

Результаты моделирования системы SiO₂ (в часах)

N	Локальный вариант	MPI-8	MPI-16	CUDA	CUDA-2
1000	3,1	0,41	0,22	0,17	0,09
4200	11,5	1,75	0,9	0,63	0,37
12000	134	21,7	10,3	7,4	4,1
100000	1456	221	87,6	77,3	39,3

В моделирование сравнивалось время, затраченное на проведение эксперимента с числом частиц в системе 105 и числом шагов 5000.

Компьютерные эксперименты проводились на кластере, состоящем из 16 вычислителей Dual Core AMD Opteron с частотой 2.21 Ghz для приложения, реализованного с использованием технологии MPI. На графическом процессоре с использованием двух видеокарт NVidia GeForce GTS 450 для приложения реализующего расчет на CUDA.

Полученные результаты дают возможность говорить о целесообразности использования графических процессоров для экспериментов в области молекулярной динамики и позволяют ускорять время расчета в 36 раз по сравнению с локальным вариантом МД-приложения.

Список литературы

1. Brown, W.M., Kohlmeier, A., Plimpton, S.J., Tharrington, A.N. Implementing molecular dynamics on hybrid high performance computers – Particle-particle particle-mesh // Computer Physics Communications Volume 183, Issue 3, March 2012, Pages 449-459
2. Harvey, M.J., De Fabritiis, G. An implementation of the smooth particle mesh Ewald method on GPU hardware // Journal of Chemical Theory and Computation Volume 5, Issue 9, September 2009, Pages 2371-2377
3. Jha, P.K., Sknepnek, R., Guerrero-García, G.I., Olvera De La Cruz, M. A graphics processing unit implementation of coulomb interaction in molecular dynamics // Journal of Chemical Theory and Computation Volume 6, Issue 10, 12 October 2010, Pages 3058-3065
4. Воронова Л.И., Григорьева М.А., Воронов В.И., Трунов А.С. Программный комплекс «MD-SLAG-MELT» для моделирования наноструктуры и свойств многокомпонентных расплавов. – Расплавы. 2013. № 4. С. 36-49.
5. Трунов А.С., Воронова Л.И., Шалабай Т.С. Метод параллельного расчета коррелированной системы n-частиц на графическом процессоре. – Современные наукоемкие технологии. 2013. № 6. С. 117-119.
6. Воронова Л.И., Трунов А.С., Воронов В.И. Разработка методов параллельного расчета коррелированной многочастичной системы на графическом процессоре. – Вестник Российского государственного гуманитарного университета. 2013. № 14. С. 236-247.
7. Трунов А.С., Воронова Л.И., Воронов В.И. Разработка методов распределения для высокопроизводительных вычислений в многочастичных системах. – Международный журнал прикладных и фундаментальных исследований. 2013. № 10-2. С. 192-194.

МОДЕЛЬ БАЛАНСИРОВКИ НАГРУЗКИ ДЛЯ ПАРАЛЛЕЛЬНОГО РАСЧЕТА СИСТЕМЫ N-ЧАСТИЦ НА ГЕТЕРОГЕННОЙ КЛАСТЕРНОЙ СИСТЕМЕ

Пилипчук П.Е., Воронова Л.И., Трунов А.С.

Филиал Российского государственного гуманитарного университета, Домодедово, Московская область,
e-mail: thelastrainbow@yandex.ru

В настоящее время в различных областях современной науки широко применяется компьютерное моделирование (КМ). В физической химии, например, КМ позволяет исследовать структуру и свойства материалов. Широко применяемый в физической химии метод молекулярной динамики (МД) позволяет определить целый комплекс свойств (структурные, термодинамические, транспортные) и исследовать взаимосвязи наноструктуры и физико-химических свойств.

Важной характеристикой при МД-моделировании является размер (количество частиц) моделируемой системы. С увеличением числа частиц до 106 линейные размеры модельного куба достигают нанометров, что позволяет получать результаты нового качества, обладающих практической значимостью. Увеличение числа частиц моделируемой системы требует серьезных временных затрат на проведение КМ, что делает невозможным проведение КМ без привлечения высокопроизводительных вычислений.

Для проведения параллельного расчета системы N-частиц широко применяется вычислительный кластер, состоящий из вычислительных узлов (вычислителей) объединенных каналами связи. В большинстве случаев рассчитываемое множество частиц разбивается на одинаковые подмножества между вычислительными узлами кластера для проведения параллельного расчета. Но возникают ситуации, когда на кластерных системах вычислители обладают разной производительностью. Это означает, что время проведения КМ будет равно времени завершения своего расчета самым слабым (по производительности) узлом кластера.

Для обеспечения одновременного завершения расчетов всеми вычислителями, разработана модель балансировки нагрузки для параллельного расчета системы N-частиц на гетерогенной кластерной системе, что позволит существенно сократить время КМ.

