

Перспективы эксафлопсных вычислений в разработке новых технологий энергетики*

В.С. Вечер, Н.Д. Кондратюк, Н.Д. Орехов, Г.Э. Норман, В.В. Писарев,
Г.С. Смирнов, В.В. Стегайлов

Объединенный институт высоких температур РАН

Развитие новых технологий в энергетике в значительной степени основано на многомасштабном моделировании, успех которого, в свою очередь, зависит от прогресса суперкомпьютерных технологий. В данной работе мы рассматриваем несколько ярких примеров задач, которые решаются нами с использованием атомистических и многомасштабных моделей на лучших суперкомпьютерах России: фазовые переходы в газовых гидратах и газоконденсатных смесях, расчет реологических свойств органических жидкостей и создание новых полимерных материалов. Данные исследования представлены в контексте примеров задач из американской программы INCITE. На основе рассмотренных примеров мы обсуждаем вопросы эффективности аппаратного обеспечения, масштабируемости алгоритмов, проблемы ввода-вывода и обработки данных.

Введение

Развитие суперкомпьютерных технологий в США и других передовых странах встало на путь радикального наращивания числа вычислительных элементов [1]. Серия систем IBM BlueGene/L [2] была с самого начала задумана для развития технологии массового параллелизма. В 2006 году, когда системы IBM BlueGene прошли начальную апробацию и появились суперкомпьютеры Cray XT3/4, для развития алгоритмов параллельного решения математических задач на этих новых системах Департамент энергетики США (DOE) расширил уже существовавшую ранее программу Innovative and Novel Computational Impact on Theory and Experiment (INCITE) [3]. Так в 2012 году на 60 проектов было выделено 1672 миллиона процессор-часов на суперкомпьютерах IBM BlueGene/P и Cray XT5. Тот факт, что в программе INCITE участвуют системы только с тороидальной топологией, свидетельствует, что данная топология рассматривается в качестве вероятного кандидата для суперкомпьютеров эксафлопсной эры [4].

Как показано на рис. 1, тематика проектов INCITE охватывает практически все области науки и техники. Исходя из доступного краткого описания, все проекты с некоторой долей условности разделены на следующие тематики: атомистические модели (Ab initio-расчеты из первых принципов и MD-молекулярная динамика), модели в рамках механики сплошных сред (CFD), астрофизические модели (Astrophys), физика плазмы (Plasma), квантовая хромодинамика (QCD), физика ядра (Nuclear), computer science (CS). Атомистические модели (Ab initio и MD) охватывают очень широкие приложения от механики ударно-волнового разрушения до химических реакций и систем биомолекул. В программу INCITE принимаются проекты с международным участием, даже без партнера из США.

Обязательным условием получения расчетного времени по программе INCITE является использование массового параллелизма. При этом задачи, связанные с одновременным запуском большого числа однотипных задач (тривиальный параллелизм для набора статистики), рассматриваются, однако не являются приоритетными [3]. От алгоритма решения

*Работа частично поддержана грантами программ фундаментальных исследований Президиума РАН № 43 (координатор ак. Бетелин В.Б.), РФФИ 13-01-12070-офи_м. Отдельные авторы частично поддержаны грантами РФФИ 13-08-01428-а (Смирнов Г.С., Стегайлов В.В.), 14-08-31550-мол_а (Смирнов Г.С., Орехов Н.Д.), 14-08-31587-мол_а (Писарев В.В.), стипендией Президента РФ для молодых учёных (Писарев В.В.). Расчеты выполнены в МСЦ РАН, на кластерах К-100 ИПМ им. М. В. Келдыша РАН и «Ломоносов» МГУ им. М. В. Ломоносова.

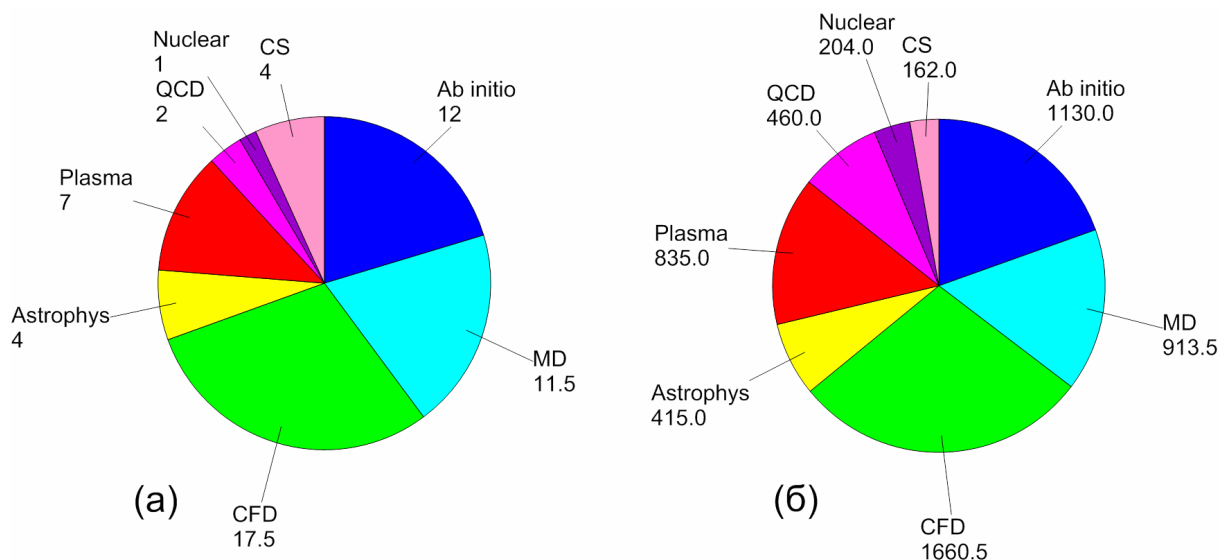


Рис. 1. Деление проектов INCITE 2014 г. по тематикам: (а) число проектов по каждой тематике и (б) выделенное на каждую тематику вычислительное время в миллионах процессоро-часов

прикладной задачи требуется демонстрировать параллельную эффективность на вычислительном поле порядка 20% используемой машины. В 2011 году это соответствовало размеру вычислительного поля порядка 20 тысяч ядер, а в 2013–14 гг. – 200 тысяч ядер!

