Министерство образования и науки Российской Федерации
Федеральное агентство по образованию
Санкт-Петербургский государственный университет

ВЫСОКОПРОЗВОДИТЕЛЬНЫЕ ПАРАЛЛЕЛЬНЫЕ ВЫЧИСЛЕНИЯ НА КЛАСТЕРНЫХ СИСТЕМАХ

Материалы шестого Международного
научно-практического семинара

Том 1

12–17 декабря 2006 г.

Издательство Санкт-Петербургского госуниверситета
Санкт-Петербург
2007

УДК 681.3.012:51
ББК 32.973.26–018.2:22
В 93

В93 **Высокопроизводительные параллельные вычисления на кластерных системах.** Материалы шестого Международного научно-практического семинара. Том 1 / Под ред. проф. **Р.Г. Стронгина.** Санкт-Петербург: Изд-во Санкт-Петербургского госуниверситета, 2007. 281 с.

Сборник сформирован по итогам научного семинара, посвященного теоретической и практической проблематике параллельных вычислений, ориентированных на использование современных многопроцессорных архитектур кластерного типа. Семинар проходил в Санкт-Петербурге 12–17 декабря 2006 года.

Вошедшие в сборник материалы семинара представляют интерес для преподавателей и научных сотрудников высших учебных заведений, а также для инженеров, аспирантов и студентов вузов.

Отв. за выпуск к.ф.-м.н. **И.М. Григорьев**

ББК 32.973.26–018.2:22

Поддержка семинара

Компания Intel Technologies

© Санкт-Петербургский госуниверситет, 2007

ФОНД ИННОВАЦИОННЫХ ПРОЕКТОВ ФАКУЛЬТЕТА «ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ» УДМУРТСКОГО ГОСУНИВЕРСИТЕТА

Исламов Г.Г., Коган Ю.В., Исламов А.Г., Лукин О.Л.
Удмуртский государственный университет, Ижевск

Фонд инновационных проектов факультета «Информационных технологий и вычислительной техники» (кратко Фонд ИП факультета ИТ и ВТ) является формой равноправного сотрудничества образования, науки и бизнеса, направленного на формирование общих целей ускоренного развития за счёт интеграции интеллекта, менеджмента, материальных и финансовых ресурсов ведущих процессов общества: образовательного, научного и производственного. Руководитель, заместитель и группа экспертов фонда назначается ректором университета из ведущих специалистов факультета ИТ и ВТ.

Участником Фонда ИП может стать любое физическое и юридическое лицо, внёсшее в кассу университета ежегодный членский взнос. Главным участником Фонда ИП выступает университет – содержатель финансовых и материальных средств фонда ИП. Из средств Фонда ИП осуществляется финансирование инновационных проектов, разрабатываемых на факультете ИТ и ВТ. Это финансирование обеспечивает персональную надбавку к бюджетной зарплате руководителей и исполнителей инновационных проектов. Величина персональной надбавки устанавливается руководителем Фонда ИП согласно принятым нормам на текущий академический год и не зависит от числа проектов, в которых участвует преподаватель или сотрудник факультета ИТ и ВТ. Участники Фонда имеют право выдвигать свои инновационные проекты, формировать заявки на специалистов ИТ, повышать квалификацию своих сотрудников на курсах, разрабатываемых по согласованию с факультетом ИТ и ВТ, имеют равный доступ к результатам всех инновационных проектов текущего академического года, осуществлять контроль за ходом выполнения инновационных проектов.

Группа экспертов Фонда ИП формирует в сентябре каждого года пакет инновационных проектов из области образования, науки и бизнеса сроком исполнения на один академический год. Текстовая часть проекта должна содержать название проекта, краткое описание, данные о руководителе проекта, резюме, введение и обзор,

формулировку проблемы, цель и задачи предлагаемого исследования, практические методы, используемые в исследовании, подробный план исследования, включающий график выполнения проекта и формы представления отчёта в Фонд ИП, ожидаемые результаты исследования (научные, организационные и технологические), состав исследовательской группы и способ управления ходом выполнения проекта. Смета расходов на аппаратное и программное обеспечение проекта утверждается группой экспертов с учётом всего пакета инновационных проектов на текущий академический год. Руководителем проекта может быть только профессор, доктор наук или доцент, кандидат наук (по специальному решению группы экспертов). Не существует иной формы отклонения инновационного проекта кроме несоответствия принятым нормам его оформления. Все правильно оформленные проекты имеют право на существование и их реализацию. Проекты, получившие одобрение большего числа участников Фонда ИП, премируются руководителем Фонда ИП в конце академического года. Курсовые и дипломные проекты на факультете ИТ и ВТ должны выполняться только в рамках инновационных проектов, одобренных Фондом ИП.

