с собственным временем исполнения пакета.

Расширение состава прикладных сервисов HPC-NASIS. Отдельным направлением исследований и разработок является адаптация технологии HPC-NASIS для решения новых классов задач как в рамках заданной предметной области (моделирование наноразмерных атомно-молекулярных систем и комплексов), так и в других предметных областях (пилотные решения). В частности, в рамках данного проекта разработан экспериментальный образец молекулярного редактора в форме сервиса в комплексе HPC-NASIS. Основная задача сервиса заключается в предоставлении возможностей визуализации и редактировании молекул и их систем. Молекулярный редактор в составе HPC-NASIS делает подготовку данных для расчетов более удобной, а также улучшает восприятие решения самой задачи расчета молекулярной структуры.

На рис. 2 представлена молекулярная поверхность большой молекулы белка, изображение построено с помощью разработанного для HPC-NASIS молекулярного редактора.

— Обоснованы возможности комплекса HPC-NASIS в области интеграции пакетов инженерного проектирования и анализа, разработано программное решение, которое позволяет строить вычислительные цепочки из нескольких сервисов, реализующих процесс исследования наноматериалов и сред с переходом от микро- к макроуровням. Для этого в состав HPC-NASIS встроен сервис, обеспечивающий доступ к возможностям свободно распространяемого пакета OpenFOAM.

— Разработан и реализован в качестве прикладного сервиса HPC-NASIS сопоставимый с лучшими мировыми аналогами параллельный алгоритм геометрической декомпозиции тетраэдральных сеток. Преимущества технологии HPC-NASIS для решения данной задачи заключаются в возможности составления экспериментальной базы разбиений различных сеток для различных вычислительных задач с использованием экспертной системы, помогающей в настройке параметров под конкретную задачу. Интеграция в комплекс HPC-NASIS различных пакетов решения вычислительных задач наравне с пакетом параллельной декомпозиции сеток позволит создавать сценарии цепочек запуска различных пакетов для решения комплексных ресурсоемких задач.

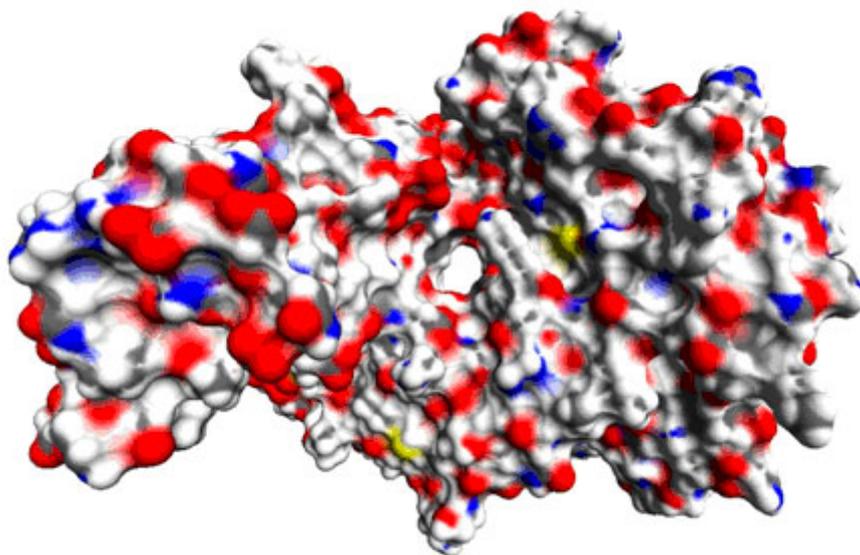


Рис. 2

— Разработан и программно реализован в качестве прикладного сервиса HPC-NASIS параллельный алгоритм оптимизации многошаговых управляющих воздействий в наносистемах. В качестве примера расчетов была выбрана задача распределения ресурсов в целях оптимизации эффективности нанотехнологического процесса, которая показала хорошую масштабируемость алгоритма.

— Рассмотрен вариант решения проблемы поиска сферических кодов с использованием ресурсов комплекса HPC-NASIS и математической среды MathCloud, выявлена возможность интеграции указанных сред в рамках решения поставленной задачи, обеспечивающей унифицированный доступ к широкому кругу доступных вычислительных платформ.

Заключение. В рамках проекта „Развитие технологии проблемно-ориентированных распределенных инструментальных сред нового поколения для компьютерного моделирования сложных систем на основе высокопроизводительного программного комплекса HPC-NASIS“ в 2010 г. работы по развитию технологии HPC-NASIS выполнялись с привлечением двенадцати молодых приглашенных специалистов (преподавателей, аспирантов и научных сотрудников). Так, исследования в области механизмов, методов и технологий адаптации сторонних программных модулей выполнены специалистами ЮУрГУ и МФТИ и ЮФУ. Работы по развитию дополнительных системных инструментов проведены специалистами ННГУ, ЮФУ и НИВЦ МГУ. Разработки в области адаптации технологии HPC-NASIS к решению новых классов прикладных задач связаны с деятельностью специалистов ИММ РАН, НИВЦ МГУ, ВГУ и ИСА РАН.

