

ОЛЕНЁВ Н.Н, ПЕЧЁНКИН Р.В, ЧЕРНЕЦОВ А.М.

**ПАРАЛЛЕЛЬНОЕ ПРОГРАММИРОВАНИЕ В
МАТЛАВ И ЕГО ПРИЛОЖЕНИЯ**



ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР ИМ. А.А ДОРОДНИЦЫНА РАН
МОСКВА ♦ 2007

УДК 512.643 + 519.86

Ответственный редактор
академик РАН А.А. Петров

Монография посвящена вопросам использования распределенных и параллельных вычислений в среде MATLAB. Изложена технология настройки кластера для использования MATLAB. Технология апробирована на стандартных задачах линейной алгебры. Предложен промышленный подход идентификации математических моделей сложных систем на основе параллельных средств MATLAB. Подход рассмотрен на примере простейшей динамической модели экономики современной России, и позволяет освоить его большому числу потенциальных пользователей. Дана экономическая интерпретация полученным результатам.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (коды проектов 07-01-00563-а, 07-01-12032-офи), Российского гуманитарного научного фонда (код проекта 06-02-91821-а/G), программы Intel Education Project 2006-2007, программы Президиума РАН № 15, программы ОМН РАН № 3, по программе государственной поддержки ведущих научных школ (код проекта НШ-5379.2006.1).

Рецензенты Г.М. Михайлов,
В.С. Молоствов

Научное издание

©Вычислительный центр им. А.А. Дородницына
Российской академии наук, 2007

Оглавление

Введение	5
I. Администрирование и конфигурирование	10
1.1 Используемые термины и обозначения	10
1.2 Требования к аппаратному и программному обеспечению	15
1.3 Установка и запуск MDCE	18
1.3.1 Работа в Windows	19
1.3.2 Работа в Unix	22
1.4 Запуск планировщика	23
1.4.1 Запуск в Windows	23
1.4.2 Запуск в Unix	24
1.5 Запуск рабочих процессов	25
1.6 Объекты-ссылки на системные процессы	27
1.7 Эквиваленты между функциями DCT и MPI	33
1.8 Использование сторонних планировщиков	34
1.9 Заключение	35
II. Программирование параллельных задач	38
2.1 Режим rmode, первая параллельная задача	39
2.2 Вычисление определенного интеграла	42

2.3 Объект "параллельная задача"	46
2.3.1 Объект "параллельное задание"	48
2.3.2 Матричное умножение	51
2.3.3 Решение СЛАУ методом Гаусса	54

III. Параллельные вычисления в моделировании экономики	65
3.1 Проблема идентификации внешних параметров модели экономики	67
3.2 Численная реализация задачи идентификации	82
3.2.1 Листинги программ	86
3.2.2 Результаты идентификации модели, их графическое представление	90
3.3 Сценарные расчеты с моделью	93
3.3.1 Базовый, он же пессимистический сценарий	93
3.3.2 Оптимистичный сценарий	95
3.3.3 Сравнение, выводы	97
Заключение	108

Введение

Данная работа представляет собой первую попытку в виде научной монографии, доступной начинающему пользователю MATLAB, описать параллельную технологию вычислений, реализованную компанией MathWorks с помощью двух взаимосвязанных пакетов расширений (приложений): MATLAB Distributed Computing toolbox [1] и MATLAB Distributed Computing Engine [2].

В настоящее время не только компания MathWorks обладает технологиями, позволяющими реализовывать параллельные вычисления. Следует упомянуть об аналогичных средствах, имеющихся в математических пакетах других фирм. Так, Maple, начиная с версии 9.5, предлагает toolbox HPC-Grid [3], который предлагает пользователю этой среды аналогичные возможности. Для пакета Mathematica имеется Parallel Computing Toolkit [4]. Также предлагается независимый toolkit Pooch MPI [5], который работает только под операционной системой MacOS, что резко снижает его область применения. Часто применяемый пакет Mathcad на момент написания работы не имеет каких-либо средств, позволяющих выполнять распределенные вычисления.