Как для задач крупномасштабного суперкомпьютерного физико-математического моделирования, так и для задач обработки больших объемов данных второй важнейшей технической задачей на пути в «эру эксафлопса» является организация параллельного ввода-вывода с сотен тысяч и миллионов ядер.

Эти цифры демонстрируют переход к новой эре использования вычислительных методов в науке и технике. В этой работе мы приводим примеры атомистических моделей для решения задач, предполагающих, в перспективе, использование вычислительных полей эксафлопсного уровня производительности. Вместе с тем, ученые не всегда представляют себе, как надо выбирать суперкомпьютер для задачи, которой они занимаются. Тем более, неясной остаётся архитектура требуемых компьютеров. Этим вопросам посвящена вторая часть настоящей работы. Рассматривается подход, позволяющий выявить, для каких задач нужны суперкомпьютеры эксафлопсного класса. Имеются в виду задачи, для решения которых требуется загрузка всего или значительной части суперкомпьютера, т.е. сотен тысяч или даже миллионов вычислительных ядер одновременно под одну задачу.

1. Атомистические модели вещества

Повышение эффективности существующих технологий и разработка новых в значительной мере связана с использованием более точных представлений о строении вещества и его свойствах на микроуровне. Соответствующие атомистические модели основываются на математическом описании межатомного и межмолекулярного взаимодействия [5]. Расчет сил, действующих на атомы, обычно занимает большую часть (более 80%) МД расчета. Задачи молекулярно-динамического (МД) моделирования являются одними из наиболее хорошо масштабируемых на параллельные вычислительные системы [6–10]. В рамках программы INCITE с задачами МД моделирования связано почти 2/3 всех проектов (Ab initio и MD на рис. 1). Далее рассмотрим несколько характерных примеров подобного рода.

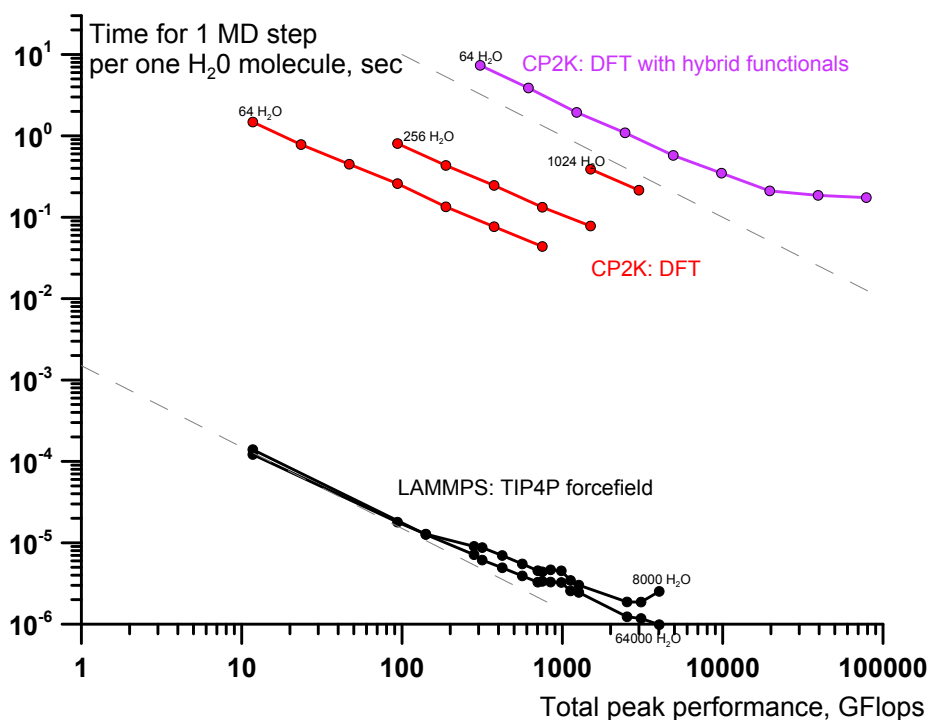


Рис. 2. Сравнение времени расчета на 1 шаг МД интегрирования в классических (TIP4P) и квантовых МД моделях при различных пиковых мощностях вычислительного поля. Приведены данные для пакетов LAMMPS и CP2K.

1.1. Вода

В связи с тем, что вода является важнейшим примером для молекулярного моделирования, имеющим отношение практически ко всем областям науки и техники, используемые для ее моделирования модели представляют большой интерес. По своей природе взаимодействие между атомами и молекулами является квантовым эффектом. Однако вычислительная сложность квантового описания даже при использовании таких подходов, как теория функционала электронной плотности (ТФП), обуславливает необходимость использования эмпирических потенциалов для классического «механического» описания межмолекулярного взаимодействия. На рис. 2 показана зависимость времени расчета на 1 шаг МД интегрирования в классической МД модели воды с 4-х точечным потенциалом TIP4P и в квантовых моделях на основе ТФП и гибридной ТФП, более точно учитывающей эффекты обменного взаимодействия. Классические МД модели воды имеют вычислительную сложность ~ 1.5 МФлоп на 1 молекулу на 1 МД шаг (Флоп — операция с плавающей точкой), которая растет с увеличением числа частиц как $O(N)$ или $O(N \log(N))$ при явном учете дальнегодействующего электростатического взаимодействия. В случае ТФП модели время расчета растет квадратично по отношению к числу молекул: $\sim 300N$ МФлоп на 1 молекулу на 1 МД шаг.

