

УДК 539.1+537.31

**AB INITIO МОЛЕКУЛЯРНАЯ ДИНАМИКА: ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ
МНОГОПРОЦЕССОРНЫХ И ГИБРИДНЫХ СУПЕРЭВМ****П. А. Жилиев¹, В. В. Стегайлов¹**

Представлен обзор распараллеливания алгоритмов ab initio молекулярной динамики на основе теории функционала электронной плотности в базисе плоских волн. Проанализированы требования к балансу вычислительной мощности узлов и коммутационной сети суперЭВМ с точки зрения достижения максимальной эффективности для примеров особенно требовательных в вычислительном отношении задач физики разогретого плотного вещества. Описана альтернативная стратегия параллелизации в вейвлетном базисе и выигрыш в производительности при использовании гибридных вычислительных систем в этом случае. Работа выполнена при поддержке гранта РФФИ (код проекта 11-01-12131-офи-м-2011). Статья рекомендована к публикации Программным комитетом Международной научной конференции “Параллельные вычислительные технологии” (ПАВТ-2012; <http://agora.guru.ru/pavt2012>).

Ключевые слова: ab initio расчеты, молекулярная динамика, параллельные вычисления, суперЭВМ, проводимость.

1. Введение. Продолжающееся непрерывное развитие теоретических и вычислительных методов атомистического моделирования на протяжении последних десятилетий обеспечивает основу средств анализа и прогноза для физики конденсированного состояния, материаловедения, химии, молекулярной биологии и нанотехнологий. Свойства материалов определяются откликом многоатомной системы на изменение внешних условий. Вообще говоря, для теоретического описания данного отклика не может быть достаточно методов кинетики и теории сплошных сред, а нужен выход на атомистический уровень описания.

Развитие атомистических моделей вещества изначально шло несколькими практически независимыми путями. Развитие методов квантовой механики позволило с 30-х гг. XX века проводить расчеты электронной структуры молекулярных систем небольшого размера. Совершенствование данных методов привело к созданию чрезвычайно развитой к настоящему времени области под названием квантовая химия. Когда в результате создания компьютеров в 50-х гг. XX века появился способ проведения больших объемов математических расчетов, одним из первых их применений стали задачи статистической механики, решение которых предполагалось получать путем непосредственного расчета динамики многочастичной системы на основе классических уравнений движения с простейшими модельными межатомными потенциалами. Подобный метод решения задач равновесной и неравновесной статистической механики получил название метода молекулярной динамики (МД).

Метод МД в настоящее время является незаменимым методом исследования конденсированного состояния [1]. Принципиальным условием его успешного использования является наличие адекватных моделей потенциалов межатомного взаимодействия. Межатомное взаимодействие определяется электронной структурой вещества, поэтому современные методы создания потенциалов основаны на квантово-механических расчетах [2, 3]. Наряду с методами квантовой химии (применимыми для изолированных молекул и кластеров), важную роль при этом сыграли методы, развитые в квантовой физике твердого тела для периодических систем. По мере роста вычислительных возможностей все чаще рассматриваются ab initio МД модели, не использующие эмпирические потенциалы.

Общей теоретической основой квантовых методов являются методы, основанные на приближении Хартри–Фока, и методы теории функционала электронной плотности (ТФП). Для представления электронной структуры в квантово-химических расчетах используются локализованные базисы [4, 5], в то время как для периодических систем, в первую очередь, используется базис плоских волн [6, 7], что приводит к совершенно разным вычислительным алгоритмам, в том числе и к разным стратегиям их распараллеливания. Преимуществом базиса плоских волн является возможность систематического повышения

¹ Объединенный институт высоких температур РАН (ОИВТ РАН), ул. Ижорская, 13/2, 125412, Москва; П. А. Жилиев, мл. научн. сотр., e-mail: PeterZhilyaev@gmail.com; В. В. Стегайлов, зав. лабораторией, e-mail: stegailov@gmail.com

точности расчетов путем плавного увеличения размера базиса. Использование быстрого преобразования Фурье обеспечивает высокую скорость расчетов в этом базисе. С использованием базиса плоских волн для ab initio МД на настоящее время доступны системы до тысяч атомов и времена до десятков пикосекунд [6–10].

Настоящая статья посвящена параллельным алгоритмам в рамках ТФП в базисе плоских волн. Заметим, что локализованные базисы также эффективно используются для ab initio МД и обеспечивают точность, сопоставимую с базисом плоских волн [11]. Перспективным направлением является использование смешанных базисов [12].

2. Теория функционала электронной плотности. В рамках теории функционала электронной плотности (ТФП) полная энергия системы взаимодействующих электронов в поле классических ядер $\{\mathbf{r}_I\}$ представляется в виде $\min_{\Psi_0} \{\langle \Psi_0 | H_e | \Psi_0 \rangle\}$ как минимум функционала Кона–Шэма

$$E^{KS}[\{\phi_i\}] = T_s[\{\phi_i\}] + \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} V_H(\mathbf{r}) n(\mathbf{r}) + E_{\text{xc}}[n] + E_{\text{ions}}(\mathbf{r}^N), \quad (1)$$

который явным образом зависит от набора вспомогательных функций (орбиталей Кона–Шэма), удовлетворяющих условию ортонормированности $\langle \phi_i | \phi_j \rangle = \delta_{ij}$.

