



ИССЛЕДОВАНИЕ ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ АЛГОРИТМА ГАУССА НА ВЫЧИСЛИТЕЛЬНОМ КЛАСТЕРЕ

**Г. А. ПОЛЯКОВ¹, К. В. ЛЫСЫХ¹
Е. Г. ТОЛСТОЛУЖСКАЯ²**

¹Белгородский государственный
национальный исследовательский
университет

²Харьковский национальный
университет имени В. Н. Каразина

e-mail:
tda_ua@pochtamt.ru
514453@bsu.edu.ru

В статье описаны этапы и представлены результаты исследования показателей эффективности алгоритма Гаусса на кластере НИУ «БелГУ».

Ключевые слова: MPI, алгоритм Гаусса, кластерные вычислительные системы (ВС).

Введение

Центральной проблемой вычислительной техники является повышение эффективности параллельного программного обеспечения суперЭВМ и Кластеров [1-3]. Проблема включает:

- а) выбор и оценку показателей эффективности параллельного выполнения задач;
- б) моделирование параллельного выполнения задач кластерами и оценку их показателей эффективности с учетом особенностей задач и конкретных требований/ограничений;
- в) выработку рекомендаций по выбору конкретных конфигураций кластера, ориентированных на конкретные задачи и требования/ограничения и совершенствованию систем параллельного программирования, направленных на повышение эффективности параллельных программ.

Постановка задачи

Исходная информация:

- математическая модель и Си – программа задачи – алгоритма Гаусса;
- параллельная ВС – Кластер НИУ «БелГУ»: для исследования использовались 4 вычислительных узла (ВУ). В одном вычислительном узле содержатся 2 процессора Intel Xeon E5-2665 с тактовой частотой 2,4 ГГц, 8 ядер в каждом процессоре, 64 ГБ оперативной памяти DDR3-1600. Топология коммутационной среды – полный граф. Скорость передачи данных между узлами организована по технологии Infiniband и составляет 10 Гбит/с;
- состав поддерживаемых при оценке эффективности факторов: число процессоров NM, значения t^o (typ) длительностей выполнения различных типов « typ » операций v ; время обмена одним сообщением t_c^0 ; фактическая топология коммутационной среды;
- состав показателей эффективности и поддерживаемых требований и ограничений.

Требуется получить для алгоритма Гаусса оценки зависимостей показателей эффективности параллельного решения задачи для различных значений характеристик конфигурации кластера и различных размеров матрицы системы линейных уравнений.

Математическая модель алгоритма Гаусса.

Система n линейных алгебраических уравнений:



$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= a_{1\ n+1} \\
 a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= a_{2\ n+1} \\
 \dots\dots\dots \\
 a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= a_{n\ n+1}
 \end{aligned}$$

Решение системы уравнений по компактной схеме заключается в последовательном определении x_n, x_{n-1}, \dots, x_1 из системы уравнений:

$$\begin{aligned}
 x_1 + c_{12}x_2 + \dots + c_{1n}x_n &= c_{1\ n+1} \\
 x_2 + \dots + c_{2n}x_n &= c_{2\ n+1} \\
 \dots\dots\dots \\
 x_n &= c_{n\ n+1}
 \end{aligned}$$

Эта система получается как результат разложения исходной матрицы на треугольные матрицы в соответствии со следующими формулами:

$$\begin{aligned}
 b_{ij} &= a_{ij} - \sum_{k=1}^{j-1} b_{ik} \cdot c_{kj}, \quad 1 \leq j, j \leq n, i \geq j; \\
 c_{ij} &= \left(a_{ij} - \sum_{k=1}^{i-1} b_{ik} \cdot c_{kj} \right) / b_{ii}, \quad 1 \leq j \leq n, 1 < j \leq n+1, i < j.
 \end{aligned}$$