Несмотря на наличие подобного рода инструментов в пакетах компьютерной алгебры Maple и Mathematica, среда

технических расчетов MATLAB является, по мнению авторов, более привычной и удобной для программирования, отладки и реализации параллельных алгоритмов.

За последние годы, различными научными коллективами, были созданы многочисленные пакеты расширений (Toolbox) MATLAB, реализующих какую-либо функциональность распределенных вычислений. С их обзором можно ознакомиться в [6]. Эти пакеты в своей реализации использовали различные способы (парадигмы):

- Использование модели передачи сообщений.
- Работа с общей памятью.
- Использование специальных процедур, с помощью которых производится обмен данных между последовательными сессиями MATLAB.
- Перекомпиляция кода MATLAB на C и создание параллельного кода уже для программ C.

В зависимости от реализации, при использовании этих пакетов расширений, требовалось иметь разное количество лицензий на запуск клиентских сессий MATLAB.

Пакеты расширений для параллельного программирования, сторонними производителями начали разрабатываться в то время, когда необходимость в параллельном программировании для MATLAB не являлась критической. С тех пор, в связи с необходимостью расчета более сложных задач, эта необходимость резко возросла. Некоторые toolbox работали только под определенную операционную систему, чаще всего - под Unix. Поэтому использование единого подхода для всех операционных систем стало необходимым. Большая часть разработанных пакетов уже не поддерживается.

Особо следует отметить пакет MPITB [7], написанный в 2000 г. и получивший достаточно широкое распространение для кластерных систем. В нем реализовано применение функций MPI в программах MATLAB. Последняя на момент написания работы версия - для MATLAB 7.0.1 R14SP1 и библиотеки LAM MPI v. 7.1.1 [8] - датируется 2005 годом. Основными его недостатками являются:

- Необходимость получения лицензии на каждое выполняемое ядро.
- Невозможность применения пакета для неоднородных вычислительных архитектур.
- Отсутствие официальной поддержки от компании MathWorks, постоянно развивающей систему MATLAB.
- Исключение работы пакета с операционными системами, отличающимися от ОС Unix.

Данный пакет был импортирован в среду Windows и реализацию MPICH2 [9]. Его дистрибутивы можно скачать с [10]. Последняя версия от 02.2006 работает с MATLAB v. 7.1.0.246 (R14) Service Pack 3, т.е. также с устаревшей версией MATLAB.

Учитывая все вышеперечисленное, реализация подобных пакетов расширений уже самой фирмой MathWorks выглядит вполне логичным шагом.

Авторы ставили перед собой цель написать такую работу, которая могла бы значительно ускорить процесс освоения технологии параллельного программирования читателем, который вообще не знаком ни с навыками параллельного программирования в системе MATLAB, ни в целом с системой.

В первом разделе описаны основные процедуры установки, настройки и конфигурирования кластерной части MATLAB Distributed Computing Engine. Этот раздел будет интересен в первую очередь системным администраторам и администраторам кластерных систем. В данном разделе авторы попытались представить информацию, содержащуюся "между строчек" в стандартной технической документации. Как показал опыт сотрудничества с другими исследовательскими группами, информации, представленной в этом разделе, достаточно для настройки высокопроизводительной вычислительной среды.

Во втором разделе описаны первоначальные сведения для создания параллельных программ с использованием средств библиотеки MPI и ее реализации на основе платформы MATLAB. Рассмотренные простые задачи позволяют пользователю убедиться в приросте производительности вычислений при увеличении числа процессоров и понять базовые принципы написания параллельных программ в MATLAB.

В третьем разделе авторы в качестве примера приложения параллельных вычислений в MATLAB для научных исследований рассматривают их применение в математическом моделировании экономических систем. Это приложение ни в коей мере не ограничивает применение параллельных вычислений в математическом моделировании сложных систем только данной областью науки. Во-первых, рассмотренные примеры легко могут быть обобщены на аналогичные задачи в других областях знаний. Во-вторых, экономические приложения к современной жизни интересны для всех любознательных людей.