Результаты проведенных исследований в целом продемонстрировали масштабируемость и востребованность технологии HPC-NASIS для создания проблемно-ориентированных платформ компьютерного моделирования в различных предметных областях; при этом

научно-технический уровень достигнутых результатов соответствует современным мировым образцам программного инструментария eScience.

Работа выполнена при поддержке ФЦП „Научные и научно-педагогические кадры инновационной России на 2009—2013 гг.“. Результаты проекта использованы в рамках выполнения государственного контракта № 16.647.12.2025 „Создание функционирующего в режиме удаленного доступа интерактивного учебно-методического комплекса для выполнения работ в области моделирования наноразмерных атомно-молекулярных структур, наноматериалов, процессов и устройств на их основе, в распределенной вычислительной среде“.

СПИСОК ЛИТЕРАТУРЫ

1. Бухановский А. В., Васильев В. Н. Современные программные комплексы компьютерного моделирования e-Science // Изв. вузов. Приборостроение. 2010. Т. 53, № 3. С. 60—64.
2. Ядро высокопроизводительного программного комплекса для квантово-механических расчетов и моделирования наноразмерных атомно-молекулярных систем и комплексов „HPC-NASIS“. Свид-во о гос. регистрации программы для ЭВМ № 20010610161 от 11.01.2010 г.
3. Бухановский А. В., Ковальчук С. В., Марьин С. В. Интеллектуальные высокопроизводительные программные комплексы моделирования сложных систем: концепция, архитектура и примеры реализации // Изв. вузов. Приборостроение. 2009. Т. 52, № 10. С. 5—24.
4. Гуськов А. А., Дунаев А. В., Подтелкин Ф. В., Бухановский А. В. Интеллектуальные технологии поддержки пользователей в области компьютерного моделирования наноразмерных структур и наноматериалов // Науч.-техн. вестн. СПбГУ ИТМО. 2010. Вып. 66. С. 120.

Сведения об авторах

- Александр Валерьевич Бухановский** — д-р техн. наук, профессор; НИИ Научно-технических компьютерных технологий Санкт-Петербургского государственного университета информационных технологий, механики и оптики; директор; E-mail: avb_mail@mail.ru
- Сергей Владимирович Марьин** — канд. техн. наук; НИИ Научно-технических компьютерных технологий Санкт-Петербургского государственного университета информационных технологий, механики и оптики; младший научный сотрудник; E-mail: sm.niinkt@gmail.com
- Константин Валерьевич Князьков** — НИИ Научно-технических компьютерных технологий Санкт-Петербургского государственного университета информационных технологий, механики и оптики; младший научный сотрудник; E-mail: constantinvk@gmail.com
- Алексей Александрович Сиднев** — Нижегородский государственный университет им. Н. И. Лобачевского, кафедра математического обеспечения ЭВМ; ассистент; E-mail: sidnev.a.a@gmail.com
- Сергей Николаевич Жабин** — Московский физико-технический институт, кафедра общей физики; ассистент; E-mail: delafrog@gmail.com
- Антон Павлович Баглий** — аспирант; Южный федеральный университет, факультет механики, математики и компьютерных наук, Ростов-на-Дону; E-mail: taccessviolation@yandex.ru
- Роман Борисович Штейнберг** — Южный федеральный университет, кафедра алгебры и дискретной математики, Ростов-на-Дону; научный сотрудник; E-mail: romanofficial@yandex.ru
- Анастасия Валерьевна Шамакина** — Южно-Уральский государственный университет, отдел распределенных вычислений Лаборатории суперкомпьютерного моделирования, Челябинск; начальник; E-mail: sham2004@bk.ru
- Вадим Владимирович Воеводин** — НИВЦ МГУ им. М. В. Ломоносова, лаборатория параллельных информационных технологий, Москва; младший научный сотрудник; E-mail: vadim@parallel.ru
- Евдокия Николаевна Головченко** — Институт прикладной математики им. М. В. Келдыша РАН, отдел № 16, Москва; младший научный сотрудник; E-mail: ge03@imamod.ru

- Роман Тимофеевич Фалалеев* — Московский физико-технический институт, кафедра информатики; зав. кафедрой; E-mail: roman@phystech.edu
- Алексей Валентинович Духанов* — канд. техн. наук; НИИ Научно-технических компьютерных технологий Санкт-Петербургского государственного университета информационных технологий, механики и оптики; E-mail: avd_mail@mail.ru
- Алексей Александрович Тарасов* — канд. физ.-мат. наук; Московский физико-технический институт, кафедра распределительных вычислений; доцент; E-mail: tarasov@isa.ru
- Лев Витальевич Шамардин* — Научно-исследовательский институт ядерной физики им. Д. В. Скобельцына, Москва; старший научный сотрудник; E-mail: shamardin@theory.sinp.msu.ru
- Алексей Игоревич Моисеенко* — ЗАО „Фирма АйТи. Информационные технологии“, Департамент технологий информационного менеджмента, Москва; директор; E-mail: AMoiseenko@it.ru

Рекомендована НИИ НКТ

Поступила в редакцию
15.05.11 г.