В результате вычислений получается матрица:

$$\begin{aligned}
 &b_{11}c_{12}c_{13} \dots c_{1n}c_{1\ n+1} \\
 &b_{21}b_{22}c_{23} \dots c_{2n}c_{2\ n+1} \\
 &\dots \\
 &b_{n1}b_{n2}b_{n3} \dots b_{nn}c_{n\ n+1}.
 \end{aligned}$$

Собственно вычисление значений $x_i, i = n, n-1, \dots, 1$, выполняется по формуле

$$x_i = c_{i\ n+1} - \sum_{k=1}^{n-i} c_{i, n-k+1} \cdot x_{n-k+1}$$

Используемые показатели эффективности

Математическое ожидание времени реализации множества P операторов произвольного алгоритма определяется выражением (1) [5,6]

$$T(P) = \sum_{\xi=1}^w p_{\xi} T_{\xi}, \quad (c), \tag{1}$$

где w – число ветвей в алгоритме, p_{ξ} – вероятность реализации ξ -й ветви; T_{ξ} – время реализации ξ -й ветви параллельного алгоритма, определяемое по формуле

$$T_{\xi} = \max_{P_j \in P(\xi)} (t_j^H + t_j), \quad (c),$$



где $P(\xi)$ – множество операторов ξ -й ветви, t_j^H и t_j – момент начала и относительная временная глубина оператора $P_j \in P(\xi)$. Уменьшение временных затрат («ускорение») за счет параллельного решения задачи определяется выражением (2)

$$DT = \frac{T_{noc}(P)}{T_{пар}(P)} \quad (\text{раз}). \quad (2)$$

В соотношении (2) $T_{noc}(P)$ и $T_{пар}(P)$ – среднее время соответственно последовательной и параллельной реализации задачи. Значение показателя эффективности распараллеливания рассчитывается на основе соотношения (3)

$$R(P) = K_T \cdot \frac{DT(NM)}{NM} + K_S * S(NM). \quad (\text{раз}) \quad (3)$$

В соотношении (3) K_T и K_S являются весовыми коэффициентами, определяющими «пользовательскую» важность учета в эффективности распараллеливания величины сокращения времени реализации алгоритма ($K_T \leq 1$) и степени загрузки оборудования параллельным алгоритмом ($K_S \leq 1$).

Этапы и результаты исследования

Основными этапами исследования показателей эффективности параллельного решения задачи являлись:

- а) выбор программистом «поддерживающей задачу» конкретной конфигурации кластера;
- б) разработка текста параллельной MPI – программы исследуемой задачи;
- в) прогон параллельной MPI – программы на «поддерживающей» конфигурации кластера и оценка значений показателей эффективности;
- г) оценка дифференциала требуемых и фактических значений показателей эффективности; стоп – при выполнении требований, иначе -переход к п.е; формальная спецификация MPI – программы с использованием структур
- д) семантико – числовой спецификации (СЧС) [5,6];
- е) синтез времяпараметризованной параллельной СЧС модели процесса, минимизирующей разности дифференциалов требуемых и фактических значений показателей эффективности [5,6];
- ж) определение из СЧС модели процесса состава и величин коррекции «поддерживающей задачу» конкретной конфигурации кластера; переход к п. в.

Результаты исследования зависимостей времени параллельной реализации алгоритма Гаусса от количества процессов и размеров матрицы системы уравнений представлены в табл. 1, табл. 2 и на рис. 2 и рис. 3.

Таблица 1

Время выполнения (сек.)

