

С. М. Абрамов, В. Ф. Заднепровский, А. А. Московский

Опыт использования СуперЭВМ для эффективного развития «прорывных технологий» (на примере нанотехнологий)

Научный руководитель: С. М. Абрамов

Аннотация. В статье анализируется зарубежный опыт использования суперЭВМ для развития «прорывных технологий». Особое внимание уделено области нанотехнологий. Показано преимущество суперкомпьютерных расчетов перед натурными экспериментами.

1. Опыт использования СуперЭВМ

1.1. Особенности использования высокопроизводительных средств для развития нанотехнологий

Особенностью свойств наночастиц является то, что их размер занимает промежуточное положение между характерными размерами твердых тел и характерными размерами отдельных атомов. Именно «промежуточный» размер во многих случаях определяет наличие у наночастиц практически важных свойств. Это обстоятельство существенно усложняет моделирование, поскольку часто бывает необходимо описывать именно наночастицы, а не индивидуальные молекулы и/или твердые тела — явно учитывать в математической модели от десятков тысяч до миллионов атомов.

Развитие нанотехнологий, создание новых материалов с заданными свойствами требует проведения большого количества дорогостоящих экспериментов, часть из которых может быть заменена более дешевыми вычислительными экспериментами. В то же время, проведение точных расчетов методами молекулярного моделирования для реалистично больших систем требует огромных затрат вычислительных мощностей, что обуславливает использование вычислительных

Представлено по тематике: *Программное обеспечение для суперЭВМ, Аппаратные компоненты и системы суперЭВМ.*

комплексов с параллельной архитектурой, в том числе суперкомпьютеров или распределенных вычислительных систем большого масштаба.

В среднесрочной перспективе (5–10 лет) средства молекулярного моделирования будут играть такую же, и даже более важную роль в нанотехнологиях, какую системы Computer-Aided Engineering (CAE) и расчетов на их основе играют в современном машиностроении. Внедрение средств молекулярного компьютерного моделирования будет иметь определяющее влияние на сокращение сроков вывода продуктов на рынок и уменьшение расходов на разработку продукта — то есть, на конкурентные преимущества тех или иных высокотехнологичных продуктов.

Математическое моделирование предоставляет широкие возможности в определении тех характеристик молекулярных систем, которые крайне тяжело или невозможно определить в эксперименте, например:

- свойства электронной структуры наносистем: сопротивление нанопроводов и т. п.;
- детали молекулярной структуры в растворе;
- детали динамики систем.

Суперкомпьютерные расчеты позволяют достаточно легко определить:

- структуру, равновесные геометрические конфигурации систем;
- спектры микроволнового диапазона электромагнитных волн (ИК);
- спектры в оптическом диапазоне;
- термодинамические характеристики;
- оценить константы скорости химических реакций;
- определить дипольный момент, поляризуемость и т. п.;
- и т. п.

При этом возможен поиск соединения или конфигурации с нужными свойствами — «виртуальный скрининг», — который может быть проведен гораздо быстрее *in silico* — с использованием вычислительных средств, — а не реальным синтезом необходимых веществ. Возможность быстро определить, какие вещества обладают нужными

свойствами, важна, в первую очередь, при разработке новых технических систем. Например, при разработке новых лекарств использование виртуального скрининга может сэкономить существенные суммы — более сотни миллионов долларов, — и сократить время разработки почти на год: «In silico methods may benefit drug discovery and development significantly by saving an average of \$130 million and 0.8 years per drug» [2]. Подобные конкурентные преимущества дают суперкомпьютерные расчеты и при проектировании наноматериалов.

Преимуществом суперкомпьютерных расчетов является и то, что экспериментальные данные всегда требуют интерпретации на основе той или иной модели, что зачастую вносит элемент неопределенности в получаемые результаты. Суперкомпьютерные же расчеты обычно позволяют взглянуть на систему с детализацией (пространственное разрешение, временное разрешение), недоступной в эксперименте. Сравните, например, изображения движения «наноавтомобилей» по поверхности золота, полученные в эксперименте при помощи электронного микроскопа (см. [3]) и при помощи расчетов методами молекулярной динамики (см. [5]) и [4]).

