

Умножение матриц в среде с распределенной памятью

С.С. Ефимов, А.В. Мартынович

Умножение матриц широко используются при выполнении многих научных и инженерных расчетов. Поэтому ускорение нахождения произведения матриц является актуальной задачей.

Многие организации и предприятия, сталкивающиеся с такими задачами, пока не имеют в распоряжении современных суперкомпьютеров, но могут выполнять расчеты с использованием кластеров невыделенных рабочих станций. Поэтому является актуальным исследование возможностей реализации алгоритмов умножения матриц в такой вычислительной среде с учетом ее технических характеристик. К ним, в частности, относятся: число узлов кластера, быстродействие процессоров, объем доступной оперативной памяти, наличие в узлах многоядерных процессоров, пропускная способность сети, характеристики сетевых карт.

Представляемые в работе результаты получены с использованием кластера факультета компьютерных наук Омского государственного университета им. Ф.М. Достоевского. Краткая характеристика кластера. Сеть: 1 Гбит/с и 100 Мбит/с; конфигурация компьютеров: P4 ~3 GHz, 1 GB и 500 Мбайт DDR; операционная система – Windows XP; количество рабочих станций: около 50-ти, из которых половина двухядерных. Имеются разные сетевые карты.

Проведено исследование параллельных вычислительных алгоритмов перемножения квадратных матриц размера от нескольких десятков строк/столбцов до нескольких тысяч. Реализованы и протестированы ленточные и блочные алгоритмы, включая алгоритмы Фокса, Кэнона, Штрассена в сочетании с блочным, алгоритм, использующий топологию гиперкуба.

Результаты аналогичных исследований, проводившихся с использованием кластера Нижегородского университета, представлены в работе [1]. Однако в ней исследованы только три метода из названных. Кроме того, по сравнению с приведенными в [1] результатами, авторам настоящей публикации удалось существенно расширить диапазон исследуемых размеров матриц: как верхний, так и нижний. Результаты в Нижегородском университете были получены с использованием кластера выделенных рабочих станций. Авторы настоящей работы добились эффективной работы алгоритмов с использованием кластера невыделенных рабочих станций, построенного на основе обычных учебных компьютерных классов.

Для исследований разработана программа с удобным интерфейсом и со всеми необходимыми встроенными возможностями, включая опрос сети, выбор метода, размера матрицы, отображение затрат времени, достигаемых ускорений и др. Все программные модули разрабатывались в среде Microsoft Visual Studio 2008 на языке C++ с использованием MPI.

Названные выше методы позволили получить ускорение больше единицы на матрицах размера от 400x400 до 5000x5000 и более при использовании одноядерных узлов. Использование двухядерных узлов позволяет получить больше единицы ускорение при уменьшении нижней границы размера матриц до 30x30.

Исследование сети кластера показало, что при размере пакетов примерно менее 10 Кбайт и более 1-10 Мбайт (в зависимости от сетевой карты) пропускная способность сети снижается. Модификация алгоритмов путем разрезки пересылаемых частей умножаемых матриц на фрагменты меньшего указанного размера позволила достичь больших ускорений.

На конференции будут представлены семейства графиков достигаемых ускорений для различных значений технических характеристик среды моделирования, в частности, скорости сети, числа ядер, доступной оперативной памяти, размера умножаемых матриц и выбранного метода умножения.

Литература

1. Гергель В. П. Теория и практика параллельных вычислений. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2007. – 423 с.
2. Kumar V., Grama A., Gupta A., Karypis G. Introduction to Parallel Computing. Second Edition. – Addison Wesley, 2003. – 856 p.