Размер матрицы	Последовательный	Параллельный (np=2)	Параллельный (np=4)	Параллельный (np=8)	Параллельный (np=16)	Параллельный (np=32)
2048x2048	13,86	7,33	3,56	1,95	1,85	1,38
4096x4096	112,38	56,08	28,8	15,45	13,6	7,74
8192x8192	885,81	445,5	227,21	121,38	106,8	55,1
16384x16384	7756,05	3522,51	1807,88	948,91	518,93	432,63

Таблица 2

Ускорение (разы)

Размер матрицы	Последовательный	Параллельный (np=2)	Параллельный (np=4)	Параллельный (np=8)	Параллельный (np=16)	Параллельный (np=32)
2048x2048	-	1,89	3,89	7,10	7,49	10,04
4096x4096	-	2,003	3,9	7,27	8,26	14,51
8192x8192	-	1,98	3,89	7,29	8,29	16,08
16384x16384	-	2,201	4,29	8,17	14,95	17,92

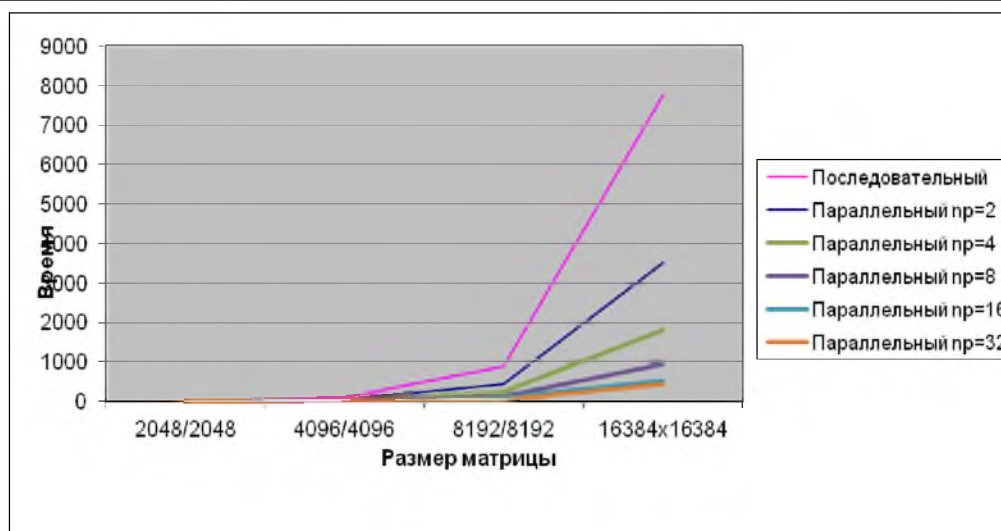


Рис.2. Зависимости времени решения задачи (сек.) от размер матрицы и количества используемых процессов

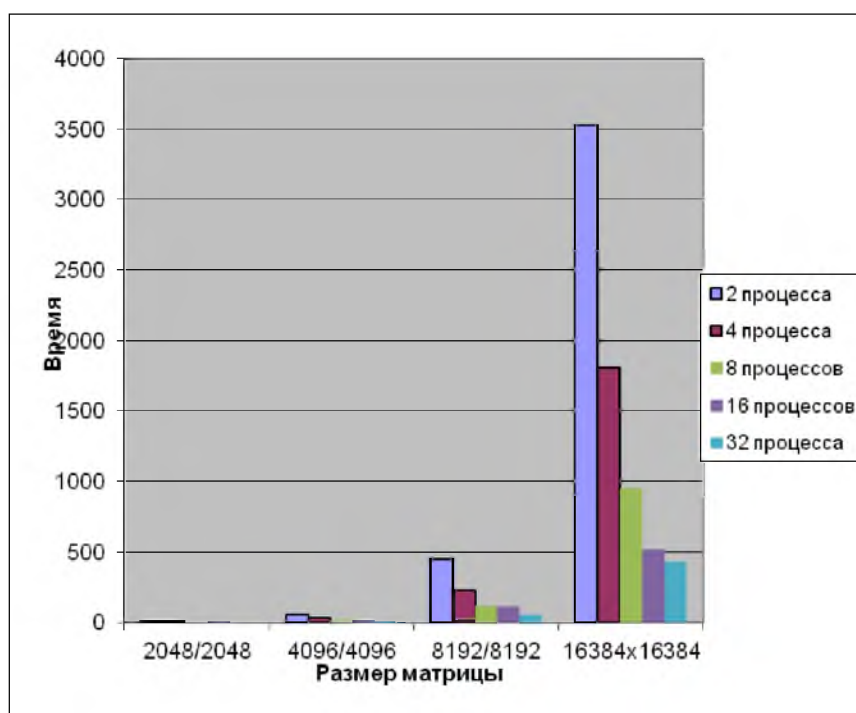


Рис.3. Гистограмма временных затрат (сек.) на выполнение алгоритма Гаусса при варьировании размеров матрицы и количества процессов

