

**Sukhinov Alexander Ivanovich**

Taganrog Institute of Technology – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: sukhinov@gmail.ru

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: +78634310599.

The Head of TIT SFedU; Dr. of Phis.-Math. Sc.; Professor.

**Chistyakov Alexander Evgenjevich**

E-mail: cheese\_05@mail.ru.

Phone: +78634371606.

The Department of Higher Mathematics; Assistant.

**Bondarenko Juliana Sergeevna**

E-mail: juliana\_bond@mail.ru.

Phone: +78634371606.

Postgraduate Student.

УДК 519.86

**А.И. Сухинов, В.К. Гадельшин, Д.С. Любомищенко**

**ТЕОРЕТИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ОЦЕНКИ ЭФФЕКТИВНОСТИ  
ПАРАЛЛЕЛЬНОЙ РЕАЛИЗАЦИЙ МЕТОДА СТОУНА**

*В работе приведены теоретические и практические оценки эффективности параллельной версии алгоритма неполного LU-разложения Стоуна (SIP) для суперпроизводительной вычислительной системы. В качестве базового подхода распараллеливания используется метод декомпозиции области моделирования на подобласти (domain decomposition). В рамках данного подхода рассматриваются: двумерная декомпозиция по данным без учета общего волнового фронта; двумерная декомпозиция по данным с учетом общего волнового фронта; одномерная декомпозиция по данным с учетом общего волнового фронта и пакетной организацией обменов.*

*Показано, что для задачи приземной аэродинамики наилучшие результаты из трех оцениваемых параллельных версий дает алгоритм одномерной декомпозиции с учетом общего волнового фронта. Приведены оценки оптимального размера пакета передачи.*

*Метод Стоуна; оценки эффективности для системы с массовым параллелизмом.*

**A.I. Sukhinov, V.K. Gadelshin, D.S. Lyubomishchenko**

**THEORETICAL AND PRACTICAL EFFICIENCY ASSESSMENT OF PARALLEL  
VERSION OF STONE METHOD**

*In the paper there are theoretical and practical efficiency assessments of parallel method of incomplete factorization of Stone (SIP) for supercomputer. Domain decomposition is used as background parallel method. 2D data decomposition without wave front, 2D decomposition with wave front and 1D data decomposition with wave front and pocket data exchange are considered. The best results among three algorithm versions of parallel realization are achieved for wave front algorithm and 1D domain decomposition. The assessment of optimal package size is given.*

*Stone Implicit Procedure (SIP); efficiency assessment for MPP system.*

Для решения задач атмосферного моделирования очень важным аспектом является выбор параллельного алгоритма. Процесс построения параллельной программы не всегда может быть ограничен простой переделкой последовательных линейных конструкций в параллельное представление. Такой подход может дать параллельный аналог алгоритма, который при масштабировании на большое коли-

чество узлов параллельной вычислительной системы будет давать незначительное преимущество или не давать преимуществ вовсе по сравнению с последовательным алгоритмом, если учесть трудозатраты на разработку параллельной версии. Также стоит отметить, что многие параллельные алгоритмы вовсе не имеют последовательных аналогов, что существенно усложняет их разработку и отладку. Поэтому перед началом создания программного средства для многопроцессорной вычислительной системы необходимо произвести теоретические оценки эффективности.

Анализ производительности параллельных алгоритмов чаще всего проводят с использованием понятий ускорения и эффективности, определяемых соответственно по формулам

$$S_n = \frac{T_s}{T_n}, \quad (1)$$

$$E_n^{tot} = \frac{T_s}{nT_n}, \quad (2)$$

где  $T_s$  – время выполнения лучшего последовательного алгоритма,  $T_n$  – время выполнения параллельного алгоритма на  $n$  процессорах.

