

В. В. Чуданов, С. А. Горейнов, А. Е. Аксенова, В. А. Первичко,  
А. А. Макаревич

## Новый метод решения CFD задач на кластерных ЭВМ петафлопсной производительности<sup>1</sup>

АННОТАЦИЯ. Описывается новый метод решения CFD задач на кластерных ЭВМ петафлопсной производительности. Метод обладает высокой эффективностью и позволяет достигать значительного быстродействия за счет прямой реализации. Продемонстрирована масштабируемость близкая к идеальной.

*Ключевые слова и фразы:* CFD, кластерные ЭВМ, петафлопсная производительность.

### Введение

В ИБРАЭ имеется многолетний опыт разработки численных методов [1] для решения CFD задач как несжимаемой, так и слабо сжимаемой жидкости [2]–[4]. При решении задач гидродинамики несжимаемой жидкости основной трудностью является решение эллиптического уравнения для задачи поправки давления (уравнения Пуассона с дивергентной правой частью), поскольку оно является крайне ресурсоемким и занимает почти 90% всего времени решаемой проблемы. Чтобы обойти трудности, возникающие при решении уравнения эллиптического типа, в качестве альтернативы предлагается использовать алгоритмы с малой сжимаемостью, что позволяет избежать трудностей решения уравнения Пуассона. Особого подхода при переходе к кластерным системам требует и решение уравнений параболического типа, возникающих в результате расщеплении по физическим процессам решаемых задач математической физики.

---

<sup>1</sup> Статья рекомендована к публикации в журнале «Программные системы: теория и приложения» Программным комитетом НСКФ-2013.

Наблюдающийся в последние десятилетия стремительный рост вычислительных мощностей оказывает существенное влияние на разрабатываемые методы и вычислительные алгоритмы решения во всех областях науки. В настоящий момент уже стали привычными кластерные системы с быстродействием в несколько терафлопс ( $10^{12}$  операций в секунду). Появились первые многопроцессорные компьютерные системы петафлопсной производительности ( $10^{15}$  операций в секунду), а к 2020 году ожидается появление многопроцессорных вычислительных систем с эксафлопсной ( $10^{18}$  операций в секунду) производительностью.

Реальное быстродействие или вычислительная эффективность многопроцессорных вычислительных систем на прикладных задачах зависит не только от успехов в области электроники, но и от эффективности алгоритмов распараллеливания и качества численных методов, применяемых для расчета физических процессов. Переход от одних разностных схем к другим может повысить реальное быстродействие на несколько порядков. Безусловно, такая оценка является очень приблизительной, она может иметь место только в том случае, если переход к новым вычислительным методам не ухудшает характеристик их эффективного распараллеливания.

Принимая во внимание сказанное выше необходимо, учитывая темпы развития вычислительных мощностей, разрабатывать современные и адаптивные вычислительные алгоритмы и расчетные солверы, позволяющие осуществлять переход со стандартных систем с общей памятью на системы с распределенной памятью и супербольшие кластерные ЭВМ.

Для решения названной задачи в ИБРАЭ разрабатываются новые методы [2], [3], [7], позволяющие получать эффективные результаты, в частности, с помощью стандартных процедур на системах с распределенной памятью, использующих MPI-технологии.

В данной работе описывается прямой метод решения эллиптических уравнений (ПМРЭУ), а также быстрый прямой солвер

(SMITH) для решения линейных систем с сепарабельной блочно-трехдиагональной матрицей для кластерных компьютеров.

## 1. Особенности метода

Существует достаточно много задач, требующих перебора большого количества комбинаций данных для выбора оптимального решения. Такие задачи очень трудоемки для памяти персонального компьютера, поэтому существуют методы ограниченного перебора (оптимизация решения задачи в процессе решения). Одним из эффективнейших методов построения и анализа алгоритмов является метод, основанный на стратегии «разделяй и властвуй». Алгоритм разделяет задачу на более мелкие подзадачи, а затем собирает решение основной задачи «снизу вверх». Метод применим только в случаях, когда подзадачи являются независимыми. Если подзадачи будут взаимозависимыми, то алгоритм будет делать лишнюю работу, решая одни и те же подзадачи по несколько раз. Метод первоначально был введен в [5]. После этого еще несколько работ были опубликованы по той же теме, включая [6].

Разбиение исходного множества на подмножества должно производиться таким образом, чтобы в каждом полученном подмножестве доминировал (властвовал) один из классов. Параметры разбиения на каждом шаге должны выбираться так, чтобы обеспечить максимальное преобладание одного из классов. Например, в случае решения линейных систем это отделяемые блоки трехдиагональных матриц.