В рамках данного подхода решение многоэлектронной квантово-механической задачи упрощается путем перехода от описания электронов в виде волновых функций к описанию в терминах одночастичной электронной плотности. Метод ТФП отличается от метода Томаса–Ферми тем, что представление электронов в виде волновых функций ϕ_i все-таки сохраняется. Это позволяет учесть обменно-корреляционные эффекты, которые не могут быть описаны в модели Томаса–Ферми.

Распределение электронной плотности рассчитывается на основе одного детерминанта Слэтера, построенного по заполненным орбиталям с учетом чисел заполнения $n(\mathbf{r}) = \sum_i^{\text{occ}} f_i |\phi_i(\mathbf{r})|^2$.

Первое слагаемое в функционале Кона–Шэма (1) соответствует кинетической энергии воображаемой системы невзаимодействующих электронов $T_s[\{\phi_i\}] = \sum_i^{\text{occ}} f_i \left\langle \phi_i \left| -\frac{1}{2} \nabla^2 \right| \phi_i \right\rangle$, состоящей из того же числа электронов и находящейся в поле того же внешнего потенциала, что и исходная система с полным взаимодействием. Второе слагаемое соответствует внешнему потенциалу $V_{\text{ext}}(\mathbf{r})$ — в большинстве случаев он соответствует сумме потенциалов классических ядер. Третье слагаемое соответствует классической электростатической энергии и рассчитывается на основе потенциала Хартри $V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$, связанного, в свою очередь, с распределением электронной плотности уравнением Пуассона $\nabla^2 V_H(\mathbf{r}) = -4\pi n(\mathbf{r})$.

Предпоследнее слагаемое в функционале Кона–Шэма представляет собой обменно-корреляционный функционал $E_{\text{xc}}[n]$, который одновременно включает в себя обменные и корреляционные эффекты. По сути, функционал $E_{\text{xc}}[n]$ представляет собой разность между точной энергией многоэлектронной системы и ее разложением по методу Кона–Шэма на три предшествующих слагаемых. Последнее слагаемое в (1) соответствует энергии взаимодействия зарядов ядер.

Путем минимизации функционала Кона–Шэма (1) при условии фиксированного полного числа электронов с учетом условия ортонормированности орбиталей получается система уравнений Кона–Шэма

$$\left\{ -\frac{1}{2} \nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + \frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})} \right\} \phi_i(\mathbf{r}) = \sum_j \Lambda_{ij} \phi_j(\mathbf{r}),$$

которая может быть записана в виде $\left\{ -\frac{1}{2} \nabla^2 + V_{KS}(\mathbf{r}) \right\} \phi_i(\mathbf{r}) = \sum_j \Lambda_{ij} \phi_j(\mathbf{r})$, или $H^{KS} \phi_i(\mathbf{r}) = \sum_j \Lambda_{ij} \phi_j(\mathbf{r})$,

если ввести одноэлектронный гамильтониан H^{KS} с локальным потенциалом $V_{KS}(\mathbf{r})$. Заметим, что данный одноэлектронный гамильтониан H^{KS} эффективным образом учитывает многочастичные эффекты за счет обменно-корреляционного потенциала $\frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})} = V_{\text{xc}}(\mathbf{r})$.

Унитарное преобразование в пространстве заполненных орбиталей дает каноническую форму системы уравнений Кона–Шэма

$$H^{KS} \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \quad (2)$$

Для того чтобы получить орбитали и распределение одноэлектронной плотности для основного состояния многоэлектронной системы, система уравнений (2) должна решаться самосогласованным образом.

Для применения ТФП в расчетах реальных систем критическим является используемое приближение для аппроксимации неизвестного обменно-корреляционного функционала (приближение локальной плотности, градиентное приближение, комбинированные подходы).

С вычислительной точки зрения вычисления электронной структуры в рамках теории функционала электронной плотности могут быть существенно облегчены путем введения в модель псевдопотенциалов. Данные псевдопотенциалы учитывают тот факт, что электроны, находящиеся на сильно связанных заполненных электронных оболочках атомов, практически не изменяют свое состояние при взаимодействии с другими атомами (при образовании и разрыве химических связей и т.п.) Таким образом, при построении моделей возможно разделение электронной на валентные, описываемые явно, и электроны сердцевины атома, описываемые эффективным псевдопотенциалом, заменяющим чисто кулоновский потенциал ядра. При конструировании псевдопотенциалов обычно накладывается условие соответствия волновых функций валентных электронов атома в полноразмерном и псевдопотенциальном представлении (рис. 1).