Рассмотрение приложений параллельных вычислений очень важно, поскольку в связи с развитием многопроцессор-

ных и многоядерных [11] архитектур вычислительных систем в настоящее время параллельное программирование планируют применять повсеместно во всех областях науки, техники и бизнеса. Стандартом де-факто для параллельных вычислений служит довольно сложный для первоначального изучения интерфейс передачи сообщений MPI [12], [13]. Приложения, разработанные компанией MathWorks для создания параллельных и распределенных программ с использованием средств библиотеки MPI и их реализация на платформе MATLAB, упрощает практическое применение параллельных вычислений на многоядерных компьютерах, кластерах и GRID-системах.

Авторы выражают *благодарность* департаменту MathWorks компании Softline за предоставленные дистрибутивы MATLAB и возможность ознакомиться и протестировать технологические возможности, заложенные в MATLAB Distributing Computing Engine.

Заключение

Данная работа является первой попыткой описания параллельного программирования в системе MATLAB. Появление новых архитектурных решений: многоядерных платформ, — подвигнуло создателей системы MATLAB и других систем уделять больше внимания параллельным аспектам этих систем, что влечет быстрые их изменения. Соответственно увеличиваются потенциальные возможности использования параллельных вычислений в различных прикладных областях. Исходя из имеющихся тенденций авторы настоящей монографии планируют ее модификацию как для учета быстрого развития технологических возможностей системы MATLAB в области параллельных вычислений, так и для учета быстрого расширения области использования параллельных вычислений.

Кроме того, в каждом из разделов остались нерассмотренными интересные и важные вопросы. В первом разделе не были рассмотрены такие вопросы, как:

- работа с произвольной реализацией MPI,
- работа MATLAB с существующими системами очередей.

Во втором разделе не был рассмотрен вопрос, связанный с распределенным хранением исходных данных некой вычислительной задачи (в том случае, если размер исходных данных не позволяет хранить ее на рабочей станции).

В третьем разделе не затронуты вычислительные задачи, связанные с пространственно распределенной эколого-экономической системой, в которой могут использоваться не только пространственно-распределенные данные, но и отличающиеся модели для экологического и экономического блоков. В таких задачах применение параллельных вычислений существенно, эти вычисления требуют проведения однотипных вычислений в огромном числе пространственно-распределенных точек.

Еще одним типом задач для приложения параллельных вычислений являются задачи нелинейной математической экономики, включающие поиск решения дифференциальных уравнений с частными производными. Эти задачи и методы их решения похожи на соответствующие задачи математической физики и пример такой задачи нужно будет рассмотреть.

Эти и некоторые другие вопросы планируется осветить в следующем издании.

Лучшие применения параллельных вычислений в моделировании экономики раскрывают возможности разработчиков моделей. Появилась возможность идентифицировать внешние параметры сложных нормативных балансовых динамических моделей экономических систем. Можно идентифицировать сложные, пространственно-распределенные модели эколого-экономических систем, чтобы исследовать на них последствия изменения климата на протекание экономических и экологических процессов в различных регионах и странах. Для этого разрабатываются специальные критерии близости и похожести для статистических и рассчитанных по модели временных рядов макропоказателей изучаемой экономической системы страны или региона [3], применяются информационные технологии создания математических моделей с помощью системы ЭКОМОД для уменьшения числа независимых внешних параметров, совершенствуются методы глобальной оптимизации, сокращающие время расчета. Построение специальной эконометрической модели упрощает изложение.

