ИПМ им.М.В.Келдыша РАН

Абрау-2016 • Труды конференции

В.М. Волохов, Д.А. Варламов,
А.В. Волохов, А.И. Прохоров

**Веб-ориентированные
вычислительные сервисы как
компоненты среды компьютерного
моделирования нанокompозитных
материалов**

Рекомендуемая форма библиографической ссылки

Волохов В.М., Варламов Д.А., Волохов А.В., Прохоров А.И. Веб-ориентированные вычислительные сервисы как компоненты среды компьютерного моделирования нанокompозитных материалов // Научный сервис в сети Интернет: труды XVIII Всероссийской научной конференции (19-24 сентября 2016 г., г. Новороссийск). — М.: ИПМ им. М.В.Келдыша, 2016. — С. 104-110. — doi:[10.20948/abrau-2016-30](https://doi.org/10.20948/abrau-2016-30)

Размещена также [презентация к докладу](#)

Веб-ориентированные вычислительные сервисы как компоненты среды компьютерного моделирования нанокompозитных материалов

В.М. Волохов¹, Д.А. Варламов^{1,2}, А.В. Волохов¹, А.И. Прохоров¹

1 Институт проблем химической физики РАН, г. Черноголовка, РФ

2 Институт экспериментальной минералогии РАН, г. Черноголовка, РФ

Аннотация. В ИПХФ РАН на основе высокопроизводительных ресурсов и комплекса программных и инфраструктурных средств создана и развивается веб-ориентированная вычислительная среда, предназначенная для решения ресурсоемких задач по компьютерному моделированию структур и свойств наноматериалов в интересах широкого круга исследователей. Проблемно-ориентированные высокоуровневые веб-интерфейсы к вычислительным сервисам позволяют решить многие проблемы проведения массовых ресурсоемких вычислительных экспериментов в области моделирования нанокompозитов и предоставить облегченный доступ к суперкомпьютерным ресурсам.

Ключевые слова: высокопроизводительные вычисления, нанокompозитные материалы, вычислительная среда, веб-интерфейсы, компьютерное моделирование

Современное состояние вычислительной химии (прежде всего, задач в области квантово-химического и молекулярного моделирования нанокompозитов и биологических структур) требует практически обязательного использования высокопроизводительных вычислительных ресурсов (параллельных, гибридных, распределенных) для решения задач различных классов и различных уровней требуемой вычислительной мощности. Ряд примеров подобных задач был приведен в нашей предыдущей работе [1].

Опыт работ авторов в области вычислительной химии [2-3], в том числе при моделировании нанокompозитов и наноструктурированных объектов [4-6] показал, что, помимо использования суперкомпьютерных центров класса «Ломоносов-1,2» или МВС-10000, весьма перспективным для целей массового применения компьютерного моделирования вещества является создание проблемно-ориентированных ресурсов среднего размера, направленных на решение определенных классов задач и при этом обладающих интуитивными, дружелюбными конечному пользователю (часто неискушенному в суперкомпьютерных расчетах) веб-интерфейсами, что позволяет во многом решить следующие проблемы:

- объединить в рамках единого интерфейса все стадии моделирования (нано)объектов – от формирования массивов исходных данных и выбора методов расчетов до получения интерактивных компьютерных моделей и баз данных результатов моделирования;
- предоставить доступ к средствам моделирования как можно более широкому кругу пользователей через организацию интуитивно понятных и дружелюбных пользователю веб-сервисов;
- вовлекать в расчеты большинство доступных вычислительных ресурсов – локальные кластеры, «гибридные» установки, ресурсные сайты распределенных сред;
- по возможности оказывать интеллектуальную поддержку пользователю при выборе стратегии моделирования и использования методов расчетов;
- соединять в рамках единого комплекса прикладные пакеты, ориентированные на разные аспекты моделирования (нано)объектов – создание своего рода «вычислительных последовательностей», при этом выбирая наиболее приемлемые варианты установки данных пакетов и наиболее подходящие пулы ресурсов для них.