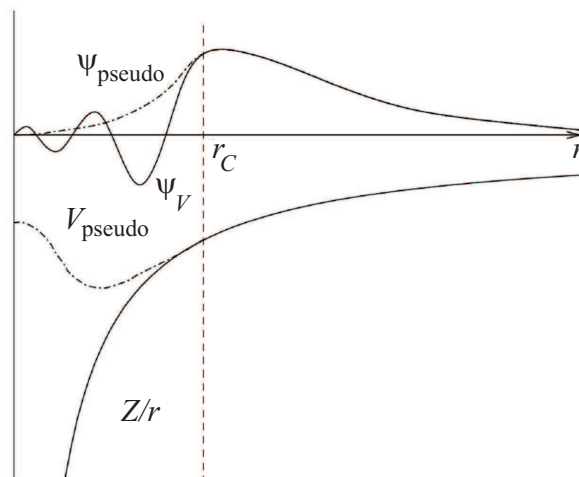


Рис. 1. Качественная схема, иллюстрирующая принцип введения псевдопотенциала для описания эффективного влияния внутренних электронов атома на валентные электроны

3. Расчеты в базисе плоских волн. Основа численного решения задачи о вычислении электронной структуры в рамках ТФП состоит в выборе базиса для представления волновых функций.

Одно из ключевых преимуществ базиса плоских волн заключается в том, что его реализация основана на использовании преобразования Фурье, для которого существуют высокоэффективные численные алгоритмы (быстрое преобразование Фурье).

Входными параметрами задачи являются геометрия системы (положения атомов, размеры расчетной ячейки), модели псевдопотенциала для атомов в модели, обрезка по энергии базиса плоских волн, параметры итерационного алгоритма нахождения самосогласованного решения системы уравнений Кона–Шэма.

Точность решения задачи о вычислении электронной структуры в рамках ТФП в первую очередь зависит от выбора модели псевдопотенциала и модели обменно-корреляционного функционала. Границы применимости существующих моделей определяются сравнением с экспериментальными данными для определенных классов веществ и соединений. Описанные в настоящей статье алгоритмы реализованы в кодах CPMD [6] и ABINIT [7].

3.1. Укрупненная блок-схема последовательного алгоритма. Введем следующие обозначения: N_{at} — число атомов, N_P — число проекторов, N_b — число электронных состояний, N_{PW} — число плоских волн, N_D — число плоских волн для представления плотности и потенциалов, N_x, N_y, N_z — число точек сетки в направлениях x, y и z , $N = N_x \times N_y \times N_z$ — полное число точек сетки.

В таблице приведены характерные значения указанных переменных для следующих двух систем: пример модели кристалла кремния с обрезкой по энергии на 13 Ry и нелокальностью псевдопотенциала s-типа и пример водной системы (значения приводятся из расчета на одну молекулу) с обрезкой по энергии на 70 Ry и нелокальностью s-типа псевдопотенциала кислорода и локальным псевдопотенциалом для водорода.

3.2. Структура представления данных. Основные величины в расчете с использованием псевдопотенциалов в базисе плоских волн либо не зависят от размера системы, либо увеличиваются линейно, либо квадратично. Примерами величин первого типа являются матрица элементарной ячейки h и значение энергии обрезки E_{cut} .

Следующие переменные растут линейно с увеличением размера системы: $r(3, N_{at})$ — координаты атомов, $v(3, N_{at})$ — скорости атомов, $f(3, N_{at})$ — силы, действующие на атомы, $g(3, N_{PW})$ — индексы

Характерные параметры моделей в рамках ТФП в базисе плоских волн

	Кремний	Вода
N_{at}	1	3
N_P	1	1
N_b	2	4
N_{PW}	53	1000
N_D	429	8000
N	1728	31250

плоских волн, $ipg(3, N_{PW})$ — отображение G-векторов (положительная часть), $img(3, N_{PW})$ — отображение G-векторов (отрицательная часть), $rhog(N_{PW})$ — плотности (n, n_c, n_{tot}) в пространстве Фурье, $vpot(N_{PW})$ — потенциалы (V_{loc}, V_{xc}, V_H) в пространстве Фурье, $n(N_x, N_y, N_z)$ — плотности (n, n_c, n_{tot}) в реальном пространстве, $v(N_x, N_y, N_z)$ — потенциалы (V_{loc}, V_{xc}, V_H) в реальном пространстве, $vps(N_D)$ — локальный псевдопотенциал, $rpc(N_D)$ — заряды ядра атомов, $pro(N_{PW})$ — проекторы для нелокальной части псевдопотенциала.

Величины vps , rpc и pro , относящиеся к псевдопотенциалу, зависят линейно от размера системы, но в то же время зависят от числа типов атомов в модели. В дальнейшем описании без потери общности будем предполагать, что в системе один тип атомов. Большая часть оперативной памяти уходит на хранение величин, зависящих квадратично от размера системы: $eigr(N_D, N_{at})$ — структурные факторы, $fnl(N_P, N_b)$ — перекрытия проекторов и состояний, $dfnl(N_P, N_b, 3)$ — производная от fnl , $smat(N_b, N_b)$ — матрицы перекрытия между состояниями, $cr(N_{PW}, N_b)$ — состояния в фурье-пространстве, $cv(N_{PW}, N_b)$ — скорости состояний в фурье-пространстве, $cf(N_{PW}, N_b)$ — силы состояний в фурье-пространстве.

Структурные факторы $eigr$ и величины cr , cv , cf , относящиеся к волновым функциям, представляют собой комплексные числа. Другие величины могут храниться в виде действительных чисел, если учитывать множитель $(-i)^l$.