Выводы

1. В настоящее время проблема разработки эффективного параллельного программного обеспечения для известных и будущих суперЭВМ стала центральной проблемой параллельной компьютерной техники.
2. Система MPI не содержит в своем составе средств разработки параллельных программ, обеспечивающих поддержку создания параллельных программ, удовлетворяющих требованиям пользователей к времени выполнения.
3. Предложенный аналитически – имитационный подход основан на сочетании методов формальной семантико – числовой спецификации и автоматического синтеза времяпараметризованных (временных) моделей процессов и использовании средств MPI для текстовой спецификации и кластерной реализации MPI-программ. Подход



поддерживает расширение состава учитываемых факторов эффективности, поддержку требований и ограничений пользователей, формализацию и автоматизацию трудоемких этапов синтеза временных моделей процессов и обеспечивает возможность разработки в сжатые сроки высокоэффективных параллельных программ.

Список литературы

1. Воеводин В.В., Воеводин Вл. В. Параллельные вычисления. – СПб.: БХВ-Петербург, 2002. – 608 с.
2. Немнюгин С.А. Параллельное программирование для многопроцессорных вычислительных систем. / С.А. Немнюгин, О.Л. Стесик. . – СПб.: БХВ-Петербург, 2002. – 400 с.
3. Корнеев В.В. Параллельное программирование в MPI М. – Ижевск: Институт компьютерных исследований, 2003
4. Гергель В.П. Теория и практика параллельных вычислений БИНОМ. Лаборатория знаний, Интернет-университет информационных технологий – ИНТУИТ.ру, 2007
5. Поляков Г.А. Технология проектирования времяпараметризованных мультипараллельных программ как стратегия развития систем параллельного проектирования. / Г.О. Поляков, Е.Г. Толстолужская. // *Радіоелектронні і комп'ютерні системи* – Х.: Національний аерокосмічний університет ім. М.С. Жуковського «Харківський авіаційний інститут», 2009. – Вип. 6(40). – С. 166-171.
6. Поляков Г.А. Синтез и анализ параллельных процессов в адаптивных времяпараметризованных вычислительных системах /Г.А.Поляков, С.И. Шматков, Е.Г. Толстолужская, Д.А. Толстолужский: монография. – Х.: ХНУ имени В.Н. Каразина, 2012. – 672 с.
7. Лысых К.В., Поляков Г.А. «Разработка фрагментированной временной параллельной модели алгоритма гаусса на основе формальных полиномов и структур семантико-числовой спецификации»// *Научные ведомости БелГУ: история, политология, экономика, информатика* – Б.: Белгородский государственный университет, 2012 г. № 19(138) 2012 – Вып. 24/1 – С. 133-136.

RESEARCH OF THE GAUSSIAN ALGORITHM PERFORMANCE INDICATORS ON A COMPUTING CLUSTER

G.A. POLYAKOV¹
K.V. LYSYKH¹
E.G. TOLSTOLUZKA²

¹⁾ *Belgorod State National
Research University, Belgorod*

²⁾ *Kharkov National University
nm. Karazin, Kharkov*

e-mail:
tda_ua@pochtamt.ru
514453@bsu.edu.ru

The paper presents the steps and results of the Gaussian algorithm performance indicators research on a cluster of Belgorod State National Research University

Keywords: MPI, Gaussian algorithm, cluster computing systems.