При этом может быть также во многом решен целый ряд проблем, специфичных для квантово-химических расчетов и молекулярного моделирования и обычно существенно ограничивающих эффективность работы исследователей:

- подавляющее большинство прикладных пакетов квантовой химии и молекулярной динамики в качестве исходных данных требует введения нескольких файлов сложного синтаксиса, содержащих описания до нескольких сот параметров (в том числе весьма неявно связанных между собой) и методов расчетов, многих сотен и даже тысяч первичных координат атомов, полевых потенциалов, определенной иерархии используемых констант, данных и вычислительных процедур, многовариантности построения моделей и т.п. Их подготовка очень трудоемка, требует детального знания синтаксиса входных файлов, тщательной проверки введенных данных и практически затруднена для пользователя (часто ошибки несогласованности данных обнаруживаются на этапе непосредственно выполнения расчетов, при уже большом затраченном времени счета);
- зачастую требуется выполнять многочисленные расчеты задач, отличающихся друг от друга лишь небольшими изменениями параметров («шаги» по температуре, сдвиги координат отдельных атомов, замена сортов атомов и т.п.). На основе хранимых данных весьма желательна возможность создавать массовые «шаблоны» для последующих проектов;
- для построения и хранения мультипараметрических моделей требуется создание и поддержка достаточно объемных структурированных баз данных;

- большие временные затраты на расчет моделей (например, до 40-50 часов на расчетный шаг на СК «Ломоносов» при необходимых 150-200 шагах для расчета транспорта единичного протона через наномембрану) требуют применения на ресурсах новейших методов проведения расчетов, основанных на использовании «гибридных» и распределенных моделей вычислений;
- прикладные пакеты квантовой химии в результате вычислений создают огромные (вплоть до многих гигабайтов) «сырые» (*raw*) файлы результатов, требующие постпроцессинга данных, генерализации, упрощения и перевода в другие форматы для наглядной интерактивной визуализации и сопряжения с другими прикладными пакетами.
- как правило, на основе результатов требуется построение 3-D моделей молекулярных структур, позволяющих проводить различные манипуляции с ними и получать наиболее эффективное представление о данных структурах. Кроме того, требуется формировать «на лету» большое количество текстово-табличной и графической информации (бинарные и более сложные диаграммы), позволяющей адекватно оценивать качество созданных моделей в зависимости от входных параметров.

Большинство этих проблем плохо решаемы непосредственно в зоне ответственности суперкомпьютерных центров ввиду «всеядности» последних и незаинтересованности их персонала, поэтому очевидна необходимость построения «внешних» проблемно-ориентированных пользовательских «шлюзов» в форме веб-ориентированных высокоуровневых интерфейсов. В области собственно химии (исключая биохимию и фармацевтику) сейчас почти отсутствуют комплексные решения, позволяющие снять эти проблемы, поэтому разработка современного комплексного компьютерного инструментария для упрощения работы с квантово-химическими и молекулярно-динамическими прикладными пакетами вкупе со значительной интенсификацией расчетов на основе суперкомпьютеров и распределенных сред способна коренным образом улучшить ситуацию в области молекулярного моделирования нанокompозитных материалов и нанообъектов на их основе.

В последние годы тенденция на создание проблемно-ориентированных вычислительных RIA-сервисов с дистанционным доступом пользователей является одним из основных трендов развития прикладных вычислений в разных областях науки и должна стать эффективным обязательным научным инструментом для российских исследователей. Развитие же программных и инфраструктурных средств для создания высокоэффективных веб- и грид-сервисов в настоящее время позволяет создавать инструменты, контролирующие все стадии процесса моделирования: от ввода первичных данных до интерактивной 3-D визуализации объектов или процессов. Подобные программно-аппаратные комплексы позволяют решать на принципиально новом вычислительном уровне многочисленные «тяжелые»

задачи в следующих областях: строение молекул и структур твердых тел, кинетика и механизмы сложных химических реакций, электрохимические превращения, описание процессов образования и модификации полимеров, биологических процессов и систем, а также в области общих проблем химической физики.