В настоящее время большинство активных групп химиков-экспериментаторов привлекают квантовохимические расчеты как для интерпретации получаемых данных, так и для дополнительного подтверждения своих выводов. (Пример публикации о работе, построенной по такому принципу — исследование магнитных свойств нанопроводов платины [6]).

1.2. Примеры использования высокопроизводительных вычислений в развитии нанотехнологий

Трудно говорить о конкретных примерах использования молекулярного моделирования, поскольку тенденция затрагивает целую отрасль знаний — при выделении отдельных «точек» слишком много остается в стороне, а использование квантовохимических расчетов — действительно массовое явление. Тем не менее, попробуем привести несколько примеров.

- (1) Потребности инженеров в знании тех или иных термодинамических свойств различных веществ, в большинстве случаев, удовлетворяются при помощи теоретических расчетов методами молекулярного моделирования [7]. Так, например, стоимость проведения экспериментального расчета теплоты протекания химической реакции для «небольших» молекул

составляет, в среднем, около \$5,000. В то время как стоимость экспериментального определения этой же величины составляет [1] около \$100,000.

- (2) Программа Департамента Энергетики США INCITE — Innovative and Novel Computational Impact on Theory and Experiment, — в начале 2008 года распределила 250 000 000 процессорочасов машинного времени в ведущих суперкомпьютерных центрах США [8] — что равносильно 28 539 годам вычислений на одном процессоре. На основании конкурса были отобраны 55 проектов, из которых 4 — направлены на развитие непосредственно нанотехнологий и еще 13 — выполняются в смежных областях: моделирование биосистем, науки о материалах, науки о жизни и т. п. Общее машинное время, выделенное проектам в рамках программы INCITE, составляет 68 513 077 процессоро-часов (7 819 лет), в том числе, непосредственно нанотехнологическим проектам — 23 100 000 процессоро-часов (2 637 лет). Примеры проектов приведены ниже (Таблица 1).
- (3) «Case studies» фирмы AcceLrys — производителя программного обеспечения для проведения расчетов методами молекулярного моделирования, около 20 проектов [9]:
 - использование углеродных нанотрубок, как электромеханических сенсоров;
 - исследование механизма роста углеродных нанотрубок в ST Microelectronics (Италия);
 - исследование взаимодействия конца углеродной нанотрубки с молекулами воды (исследовательский центр компании Моторола);
 - описание механизма формирования нанокompозитных материалов на основе полимерных глин в ходе самокатализируемой интерколирующей полимеризации.
- (4) «Case study» компании Cray — моделирование свойств наночастиц PtFe, описание термодинамических характеристик для создания новых материалов — основы для запоминающих устройств нового поколения, Национальная лаборатория в Оак-Ридже [10].
- (5) Проекты, выполняемые в нанотехнологическом центре Molecular Foundry (при национальной лаборатории Лоуренса в Беркли, США) — примеры комплексного подхода. Molecular

Foundry — «центр коллективного пользования» в области нанотехнологий. Центр предоставляет свои ресурсы (доступ к многочисленным приборам, а также тесные связи с другими центрами, таким как синхротронный источник света, центр электронной микроскопии и другие) для выполнения проектов, выбранных на конкурсной основе. Список поддерживаемых проектов [11] 21.01.2008 содержал несколько десятков тем исследований. Вычислительную поддержку проектам оказывает National Energy Research Scientific Computing Center (NESRC) — один из мощнейших суперкомпьютерных центров в мире. Теоретическое подразделение центра [12] решает следующие задачи:

- расчеты методами *ab initio* структурных, электронных и спиновых свойств;
 - точные расчеты электронных свойств и спектров при помощи GW-приближения;
 - расчеты спектров ультрафиолетовой области при помощи уравнения Бете-Салпетера временной теории функционала плотности (*time dependent density functional theory*);
 - расчеты спектров в рентгеновской области из первых принципов;
 - расчеты линейных и нелинейных транспортных свойств с целью исследования проводимости и вольт-амперных характеристик наноразмерных электрических соединений;
 - моделирование из первых принципов и на основе эмпирических силовых полей механических, динамических и морфологических свойств.
- (6) Моделирование гибридным методом¹ окисления поверхности кремния кислородом [13] потребовали усилий 6 суперкомпьютерных центров (США и Японии) и около 150 000 часов процессорного времени (17 CPU×лет), позволили пояснить детали процесса окисления поверхности кремния при бомбардировке кислородом.
- (7) Создаваемый вычислительный центр нанотехнологических инноваций CCNI² будет обладать общей мощностью около

¹Включая квантовохимические и молекулярно-механические расчеты.