3.3. Вычислительные алгоритмы. Большая часть вычислений в коде, основанном на представлении плоских волн, производится в нескольких базовых алгоритмах. Эти алгоритмы могут быть описаны на псевдокоде. Там, где это возможно, применяются стандартные процедуры линейной алгебры (библиотека BLAS). Первый алгоритм — расчет структурных факторов:

```

MODULE StructureFactor
  FOR i=1:Nat
    s(1:3) = 2 * PI * MATMUL[html(1:3,1:3), r(1:3,i)]
    dp(1:3) = CMLPX[COS[s(1:3)], SIN[s(1:3)]]
    dm(1:3) = CONJG[dp(1:3)]
    e(0,1:3,i) = 1
  FOR k=1:gmax
    e(k,1:3,i) = e(k-1,1:3,i) * dp
    e(-k,1:3,i) = e(-k+1,1:3,i) * dm
  END
  FOR j=0:ND
    eigr(j,i) = e(g(1,j),1,i) * e(g(2,j),2,i) * e(g(3,j),3,i)
  END
END

```

Алгоритм расчета перекрытий проекторов нелокальной части псевдопотенциала и волновых функций в фурье-пространстве масштабируется по времени исполнения как $N_P N_b N_{PW}$. Два следующих алгоритма проводят преобразование зарядовой плотности из фурье-пространства в прямое пространство и обратно:

```

MODULE INVFFT
  scr(1:Nx,1:Ny,1:Nz) = 0
  FOR i=1:ND
    scr(ipg(1,i),ipg(2,i),ipg(3,i)) = rhog(i)
    scr(img(1,i),img(2,i),img(3,i)) = CONJG[rhog(i)]
  END
  CALL FFT3D("INV",scr)
  n(1:Nx,1:Ny,1:Nz) = REAL[scr(1:Nx,1:Ny,1:Nz)]

MODULE FWFFT
  scr(1:Nx,1:Ny,1:Nz) = n(1:Nx,1:Ny,1:Nz)
  CALL FFT3D("FW",scr)
  FOR i=1:ND
    rhog(i) = scr(ipg(1,i),ipg(2,i),ipg(3,i))
  END

```

На рис. 2 иллюстрируются алгоритмы расчета плотности и применение локального потенциала. Расчет плотности требует N_b трехмерных преобразований Фурье. Для применения локального потенциала

необходимо одно фурье-преобразование на каждое состояние. Число операций в этих алгоритмах пропорционально $N_b N \log(N)$. В большинстве приложений эти алгоритмы занимают больше всего вычислительного времени.

3.4. Принципы работы параллельного алгоритма. При распараллеливании алгоритмов для систем с распределенной памятью в рамках представленного метода расчета электронной структуры, основанного на базисе плоских волн, главное внимание уделяется минимизации обменов между вычислительными узлами.

В этой связи крайне важно распределение данных по вычислительным узлам. В первую очередь это относится к информации о волновых функциях. Распределение данных является статическим, т.е. оно не изменяется в процессе расчета. Во всех частях программы, где невозможна параллелизация по плоским волнам, используется параллелизация по числу атомов и состояниям.

На каждый вычислительный узел p приписывается определенное число плоских волн, атомов, электронных состояний и точек сетки в реальном пространстве: N_{at} — число атомов, N_P^p — число проекторов, N_b^p — число электронных состояний, N_{PW}^p — число плоских волн, N_D^p — число плоских волн для представления плотности и потенциалов, N_x^p, N_y^p, N_z^p — число точек сетки в направлениях x, y и z , полное число точек сетки $N^p = N_x^p \times N_y^p \times N_z^p$.

Сетка по пространству распределяется по процессорам только в направлении x (рис. 3). Это связано с производительностью алгоритма фурье-преобразования. Конкретные алгоритмы, основанные на распределении указанных данных, незначительно влияют на общую эффективность распараллеливания кода.

Следующие структуры данных реплицируются на всех вычислительных узлах: $r(3, N_{at})$ — координаты атомов, $v(3, N_{at})$ — скорости атомов, $f(3, N_{at})$ — силы, действующие на атомы, $fnl(N_P, N_b)$ — перекрытия проекторов и состояний, $smat(N_b, N_b)$ — матрицы перекрытия между состояниями.

Следующие структуры данных распределяются по процессорам: $g(3, N_{PW}^p)$ — индексы плоских волн, $ipg(3, N_{PW}^p)$ — отображение G-векторов (положительная часть), $img(3, N_{PW}^p)$ — отображение G-векторов (отрицательная часть), $rhog(N_{PW}^p)$ — плотности (n, n_c, n_{tot}) в пространстве Фурье, а также $v_{pot}(N_{PW}^p)$ — потенциалы (V_{loc}, V_{xc}, V_H) в пространстве Фурье, $n(N_x^p, N_y^p, N_z^p)$ — плотности (n, n_c, n_{tot}) в реальном пространстве, $v(N_x^p, N_y^p, N_z^p)$ — потенциалы (V_{loc}, V_{xc}, V_H) в реальном пространстве, $vps(N_D^p)$ — локальный псевдопотенциал, $rpc(N_D^p)$ — заряды ядра атомов, $pro(N_{PW}^p)$ — проекторы для нелокальной части псевдопотенциала, $eigr(N_D^p, N_{at})$ — структурные факторы, $dfnl(N_P, N_b^p, 3)$ — производная от fnl , $cr(N_{PW}^p, N_b)$ — состояния в фурье-пространстве, $cv(N_{PW}^p, N_b)$ — скорости состояний в фурье-пространстве, $cf(N_{PW}^p, N_b)$ — силы состояний в фурье-пространстве.