В рамках этого Фонда планируется продолжить работы в области высокопроизводительных вычислений на многопроцессорных комплексах, разработки параллельных алгоритмов для решения задач управления детерминированными и случайными процессами [5], изучения математической модели банковского скоринга [7] и исследования функции электронной плотности на основе обобщенного уравнения Шрёдингера [8]. Ниже приводится краткое описание нового проекта на тему *«Параллельный алгоритм одновременного нахождения решений и областей устойчивости для симметричной пары прямой и двойственной задач линейного программирования»*.

Вычислительный ресурс – пакет прикладных программ, ориентированных на решение определенного класса научно-технических и производственных задач – может предоставляться пользователям в самой разнообразной форме. Как правило, это простая передача пакета и документации к нему на правах платного или бесплатного программного обеспечения. Мы считаем, что в случае сложных и оригинальных разработок целесообразна, по многим причинам, регистрация вычислительного ресурса в сети Internet и организация доступа к нему средствами инструментария Globus Toolkit 4 [1]. Положительный мировой опыт применения предыдущих версий

этого инструментария, а также расширенные возможности четвёртой и последующих версий Globus Toolkit, указывают нам новые способы доставки вычислительных ресурсов пользователям через grid-службу. Важным ресурсом может служить пакет программ для решения задач линейного программирования большой размерности. Прикладное значение этих задач раскрывалось во многих работах (см., например, [2-5]). Новый интерес к оптимизационным задачам такого класса возник в связи с быстрым распространением в мире кластерных технологий организации высокопроизводительных вычислений. Globus-машина (сервер) в составе вычислительного кластера обеспечивает доступ к вычислительным ресурсам этого кластера. Пользователи, имеющие globus-машину (клиента), через сеть Internet получают возможность затребовать вычислительный ресурс через globus-машину (сервер).

В учебном пособии [6] на основе идей монографии [2] разработан последовательный алгоритм в форме симплекс-таблиц одновременного нахождения решений и областей устойчивости для симметричной пары прямой и двойственной задач линейного программирования. В настоящей работе предлагается параллельная реализация этого алгоритма, использующая методы распределённых вычислений, применённые в отчёте [4]. Толчком к инициированию нашего исследования послужили важные результаты, полученные в [4] относительно эффективности параллельной реализации симплекс-метода в виде симплекс-таблиц. Вычислительный эксперимент показал, что эта реализация выигрывает в сравнении с методом обратной матрицы. Преимущество же нашего параллельного алгоритма перед алгоритмом работы [4] заключается в том, что он сразу даёт решение четырёх задач: находятся оптимальные планы симметричной пары прямой и двойственной задач линейного программирования, а также области устойчивости этих планов. Значимость этих задач для экономико-математического анализа линейных моделей производственных процессов отмечается в монографии [2]. Исходные структуры данных параллельного алгоритма симплекс-метода в виде симплекс-таблиц соответствуют структурам, разработанным в учебном пособии [6]. Подходы к распределению столбцов информационной структуры по процессорам, а также правила выбора столбца, вводимого в базис, заимствованы из работы [4]. Следует отметить, что в указанной работе проведён теоретический анализ времени выполнения для последовательного и

параллельного варианта метода симплекс-таблиц, метода обратной матрицы, а также дано сравнение этих методов по эффективности. Подобный анализ модели производительности для взаимодействующих процессоров можно провести и в нашем случае. Это позволяет указать оптимальное число процессоров, которое минимизирует общее время выполнения как функцию размерности, структуры и степени разрежённости матрицы задачи линейного программирования.