²Computational Center for Nanotechnology Innovations, [14]

100 триллионов операций с плавающей точкой в секунду (100 Tflops). Программа исследований в CCNI ориентирована на моделирование (из первых принципов) нанoeлектронных устройств.

Таблица 1: Избранные проекты, выполненные по программе Департамента Энергетики США INCITE

Проект (название)	Классификация	Машинное время CPU×час (CPU×лет)	СуперЭВМ
Computational Protein Structure Prediction and Protein Design	Химия/биология	12 000 000 (1 369)	IBM Blue Gene/P
Modeling the Rheological Properties of Concrete	Науки о материалах	750 000 (86)	IBM Blue Gene/P
An Integrated Approach to the Rational Design of Chemical Catalysts	Квантовая химия	10 000 000 (1 141)	Cray XT4
Understanding the electronic structure of novel electronic materials using many-body perturbation theory	Квантовая химия	124 615 (14)	Системы суперкомпьютерного центра NERC
Molecular simulation of complex chemical Systems	Вычислительная химия	1 500 000 (171)	Cray XT4
Predictive and accurate Monte Carlo based simulations for Mott insulators, cuprate superconductors, and nanoscale systems	Науки о материалах, нанотехнологии	10 000 000 (1 141)	Cray XT4
Cellulosic Ethanol: Physical Basis of Recalcitrance to Hydrolysis of Lignocellulosic Bio-mass	Биофизика	3 500 000 (399)	Cray XT4
Kinetics and Thermodynamics of Metal and Complex Hydride Nanoparticles	Науки о материалах, нанотехнологии	1 000 000 (114)	IBM Blue Gene/P

Проект (название)	Классификация	Машинное время CPU×час (CPU×лет)	СуперЭВМ
Electronic, Lattice, and Mechanical Properties of Novel Nano-Structured Bulk Materials	Науки о материалах, нанотехнологии	10 000 000 (1 141)	Cray XT4
Molecular simulations of surfactant assisted aqueous foam formation	Вычислительная химия	4 000 000 (456)	IBM Blue Gene/P
Water in confined states	Физическая химия	6 000 000 (685)	IBM Blue Gene/P
Reactions of lithium carbenoids, lithium enolates, and mixed aggregates	Молекулярное моделирование	138 462 (16)	Системы суперкомпьютерного центра NERC
Bose-Einstein condensation vs. quantum localization in quantum magnets	Науки о материалах	1 200 000 (137)	Cray XT4
Gating Mechanism of Membrane Proteins	Науки о жизни	1 500 000 (171) и 3 500 000 (399)	IBM Blue Gene/P и Cray XT4
Simulation and modeling of synuclein-based 'protofibril structures' as a means of understanding the molecular basis of Parkinson's Disease	Науки о жизни	1 200 000 (137)	IBM Blue Gene/P
Linear Scale Electronic Structure Calculations for Nanostructures	Науки о материалах, квантовая химия, нанотехнологии	2 100 000 (240)	Cray XT4

1.3. Классификация современных высокопроизводительных вычислительных средств

Высокопроизводительные вычисления могут осуществляться при помощи самых различных аппаратных средств. Для нужд данного документа можно предложить достаточно грубую классификацию, приведенную ниже (Таблица 2).

Преимуществом суперкомпьютеров является возможность решения задач, требующих интенсивных и/или частых обменов между процессорами.

Специализированные процессоры могут обеспечить уникальные соотношения «потребляемая мощность/производительность», но обладают рядом недостатков: такие решения применимы лишь к узкому кругу задач и требуют существенных затрат на разработку аппаратуры и программного обеспечения в течение достаточно долгого времени. Тем не менее, в области молекулярного моделирования известно минимум два таких проекта: Protein Explorer/MD Grape в Японии, Anton в США.

Как представляется, в настоящее время именно суперкомпьютеры, построенные по кластерной технологии, являются оптимальным выбором аппаратных средств на первом этапе создания предполагаемого СЦКП³, поскольку позволяют решать широкий круг задач, связанных не только с молекулярным моделированием, но и инженерными расчетами, оптимизацией свойств технических систем и т. п.