При распределении плоских волн по вычислительным узлам должны выполняться несколько условий. Все узлы должны иметь приблизительно равное число плоских волн. Если плоская волна, соответствующая обрезке волной функции, находится на некотором узле, то та же плоская волна должна быть на

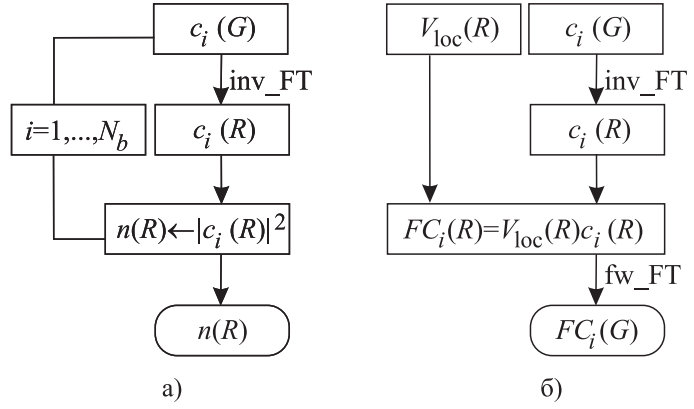


Рис. 2. Схема алгоритмов расчета плотности (а) и применения локального потенциала для расчета сил, действующих на волновую функцию (б)

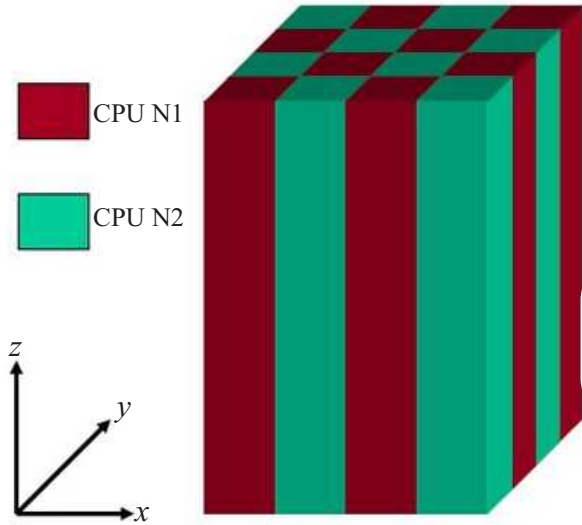


Рис. 3. Распределение данных для проведения 3D FFT. Для каждой волновой функции ее коэффициенты распределяются по G-векторам в направлении z (на рисунке пример для 2 процессоров)

том же процессоре для обрезки плотности. Распределение плоских волн должно быть организовано таким образом, чтобы в начале и конце трехмерного преобразования Фурье не требовалось дополнительных обменов данными. Специальное внимание должно уделяться вычислительному узлу (p_0), на котором находится компонента $G = 0$, в связи с особенностями расчета перекрытий с ее участием. Главными для распараллеливания алгоритмов являются три коммуникационные процедуры. Все они — коллективные, т.е. задействуют все вычислительные узлы, что подразумевает проведение синхронизации при выполнении этих процедур. Этими процедурами являются Broadcast, GlobalSum и MatrixTranspose. Процедура Broadcast пересылает данные с одного узла (px) на все остальные $x^p \leftarrow x^{px}$.

Процедура GlobalSum заменяет элемент данных на каждом узле суммой этой величины по всем узлам $x^p \leftarrow \sum_p x^p$.

Процедура MatrixTranspose производит транспонирование матрицы $x(p, :) \leftarrow x(:, p)$.

На параллельной вычислительной системе, состоящей из P процессоров, с характерным временем задержки при коммуникации (латентностью) t_L и пропускной способностью B , на выполнение одного запуска указанных процедур уходит следующее время: Broadcast — $\log_2[P]\{t_L + N/B\}$, GlobalSum — $\log_2[P]\{t_L + N/B\}$, MatrixTranspose — $Pt_L + N/(PB)$.

С учетом описанных правил распределения данных параллелизация большинства алгоритмов достаточно проста.

Процедуры, для распараллеливания которых требуется большое число изменений, содержат как составную часть преобразование Фурье. Вводится процедура *marxy*, которая обеспечивает непрерывное распределение данных в памяти на каждом вычислительном узле:

```

MODULE INVFFT
scr(1:Nx,1:NDpencil) = 0
FOR i=1:NpD
  scr1(ipg(1,i),mapxy(ipg(2,i),ipg(3,i))) = rhog(i)
  scr1(img(1,i),mapxy(img(2,i),img(3,i))) = CONJG[rhog(i)]
END
CALL ParallelFFT3D("INV",scr1,scr2)
n(1:Npx,1:Ny,1:Nz) = REAL[scr2(1:Npx,1:Ny,1:Nz)]

MODULE FWFFT
scr2(1:Npx,1:Ny,1:Nz) = n(1:Npx,1:Ny,1:Nz)
CALL ParallelFFT3D("FW",scr1,scr2)
FOR i=1:NpD
  rhog(i) = scr1(ipg(1,i),mapxy(ipg(2,i),ipg(3,i)))
END