В общем, стратегия «разделяй и властвуй» сводится к трем основным шагам. Во-первых, к разделению задачи на подзадачи, как правило, меньшего размера. Во-вторых, решению каждой из подзадач (напрямую, если они достаточно небольшого объема — иначе рекурсивно, разбивая на меньшие части). В-третьих, объединению полученных решений подзадач.

Существует несколько пакетов, находящихся в свободном доступе, которые принадлежат описываемому классу солверов и позволяют решать эллиптические уравнения, дискретизованные на регулярной трехмерной сетке [6], [9], [10]. Наше внимание привлек-

ли подпрограммы, реализующие так называемый быстрый прямой метод [6] для двумерного и трехмерного случая (солверы dc2d и dc3d) — по объему кода и по быстродействию они оказались привлекательнее как многосеточного метода SMG в реализации HYPRE [9], так и алгебраического аггломерационного метода AMG в реализации DUNE [10]. Мы провели анализ данных подпрограмм и пришли к выводу, что они имеют ряд ограничений, которые мы и попытались устранить [4], [7] в нашей версии прямого метода решения эллиптических уравнений, получившего сокращенное название ПМРЭУ. Далее, мы будем говорить только о трехмерной версии.

Сначала перечислим обнаруженные нами ограничения:

- большие вычислительные погрешности в случае сеток с переменными шагами;
- невозможность счета при количестве процессоров пр свыше 256;
- невозможность счета при  $np_1 \neq np_2$ .

Алгоритм ПМРЭУ был реализован в виде солвера SMITH, который в отличие от солвера dc3d позволяет:

- кроме дискретного аналога задачи

$$cu + \operatorname{div} \begin{pmatrix} \kappa_1(x) \\ \kappa_2(y) \\ \kappa_3(z) \end{pmatrix} \operatorname{grad} u = f$$

решать и дискретный аналог  $\operatorname{div}(\kappa_1(x)\kappa_2(y)\kappa_3(z)\operatorname{gradu}) = f$  ;

- использовать условия Дирихле или Неймана общего вида;
- аппроксимация краевого условия Неймана выполнена со вторым порядком точности вместо первого;
- вместо сетки процессоров  $2^k \times 2^k \times 1$  нами используется сетка  $2^k \times 2^j \times 1$ , где  $k$  и  $j$  не связаны друг с другом.

ПМРЭУ предназначен для решения линейных систем с сепарабельными ленточными матрицами вида:

$$A = A_1 \times K_2 \times K_3 + K_1 \times A_2 \times K_3 + K_1 \times K_2 \times A_3,$$

где  $A \times B = [a_{kj} B_{-kj}]$  означает кронекерово произведение матриц,  $A_j$  — симметричные трехдиагональные, а  $K_j$  — диагональные матрицы.

Сведение простейшей дискретизации уравнения

$$\operatorname{div}(\kappa_1(x)\kappa_2(y)\kappa_3(z)\operatorname{grad} u) = f,$$

рассматриваемого на бруске с граничными условиями Дирихле и Неймана к виду  $A = A_1 \times K_2 \times K_3 + K_1 \times A_2 \times K_3 + K_1 \times K_2 \times A_3$  в данной работе приведено не будет. Авторами запланирована отдельная публикация разработанного подхода в журнале «Известия Российской академии наук. Энергетика».

## 2. Особенности параллелизации

Параллелизация основного алгоритма состоит в разбиении полос всех уровней по вычислительным ядрам. Единственная техническая проблема состоит в передаче данных Дирихле между соседними полосами. Основная трудность параллелизации заключается в эффективной реализации частичных задач. А именно, все три указанных выше шага требуют разбиения по вычислительным ядрам; при этом первый и третий шаги требуют обменов, тогда как решения систем на втором шаге не зависят друг от друга.

И при разбиении полос, и при решении частичных задач необходимо учитывать, что элементом решения трехмерной задачи является задача двумерная, так что между выделенным числом ядер на полосе, соответствующую трехмерной задаче, и числом ядер на «двумерную» полосу будет жесткая связь.

Пусть число процессоров по одной оси,  $np_i$ , есть степень двойки. Локальное число полос (locs) на уровне  $\lambda$  тогда будет

$$\operatorname{locs} = 4^{\lambda-1} / np_i.$$

Минимальное условие баланса работы состоит в том, чтобы на последнем уровне не было не занятых работой ядер. Поскольку последний уровень может быть нерегулярным, это означает, что

$l_{ocs}$  не должно быть менее 4. В силу следующей леммы [7], неравенство  $8np_i \leq np_i + 1$  гарантирует нам  $l_{ocs} \geq 4$ .