1.4. Математическое моделирование в химии

Описание свойств веществ, важных для различных приложений (наноэлектроника, наноматериалы, наноустройства, нанобиосистемы и т. п.) достигается применением различных методов математического моделирования. Используемые в этих расчетах алгоритмы предъявляют различные требования к аппаратным средствам — Таблица 3.

Следует обратить внимание, что большинство методов требуют для проведения высокопроизводительных расчетов именно суперкомпьютеров, чьи аппаратные средства позволяют осуществлять тесную (высокоскоростную и с низкой задержкой) связь между индивидуальными процессорами, участвующими в параллельном счете.

Очень часто невозможно рассчитать интересующие исследователей и инженеров свойства молекул, иначе как привлекая методы

³Суперкомпьютерного центра коллективного пользования.

Обозначение	Отличительные особенности	Комментарий
Вычислительные фермы	Некоторое число компьютеров, связанных между собой в единую сеть при помощи технологий семейства Ethernet	Низкое отношение «Стоимость/производительность»
Грид-системы	Объединение нескольких высокопроизводительных установок, принадлежащих разным организациям и находящихся в разных географических точках	Требует организационных мероприятий по развертыванию
Суперкомпьютеры (вычислительные кластеры)	Большое число процессоров, объединенных высокоскоростной сетью (интерконнектом) и/или имеющих общее адресное пространство (общую память). Способны решать задачи, требующие интенсивных обменов данными между отдельными процессорами	Отношение «Стоимость/производительность» выше, чем у вычислительных ферм за счет необходимости применять высокоскоростные сети и другие специальные решения
«Ускорители»	Использование реконфигурируемых вычислителей, графических процессоров и т. п. Трудностью в использовании данного типа решений является необходимость адаптации прикладного программного обеспечения	Отношение «Стоимость/производительность» может быть существенно ниже, чем у других систем. Сложность программирования, не универсальность
Специализированные вычислители	Использование для расчетов процессоров, ориентированных на решение определенных классов задач	Отношение «Стоимость/производительность» очень высокое. Высок риск отстать от возрастающих возможностей массовых процессоров

ТАБЛИЦА 2. Классификация систем высокопроизводительных вычислений

квантовой химии. Это относится, в первую очередь к спектрам, электронной структуре. Квантовая химия является хорошо разработанной областью теоретической химии, работы в этой области ведутся с

Моделируемые характеристики	Методы решения	Необходимые аппаратные средства
Механические свойства	Методы Монте-Карло	Вычислительные фермы
Механические свойства	Молекулярная динамика	Суперкомпьютеры, грид-системы
Механические свойства	Молекулярная динамика с разрывом и образованием химических связей	Суперкомпьютеры
Электронные свойства (электропроводность и т. п.)	Моделирование электронной структуры вещества методами квантовой химии (например, методы функционала плотности)	Суперкомпьютеры для моделирования свойств наночастиц требуются суперкомпьютеры до 100 Tflops и более
Электронно-оптические свойства (спектры поглощения, люминесценции и т. п.)	Моделирование энергетических спектров частиц методами квантовой химии (например, методами многоконfigurационного самосогласованного поля)	Суперкомпьютеры, для моделирования свойств наночастиц требуются суперкомпьютеры до 100 Tflops и более
Транспортные свойства (скорость диффузии и т. п.)	Методы Монте-Карло	Вычислительные фермы, грид-системы
Транспортные свойства (скорость диффузии и т. п.)	Молекулярная динамика	Суперкомпьютеры, грид-системы
Транспортные свойства (скорость диффузии и т. п.)	Молекулярная динамика с разрывом и образованием химических связей	Суперкомпьютеры
Химические свойства протекания реакций (скорости реакций) в окружении (растворе, матрице, белке)	Построение энергетических профилей реакций на основе квантовохимических расчетов или расчетов гибридными методами	Суперкомпьютеры, для проведения расчетов

ТАБЛИЦА 3. Требования к системам высокопроизводительных вычислений и их зависимость от решаемых задач