```

Из-за отображения направлений y и z на одно направление при преобразовании Фурье массивы ввода и вывода параллельного преобразования Фурье должны иметь различный вид. Процедура трехмерного параллельного фурье-преобразования может быть сконструирована из нескольких одномерных фурье-преобразований и параллельных транспонирований матрицы:

```

MODULE ParallelFFT3D(tag,a,b)
IF (tag == "INV") THEN
  CALL MLTFFT1D(a)
  CALL ParallelTranspose("INV",b,a)
  CALL MLTFFT2D(b)
ELSE
  CALL MLTFFT2D(b)
  CALL ParallelTranspose("FW",b,a)
  CALL MLTFFT1D(a)
END IF

```

4. Тесты параллельной производительности. На рис. 4–6 представлены результаты тестов на суперкомпьютерах МФТИ-60, BlueGene ВМК МГУ и “Ломоносов” НИВЦ МГУ. На рис. 4 показаны данные по ускорению расчетов молекулы N-метилформамида. Результаты тестов показывают, что даже расчеты такой небольшой системы могут эффективно распараллеливаться на десятки процессоров.

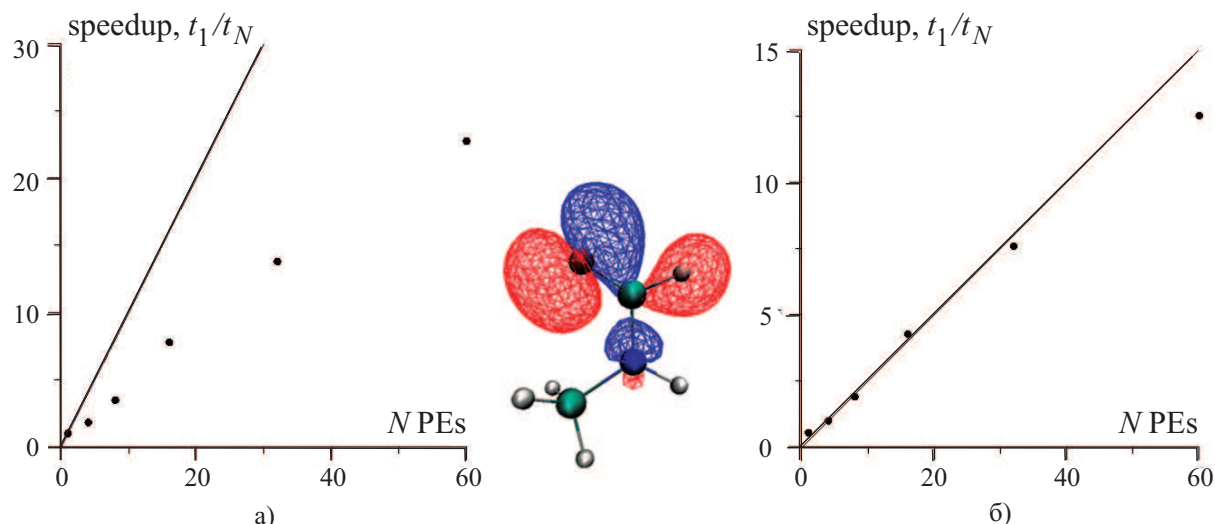


Рис. 4. Относительное ускорение расчетов на параллельной вычислительной системе для расчета электронной структуры восьмиатомной молекулы N-метилформамида (пакет CPMD): а) рост эффективности по отношению к однопроцессорному варианту расчета, б) рост эффективности по отношению к одному полному узлу вычислительного кластера, имеющему 4 вычислительных ядра. Расчеты проведены на кластере МФТИ-60.

Сплошная линия показывает случай стопроцентного распараллеливания. В центре показана визуализация высшей заполненной орбитали (HOMO) основного состояния S_0 N-метилформамида C_2H_5NO

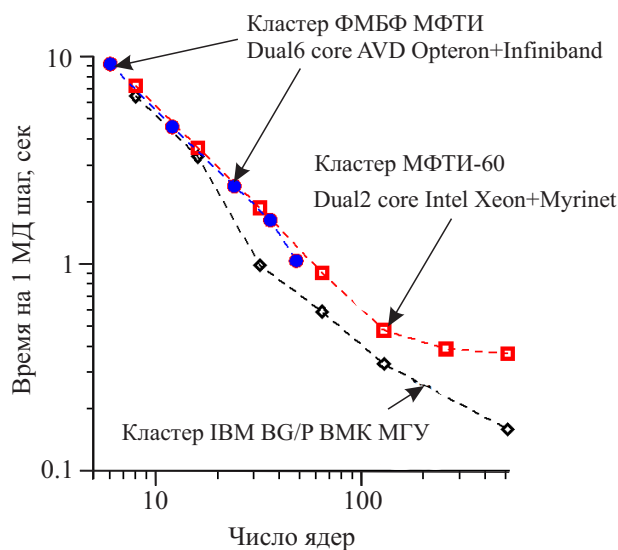


Рис. 5. Результаты тестирования производительности кода CPMD на различных параллельных системах на стандартном примере системы с 32 молекулами воды

