

МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ СВОЙСТВ ТВЕРДЫХ ТЕЛ С ПОМОЩЬЮ СОВРЕМЕННЫХ ПАКЕТОВ ПРОГРАММ НА СОВРЕМЕННЫХ СУПЕРКОМПЬЮТЕРАХ

Ю.С. Митрохин, С.А. Мельчуков, М.А. Ключков
Удмуртский государственный университет, г. Ижевск
Тел.: (3412) 91-60-89, e-mail: mit@uni.udm.ru

Современные методы моделирования в науке и технике требуют от исследователя решения целого ряда довольно сложных проблем. Это, прежде всего, наличие параллельных вычислительных систем или возможность выхода на такие системы по сети Интернет. Во-вторых, наличие на этих системах необходимых пакетов программ или возможность установки таких пакетов обычному пользователю. Эта задача может быть связана с необходимостью иметь права администратора, что не допускается на удаленных системах. Часто бывает, что на суперкомпьютере могут стоять не те трансляторы и библиотеки, которые требуются пользователю. Следующее условие успешной работы на удаленном суперкомпьютере – это наличие достаточной скорости передачи большого объема информации для работы в режиме on-line. Используемые пакеты программ должны быть хорошо документированы, а пользователь должен иметь достаточно высокую квалификацию, как в предметной области, так и в области программирования и вычислительной техники. Важное значение имеет загрузка вычислительной установки и политика администрации вычислительного центра по отношению к пользователю. Опуская вопросы финансирования, можно сделать вывод, что моделирование на современных удаленных суперкомпьютерах – довольно сложная задача, которая предполагает наличие достаточного опыта у пользователя и наличие необходимого программного обеспечения на вычислительном центре.

В настоящее время создано много хорошо документированных пакетов для моделирования атомной и электронной структур молекул, наноструктур и кристаллов. В таблице 1 приведен список пакетов программ установленных на суперкомпьютере Blue Gene в лаборатории им. Лоуренса Ливермора. Этот компьютер некоторое время возглавлял список Top 500 суперкомпьютеров, а в настоящее время находится в верхней части этого списка. Недавно в МГУ был установлен такой компьютер с 40000 ядер. Это позволяет выполнять в России работы по моделированию на современном международном уровне.

Во многих ведущих университетах и академических институтах сейчас появились достаточно мощные параллельные кластеры, которые хоть и уступают ресурсам МГУ, но зато они более доступны и во многих случаях достаточны для решения задач пользователя. В Удмуртском университете также имеется небольшой кластер, который, в основном, используется в учебном процессе. Для моделирования больших задач мы используем кластер umt (1664 ядер), установленный в Институте математики и механики ИММ УрО РАН в г. Екатеринбурге. Этот кластер имеет все современные трансляторы и другой системный софт необходимый для установки прикладных программ пользователя. Здесь следует отметить, что пакеты программ пользователя должны быть в исходных кодах. Обычно они написаны на языке Fortran, C или C++. Если разработчик пакета предусмотрел возможность распараллеливания программы на возможных платформах, то тогда такой пакет устанавливается относительно легко. Если нет, то от пользователя потребуются значительные усилия, или вообще придется отказаться от установки такого пакета. Однако, даже при успешной установке пакета, нужно еще убедиться, что он будет правильно работать в параллельном режиме. Кластер umt работает под управлением операционной системы Linux, что исключает установку пакетов, разработанных под Windows.

Таблица 2. Список некоторых пакетов установленных на BG/L (Blue Gene)

Область применения	Имя пакета	Масштабируемость (%)
Классическая Молекулярная динамика	LAMMPS	55
	DL_POLY	50
	NAMD	45
	GROMACS	52
	ddcMD	28
Первопринципная (<i>ab initio</i>) молекулярная динамика	CPMD	50
	VASP	65