Получим теоретические оценки эффективности параллельных реализаций метода SIP с учетом основных параметров параллельной вычислительной системы. В ТТИ ЮФУ установлен НРС-кластер, представляющий собой систему с распределенной памятью из 128 узлов, объединенных сетью InfiniBand. Каждый узел кластера является системой с общей памятью (SMP-системой) с 16-ю вычислителями и объемом ОЗУ 32 Гбайт. Фактическая пропускная способность и латентность сети Infiniband для узлов в разных корзинах (худший случай):

- ◆ латентность:  $7,761 \cdot 10^{-3}$  с;
- ◆ среднее время одной операции с плавающей точкой:  $\tau = 5,376 \cdot 10^{-7}$  с;
- ◆ среднее время пересылки одного элемента массива:  $\tau_{send} = 8,595 \cdot 10^{-9}$  с;
- ◆ пропускная способность: 1224,38 Гбайт/с;
- ◆ пиковая производительность одного вычислительного ядра узла кластера: 9,2 ГФлопс.

В работе рассматривается одномерная и двумерная топологии процессоров. Количество процессоров в параллельной вычислительной системе равно  $P$ . Для двумерной топологии  $P = P_x \cdot P_y$ , где  $P_x$  и  $P_y$  равны количеству процессоров соответственно в направлениях  $X$  и  $Y$ . Для определенности будем считать, что область моделирования имеет квадратное основание и одинаковое количество контрольных объемов (КО)  $N_h$  в горизонтальных направлениях и  $N_v$  – количество КО в вертикальном направлении.

Подсчитаем время  $T_s$  работы одной итерации последовательного алгоритма SIP. Это время получается суммированием временных затрат на каждый из этапов алгоритма, т.е. декомпозицию, прямую подстановку, обратную подстановку:

$$T_s = T_{dec} + T_{frw} + T_{bkw}, \quad (3)$$

$$T_{dec} = N_h^2 N_v \tau \cdot 28, T_{frw} = N_h^2 N_v \tau \cdot 17, T_{bkw} = N_h^2 N_v \tau \cdot 5, \quad (4)$$

$$T_s = 50 N_h^2 N_v \tau. \quad (5)$$

Подсчитаем время  $T_p^1$  работы одной итерации параллельного алгоритма без учета общего волнового фронта. Это время складывается из времен, необходимых на осуществление трех этапов алгоритма *SIP*, и времени на синхронизацию данных на границе:

$$T_p^1 = T_{calc}^1 + T_{sync}^1, \quad (6)$$

где

$$T_{calc}^1 = T_{dec} + T_{frw} + T_{bkw} = 50 N_h^2 N_v \tau. \quad (7)$$

Время синхронизации равно

$$T_{sync}^1 = T_{lat}^1 + T_{send}^1, \quad (8)$$

где  $T_{lat}^1$  – время, необходимое на организацию обмена;  $T_{send}^1$  – время на пересылку.

Процесс синхронизации проходит вначале в направлении  $\mathcal{X}$ , а затем в направлении  $\mathcal{Y}$  двумерной топологии. Рассмотрим направление  $\mathcal{X}$ , а для направления  $\mathcal{Y}$  процессы обменов протекают аналогично. Синхронизация в направлении  $\mathcal{X}$  протекает в два этапа: сначала осуществляются обмены нечетных и четных узлов, затем четных и нечетных. На каждом этапе тратится время на организацию двух обменов  $2\tau_{lat}$  и двух пересылок двух плоскостей данных синхронизации

$2 \left( \frac{2N_h N_v}{p_y} \tau_{send} \right)$ . Таким образом, на синхронизацию в направлении  $\mathcal{X}$  затрачивается время

$$2 \left[ 2\tau_{lat} + 2 \left( \frac{2N_h N_v}{p_y} \tau_{send} \right) \right], \quad (9)$$

а в направлении  $\mathcal{Y}$  – время

$$2 \left[ 2\tau_{lat} + 2 \left( \frac{2N_h N_v}{p_x} \tau_{send} \right) \right]. \quad (10)$$