В дополнение к работе [5] рассмотрим однородную цепь Маркова с переходной матрицей L порядка n, в которой строка $x(t) = (x_1(t),...,x_n(t))$ представляет безусловные вероятности системы в фиксированный момент времени $t = 0,1,...,N$. Эволюция системы описывается равенством $x(t+1) = x(t)\,L$, где $t = 0,1,...,N-1$ и начальное состояние $x(0)$ выбирается из условия выполнения конечной системы неравенств

$$x(t_k)\,c^k \geq \beta_k\,, k = 1,...,m\,. \qquad (1)$$

Здесь c^k представляет собой столбец «доходов», которые обеспечивает система в момент времени $t_k \in \{1,...,N\}$ $(t_{k-1} < t_k, t_0 = 0)$ и все «пороговые» значения $\beta_k > 0$. Нас интересует случай, когда величины n и m достаточно велики $(n \approx 2^{16}, m \approx 2^{10})$. Условия разрешимости относительно начальных безусловных вероятностей $x(0) = (x_1(0),...,x_n(0))$ системы аналогичны тем, что получены в работе [5] и формулируются в терминах следующих двух расчётных задач большой размерности. Первая задача, требующая высокой точности вычислений, состоит в нахождении m столбцов $y^k(t) = (y^k{}_1(t),...,y^k{}_n(t))^T$ - решений сопряжённой разностной системы $y(t+1) = Ly(t), t = 1,...,N$ с начальным условием $y(0) = c^k$ $(k = 1,...,m)$. Вторая расчётная задача состоит в отыскании максимального значения следующей экстремальной задачи

$$\sum_{k=1}^{m} \beta_k\,\lambda_k \to \max,\ \sum_{k=1}^{m} y^k(t_k)\,\lambda_k \leq e, \lambda_k \geq 0, = 1,...,m, \qquad (2)$$

где e - столбец, состоящий из n единиц (сравнение векторов-столбцов покоординатное). Аналогично работе [5] показывается, что задача (1) разрешима тогда и только тогда, когда максимальное значение целевой функции задачи (2) больше нуля и не превосходит единицы и после нормировки решение $q = (q_1, \ldots, q_n)$ двойственной к (2) экстремальной задачи

$$q\,e \to \min, q_i \geq 0, i = 1, \ldots, n, q\,y^k(t_k) \geq \beta_k, k = 1, \ldots, m \qquad (3)$$

даёт одно из начальных состояний $x(0) = q$, обеспечивающее выполнение системы неравенств (1). Для нахождения решений этой пары симметричных двойственных задач и их областей устойчивости мы применяем параллельную реализацию алгоритма симплекс-метода, описанного в [6]. Первая симплекс-таблица предложенного в [6] алгоритма одновременного решения задач (2) и (3) имеет вид

$$\begin{pmatrix}
 & & & \lambda_1 & \cdots & \lambda_m & & & \\
 & & & q_{n+1} & \cdots & q_{n+m} & & & \\
 & & 0 & -\beta_1 & \cdots & -\beta_m & & & \\
q_1 & \lambda_{m+1} & 1 & y_1^1(t_1) & \cdots & y_1^m(t_m) & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
q_n & \lambda_{m+n} & 1 & y_n^1(t_1) & \cdots & y_n^m(t_m) & 0 & \cdots & 1 \\
 & & & 1 & \cdots & 0 & & & \\
 & & & \vdots & \ddots & \vdots & & & \\
 & & & 0 & \cdots & 1 & & &
\end{pmatrix},$$

где справа и снизу приписаны единичные матрицы размеров соответственно n и m. В пособии [6] указано правило преобразования симплекс-таблиц такой структуры, условие окончания алгоритма, правило нахождения экстремальных значений, оптимальных планов и их областей устойчивости, а также рассмотрены иллюстративные примеры.