Гидраты природных газов или газовые гидраты — это нестехиометрические соединения включения, в которых молекулы газа (молекулы-гости) заключены в полостях трехмерной решетки из молекул воды (каркас хозяина). Стабилизация водных клатратных каркасов, термодинамически менее стабильных, чем лед или жидкая вода при тех же условиях, обеспечивается за счет ван-дер-ваальсовых взаимодействий гость-хозяин.

Месторождения природных газовых гидратов встречаются на шельфе и в зоне вечной мерзлоты. Они активно изучаются уже более сорока лет, так как содержат большое ко-

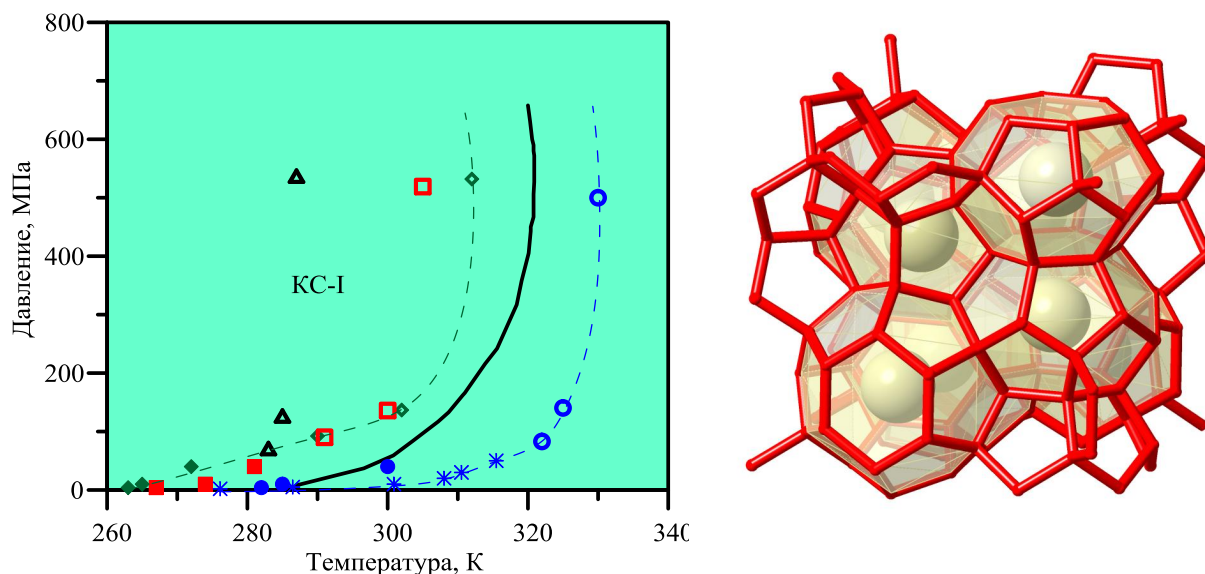


Рис. 3. Фазовая диаграмма гидрата метана (см. [11]). Черной кривой показана экспериментальная линия плавления [21]. Заполненные синие, красные и зеленые символы и синие снежинки — литературные данные. Незаполненные символы — наши результаты для тех же моделей межмолекулярного взаимодействия. Синие символы — модель воды TIP4P/Ice; красные — TIP4P/2005 ($\chi=1$); зеленые — TIP4P/2005 ($\chi=1.07$); черные — SPC/E. Область между пунктирными линиями (интерполяциями результатов расчетов в рамках моделей TIP4P/Ice и TIP4P/2005 ($\chi=1.07$)) иллюстрирует текущую точность определения кривой плавления по результатам моделирования. Справа показана визуализация молекулярной модели гидрата метана KS-I структуры. Красным показаны направления водородных связей между молекулами воды в клатратном каркасе (сами молекулы воды не показаны), шары соответствуют молекулам метана.

личество метана и рассматриваются как перспективный источник топлива. Вместе с тем газовые гидраты, существующие при температурах и давлениях, близких к условиям их разложения, представляют экологическую опасность - в случае смещения теплового равновесия освободившийся метан может внести значительный вклад в парниковый эффект. Поэтому знание динамических, термодинамических и механических свойств газовых гидратов важно для оценки их поведения при разработке газогидратных месторождений.

За последние годы было проведено множество теоретических, экспериментальных и компьютерных исследований газовых гидратов. Проведенные исследования месторождений на Аляске, в Канаде и Японии показывают, что добыча метана из природных гидратов может быть выполнена с использованием существующих технологий. Более того, около побережья Японии недавно была произведена первая в мире пробная добыча газа из месторождения в желобе Нанкай.

Молекулярная динамика позволяет исследовать свойства вещества на микроскопическом уровне, поэтому ее развитие привело к пониманию ряда фундаментальных свойств газовых гидратов. МД модели для газовых гидратов существенным образом зависят от точности описания взаимодействия между молекулами воды. Точность таких моделей сильно возросла за последнее десятилетие. Наиболее сложные модели, учитывающие поляризацию, лучше описывают структурные свойства и растворимость молекул газа, тогда как модели без поляризации лучше подходят для описания фазовой диаграммы [11,12] (см. рис. 3).