*Лемма:* Пусть длина исходной полосы равна  $n_i + 1$ , где  $n_i \in \mathbb{N}$ . Пусть  $\lambda$  — последний уровень, то есть такой, на котором  $2 \leq x_\lambda \leq 3$  или  $3 \leq x_\lambda \leq 4$ . Тогда справедливы оценки

$$n_i + 1 \leq 4^\lambda, \quad n_i \geq 4^{\lambda-1}.$$

Использованы обозначения:  $x$  — длина полосы,  $x = kq + r$ , где  $0 \leq r < k$ ,  $k$  — число частей, на которое разбивается полоса,  $k \geq 1$ .

Далее, предположим, что разбиение  $n_i + 1$  степеней свободы по  $np_i$  ядрам происходит по целым полосам последнего ( $n_{lev}$ ) уровня — из условия наилучшего баланса. На последнем уровне работа состоит в решении подзадач, размерность которых по разбиваемой оси составляет 1, 2 или 3 (длина полосы без единицы). Поэтому естественно ограничить число ядер для такой подзадачи малым числом. Мы используем ограничение: число ядер по оси  $i$  на полосу ( $nps$ ) для последнего уровня не более 1. В силу леммы, неравенство

$$4np_i \leq n_i + 1$$

гарантирует нам  $nps = np_i / 4^{\lambda-1} \leq 1$ .

Детали параллельной реализации кода SMITH подробно изложены в работе [7]. Приведем здесь только оценку времени параллельного счета.

*Теорема:* Пусть модель сети такова, что пересылки по разным парам адресов не мешают друг другу, и время пересылки сообщения длины  $m$  равно  $\alpha + \beta m$ . Пусть также время выполнения одного флопа равно  $\varphi$ , размерность процессорной сетки  $np_x \times np_y \times \dots$ , и размерность расчетной сетки  $n_x \times n_y \times \dots$ , тогда параллельное время счета нашей реализации быстрого прямого метода (SMITH) для задачи в  $d$  измерениях допускает оценку

$$T_d^{(x,y,\dots,z)} \leq C_1(\alpha + \beta n_y \cdots n_z) \log^2 np_x +$$

$$C_2 \frac{n_x \log n_x}{np_x} T_{d-1}^{(y,\dots,z)} +$$

$$C_3 n_y \cdots n_z \left( \frac{n_x}{np_x} + \log np_x \right).$$

В частности, на кубических сетках  $n \times n \times n$  с общим числом ядер  $N_p$ , равномерно распределенным по осям, имеем оценку

$$T_3 \leq (C_4 \alpha + C_5 n^2 \beta) \log N_p + C_6 \left( \frac{n^3}{N_p^{1/3}} + \frac{n^3 \log n}{N_p} \right).$$

### 3. Основные результаты

В данном разделе приведены результаты расчета тестовой задачи естественной конвекции в полости с разнотемпературными стенками (международная база данных ERCOFTAC) с помощью кода SMITH. Масштабируемость кода SMITH сравнивалась в двух формах, «слабой» и «сильной». Здесь мы приводим результаты только сильной масштабируемости. Более детально этот вопрос обсуждается в [4].

Сильная масштабируемость характеризуется числом

$$\frac{p_0 T(N, p_0)}{p T(N, p)}.$$

График, иллюстрирующий динамику изменения времени расчета в секундах в зависимости от количества процессоров при сильной масштабируемости, полученный на суперкомпьютере BlueGene/Q (ANL) представлен на рисунке Рис. 1 и демонстрирует идеальную масштабируемость вплоть до 256 тысяч процессоров. Что позволяет говорить о перспективности развитого подхода при наличии стабильно работающего оборудования.

С помощью разработанного метода были также проведены CFD расчеты слепого теста OECD/NEA-MATIS-H (KAERI, Корея, CFD4NRS4) на кластере BlueGene/Q(ANL). MATIS-H — тестовая установка с холодным циклом, используется в качестве гидравлического теста в связке стержней с разными типами дистанциони-

рующих решеток (ДР) в условиях нормального давления и температуры.