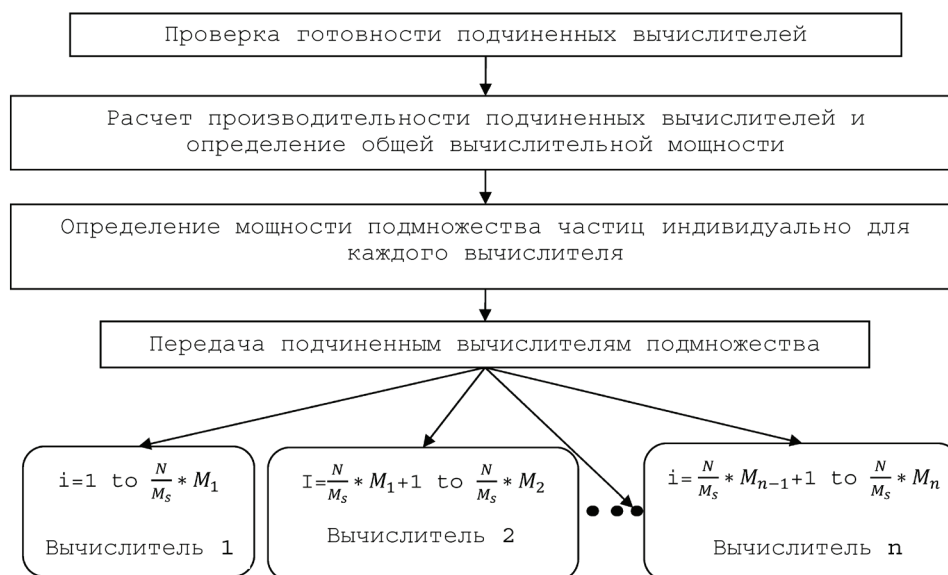
Для определения размеров подмножеств частиц используется процентная доля производительности каждого из вычислительных узлов кластера, в соответствии с которой рассчитывается число частиц, направляемое вычислителю для расчета. Формула для вычисления мощности подмножества частиц, направляемого каждому вычислителю в зависимости от его производительности:

$$k_i = \frac{N}{M_s} M_i \quad (1)$$

где k_i – подмножество частиц направляемых для расчета вычислителю P_i , N – множество частиц в системе, M_i – производительность i -го вычислителя, M_s – суммарная мощность всех вычислителей в системе.

На рисунке представлен алгоритм распределения нагрузки между вычислителями в соответствии с мощностями каждого вычислителя.

Определение мощности подмножества дескрипторов производится в тестовом режиме до начала моделирования системы N-частиц. Главный вычислитель проверяет готовность всех подчиненных вычислителей и передает для тестового расчета одинаковые подмножества дескрипторов $\{Dls(i)\}$ и $\{Dlv(i)\}$. С учетом времени расчета каждым вычислителем своего подмножества дескрипторов главный вычислитель определяет мощность подмножества для каждого вычислителя по формуле (1). Вновь сформированные подмножества будут передаваться вычислителям во время процесса моделирования.



Алгоритм распределения нагрузки между вычислителями в соответствии с производительностью каждого вычислителя.

Для оценки эффективности разработанной модели распределенных вычислителей проведено тестовое моделирование системы оксид кремния SiO₂ в программном комплексе ИИС «MD-SLAG-MELT»[5]. Тестирование проходило на 16 вычислителях AMD и Intel разной производительной мощности. В моделировании сравнивалось время, затраченное на проведение эксперимента с числом частиц в системе 105 и числом шагов 5000. Результаты компьютерных экспериментов показаны в таблице.

Результаты тестового моделирования системы SiO₂ (в часах)

N	1000	4200	12000	100000
Метод загрузки				
Равномерная загрузка	0,9	1,75	3,7	11,5
Загрузка по вычислительным мощностям	0,42	0,81	2,9	9,4

Таким образом, использования модели балансировки нагрузки позволит в равной степени задействовать все вычислители на максимальную мощность в течении всего времени выполнения моделирования системы N-частиц. Данный подход обеспечивает наиболее равномерное распределение нагрузки между вычислителями, т.к. позволяет гибко регулировать объемы обрабатываемых подмножеств и обеспечивает эффективность работы всей системы приближенную к оптимальной, что не позволяют обеспечить ранее описанные методы для гетерогенной вычислительной среды. Кроме того, данная модель позволяет обеспечить оптимальную эффективность использования вычислителей, как для гетерогенной, так и для однородной среды.

Список литературы

1. Воронова Л.И., Григорьева М.А., Воронов В.И., Трунов А.С. Программный комплекс «MD-SLAG-MELT» для моделирования nanoструктуры и свойств многокомпонентных расплавов. – Расплавы. 2013. № 4. С. 36-49.
2. Трунов А.С., Воронова Л.И., Шалабай Т.С. Метод параллельного расчета коррелированной системы n-частиц на графическом

процессоре. – Современные наукоемкие технологии. 2013. № 6. С. 117-119.

3. Воронова Л.И., Трунов А.С., Воронов В.И. Разработка методов параллельного расчета коррелированной многочастичной системы на графическом процессоре. – Вестник Российского государственного гуманитарного университета. 2013. № 14. С. 236-247.

4. Трунов А.С., Воронова Л.И., Воронов В.И. Разработка методов распределения для высокопроизводительных вычислений в многочастичных системах. – Международный журнал прикладных и фундаментальных исследований. 2013. № 10-2. С. 192-194.

АВТОМАТИЗАЦИЯ РАНЖИРОВАНИЯ ТИПОВ ВИЗУАЛИЗАЦИИ В DATA MINING ДЛЯ НЕРЕЛЯЦИОННЫХ БАЗ ДАННЫХ

Хомутова Е.В., Воронова Л.И.

Московский технический университет связи и информатики, Москва, e-mail: evenie@mail.ru

Основным инструментом практически любого бизнеса в современном мире является анализ информации. Без анализа данных невозможно провести оценку выбора пользователя («пользовательской корзины»), произвести сегментацию рынка и прогнозирование, оценить риски и разработать стратегию дальнейшего развития. Наиболее полно решить все перечисленные выше задачи способны программные продукты, реализующие методы технологии Data Mining. В статье рассмотрены варианты автоматизации одного из модулей Data Mining – визуализации результатов анализа данных для нереляционных баз данных.

Data Mining – (рус. добыча данных, интеллектуальный анализ данных, глубинный анализ данных) – собирательное название, используемое для обозначения совокупности методов обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности [1]. В более узком смысле, Data Mining – это интеллектуальный анализ данных (ИАД), направленный на выявление скрытых, неявных, ранее неизвестных закономерностей.