При изучении рассмотренной в разделе III простой эконометрической модели современной экономики России получены интересные результаты. Дана оценка величины эффективного капитала и его динамики. Интересно то, что параметр μ , средний темп выбытия старых производственных фондов, принимает отрицательное значение, $\mu < 0$. В рамках рассмотренной модели это означает, что часть доставшихся от советского времени производственных фондов вовлекается в процесс производства с большим темпом, чем амортизируются производственные фонды, уже вовлеченные в процесс производства. Но такой процесс не может продолжаться бесконечно долго. Объем неиспользуемых фондов сокращается и наступит момент, когда больше не выгодно будет вовлекать в процесс воспроизводства старые, доставшиеся с советских времен фонды. Рост труда после 2008 года также не может продолжаться из-за демографических проблем. Это описывает базовый пессимистический сценарий будущего развития экономики России на основе сложившихся в ней экономических структур.

Рассмотрен и оптимистический сценарий, в котором предполагается, что уже начался рост за счет инновационной экономики. В модели это выражается в повышенной отдаче от вложенных производственных факторов. В инновационной экономике труд увеличивает свое качество, поэтому можно считать, что количество простого труда, учитываемого в производственной функции в качестве одного из факторов, возрастает даже при уменьшении числа занятых. Хотелось бы надеяться на лучшее, но такие выводы делать преждевременно. Для более полного учета процессов, происходящих в реальной экономике, и для более обоснованных рекомендаций лицам, принимающим решения, надо строить и исследовать более сложные модели системного анализа развивающейся экономики.

Многие, пока не поддающиеся решению задачи математической и прикладной экономики могут быть решены благодаря использованию современных суперкомпьютеров, так же как и слож-

ные математические задачи [27].

Использование высокопроизводительных вычислений расширяет область применения и исследования для математических методов. Так, значительное увеличение скорости оценки параметров математической модели сложной системы можно достичь за счет применения направленного перебора, реализованного в численных методах глобальной оптимизации [25].

Web-страничка данной книги находится по адресу:
<http://www.ccas.ru/olenev/parmatlab.html>

Новости и дальнейшие планы авторов книги можно узнать по указанным ниже Web-адресам их домашних страниц.
<http://www.ccas.ru/olenev/>
<http://pechonkin.pochta.ru>
<http://www.ccas.ru/chernetsov/>

Замеченные опечатки, пожелания, комментарии, просьба высылать на указанные ниже адреса электронной почты, для нас они будут очень ценными.

Оленев Николай Николаевич	olenev@ccas.ru
Печенкин Руслан Викторович	ruslan.pechenkin@mail.ru
Чернецов Андрей Михайлович	an@ccas.ru

Литература

1. *Distributed Computing Toolbox For Use with MATLAB*. 2004-2006 by The MathWorks, Inc.
http://www.mathworks.com/access/helpdesk/help/pdf_doc/distcomp/distcomp.pdf
2. *MATLAB Distributed Computing Engine For Use with MATLAB*. 2004-2006 by The MathWorks, Inc.
http://www.mathworks.com/access/helpdesk/help/pdf_doc/mdce/mdce.pdf
3. *HPC-Grid for Maple*. <http://www.maplesoft.com/products/toolboxes/HPCgrid/index.aspx>
4. *Mathematica Parallel Toolkit*. <http://documents.wolfram.com/applications/parallel/>
5. *PoochMPI Toolkit for Mathematica*
<http://daugerresearch.com/pooch/mathematica.shtml>
6. *Parallel MATLAB Survey*.
<http://www.interactivesupercomputing.com/reference/parallelMatlabsurvey.php>
7. *MPI Toolbox for MATLAB*.
http://atc.ugr.es/javier-bin/mpitb_eng