Обозначим через P матрицу порядка n, стоящую справа в последней симплекс-таблице, а через Q матрицу порядка m, стоящую снизу в этой таблице. Пусть, далее, $\Delta = (\Delta_1, \ldots, \Delta_m)$ -

строка, стоящая в последней симплекс-таблице в положении строки $(-\beta_1,\ldots,-\beta_m)$ первой симплекс-таблицы, а нулевые координаты этой строки имеют номера $j_1,\ldots,j_t\ (1\le t\le m)$; $U=(u_1,\ldots,u_n)^T$ - столбец, стоящий в последней симплекс-таблице в положении столбца из единиц e первой симплекс-таблицы, а нулевые координаты этого столбца имеют номера $i_1,\ldots,i_s\ (1\le s\le n)$. Заметим, что величины Δ и U определяют соответственно решения двойственной задачи (3) и прямой задачи (2), а экстремальное значение целевых функций этих задач находится в последней симплекс-таблице на месте «жирного» нуля в первой симплекс-таблице.

Имеет место следующее утверждение, частная формулировка которого дана в учебном пособии [6] для невырожденных задач линейного программирования.

Теорема. *Область устойчивости прямой задачи (2) даётся системой неравенств*

$$\delta\beta\cdot Q\le\Delta+\sum_{k=1}^{s}\nu_k\,\tilde{y}_{i_k},$$

где $\delta\beta=(\delta\beta_1,\ldots,\delta\beta_m)$, *а* $\nu_k\ge 0, k=1,\ldots,s$ *произвольно заданы и* $\tilde{y}_{i_k}$ *есть* i_k *–я строка основной матрицы последней симплекс-таблицы. Область устойчивости двойственной задачи (3) описывается системой неравенств*

$$P\cdot\delta e+U\ge\sum_{k=1}^{ts}\mu_k\,\tilde{y}^{j_k},$$

где $\delta e=(\delta e_1,\ldots,\delta e_n)^T$, *а* $\mu_k\ge 0, k=1,\ldots,t$ *произвольно заданы и* $\tilde{y}^{j_k}$ *есть* j_k *–ый столбец основной матрицы последней симплекс-таблицы.*

Литература

1. Foster I. Globus Toolkit Version 4: Software for Service – Oriented Systems // NPC 2005, LNCS 3779, pp. 2-13, 2005.

2. Бункин В.А., Курицкий Б.Я., Сокуренко Ю.А. Решение задач оптимизации в управлении машиностроительным производством. Л.: Машиностроение, 1976. – 232 с.
3. Булавский В.А., Звягина Р.А., Яковлева М.А. Численные методы линейного программирования. М.: Наука, 1977. – 368 с.
4. Yarmish G. A distributed implementation of the simplex-method. Technical Report TR-CIS-2001-04, Polytechnic University, Brooklyn.
5. Исламов Г.Г. Высокопроизводительные вычисления и параллельное программирование для марковского процесса с конечным числом состояний и непрерывным временем // Труды XIII Всероссийской научно-методической конференции «Телематика'2006», Санкт-Петербург. - Т. 1, С. 219-220.
6. Исламов Г.Г. Принципы оптимальности // Учебное пособие, Ижевск, Изд-во УдГУ, 1998. – 124 с.
7. Пальянов А.Л. , Исламов Г.Г. Математическая модель банковского скоринга // Технологии Microsoft в теории и практике программирования. Материалы конференции / Под ред. проф. Р.Г. Стронгина. Нижний Новгород: Изд-во НГГУ, 2006. С. 235-236.
8. Исламов Г.Г. Об одном обобщении уравнения Шрёдингера // Международная научная конференция «75 лет высшему образованию в Удмуртии»: Материалы конференции: Ч. 2. Естественные науки. Ижевск, 2006. С. 5-6.

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ СОЗДАНИЯ ДВУХПОТОЧНОГО КОНВЕЙЕРА ПАРСЕРА XERCES ДЛЯ ПОЛНОГО ИСПОЛЬЗОВАНИЯ ВОЗМОЖНОСТЕЙ ДВУЯДЕРНОЙ АРХИТЕКТУРЫ

Кабардин И.К.

Лаборатория высокопроизводительных вычислительных систем, Новосибирский Государственный Университет, г. Новосибирск

Введение:

Ключевая задача для всех, кто делает процессоры, в том числе и для Интел – выдавать большую производительность. В течении длительного периода времени производительность повышали за счет