40-х годов XX-го века. За работы в этой области было присуждено четыре нобелевских премии: Уолтеру Кону и Джону Поплу (1998), Рудольфу Маркусу (1992), Роберту Милликену (1966), Лайнусу Полингу (1954). Наиболее точными и общими методами расчета являются т. н. расчеты из первых принципов (*ab initio*), которые, в отличие от полуэмпирических и эмпирических методов, опираются лишь на (приближенное) решение уравнений Шредингера — основного уравнения квантовой механики, — для системы, состоящей из ядер и электронов. Методы *ab initio* являются и наиболее затратными в вычислительном смысле, поскольку требуют решения сложных математических уравнений. В случае наночастиц, моделирование дополнительно осложняется большим размером системы (тысячи и миллионы атомов). Тем не менее, при помощи ряда приближений удается находить решения и для наночастиц, описывать их свойства. Используемый при этом математический аппарат (в первую очередь это аппарат линейной алгебры плотных матриц) позволяет обеспечить эффективное распараллеливание при условии достаточно высокоскоростных связей между процессорами, участвующими в параллельном счете.

Арсенал средств молекулярного моделирования не ограничивается методами *ab initio*, но включает в себя также и средства молекулярной динамики и другие.

В случае больших систем и в зависимости от точности приближений, расчеты становятся настолько сложными, что для их проведения требуются самые мощные суперкомпьютеры.

Так, например, в 2005 году так называемый «Приз Гордона Белла» за наивысшую производительность, достигнутую приложением на суперкомпьютере, был получен группой ученых из Национальной Лаборатории Лоуренса в Ливерморе за моделирование [15] застывания расплава тантала. На данной задаче была использована (достигнута) реальная производительность 100 триллионов операций над числами с плавающей запятой в секунду. Моделировались системы, включающие до 500 миллионов атомов.

В 2006 году данный приз был выдан за моделирование динамики кластеров молибдена (около 1000 атомов) из «первых принципов» — производительность составила около 200 триллионов операций в секунду. Если бы аналогичный расчет проводился на обычном ПК, он занял бы несколько тысяч лет.

Благодарности

Данная работа была частично поддержана и выполнялась в рамках следующих проектов:

- Проект 07–07–12038 офи Российского Фонда Фундаментальных Исследований «Метапрограммирование на основе шаблонных классов С++ как средство создания высокопроизводительных распределённых приложений».
- Программа Союзного государства «Разработка и использование программно-аппаратных средств Грид-технологий перспективных высокопроизводительных (суперкомпьютерных) вычислительных систем семейства «СКИФ», шифр «СКИФ-ГРИД».
- Программа фундаментальных научных исследований ОНИТ РАН «Оптимизация вычислительных архитектур под конкретные классы задач, информационная безопасность сетевых технологий», проекты:
1.3 «Разработка и реализация языков T# и T++ и соответствующих им средств для эффективной поддержки высокопроизводительного параллельного счета»;
2.3 «Сравнительное исследование параметров и характеристик перспективных аппаратно-программных архитектур кластерных мультипроцессорных систем»;
- Программа фундаментальных исследований Президиума РАН «Разработка фундаментальных основ создания научной распределенной информационно-вычислительной среды на основе технологий GRID». Проект 4.2 «Функционально-ориентированные суперструктуры как эффективное средство для построения высокопроизводительных распределенных приложений и сервисов». Проект 4.5 «Разработка технологий интеграции разнородных, географически распределенных данных в GRID-системах».
- Программа Президиума РАН «Поддержка инноваций и разработок», проект ИПС РАН «Технология и система параллельного программирования OpenTS»;
- Государственный контракт № 02.514.11.4034 с Федеральным агентством по науке и инновациям по теме: «Эффективные методы создания высокопроизводительных параллельных программ на языке С++ и его диалекте T++ для многоядерных ЭВМ, SMP, кластерных и распределенных систем»;
- Программа Союзного государства (шифр «ТРИАДА») «Развитие и внедрение в государствах-участниках Союзного государства наукоемких компьютерных технологий на базе мультипроцессорных вычислительных систем», проекты:
«Разработка средств параллельной обработки изображений дистанционного зондирования Земли с динамическим распределением нагрузки. Разработка прототипа распределенного архива изображений с единым каталогом информации», проекты;
«Создание прототипа системы эффективного сервера приложений для ВМВС»;

«Разработка принципов эффективного отказоустойчивого счета T++-приложений на априорно неустойчивых (мета-)кластерных конфигурациях и их программная реализация».