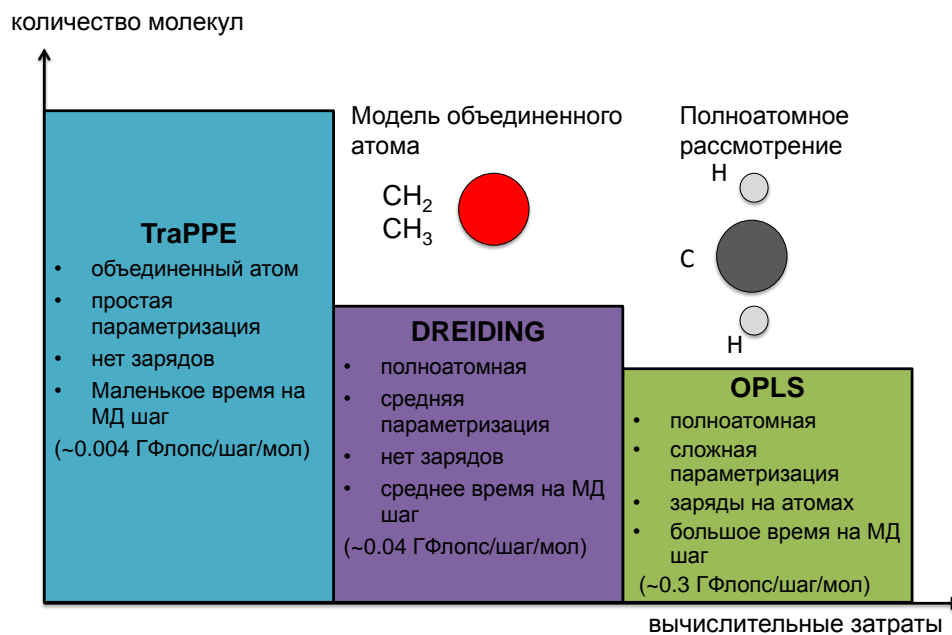


Рис. 4. Сравнение различных моделей молекулярного взаимодействия для органических жидкостей.

1.2. Органические жидкости

Модели межатомных потенциалов для органических жидкостей представляют собой хороший пример иерархии огрубления модельных представлений для расширения возможностей прямого численного счета.

Ключевым фактором для достоверности результатов атомистического моделирования является выбор потенциала взаимодействия в модельной системе. Моделирование органических соединений чаще всего производится в рамках концепции “силового поля” (англ. forcefield), когда суммарная энергия E_{mol} взаимодействия молекул вещества представляется в виде суммы вкладов ненаправленного парного взаимодействия E_{pair} , энергии напряжения связей E_{bond} и углов E_{angle} в молекулах и энергии вращения по двугранным углам E_{dihed} . Различные параметризации этих взаимодействий позволяют моделировать вещества от простейших углеводов до белков и ДНК.

Для углеводов разработан ряд различных наборов параметров силового поля, из них наиболее широко применяются наборы OPLS, DREIDING и TraPPE. Параметризация межатомных потенциалов OPLS позволяет точнее воспроизвести теплофизические свойства жидкой фазы, а параметризация TraPPE предназначена для точного предсказания фазовых диаграмм. DREIDING занимает промежуточное положение (рис. 4). Эти параметризации позволяют с высокой точностью описывать наблюдаемые фазовые диаграммы чистых углеводов и их смесей (рис. 5).

Особый интерес представляет расчет методом МК свойств жидкостей, взаимодействующих с поверхностью, например, в порах. Потенциал TraPPE позволяет воспроизвести экспериментальные теплоты адсорбции углеводов на силикатных материалах [13], таким образом, представляется целесообразным его использование для расчета фазовых диаграмм газов в порах горных пород. В силу необходимости моделирования системы на атомарном уровне метод ограничен размерами систем до нескольких десятков нанометров, чего, однако, обычно оказывается достаточно для расчета фазовых диаграмм и равновесных плотностей веществ.

Метод МД, в силу универсальности подхода (прямое интегрирование уравнений дви-

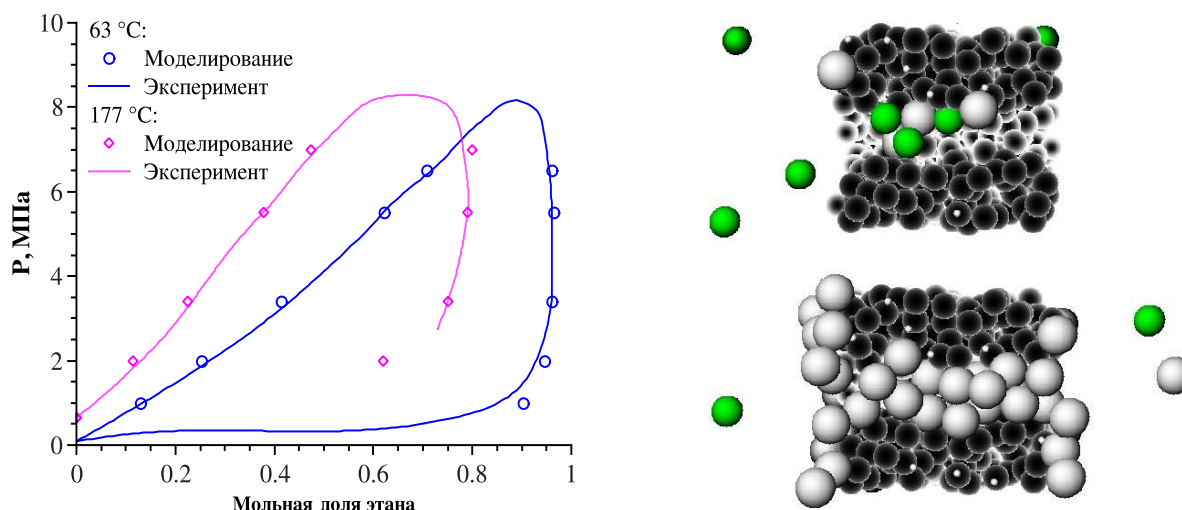


Рис. 5. (Слева) Фазовые диаграммы смеси этан-н-гептан при температурах 63 С и 177 С. Сопоставление атомистического моделирования с потенциалом TraPPE [14] с экспериментом. (Справа) Течение смеси этан-азот через пору в мембране из аморфного оксида кремния. Иллюстрация из работы [15]. Диаметр поры 0,8 нм, длина 2,85 нм, перепад давления 0,1 МПа. Течение при -13 С смесей этан-азот с мольной долей этана 5% (сверху) и 90% (снизу). Серым цветом обозначены молекулы этана, зеленым — молекулы азота, черным — атомы кремния мембраны.