Опыт участия авторов в течение 2004-2013 годов в ряде российских грид-проектов [2, 3], как-то: EGEE-RDIG и EGI-RU-NGI, СКИФ-Грид, ГридННС («Национальная нанотехнологическая Сеть»), «Развитие пилотной зоны российской грид-системы для высокопроизводительных вычислений...», различных программ Президиума РАН, ориентированных в том числе на создание проблемно-ориентированных прикладных веб-сервисов в грид-вычислениях, позволил сформировать достаточно оптимальную структуру [1] прикладной компьютерной среды (программно-аппаратного комплекса), направленной (в нашем случае) на эффективное проведение широкомасштабных вычислений в области моделирования наноматериалов через веб-сервисы, обеспечивающие дистанционный доступ (рис.1)

Рис. 1. Структура среды компьютерного моделирования нанокompозитных материалов на базе высокопроизводительных вычислительных ресурсов и веб-ориентированных вычислительных сервисов (интерфейсов) к ним [1]

Подобный комплекс способен работать как в «одиночном» режиме (с вовлечением локальных ресурсов – кластеров, «гибридных» GPU- или мультиядерных установок, многоядерных высокопроизводительных станций), так и с удаленными ресурсами (при условии соблюдения достаточных мер безопасности) и (при установке промежуточного Grid-middleware – шлюзы, ресурсные грид-сайты, грид-хранилища) в составе грид-полигонов. При этом доступ к ресурсам, а также управление комплексом осуществляется через высокоуровневые, дружелюбные конечному пользователю веб-интерфейсы. Среда объединяет следующие основные компоненты: (а) вычислительные сервисы, созданные на высокопроизводительных ресурсах на основе прикладных квантово-химических и молекулярно-динамических пакетов (пакеты VASP, CPMD), в том числе с применением «гибридных» технологий расчетов на основе GPU- и мультиядерных ускорителей; (б) интернет-ориентированный инструментарий для пользовательского доступа к этим сервисам в виде комплекса высокоуровневых веб-интерфейсов для формирования и запуска задач; (в) средства размещения, хранения и обработки первичных данных и полученных результатов; (г) средства визуализации смоделированных структур и процессов и обработки полученных данных.

В состав высокоуровневых веб-сервисов помимо стандартных приемов дистанционной работы с прикладными пакетами вводятся дружелюбные пользователю средства по формированию и тестированию сложных конфигурационных файлов и массивов входящих данных, инструменты анализа и предварительной обработки полученных результатов. Также доступны средства запуска и мониторинга выполняемых заданий (включая упрощенный мониторинг ресурсов).

В настоящее время в ИПХФ РАН реализован работоспособный вариант подобной системы, основанный на локальных ресурсах (кластер на 176 двухпроцессорных узлов HP Proliant производительностью до 15 Тфлопс и «гибридная» установка на базе GPU Tesla C2075 – 2,6 Тфлопс), с поддержкой прикладных пакетов VASP (Vienna Ab initio Simulation Package, <https://www.vasp.at>) и CPMD (Car-Parrinello Molecular Dynamics, <http://cpmd.org>). Система многоуровневых веб-интерфейсов позволяет авторизовать пользователя комплекса, проводить мониторинг доступных ресурсов, подготовить проект вычислительного задания (включая создание и редакцию начальных данных и конфигурационных файлов, в том числе на основе имеющихся шаблонов задач и предыдущих проектах); запустить прикладной пакет в инфраструктуре комплекса на выбранном ресурсе; вести мониторинг выполнения задания (включая останов и перезапуск); по завершении – получить результаты счета, провести их пост-обработку (генерализацию, архивацию, конвертацию в другие форматы) и отправку в хранилище, а также обеспечить визуализацию смоделированных объектов и структур.

Основным подходом к решению поставленной задачи является комбинированное использование большинства современных вычислительных технологий для компьютерного моделирования в области квантовой химии и молекулярной динамики в рамках единой компьютерной среды, дружественной конечному пользователю-химику, позволяющей эффективно проводить расчеты в удаленном режиме. Создаваемая вычислительная среда позволяет более эффективно решать "тяжелые", многостадийные задачи в областях молекулярного моделирования наноматериалов и наноструктур, делая их доступными более широкому кругу исследователей. Модульная структура комплекса (вычислительной среды) позволяет достаточно просто вводить новые методы расчетов, выбирать различные варианты сочетания вычислительных сервисов, вводить новые функциональные возможности (в том числе новые прикладные пакеты) без ущерба общей работоспособности комплекса.