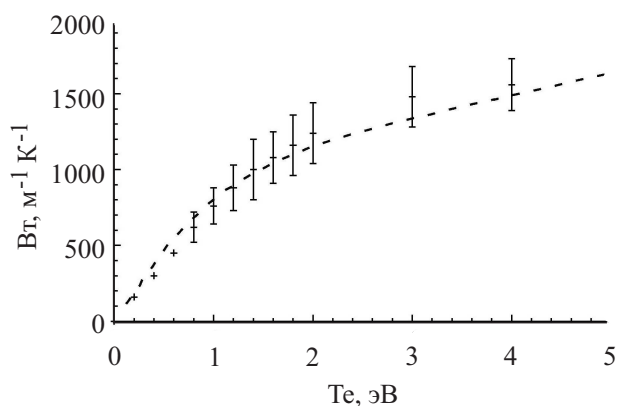


Рис. 6. Теплопроводность жидкого алюминия ($T_i = 2000$ K) в двухтемпературном случае. Точки — данная работа, пунктир — расчет по кинетической модели [14]

Другим стандартным тестом параллельной эффективности является расчет системы 32 молекул воды (рис. 5). Результаты тестов демонстрируют существенное преимущество наличия интерконнекта с высокой связностью для ускорения коллективных операций. Наличие подобного интерконнекта на суперкомпьютере BlueGene BMK МГУ обеспечивает эффективное масштабирование задачи до 512 вычислительных ядер (кроме того, значение имеет баланс скорости интерконнекта и вычислительной производительности процессоров). В то же время, падение параллельной эффективности на кластере МФТИ-60 (интерконнект Myrinet) происходит уже на 100 ядрах.

В качестве еще одного примера рассмотрим задачу расчета коэффициента электронной теплопроводности жидкого алюминия в двухтемпературном случае, когда температура электронной подсистемы (T_e) много больше температуры решетки (T_i). Мотивацией данной проблемы является описание свойств ме-

талла, находящегося в разогретом плотном состоянии, для моделирования лазерной абляции. Для этого необходимы численные значения кинетических коэффициентов металла с горячими электронами [13]. На рис. 6 представлены результаты расчета коэффициента электронной теплопроводности на основе формулы Кубо–Гринвуда для проводимости алюминия. Полученная зависимость коэффициента теплопроводности от электронной температуры хорошо согласуется с результатами кинетической модели [14]. Точность МД-расчетов (погрешности на графике) определяется усреднением по независимым конфигурациям, которые получаются путем расчета МД-траектории достаточной длины. Специфика данного типа расчетов состоит в том, что сходимость по числу атомов и по числу k -точек, покрывающих зону Бриллюэна, является достаточно медленной и для получения надежных данных необходимо использовать системы нескольких сотен атомов с числом k -точек вплоть до 10^3 , что предъявляет требования к необходимому суммарному объему оперативной памяти. Возможность расчетов электронной теплопроводности на основе *ab initio* МД и их точность существенным образом зависят от наличия эффективных параллельных алгоритмов и высокопроизводительных систем.

На рис. 7 показаны результаты тестирования масштабируемости расчетов электронной теплопроводности на основе первопринципной молекулярной динамики с использованием ТФП в базисе плоских волн (пакет ABINIT). Результаты демонстрируют очень хорошее ускорение расчетов вплоть до 100–200 ядер на кластере МФТИ-60 и кластере “Ломоносов” НИВЦ МГУ.

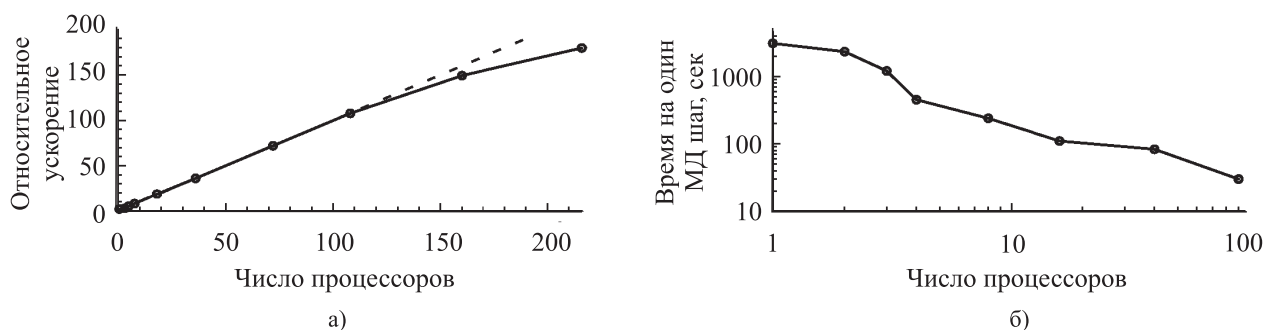


Рис. 7. (а) Относительное ускорение расчетов на параллельной вычислительной системе для расчета электронной структуры жидкого алюминия 108 атомов (пакет ABINIT). Расчеты проведены на кластере МФТИ-60. Сплошная линия показывает случай стопроцентного распараллеливания. (б) Результаты тестирования производительности кода ABINIT на суперкомпьютере “Ломоносов” НИВЦ МГУ на примере жидкого алюминия 108 атомов