жения) позволяет вычислить не только фазовые диаграммы материала с заданным потенциалом взаимодействия, так и фазовые проницаемости различных сред (рис. 5).

1.3. Полимеры

В последнее десятилетие в индустрии конструкционных материалов наблюдается заметный рост интереса к сфере полимерных нанокомпозитов. В первую очередь это обусловлено их необычными механическими свойствами при сравнительно небольшой плотности и высокой коррозионной стойкости. В качестве наполнителя в полимерной матрице могут выступать углеродные волокна, нанотрубки, фуллерены, металлические или иные наночастицы.

Свойства полимерных нанокомпозитов существенным и не вполне очевидным образом зависят от целого ряда параметров: типа, концентрации, ориентации, упругих свойств наполнителя, пластических и термических свойств самой полимерной матрицы. Описание подобной многопараметрической задачи на микроуровне с использованием сугубо экспериментальных методов крайне затруднительно и пока во многом носит эмпирический характер. Свойства полимерных цепей, находящихся в контакте с поверхностью наполнителя, неизбежно отличаются от свойств чистого полимера. В случае нановключений, ввиду большой величины площади их эффективной поверхности, уже при малых концентрациях (порядка нескольких массовых процентов), область интерфейса может охватить существенную часть полимерной матрицы. При этом на свойства материала влияет множество параметров — форма и размер наночастиц, распределение частиц в полимерной матрице, тип и сила взаимодействия полимерной матрицы с наночастицами, температурный режим, запутывание полимерных цепочек и т.д.

Одной из основных особенностей полимерных систем является их четко выраженная иерархическая структура. Последовательный подход описания при этом должен охватывать ряд пространственных масштабов начиная от размеров отдельного мономера (субнанометровый масштаб) до структуры цельной молекулы (десятки-сотни нанометров) и конфигураций, образующихся в системе многих молекул (субмикронные масштабы) (рис. 6). Аналогичная иерархия масштабов может быть построена и для шкалы времен протекания

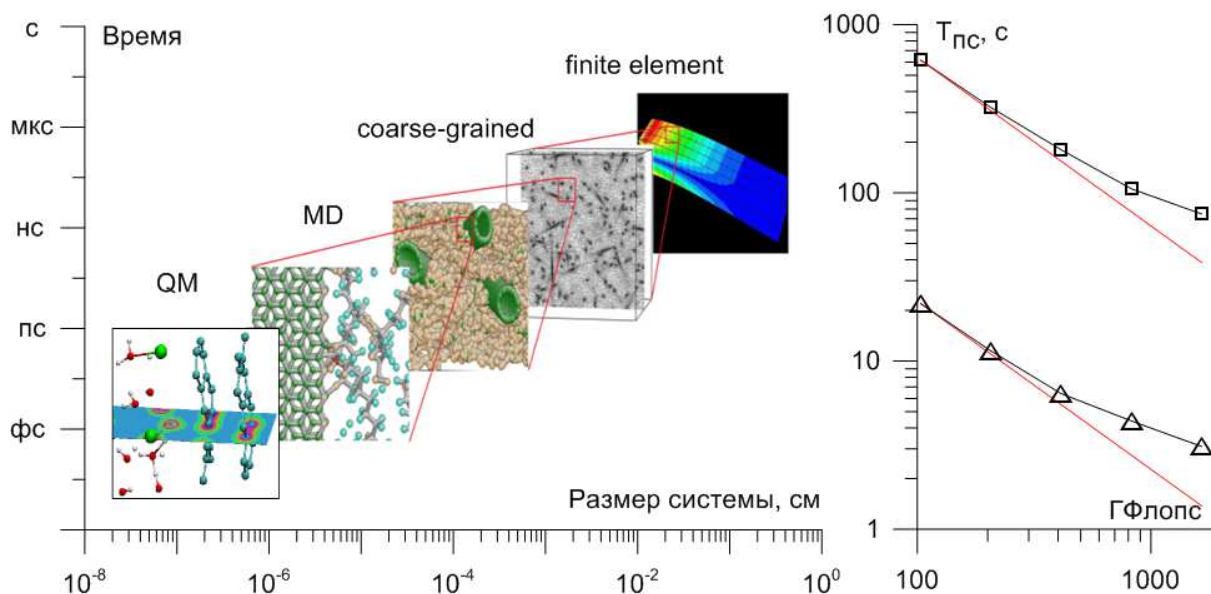


Рис. 6. а) Принципиальная схема многомасштабного моделирования полимерных композитов. б) зависимость вычислительного времени, требуемого на расчёт траектории длительностью в одну пикосекунду (T_{ps}), от величины задействованных вычислительных ресурсов. Квадраты соответствуют полноатомной модели, треугольники – огрубленной. Красными линиями обозначено идеальное масштабирование задач.

ния процессов в высокомолекулярных системах. Если характерное время переориентации отдельного мономера составляет $t \sim 10^{-11}$ с, то время релаксации цепи, содержащей N звеньев, согласно модели Рауза, будет составлять $T = tN^2$. В случае цепи из $N=1000$ звеньев характерное время релаксации макромолекулы может достигать 10^{-5} с. В связи с этим даже МД-расчеты с участием достаточно простого полимера, как, например, полиэтилен, требуют задания начальной конфигурации, максимально близкой к равновесной при рассматриваемых условиях. Более сложные конформационные перестройки полимерных систем как, например, стеклование или кристаллизация из расплава, как правило, лежат далеко за пределами временных масштабов, доступных МД.