Тогда общее время синхронизации равно

$$T_{sync}^1 = 8\tau_{lat} + 4 \left( \frac{2N_h N_v}{p_y} \tau_{send} \right) + 4 \left( \frac{2N_h N_v}{p_x} \tau_{send} \right). \quad (11)$$

Время работы одной итерации алгоритма двумерной декомпозиции без учета общего волнового фронта равно

$$T_p^1 = 50 \frac{N_h}{p} \frac{N_h}{p} N_v \tau + 8\tau_{lat} + 8N_h N_v \left( \frac{1}{p_y} + \frac{1}{p_x} \right) \tau_{send}. \quad (12)$$

Время работы алгоритма с учетом волнового фронта так же как и в последовательном случае складывается из временных затрат на три этапа алгоритма SIP:

$$T_p^2 = T_{dec}^2 + T_{frw}^2 + T_{bkw}^2. \quad (13)$$

Каждый из этапов проходит  $(p_x + p_y - 1)$ -стадий и на  $(p_x + p_y - 2)$ -стадиях требуется осуществлять синхронизационные обмены.

Таким образом:

$$T_{dec}^2 = (p_x + p_y - 1)T_{dec\_calc}^2 + (p_x + p_y - 2)T_{dec\_sync}^2, \quad (14)$$

$$T_{frw}^2 = (p_x + p_y - 1)T_{frw\_calc}^2 + (p_x + p_y - 2)T_{frw\_sync}^2, \quad (15)$$

$$T_{bkw}^2 = (p_x + p_y - 1)T_{bkw\_calc}^2 + (p_x + p_y - 2)T_{bkw\_sync}^2. \quad (16)$$

Времена, необходимые на осуществление процесса расчета для этапов декомпозиции, прямой и обратной подстановок, запишутся соответственно как

$$T_{dec\_calc}^2 = 28 \frac{N_h}{p_x} \frac{N_h}{p_y} N_v \tau, \quad (17)$$

$$T_{frw\_calc}^2 = 17 \frac{N_h}{p_x} \frac{N_h}{p_y} N_v \tau, \quad (18)$$

$$T_{bkw\_calc}^2 = 5 \frac{N_h}{p_x} \frac{N_h}{p_y} N_v \tau. \quad (19)$$

На этапе декомпозиции и прямой подстановки необходимо передавать одну плоскость синхронизационных данных в каждом направлении, а на этапе обратной подстановки – две плоскости в каждом направлении. Как и прежде, время синхронизации по каждому направлению складывается из времени на организацию соединения и времени на передачу данных. Таким образом, времена синхронизации равны

$$T_{dec\_sync}^2 = T_{frw\_sync}^2 = 2\tau_{lat} + \frac{N_h N_v}{p_y} \tau_{send} + \frac{N_h N_v}{p_x} \tau_{send}, \quad (20)$$

$$T_{bkw\_sync}^2 = 2\tau_{lat} + \frac{2N_h N_v}{p_y} \tau_{send} + \frac{2N_h N_v}{p_x} \tau_{send}. \quad (21)$$

В итоге время работы алгоритма волнового фронта на  $p$ -процессорах имеет вид

$$T_p^2 = 50 \frac{N_h}{p_x} \frac{N_h}{p_y} N_v (p_x + p_y - 1) \tau + (p_x + p_y - 2) \left[ 6\tau_{lat} + 4N_h N_v \left( \frac{1}{p_y} + \frac{1}{p_x} \right) \tau_{send} \right]. \quad (22)$$