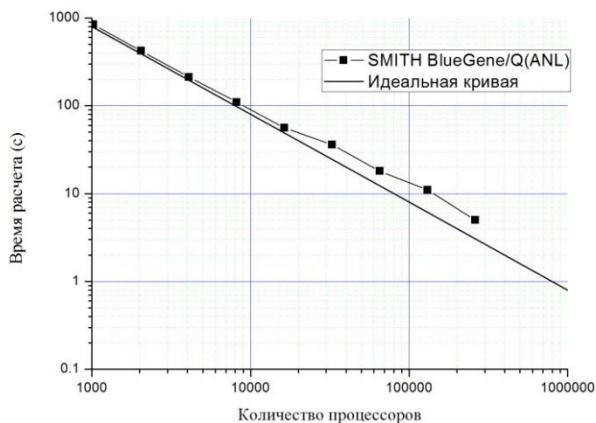


Рис. 1. Расчет эксперимента ERCOFTAC с помощью SMITH-солвера на кластере BlueGene/Q(ANL)

На рисунке Рис. 2 показана ДР split типа и результаты расчета усредненной скорости  $V$  на удалении от ДР 0.5 диаметра вниз по течению при  $y = 16.56$  мм в сравнении с экспериментальными данными.

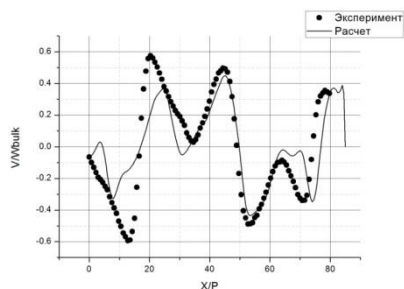
На рисунке Рис. 3 показана ДР split типа вместе со стрижнями и результаты расчета усредненной  $gms$  скорости  $V$  на удалении от ДР 0.5 диаметра вниз по течению при  $y = 16.56$  мм в сравнении с экспериментальными данными. Как можно видеть из рисунков, совпадение в обоих случаях очень хорошее.

По результатам теста институтом KAERI (Корея) в июле этого был выпущен отчет [8], где в рейтинге кодов при расчете эксперимента MATIS-H код CONV-3D (ИБРАЭ) занял общее 4 место. Для сравнения американский код Nek5000 (ANL) занял общее 3 место.



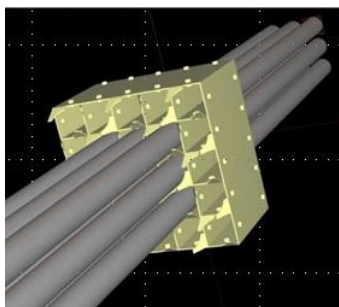


DP split типа

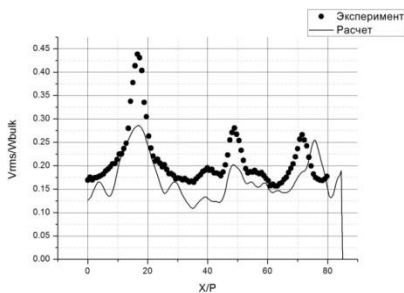


$V/W_{bulk}$

Рис. 2. Результаты расчета OECD/NEA-MATIS-H теста на кластере BlueGene/Q(ANL)



DP + стержни



$V_{rms}/W_{bulk}$

Рис. 3. Результаты расчета OECD/NEA-MATIS-H теста на кластере BlueGene/Q(ANL)

## Закключение

Разработан новый метод и на его основе создан и протестирован быстрый прямой солвер для решения линейных систем с отдельным блоком трехдиагональных матриц для кластерных компьютеров.

Солвер обладает высокой эффективностью и позволяет достигать значительного быстродействия за счет прямой реализации. Продемонстрирована масштабируемость близкая к идеальной.

Требуется проведение дальнейших исследований разработанного алгоритма, а также кода SMITH.