	CASTEP	65
	SIESTA	55
	Qbox	85
Квантовая химия	AMBER8	45
	GAMESS	60

В лаборатории параллельных вычислений УдГУ проводятся работы по моделированию атомной и электронной структуры металлов и сплавов на основе никеля титана и алюминия. Эти работы выполняются совместно Институтом физики металлов для авиакосмической промышленности. Для моделирования таких систем используются пакеты программ на основе метода первопринципного (*ab initio*) псевдопотенциала [1-3]. Этот метод в настоящее время стал одним из основных методов расчета и моделирования. Это связано с тем, что базис плоских волн, на котором основан метод псевдопотенциала, является наиболее простым и удобным базисом, который применим для любых структур, как упорядоченных, так и неупорядоченных. В нем легко контролировать сходимость с помощью одного основного параметра – радиуса обрезания плоских волн. В таблице 2 приведен список пакетов, которые мы используем для моделирования. Пакеты VASP (2000\$) и Wien2k (500\$) платные, пакеты SIESTA, CPMD и DL_POLY требуют регистрации, остальные пакеты бесплатные и не требуют регистрации.

При моделировании атомной структуры методами классической молекулярной динамики (MD) [4], число атомов в системе может колебаться в пределах от 10^3 до 10^6 , а число итераций (шагов по времени) может превышать 10^6 . При моделировании методами первопринципной молекулярной динамики (*ab initio* MD) [5-8], число атомов в суперячейке обычно не превышает 100, а число итераций составляет несколько тысяч. При переходе с 4-х ядерных «домашних» компьютеров на большие параллельные кластеры появляется возможность увеличить число атомов в моделируемой (*ab initio* MD) системе со 100 до 1000, а число итераций до 10^4 . Однако при этом нужно учитывать, что чем больше атомов в системе, тем больше размеры файлов которые нужно записывать на диск во время моделирования. Это приводит к тому, что если на кластере нет параллельной файловой системы, то основное время тратится на пересылку по сети больших файлов в десятки гигабайт, что резко снижает эффективность расчета и сильно перегружает сетевую систему. Часто это приводит к тому, что один пользователь может «подвесить» всех остальных пользователей. Такой случай имел место у нас при использовании пакета PWSCF. К счастью, в пакете имелась возможность уменьшить величину записываемых на диск файлов за счет увеличения оперативной памяти. Такую возможность нам подсказал один из разработчиков пакета PWSCF Дмитрий Коротин, за что мы хотим его поблагодарить.

Таблица 2. Список пакетов используемых в УдГУ для моделирования и расчетов

Область применения	Имя пакета	Доступность
Классическая Молекулярная динамика	LAMMPS	Free
	DL_POLY	Registration
Первопринципная (<i>ab initio</i>) молекулярная динамика	VASP	2000\$
	CPMD	Registration
	SIESTA	Registration
	Qbox	Free
	FHI98MD	Free
Квантовая химия	GAMESS	Free
	PARSEC	Free
Зонные расчеты твердых тел	ABINIT	Free
	PWSCF	Free

	Wien2k	500\$
	LMTO-ASA	Registration

Степень эффективности пакета при увеличении числа процессоров (масштабируемость) сильно зависит от разработчиков пакета. Среди используемых нами пакетов наибольшую масштабируемость имел пакет Qbox. Он написан на языке C++ с максимальным использованием архитектуры Blue Gene. На кластере umt он также оказался самым эффективным. В нем имеется три уровня распараллеливания. Первый уровень – по числу k-точек в зоне Бриллюэна (BZ). Число этих точек определяет точность трехмерного интегрирования по BZ. Оно может достигать десятков и сотен при вычислении плотности состояний. Второй уровень – это суммирование разложения волновой функции по плоским волнам – векторам обратной решетки. Это число может достигать десятков тысяч. Третий уровень – это распараллеливание по энергетическим зонам. Это число колеблется от десятков до сотен. Наиболее эффективной системой при трансляции и установке пакета Qbox на кластере umt была система OpenMPI. Здесь следует заметить, что не все пакеты допускают использование систем OpenMPI.

При моделировании больших систем важно иметь возможность записывать на диск состояния системы в определенные моменты времени (создавать контрольные точки). В MD этот вопрос решается просто, так как там почти всегда пишутся на диск файлы траекторий (координаты, скорости атомов и силы, действующие на них). В пакете CPMD предусмотрена запись файла RESTART через заданное пользователем число шагов. При аварийном окончании задания или при истечении заданного времени решения, дальнейшее моделирование можно начать с последнего файла RESTART. В других случаях, когда нет записи контрольных точек, нужно рассчитывать время решения задачи, чтобы уложиться в максимально разрешенный период времени, который определен администрацией вычислительного центра. Такой случай имел место у нас при расчете упругих констант в пакете ABINIT.