5. Вейвлетный базис и выигрыш в производительности при использовании гибридных вычислительных систем. Среди современных высокопроизводительных систем в последнее время все чаще используются гибридные системы, включающие в себя графические ускорители (ГПУ — графическое процессорное устройство). Этому во многом способствовало появление новых сред программирования, таких как CUDA и OpenCL, предназначенных для создания и выполнения на ГПУ программ, не связанных с обработкой изображений. Гибридные системы на основе ГПУ зачастую показывают быстрое действие, в десятки или сотни раз превышающее быстрое действие систем на традиционных процессорах (ЦПУ — центральное процессорное устройство). В то же время создание программ, эффективно использующих вычислительные возможности ГПУ, — это достаточно трудоемкий процесс, требующий учета специфики аппаратной архитектуры, без которого производительность на конкретной задаче может оказаться существенно меньше теоретического предела для данного устройства.

В работе [15] представлена реализация полного кода для ТФП-расчетов электронной структуры на гибридных параллельных системах. Эта реализация осуществлена с использованием свободного программного кода на основе вейвлетов Добеши. Код показывает очень хорошую параллельную эффективность и обладает устойчивой сходимостью. В частности, код может работать на многих узлах, которые могут иметь или не иметь GPU. Показано, что с двойной точностью расчетов можно добиться значительного ускорения: до 20 раз для некоторых операций и до 6 раз для полного набора операций типичного расчета в рамках ТФП.

6. Заключение. Изложена математическая формулировка теории функционала электронной плотности в базисе плоских волн для расчетов электронной структуры и первопринципной молекулярной динамики. Описаны последовательные версии алгоритмов и их параллельные версии. Приведены результаты тестовых расчетов различных систем на кластерах МФТИ-60, BlueGene/BMK МГУ и “Ломоносов” НИВЦ МГУ. Кратко обсуждается альтернативная стратегия распараллеливания в вейвлетном базисе и

ускорение на гибридных архитектурах.

Приведенные примеры модельных задач (изолированная молекула, система молекул воды, металл в разогретом плотном состоянии) демонстрируют эффективное масштабирование данных алгоритмов на параллельных вычислительных системах без существенного снижения эффективности распараллеливания на 100–500 ядрах.

СПИСОК ЛИТЕРАТУРЫ

1. Янилкин А.В., Жиляев П.А., Куксин А.Ю. и др. Применение суперкомпьютеров для молекулярно-динамического моделирования процессов в конденсированных средах // Вычислительные методы и программирование. 2010. **11**, № 1. 115–120.
2. Van Duin A.C.T. et al. ReaxFF: A Reactive Force Field for hydrocarbons // J. Phys. Chem. A. 2001. **105**. 9396–9409.
3. Smirnova D.E., Starikov S.V., Stegailov V.V. Interatomic potential for uranium in a wide range of pressures and temperatures // J. Phys.: Cond. Mat. 2012. **24**. 015702.
4. Laikov D.N., Ustynyuk Yu.A. PRIRODA-04: a quantum chemical program suite. New possibilities in the study of molecular systems with the application of parallel computing // Russian Chemical Bulletin (International Edition). 2005. **54**, N 3. 820–826.
5. Granovsky A.A. Extended multi-configuration quasi-degenerate perturbation theory: the new approach to multi-state multi-reference perturbation theory // J. Chem. Phys. 2011. **134**. 214113.
6. Hutter J., Curioni A. Car-Parrinello molecular dynamics on massively parallel computers // Chem. Phys. Chem. 2005. **6**. 1788–1793.
7. Gonze X. et al. ABINIT: First-principles approach to material and nanosystem properties // Computer Physics Communications. 2009. **180**. 2582–2615.
8. Gygi F. et al. Practical algorithms to facilitate large-scale first-principles molecular dynamics // J. Phys.: Conf. Ser. 2009. **180**. 012074.
9. Goedecker S. et al. An efficient 3-dim FFT for plane wave electronic structure calculations on massively parallel machines composed of multiprocessor nodes // Computer Physics Communications. 2003. **154**. 105–110.
10. Bottin F. et al. Large-scale ab initio calculations based on three levels of parallelization // Computational Material Science. 2008. **42**. 329–336.
11. Gusso M. Study on the maximum accuracy of the pseudopotential density functional method with localized atomic orbitals versus plane-wave basis sets // J. Chem. Phys. 2008. **128**. 044102.
12. Lippert G., Hutter J., Parrinello M. A hybrid Gaussian and plane wave density functional scheme // Mol. Phys. 1997. **92**. 477–487.
13. Стариков С.В., Стегайлов В.В., Норман Г.Э., Фортков В.Е. и др. Лазерная абляция золота: эксперимент и атомистическое моделирование // Письма в ЖЭТФ. 2011. **93**. 719–725.
14. Иногамов Н.А., Петров Ю.В. Теплопроводность металлов с горячими электронами // ЖЭТФ. 2010. **137**, № 3. 505–529.
15. Genovese L. et al. Density functional theory calculation on many-cores hybrid central processing unit-graphic processing unit architectures // J. Chem. Phys. 2009. **131**. 034103.

Поступила в редакцию
24.02.2012