### Список литературы

- [1] Chudanov V. V., Churbanov A. G., Vabishchevich P. N., Strizhov V. F., *A numerical study on natural convection of a heat-generating fluid in rectangular enclosures* // Int.J.Heat Mass Transfer. 1994. № 37(18), p. 2969–2984.
- [2] Чуданов В. В., Аксенова А. Е., Первичко В. А. *Методы прямого численного моделирования турбулентности с использованием DNS и LES подходов в задачах теплогидравлики ТВС* // Известия Академии Наук, серия Энергетика, 2007. № 6, с. 47–57.
- [3] Чуданов В. В., Аксенова А. Е., и др. *Методы вычислительной гидродинамики для анализа безопасности объектов ТЭК* // Труды ИБРАЭ РАН / Под. общей ред. чл.-корр. РАН Л.А. Большова. 2008. Вып. 3. 210 с.: ил. М: Наука.
- [4] Чуданов В.В., *CFD код CONV-3D для расчета теплогидродинамики в элементах ЯЭУ с использованием супер-ЭВМ* // Межд. Науч.-тех. конференция «Инновационные проекты и технологии ядерной энергетики», 27–29 ноября 2012 года, г. Москва. ОАО НИКИЭТ. Секция 5.
- [5] Vassilevskii P. S. *Fast algorithm for solving a linear algebraic problem with separable variables* // С. r. Acad. Bulg. Sci. 1984. № 37, pp. 305–308.
- [6] Kuznetsov Yu. A. and Rossi T. *Fast direct method for solving algebraic systems with separable symmetric band matrices.* // East-West J. Numer. Math. 1996. № 4, pp. 53–68.
- [7] Горейнов С.А. *Быстрый прямой метод для сепарабельных ленточных матриц на трехмерной сетке процессоров* // Вычислительные методы и программирование. 2014. 12 с. В печати.
- [8] *Report of the OECD/NEA KAERI Rod Bundle CFD Benchmark Exercise.* // KAERI. 2013. № NEA/CSNI/R(2013)5, pp.70–75.

- [9] Brown P. N., Falgout R. D., Jones J. E. *Semicoarsening multigrid on distributed-memory machines.* // SIAM J. Sci. Comput. 2000. V. 21. pp. 1823–1834. + <http://www.hypre.org/>
- [10] Blatt M., Bastian P. *On the generic parallelization of iterative solvers for the finite element method* // Int. J. Comp. Sci. and Engrg. 2008. V. 4. № 1. pp. 56–69. + <http://www.dune.org/>

Рекомендовал к публикации

Программный комитет

Второго национального суперкомпьютерного форума НСКФ-2013

*Об авторах:*



**Владимир Васильевич Чуданов**

Кандидат физико-математических наук, старший научный сотрудник, заведующий лабораторией Вычислительной тепло и гидродинамики института проблем безопасного развития атомной энергетики (ИБРАЭ РАН)

*e-mail:*

[chud@ibrae.ac.ru](mailto:chud@ibrae.ac.ru)



**Сергей Анатольевич Горейнов**

Кандидат физико-математических наук, научный сотрудник ИБРАЭ РАН, старший научный сотрудник ИВМ РАН

*e-mail:*

[sergei@bach.inm.ra.ru](mailto:sergei@bach.inm.ra.ru)



**Анна Евгеньевна Аксенова**

Кандидат физико-математических наук, старший научный сотрудник лаборатории Вычислительной тепло и гидродинамики института проблем безопасного развития атомной энергетики (ИБРАЭ РАН)

*e-mail:*

[aks@ibrae.ac.ru](mailto:aks@ibrae.ac.ru)

**Валерий Алексеевич Первичко**

Старший научный сотрудник лаборатории Вычислительной тепло и гидродинамики института проблем безопасного развития атомной энергетики (ИБРАЭ РАН)

*e-mail:* [valper@ibrae.ac.ru](mailto:valper@ibrae.ac.ru)

**Артём Андреевич Макаревич**

Младший научный сотрудник лаборатории Вычислительной тепло и гидродинамики института проблем безопасного развития атомной энергетики (ИБРАЭ РАН)

*e-mail:* [amak@ibrae.ac.ru](mailto:amak@ibrae.ac.ru)

*Образец ссылки на публикацию:*

В. В. Чуданов, С. А. Горейнов, А. Е. Аксенова, В. А. Первичко, А. А. Макаревич. *Новый метод решения CFD задач на кластерных ЭВМ петафлопсной производительности* // Программные системы: теория и приложения: электрон. научн. журн. 2013. Т. 4, № 4(18), с. 43–54.

URL: [http://psta.psirras.ru/read/sta2013\\_5\\_43-54.pdf](http://psta.psirras.ru/read/sta2013_5_43-54.pdf)

V. V. Chudanov, S. A. Goreinov, A. E. Aksenova, V. A. Pervichko, A. A. Makarevich. *New method for solving of CFD problems at clustered computers petascale performance.*

ABSTRACT. A new method for solving of CFD problems on the clustered petascale computers is described. The method is highly effective and can achieve significant performance under direct realization. A scalability this method close to ideal was demonstrated. (*In Russian.*)

*Key Words and Phrases:* CFD, clustered computers, petascale performance.