Рассмотрим алгоритм одномерной декомпозиции по данным с учетом общего волнового фронта и пакетной организацией обменов. Время работы алгоритма, как и в случае двумерной декомпозиции с учетом общего волнового фронта, имеет вид

$$T_p^3 = T_{dec}^3 + T_{frw}^3 + T_{bkw}^3, \quad (23)$$

где

$$T_{dec}^3 = (p_x + p_y - 1)T_{dec\_calc}^3 + (p_x + p_y - 2)T_{dec\_sync}^3, \quad (24)$$

$$T_{frw}^3 = (p_x + p_y - 1)T_{frw\_calc}^3 + (p_x + p_y - 2)T_{frw\_sync}^3, \quad (25)$$

$$T_{bkw}^3 = (p_x + p_y - 1)T_{bkw\_calc}^3 + (p_x + p_y - 2)T_{bkw\_sync}^3. \quad (26)$$

Пусть  $m$  – количество пакетов синхронизации для каждого вычислителя, тогда времена, необходимые на осуществление процесса расчета для этапов декомпозиции, прямой и обратной подстановок, запишутся соответственно как

$$T_{dec\_calc}^3 = 28 \frac{N_h N_h}{p m} N_v \tau, \quad (27)$$

$$T_{frw\_calc}^3 = 17 \frac{N_h N_h}{p m} N_v \tau, \quad (28)$$

$$T_{bkw\_calc}^3 = 5 \frac{N_h N_h}{p m} N_v \tau. \quad (29)$$

Так как декомпозиция осуществляется по одному направлению, то и синхронизация проходит только в одном направлении. Времена синхронизации для трех этапов равны

$$T_{dec\_sync}^3 = T_{frw\_sync}^3 = \tau_{lat} + \frac{N_h}{m} N_v \tau_{send}, \quad (30)$$

$$T_{bkw\_sync}^3 = \tau_{lat} + \frac{2N_h}{m} N_v \tau_{send}. \quad (31)$$

Время на осуществление одной операции:

$$T_p^3 = 50 \frac{N_h N_h}{p m} N_v (p + m - 1) \tau + (p + m - 2) \left[ 3\tau_{lat} + 4 \frac{N_h}{m} N_v \tau_{send} \right]. \quad (32)$$

Запишем формулы для теоретических оценок эффективности алгоритмов [5,6]:

$$E_p^1 = \frac{T_s}{pT_p^1} = \frac{50N_h^2 N_v \tau}{p_x p_y \cdot \left[ 50 \frac{N_h N_h}{p_x p_y} N_v \tau + 8\tau_{lat} + 8N_h N_v \left( \frac{1}{p_y} + \frac{1}{p_x} \right) \tau_{send} \right]}, \quad (33)$$

$$\begin{aligned}
 E_p^2 &= \frac{T_s}{pT_p^2} = \\
 &= \frac{50N_h^2N_v\tau}{p_x p_y \cdot \left[ 50 \frac{N_h}{p_x} \frac{N_h}{p_y} N_v (p_x + p_y - 1)\tau + \right.} \\
 &\quad \left. + (p_x + p_y - 2) \left( 6\tau_{lat} + 4N_h N_v \left( \frac{1}{p_y} + \frac{1}{p_x} \right) \tau_{send} \right) \right]}, \\
 E_p^3 &= \frac{T_s}{pT_p^3} = \\
 &= \frac{50N_h^2N_v\tau}{p \cdot \left[ 50 \frac{N_h}{p} \frac{N_h}{m} N_v (p + m - 1)\tau + (p + m - 2) \left( 3\tau_{lat} + 4 \frac{N_h}{m} N_v \tau_{send} \right) \right]} \cdot (35)
 \end{aligned}$$

Построим графики теоретических оценок эффективности с учетом того, что  $N_h = 2000$ ,  $N_v = 40$ ,  $p_x = p_y = \sqrt{p}$  (рис. 1).