Особая благодарность—М. Г. Химшиашвили за ее самоотверженный труд по корректуре и верстке статьи в программе L^AT_EX.

Список литературы

- [1] Cummings P. T. *Molecular Modeling and Simulation: Emerging Tools for Physical Properties Prediction*, June 25 (Доступно как: <http://forum2000.boulder.nist.gov/Forum-Cummings.ppt>). ↑1
- [2] Seifert M. H. J., Wolf K., Vitt D. «Virtual high-throughput in silico screening». — вып. 4. — Т. 1. — BIOSILICO, 4 сентября 2003, 143–149 с. ↑1.1
- [3] James M. Tour's Web Pages: Nano Car: Веб-сайт (Доступно как: <http://tourservr.rice.edu/movies/>). ↑1.1
- [4] You Tube: Nanocar-trimer: Веб-сайт (Доступно как: <http://youtube.com/watch?v=SDnIOLMyYwA>). ↑1.1
- [5] You Tube: Nanocar-trimer-2: Веб-сайт (Доступно как: <http://youtube.com/watch?v=OBAJNpEwpb8>). ↑1.1
- [6] Smogunov A., Dal Corso A., Delin A., Weht R., Tosatti E. «*Colossal magnetic anisotropy of monatomic free and deposited platinum nanowires*» // Nature Nanotechnology. — Т. 3, № 1, January, 2008 (доступно как: <http://www.nature.com/nano/journal/v3/n1/pdf/nnano.2007.419.pdf>), с. 22–25. ↑1.1
- [7] Irikura K. K., Frurip D. J. *Computational Thermochemistry* // Computational Thermochemistry: Prediction and Estimation of Molecular Thermodynamics ACS Symp. ред. Irikura K. K., Frurip D. J. Т. Ser. 677. — ACS: Washington DC, 1998. ↑1
- [8] DOE Awards 265 Million Hours of Supercomputing Time to Advance Leading Scientific Research Projects: U.S. Department of Energy: Веб-сайт (Доступно как: <http://www.doe.gov/news/5849.htm>). ↑2
- [9] Case Studies for Materials Science Research: Веб-сайт: Accelrys Software, Inc (Доступно как: <http://www.accelrys.com/reference/cases/application/nano.html>). ↑3
- [10] *Understanding FePt Nanoparticle Behavior* (Доступно как: http://www.cray.com/downloads/science/ornl_lsms_flyer_p.pdf). ↑4
- [11] The molecular foundry. A Nanoscince: Веб-сайт: Lawrence Berkeley National Laboratory (Материал доступен по адресу: http://foundry.lbl.gov/user_program/4_res_proj.html). ↑5
- [12] The molecular foundry. A Nanoscince. Theory of Nanostructured materials Facility: Current Capabilities: Веб-сайт: Lawrence Berkeley National Laboratory (Доступно как: http://foundry.lbl.gov/user_program/theory_capabilities.html). ↑5

- [13] Sekiguchi S., Takemiya H., Vashishta P., Nakano A., Tanaka Yo., Ogata Sh., Kalia R. K. Sustainable Adaptive Grid Supercomputing: Multiscale Simulation of Semiconductor Processing across the Pacific: SC06: International Conference for High Performance Computing Networking and Storage: Веб-сайт. — Тампа, USA, November 11 (http://sc06.supercomputing.org/schedule/event_detail.php?evid=9174). ↑6
- [14] The World's Most Powerful University-Based Supercomputing Center: Веб-сайт: CCNI: Computational Center for Nanotechnology Innovations (<http://www.rpi.edu/research/ccni/>). ↑2
- [15] Breakthrough 3D simulations win prestigious 2005 Gordon Bell Prize: Public Affairs: Веб-сайт: Lawrence Livermore National Laboratory (Документ доступен по адресу: https://publicaffairs.llnl.gov/news/news_releases/2005/NR-05-11-12.html). ↑1.4

152020, Переславль-Залесский, м. Ботик, Институт программных систем РАН.

S. M. Abramov, V. F. Zadneprovsky, A. A. Moskovsky. *Experience of the use for supercomputers for effective development by nanotechnology.* (in Russian.)

ABSTRACT. In this paper, we analyze high-performance computing assistance in new “disruptive technologies” development, with special attention to nanotechnology field. Advantages of computer simulations over experimental studies are advocated.

А. А. Московский