8. *LAM MPI Homepage.*
www.lam-mpi.org
9. *MPICH2 Homepage.*
<http://www-unix.mcs.anl.gov/mpi/mpich2/>
10. *MPITB for Windows.*
<http://www.wakun.com/download.htm>
11. *Akhter Sh., Roberts J.* Multi-Core Programming. Increasing Performance through Software Multi-threading. Intel Press. 2006. 344 p.
12. *Оленев Н.Н.* Основы параллельного программирования в системе MPI. М.: ВЦ РАН. 2005. 80 с.
13. *Стандарт MPI-2.* <http://www.mpi-forum.org/docs/mpi-20-html/mpi2-report.html>
14. *OpenPBS.* <http://www.openpbs.org>
15. *Sun Grid Engine.* <http://gridengine.sunsource.net>
16. *Голуб Дж., Ван Лоун Ч.* Матричные вычисления. М.: Мир, 1999. 548 с.
17. *Г.М. Михайлов, М.А. Копытов, Ю.П.Рогов, О.И. Самоваров, А.М. Чернецов.* Параллельные вычислительные системы в локальной сети ВЦ РАН. - М.: ВЦ РАН, 2003. 75 с.
18. *Г.М. Михайлов, Н.Н. Оленев, А.А. Петров, Ю.П. Рогов, А.М. Чернецов.* Опыт использования кластера ВЦ РАН в образовательных целях.//Высокопроизводительные параллельные вычисления на кластерных системах. Мат. V межд. научно-практ. сем. и Всерос. молод. шк. Нижний Новгород, 2005. С. 170-175.

19. *Bug Reports - 306262.* <http://www.mathworks.com/support/bugreports/details.html?rp=306262>
20. *Оленев Н.Н.* Параллельные вычисления для идентификации параметров в моделях экономики.//Высокопроизводительные параллельные вычисления на кластерных системах. Мат. IV межд. научно-практ. сем. и Всерос. молод. шк. / Под ред. чл.-корр. РАН В.А. Сойфера, Самара, 2004. - 280 с. С. 204-209.
21. *Бурнаев Е.В., Оленев Н.Н.* Меры близости на основе вейвлет коэффициентов для сравнения статистических и расчетных временных рядов.//Межвуз. сб. научн. и научно-методич. тр. за 2005 год (10 выпуск). Киров: Изд-во ВятГУ, 2006. С. 41-51.
22. *Оленев Н.Н.* Параллельные вычисления для оценки параметров динамической многосекторной балансовой модели региональной экономики.//Научный сервис в сети ИНТЕРНЕТ: технологии параллельного программирования: Тр. Всерос. научн. конф. - М.: Изд-во МГУ, 2006. - 303 с. С. 36-37.
23. *Оленев Н.Н.* Параллельные вычисления в математическом моделировании региональной экономики.//Параллельные вычислительные технологии - 2007. Тр. I междун. научн. конф. Челябинск: Изд-во ЮУрГУ, 2007. Т.1. С.140-151.
24. *Тейл Г.* Экономические прогнозы и принятие решений. М., 1971. 488 с.
25. *Гергель В.П., Стронгин Р.Г.* Параллельные методы вычисления для поиска глобально-оптимальных решений.//Высокопроизводительные параллельные вычисления на кластерных системах. Мат. 4-го межд. науч.-практ. сем. Самара, 2004. С. 54-59.
26. *Интрилигатор М.* Математические методы оптимизации и экономическая теория. - М.: Айрис-пресс, 2002. - 576 с.

27. *Нивергельт Ю., Фэррар Дж., Рейнголд Э.* Машинный подход к решению математических задач. - М.: Мир, 1977. - 352 с.
28. *Петров А.А., Поспелов И.Г., Шананин А.А.* Опыт математического моделирования экономики. - М.: Энергоатомиздат, 1996. - 544 с.
29. *Olenev N.* Production Function of Skilled and Unskilled Labour in a Model of a Non-Growing Russian Economy.//International Labour Market Conference Proceedings. - Aberdeen: Robert Gordon University, 1999. PP. 560-575.
30. *Burnaev E.V., Olenev N.N., Starikov A.S.* Parameter estimation of a macroeconomic model // Proceedings of the Vth Moscow Intern. Conf. on Oper. Res. (ORM2007), dedic. to the outstanding Russian scientist Nikita N. Moiseev 90th birthday. 2007. Moscow. PP.71-73.
31. *Weitzman M. L.* Soviet Postwar Economic Growth and Capital-Labor Substitution. The American Economic Review, Vol. 60, No. 4. (Sep., 1970). PP. 676-692.