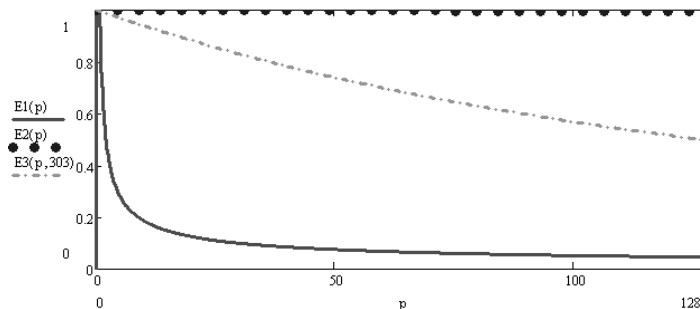


Рис. 1. Результаты теоретических оценок эффективности параллельных версий алгоритма SIP

Из результатов теоретических оценок эффективности становится ясно, что алгоритм двумерной декомпозиции по данным с учетом общего волнового фронта можно применять только на небольшом (<16) количестве вычислителей. Алгоритм без учета общего волнового фронта дает отличную асимптотическую эффективность, однако в теоретических оценках не было учтено ограничение на пропускную способность сети многопроцессорной системы, которое на практике является естественным ограничителем алгоритма. Также в алгоритме без учета общего волнового фронта сходимость итерационного процесса может быть существенно снижена за счет отсутствия необходимых обменов на этапах декомпозиции, прямой и обратной подстановки. Одномерная декомпозиция дает выигрыш по сравнению с двумерным алгоритмом волнового фронта за счет сокращения в 2 раза временных издержек на передачу синхронизационных данных на границе. Поэтому целесообразным является реализация алгоритма двумерной декомпозиции по данным без учета волнового фронта и алгоритма одномерной декомпозиции по данным с пакетной передачей сообщений.

Найдем оптимальное значение параметра разбиения  $m$  в алгоритме одномерной декомпозиции при заданном количестве вычислителей  $p$ . Для этого нужно минимизировать время  $T_p^3$ :

$$T_p^3 = 3\tau_{lat}m + \frac{1}{m} \left[ 50(p-1) \frac{N_h^2}{p} N_v \tau + 4(p-2) N_h N_v \tau_{send} \right] + 50 \frac{N_h^2}{p} N_v \tau + 4 \frac{N_h}{1} N_v \tau_{send} + 3(p-2) \tau_{lat}. \quad (36)$$

Нетрудно видеть, что минимум достигается при

$$m = \sqrt{\frac{50(p-1) \frac{N_h^2}{p} N_v \tau + 4(p-2) N_h N_v \tau_{send}}{3\tau_{lat}}}. \quad (37)$$

В результате для  $p=128$  оптимальное значение  $l=6$ , т.е. процесс синхронизации необходимо осуществлять после обработки шести строк каждым процессором.

В табл. 1 приведены временные затраты и практические замеры эффективности алгоритмов.

Таблица 1

**Время выполнения одной итерации и эффективность**

Количество процессоров	Время $T_p^3$ , с.	$E_p^3$
2	8,3823	0,743
4	4,7790	0,651
8	2,6520	0,587
16	1,4100	0,552
32	0,7913	0,492
64	0,4730	0,411
128	0,2930	0,332

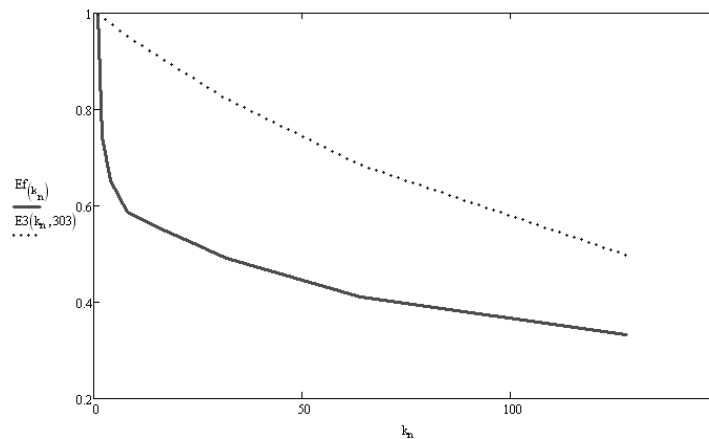


Рис. 2. Теоретическая (прерывистая) и практическая (сплошная) кривые эффективности