Активное использование МД для задач моделирования нанокомпозитов началось немногим более десятилетия назад. Первые работы в этой области [16] были нацелены на изучение аспектов взаимодействия полимера с участками поверхности наполнителя, поскольку вычислительные возможности не позволяли размещать в расчетной ячейке наночастицы характерных размеров. Классическая полноатомная МД, будучи достаточно ресурсоемким методом, накладывает серьезные ограничения на линейные размеры исследуемой системы, которые обычно не превышают сотен нанометров, и на времена расчета. На помощь в этом случае приходят методы огрубленной (т.н. coarse-grained) молекулярной динамики (см., например, [17]). В огрубленной модели составные части полимеров и полимерных композитов представляются в упрощенной геометрии. Основной идеей метода является объединение нескольких маленьких частиц (групп атомов) в один большой блок (супер-атом) и использование общих силовых постоянных и геометрических параметров, основанных на простых соображениях гибридизации. При таком упрощении функциональной формы описания межатомных взаимодействий важно следить за сохранением физических свойств материалов. Для сложных полимерных систем не существует универсального алгоритма задания структуры супер-атома (т.н. mapping). Для каждой конкретной задачи методика огрубления атомарной структуры (определение групп атомов, которые в дальнейшем сформируют супер-атомы), как правило, выявляется путём рассмотрения нескольких наиболее



Рис. 7. Основные части МД алгоритма на GPU ускорителе.

уместных вариантов. От выбора варианта огрубления напрямую зависит описательная способность создаваемой упрощённой модели. Среди стохастических методов описания эволюции атомистических систем стоит также выделить метод диссипативной динамики частиц (DPD). DPD, так же как и классическая МД, основан на интегрировании систем уравнений движения Ньютона. Благодаря меньшей вычислительной сложности он позволяет рассматривать крупные системы, включающие в себя группы из наночастиц субмикронного размера, и широко используются для изучения процессов агломерации наночастиц и их влияния на свойства полимерных матриц [17].

2. Использование специализированных GPU ускорителей

Метод молекулярной динамики зачастую требует выполнения сложных вычислений с значительным объемом данных. С развитием графических ускорителей, их вычислительные мощности все чаще стали применяться не только для визуализации, но и для выполнения трудоемких расчётов. В то время, как традиционные процессоры превосходно выполняют сложные манипуляции над малым набором данных, графические ускорители изначально были ориентированы на параллельное выполнение простых инструкций сразу над большим блоком данных. Это достигается выполнением одной инструкции сразу в нескольких параллельных потоках выполнения (SIMD).

Безусловно, подобная архитектура существенно повышает производительность в задачах, в которых изначально можно выделить высокую степень параллелизма, частую связанных с операциями над матрицами и векторами. К подобным задачам можно отнести, например линейную алгебру над плотными матрицами, выравнивание последовательностей в биоинформатике или методы машинного обучения на базе искусственных нейронных сетей. Однако, однозначно отнести в эту группу метод молекулярной динамики не представляется возможным, т.к. многие вычислительные приемы в МД создавались в эпоху господства последовательных вычислений. Потому успешное применение подобной архитектуры для расчётов методом молекулярной динамики требует существенной адаптации алгоритмов.

Для максимального использования GPU алгоритм не только должен быть высокопараллельным, но и кроме того, та часть алгоритма, которая выполняется на GPU, часто называемая ядром (kernel), должна быть маленькой, насколько возможно — ядро должно уложиться в число регистров вычислительных единиц GPU. Достичь это в реальных алгоритмах не всегда просто [18].

Кроме того, есть особенности и с организацией работы с памятью. Достаточно быстрой памятью можно считать лишь разделяемую память, но её размер в современных GPU-ускорителях составляет порядка 16 Кбайт на вычислительную единицу. К сожалению, не всегда удается сохранить результаты текущих вычислений в разделяемой памяти, что особенно характерно для расчётов дальнедействующих потенциалов. Глобальная память графического ускорителя куда больше (несколько Гбайт), но скорость доступа туда не позволяет всем потокам исполнения работать с ней одновременно, что может приводить к простаиваниям потоков и, как следствие, к потере производительности.

Существуют и фундаментальные проблемы обеспечения высокой степени параллелизма для GPU расчетов МД задач. Принцип SIMD предполагает выполнение одинаковых инструкций для группы потоков. Однако появление в алгоритме ветвлений существенно ухудшает ситуацию. Если в группе потоков разные потоки пошли по разным ветвям алгоритма, то дальнейшие инструкции для них будут уже различными, что нарушает принцип SIMD. Типичным решением является выполнение обеих ветвей алгоритма, что приводит к потерям в производительности.

Ещё одной из проблем является сортировка — практически вездесущая операция, в том числе широко применяемая и в молекулярной динамике при составлении списков соседей и сил. В то время, как на обычных процессорах решение данной задачи является скорее проблемой выбора алгоритма, на GPU создание быстрого и эффективного алгоритма несколько затруднено, ввиду массивно-параллельной архитектуры графических ускорителей.

Таким образом, достижение полной загрузки всех ядер графического ускорителя — задача вполне нетривиальная. Такая же ситуация имеет место и в других прикладных областях суперкомпьютерных вычислений. Несмотря на значительный прогресс в адаптации кодов для использования GPU ускорителей, их эффективность пока не является предметом исследований. В списке Топ-500 ноября 2014 года в верхней десятке присутствовали 3 системы с Nvidia GPU. Но в то же время в этом списке 7 самых мощных суперкомпьютеров Германии используют только обычные процессоры и только 8-я немецкая система (номер 77) имеет ускорители Nvidia. Эти примеры можно рассматривать как определенную иллюстрацию объемов текущего использования GPU ускорителей (энтузиазм в США и определенная осторожность в Европе).