После успешного выполнения задач по моделированию и расчету атомной и электронной структуры, возникает вопрос о визуализации полученных результатов. Это вопрос решается также с помощью применения различных пакетов программ для визуализации результатов. Обычно в каждом большом пакете по моделированию и расчету имеется свой пакет визуализации, который привязан к формату получаемых файлов. Таких пакетов в настоящее время насчитывается больше сотни. Однако, для задач в области физики твердого тела и квантовой химии в последнее время все большую популярность приобретает пакет XCrySDen [9]. Он первоначально создавался для пакета WENA95, а затем был использован во многих других пакетах. Мы используем его для получения 3D-картинок расположения атомов в суперячейке. Этот пакет поставляется как в виде загрузочных модулей, так и в исходных кодах. Однако, загрузочные модули, как правило, не работают, а трансляция и установка пакета из исходных модулей, довольно сложная задача. Для получения 2D-графиков широко используется пакет Gnuplot. Основное его достоинство состоит в том, что он есть практически на всех вычислительных системах, хорошо документирован и имеет подробный help.

Таблица 3. Системы визуализации результатов моделирования

Пакеты	Описание	Статус
XCrySDen	Мультиплатформенная визуализация и анализ данных	Open Source
VMD	Система 3D компьютерной графики	Open Source
Gnuplot	Системный софт для 1D-, 2D-, 3D-графики	Open Source

Среди других пакетов 3D-графики следует отметить пакет VMD [10]. Он создан в Иллинойском университете для биофизики. Это очень мощный пакет с большими возможностями и большой базой данных. Он поставляется в загрузочных модулях и очень легко устанавливается на большинстве платформ и в различных операционных системах: Linux, MacOS, Windows. Большим достоинством пакета VMD является то, что он позволяет писать сценарии и скрипты на языке Tcl/Tk. Это позволяет выполнять много рутинных операций в пакетном режиме, а некоторые сложные операции можно выполнить только программным путем с использованием довольно сложных сценариев на языке Tcl/Tk. Здесь, однако, нужно отметить, что создание таких сценариев требует от пользователя определенного опыта и квалификации. Наш опыт говорит о том, что прежде чем писать свой сценарий, лучше поискать что-нибудь похожее в Интернете. Для пакета VMD написано много скриптов для различных задач моделирования.

В заключение можно сказать, что успех моделирования сложных физических, химических и других систем, во многом зависит от наличия пакета программ для расчетов и визуализации. Такие пакеты стали доступны в настоящее время. Их можно скачать в Интернете. Выход на суперкомпьютер тоже перестал быть проблемой. И последнее условие это необходимая квалификация пользователя. Это условие выполняется, если в университете или институте имеется школа по подготовке таких специалистов. В УдГУ такая школа имеется на кафедре и лаборатории параллельных вычислений, которые существуют уже более 10 лет.

Литература

1. D.R.Hamman, M.Schluter, C.Chiang, Phys. Rev. Lett, 43, 1494 (1979).
2. G.B.Bachelet, D.R.Hamman, M.Schluter, Phys. Rev. B26, 4199 (1982).
3. Vanderbilt D, Phys. Rev. B41 7892 (1990)
4. B.J.Alder, T.E.Wainwright, J. Chem. Phys., 27, 1208 (1957).
5. R.Car, M.Parinello, Phys. Rev. Let., 55, 2471 (1985).
6. W.Kohn, L.J.Sham, Phys. Rev. 140, A1133 (1965).
7. G.Kresse, J.Futhmuller, Comp. Mat. Sci., 6, 15 (1996).
8. G.Kresse, J.Furthmuller, Phys. Rev. B, 54, 11169 (1996).
9. Kokalj, J. Mol. Graphics Modelling, 1999, Vol. 17, 176-179.
10. <http://www.ks.uiuc.edu/Research/vmd/>