Как и ожидалось, результаты практических замеров показали существенное ухудшение характеристик алгоритма (рис. 2). Это связано с тем, что в практических оценках не учитываются ограничения на пропускную способность сети, время на выделение, удаление и доступа к памяти массивов.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Сухинов А.И., Гадельшин В.К., Любомищенко Д.С.* Математическое моделирование поля ветровых течений в условиях городской застройки с применением SIP-метода // Известия ЮФУ. Технические науки. – 2009. – № 8 (97). – С. 30-37.
2. *Сухинов А.И., Гадельшин В.К., Любомищенко Д.С.* Математическая модель распространения вредных выбросов от автотранспортных средств на основе метода контрольного объема и ее параллельная реализация на кластере распределенных вычислений // Известия ЮФУ. Технические науки. – 2009. – № 2 (91). – С. 8-14.
3. *Ferziger J., Peric M.* Computational Methods for Fluid Dynamics - 3., rev. ed. – Berlin; Heidelberg; New York; Barcelona; Hong Kong; London; Milan; Paris; Tokyo: Springer, 2002. – 423 p.
4. *Stone H.L.* Iterative Solution of Implicit Approximations of Multidimensional Partial Differential Equations, SIAM J. Numerical Anal. – 1968. – Vol. 5. – P. 530-558.
5. *Гергель В.П.* Теория и практика параллельных вычислений: Учебное пособие. – М.: Интернет-Университет Информационных Технологий; БИНОМ. Лаборатория знаний, 2007. – 423 с.
6. *Корнеев В.В.* Параллельные вычислительные системы. – М.: Ноледж, 1999. – 320 с.
7. Параллельная версия алгоритма неполного  $LU$ -разложения Стоуна [Электронный ресурс] Parallel Versions of Stone's Strongly Implicit Algorithm; авт. J.S. Reeve, A.D. Scurr and J.H. Merlin / Department of Electronics and Computer Science University of Southampton. – Электрон. дан. – Southampton, [2001]. – Режим доступа: <http://eprints.ecs.soton.ac.uk/7614/>, свободный. Дата доступа: 25.10.2010. Загл. с экрана. – Яз. англ.
8. *Сухинов А.И.* Двумерные схемы расщепления и некоторые их приложения. – М.: МАКС Пресс, 2005. – 408 с.

Статью рекомендовал к опубликованию д.т.н., профессор Н.Н. Чернов.

**Сухинов Александр Иванович**

Технологический институт федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет» в г. Таганроге.

E-mail: [sukhinov@gmail.ru](mailto:sukhinov@gmail.ru).

347928, г. Таганрог, пер. Некрасовский, 44.

Тел.: 88634310599; 89281021106.

Руководитель ТТИ ЮФУ; д.ф.-м.н.; профессор.

**Гадельшин Валерий Камельянович**

Тел.: 88634601461.

Кафедра высшей математики; к.т.н.; доцент.

**Любомищенко Денис Сергеевич**

E-mail: [dexusint@gmail.com](mailto:dexusint@gmail.com).

Тел.: 88634601219.

Кафедра высшей математики; ассистент.

**Sukhinov Alexander Ivanovich**

Taganrog Institute of Technology – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”.

E-mail: [sukhinov@gmail.ru](mailto:sukhinov@gmail.ru)

44, Nekrasovskiy, Taganrog, 347928, Russia.

Phone: +78634310599.

The Head of TIT SFedU; Dr. of Phis.-Math. Sc.; Professor.



**Gadelshin Valeriy Kamel'ianovich**

Phone: +78634601461.

The Department of Higher Mathematics; Cand. of Eng. Sc.; Associate Professor.

**Lyubomishenko Denis Sergeevich**

E-mail: dexusint@gmail.com.

Phone: +78634601219.

The Department of Higher Mathematics; Assistant.