3. Перспективы развития суперкомпьютеров за рубежом

Министерство энергетики США (DOE) запустило программу CORAL [19]. Эта аббревиатура раскрывается как Collaboration трёх национальных лабораторий Oak Ridge (ORNL), Argonne (ANL) и Livermore (LLNL). В рамках этой программы DOE объявило конкурс на три системы преэксафлопсного класса, Summit (ORNL, 150-300 петафлопс), Aurora (ANL, 180 петафлопс), и Sierra (LLNL, 100+ петафлопс). Суперкомпьютеры должны в 2017 году прийти на смену Titan (ORNL, 27 петафлопс), Mira (ANL, 10 петафлопс) и Sequoia (LLNL, 20 петафлопс). В экспертную группу включены также на равных представители Los Alamos NL и Sandia NL.

Цель программы CORAL — обеспечить удержание лидерства США в развитии суперкомпьютеров. Уже решено, что request for proposal (RFP) выпускает LLNL. На основании конкурса заключаются два разовых договора на инжиниринг и три контракта на покупку суперкомпьютеров. Оба договора на инжиниринг управляются совместно тремя лабораториями. Для уменьшения рисков и чтобы удовлетворить широкому набору требований предусмотрен выбор двух путей разработки архитектуры суперкомпьютеров, уже объявлено, что Summit и Sierra будут работать на процессорах IBM POWER9 и видеокартах NVIDIA Volta, Aurora — на новом поколении сопроцессоров Intel Xeon Phi, при этом Intel будет вести разработку совместно с Cray.

На этапе подготовки решений все предложения докладываются на заседаниях Advanced

Scientific Computing Advisory Committee (ASCAC) Министерства энергетики США. В состав ASCAC входят 9 профессоров 9 ведущих университетов США, в том числе председатель ASCAC, и по одному представителю Microsoft, Google, LLNL, SNL и Pittsburgh Supercomputing Center. Всего 14 человек.

Представителей лабораторий-заказчиков здесь нет.

Поставлены цели, которым должна удовлетворять производительность новой системы, и перечислены тестовые приложения, по которым проверяется достижение этих целей.

Для выделенного режима, т.е. когда масштабируемость под одну задачу охватывает до 90% системы, поставлена цель добиться $S > 4.0$ при увеличении размерности задачи: числа атомов, узлов сетки и др., где S — отношение производительности новой системы к производительности текущей системы. Тестовыми приложениями выбраны: первопринципные расчеты основного состояния (LSMS), квантовая молекулярная динамика (QBOX), гравитационная задача N тел (НАСС), термогидромеханика (Nekbone). В скобках указаны рекомендуемые пакеты программ.

Для многозадачного режима, занимающего всю систему при одновременном запуске не менее 24 задач, цель поставлено $S > 6.0$. Тестовыми приложениями выбраны: квантовый метод Монте-Карло (QMCPACK), классическая молекулярная динамика (NAMD), динамика атмосферы (CAM-SE), многосеточный солвер (AMG2013), метод конечных элементов (MiniFE), гидродинамика ударно-волновых явлений (LULESH), перенос излучения (неструктурированные сетки) (UMT2013), перенос излучения (структурированные сетки) (SNAP), метод Монте-Карло (задачи переноса) (MCB).

Требования по LINPACK и месту в Топ-500 вообще не фигурируют.

Кроме этого, в ORNL выбрали 13 широко использующихся программных кодов из различных областей для портирования и оптимизации приложений на новый гибридный суперкомпьютер Summit. Разработчики получают ранний доступ к суперкомпьютеру и техническую поддержку от IBM и NVIDIA.

После определения победителей конкурса предусмотрено многолетнее сотрудничество между лабораторией-заказчиком и компанией-победительницей с целью достижения наивысшей производительности. Каждая из трёх лабораторий самостоятельно проводит переговоры по заключению и управлению контрактом на покупку суперкомпьютера и может выставлять свои специальные требования к контракту.

4. Заключение

Рассмотрены примеры задач, обладающих масштабируемостью на суперкомпьютеры эксафлопсного класса, для перспективных задач энергетики, физики, химии и других естественных наук. Развивается подход, который позволяет определить, может ли решение той или иной задачи загрузить одновременно сотни тысяч или даже миллионы вычислительных ядер в одном расчете и нужно ли это. Основная идея заключается, с одной стороны, в проверке масштабируемости алгоритмов при переходе к большему числу вычислительных ядер, а, с другой, в теоретической оценке размеров физической системы, рассмотрение которой требуется для решения поставленной задачи. Развиваемый подход позволяет установить перспективу нарастания сложности задач, принадлежащих к одному научному направлению, возможность решения которых будет открываться с прогрессом развития суперкомпьютеров.

Сегодня для удержания лидерства США в развитии суперкомпьютеров Министерство энергетики США (DOE) запустило на 2013-2017 годы программу CORAL. Поставлены цели, которым должна удовлетворять производительность новых систем, и перечислены тестовые приложения, ориентированные на решение ясно определенного списка прикладных задач, по которым проверяется достижение этих целей.

Пример текущего распределения суперкомпьютеров в Германии показывает осторож-

ность, с которой идет создание больших гибридных систем на GPU-ускорителях в Европе. Несмотря на активное развитие новых технологий, наибольшую эффективность для решения задач имеют обычные процессоры.

В данной ситуации отношение к развитию суперкомпьютеров в России требует определенного переосмысления. Главным является переход к ситуации, когда заказчиками на разработку новых суперкомпьютеров станут квалифицированные пользователи, владеющие задачами, которые могут эти суперкомпьютеры эффективно загрузить.

В действительности, эта весьма давняя традиция — два примера из практики IBM прошедших десятилетий. При разработке архитектуры суперкомпьютеров IBM BlueGene/L задачи молекулярного моделирования рассматривались в качестве главного приоритета [2]. А в 1970-ые годы Энрико Клементи, разработчик квантово-химической программы IBMOL, публиковавший тогда десятки статей в лучших физических журналах, тестировал на своих задачах новые модели IBM в процессе их доведения.