С использованием комплекса проведено более 200 вычислительных экспериментов в области компьютерного моделирования компонентов низкотемпературных топливных элементов [4-5] (нанокомпозитные электроды на основе диоксидов переходных металлов и платины и мембраны на основе сложных гетерополикислот) и литий-ионных источников тока новых типов [6] (композитные структуры на основе углерода и кремния – нанотрубки, наностержни, структурные кластеры кремния размерностью до 300 атомов, процессы взаимодействия наноструктур с твердыми электролитами, поведение их в процессах зарядки/разрядки). Примеры смоделированных наноструктур приведены на рис.2

Рис. 2. Примеры результатов компьютерного моделирования с использованием веб-ориентированных вычислительных сервисов: (а) модель динамики движения водорода по кластеру платины Pt_{19} на поверхность диоксида олова $(SnO_2)_{64}$ [5]; (б) компьютерная модель агрегата-«коралла» $Si_{16}C_{19}Li_{16}$ [6]

Проведение работ по данной тематике поддержано грантом РФФИ № 15-07-07867-а, рук. Волохов В.М.).

Литература

1. В.М. Волохов, Д.А. Варламов, А.В. Волохов, А.И. Прохоров, Г.А. Покатович Построение среды компьютерного моделирования нанокompозитных материалов на базе комплекса вычислительных сервисов и высокоуровневых веб-интерфейсов // Научный сервис в сети Интернет: труды XVII Всероссийской научной конференции (21-26 сентября 2015 г., г.Новороссийск). – М.: ИПМ им. М.В.Келдыша, 2015. с. 63-68 URL: <http://keldysh.ru/abrau/2015/proc.pdf#page=63>
2. В.М.Волохов, Д.А.Варламов, А.В.Волохов, А.В.Пивушков, Г.А.Покатович, Н.Ф.Сурков Грид-сервисы в вычислительной химии: достижения и перспективы // "Вестник Уфимского Государственного авиационно-технического университета. Серия «Управление, вычислительная техника и информатика»", 2011, т. 15, № 5 (45), с.161-169
3. Пивушков А.В., Волохов В.М., Варламов Д.А., Волохов А.В. Применение новых вычислительных технологий для повышения эффективности расчетов в грид-средах // "Вестник Уфимского Государственного авиационно-технического университета. Серия «Управление, вычислительная техника и информатика»", 2013, т.17, № 2 (55), с.117-124
4. В.М. Волохов, Д.А. Варламов, Т.С. Зюбина, А.С. Зюбин, Г.А. Покатович, А.В. Волохов Квантово-химическое моделирование процессов в низкотемпературных электрохимических топливных элементах // Альманах «Суперкомпьютерные технологии в науке, образовании и промышленности», (под ред. акад. В.А.Садовниченко, акад. Г.И.Савина, чл.-корр. Вл.В.Воеводина). – М., изд-во Московского университета, 2013, вып.5, с.172-176
5. В.М. Волохов, Д.А. Варламов, Т.С. Зюбина, А.С. Зюбин, А.В. Волохов, Г.А. Покатович Квантово-химическое моделирование наноструктурированных компонентов низкотемпературных электрохимических топливных элементов // "Параллельные вычислительные технологии (ПаВТ'2014)", труды международной научной конференции (1–3 апреля 2014 г., г. Ростов-на-Дону). Челябинск: Издательский центр ЮУрГУ, 2014, (389 с.), с.235-242
6. В.М. Волохов, Д.А. Варламов, Т.С. Зюбина, А.С. Зюбин, А.В. Волохов, Г.А. Покатович Суперкомпьютерное моделирование транспортных и энергетических процессов в нанокompозитных материалах на основе углерода и кремния // Параллельные вычислительные технологии (ПаВТ'2016): труды X международной научной конференции (28 марта – 1 апреля 2016 г., г. Архангельск, САФУ). Челябинск: Издательский центр ЮУрГУ, 2016. с.105-117