Литература

1. Sadovnichy V., Tikhonravov A., Voevodin V., Opanasenko V. “Lomonosov”: supercomputing at Moscow State University // Contemporary High Performance Computing: From Petascale toward Exascale: Vetter J.S. Ed. CRC Press: Boca Raton, FL, 2013. С. 283—307.
2. IBM Journal of Research and Development, 2005:
URL: <http://www.research.ibm.com/journal/rd49-23.html> (дата обращения: 14.06.2015).
3. INCITE program: URL: <http://www.doeleadershipcomputing.org/> (дата обращения: 14.06.2015).
4. Стегайлов В.В., Норман Г.Э. Проблемы развития суперкомпьютерной отрасли в России: взгляд пользователя высокопроизводительных систем // Программные системы: теория и приложения: электрон. научн. журн. 2014. Т. 5, N. 1(19). С. 111–152.
URL: http://psta.psiras.ru/read/psta2014_1_111-152.pdf (дата обращения: 14.06.2015).
5. Норман Г.Э., Стегайлов В.В. Стохастическая теория метода классической молекулярной динамики // Математическое моделирование. 2012. Т. 24, N. 6. С. 3–44.
6. Янилкин А.В., Жилиев П.А., Куксин А.Ю., Норман Г.Э., Писарев В.В., Стегайлов В.В. Применение суперкомпьютеров для молекулярно-динамического моделирования процессов в конденсированных средах // Вычислительные методы и программирование. 2010. Т. 11, С. 111–116.
7. Жилиев П.А., Стегайлов В.В. Ab initio молекулярная динамика: перспективы использования многопроцессорных и гибридных суперЭВМ // Вычислительные методы и программирование. 2012. Т. 13, С. 37–45.
8. Eckhardt W. et al. 591 TFLOPS multi-trillion particles simulation on SuperMUC // Supercomputing. Vol. 7905, Springer Berlin Heidelberg, 2013. P. 1–12.
9. Куксин А.Ю., Ланкин А.В., Морозов И.В., Норман Г.Э., Орехов Н.Д., Писарев В.В., Смирнов Г.С., Стариков С.В., Стегайлов В.В., Тимофеев А.В. ЗАЧЕМ и КАКИЕ нужны суперкомпьютеры эксафлопсного класса? Предсказательное моделирование свойств и многомасштабных процессов в материаловедении // Программные системы: теория и приложения: электрон. научн. журн. 2014. Т. 5, N. 1(19). С. 191–244.
URL: http://psta.psiras.ru/read/psta2014_1_191-244.pdf (дата обращения: 14.06.2015).

10. Подрыга В.О., Поляков С.В., Пузырьков Д.В. Суперкомпьютерное молекулярное моделирование термодинамического равновесия в микросистемах газ-металл // Вычислительные методы и программирование. 2015. Т. 16, N. 1. С. 123–138.
11. Smirnov G.S., Stegailov V.V. Melting and superheating of sI methane hydrate: Molecular dynamics study // Journal of Chemical Physics. 2012. V. 136, N. 4, P. 044523.
12. Smirnov G.S., Stegailov V.V. Toward Determination of the New Hydrogen Hydrate Clathrate Structures // Journal of Physical Chemistry Letters. 2013. V. 4, P. 3560-3564.
13. Bai P., Tsapatsis M., Siepmann J.I. TraPPE-zeo: Transferable potentials for phase equilibria force field for all-silica zeolites // J. Phys. Chem. C. 2013. V. 117, P. 24375-24387.
14. Chen B., Siepmann J.I. Transferable potentials for phase equilibria. 3. Explicit-hydrogen description of normal alkanes // J. Phys. Chem. B. 1999. V. 103, P. 5370-5379.
15. Yoshioka T., Tanaka J., Tusru T., Asaeda M. Transport Properties of Condensable and Non-condensable Gas Mixtures through Microporous Silica Membranes Studied with Molecular Dynamics Simulation // Asian Pacific Confederation of Chemical Engineering congress program and abstracts. 2004. P. 741.
16. Frankland S.J.V et al. The stress-strain behavior of polymer-nanotube composites from molecular dynamics simulation // Compos. Sci. Technol. 2003. V. 63, N. 11, P. 1655-1661.
17. Gavrilov A. A., Chertovich A. V., Khalatur P. G., Khokhlov A. R. Effect of nanotube size on the mechanical properties of elastomeric composites // Soft Matter. 2013. V. 9, N. 15. P. 4067.
18. Morozov I.V., Kazennov A.M., Bystryi R.G., Norman G.E., Pisarev V.V., Stegailov V.V. Molecular dynamics simulations of the relaxation processes in the condensed matter on GPUs // Computer Physics Communications. 2011. V. 182, P. 1974–1978.
19. Collaboration of Oak Ridge, Argonne, and Livermore (CORAL):
URL: <http://energy.gov/downloads/fact-sheet-collaboration-oak-ridge-argonne-and-livermore-coral> (дата обращения: 14.06.2015).

Perspectives of exascale computing for novel energy technologies

Vyacheslav Vecher, Nikolay Kondratyuk, Genri Norman, Nikita Orekhov, Vasiliy Pisarev, Grigory Smirnov and Vladimir Stegailov

Keywords: energy technologies, multiscale models, molecular dynamics

The development of novel energy technologies essentially relies upon multiscale modelling and simulation which success crucially depends on the progress in supercomputing. In this talk we review several vivid examples of the problems that we solve using atomistic and multiscale models and the best supercomputers in Russia: properties gas hydrates and gas condensate, rheology of organic liquids and development of polymer materials. The wider context is presented as well with the corresponding examples from the DOE INCITE program. On the basis of the case studies considered we discuss the issues of hardware efficiency, software scalability, input-output problems and data processing.