

# ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ С АЛГОРИТМАМИ ПАРАЛЛЕЛЬНОЙ ДЕКОМПОЗИЦИИ БОЛЬШИХ СЕТОК

Е.Н. Головченко, Е.Ю. Дорофеева, В.А. Гасилов, М.В. Якобовский

Задача декомпозиции сетки возникает во многих приложениях вычислительных гидрогазодинамики и механики. Развитие высокопроизводительных технологий сделало возможным решение больших задач, на расчет которых на одном процессоре не хватает памяти, или вычислительных ресурсов. В рассматриваемых параллельных приложениях зачастую используется метод геометрического параллелизма, при котором сетка, аппроксимирующая расчетную область, распределяется между процессорами по геометрическому признаку. В дальнейшем каждый процессор обрабатывает свою часть сетки. Эффективность работы процессоров определяется тем, насколько равномерно распределена сетка по процессорам и насколько минимизированы затраты на передачу данных между процессорами. Объем передач данных между процессорами зависит от числа связей между доменами (частями сеток), распределенными по процессорам. Декомпозиция регулярных сеток намного проще декомпозиции нерегулярных сеток, однако, нерегулярные сетки, в частности треугольные и тетраэдральные, лучше аппроксимируют области сложной геометрической формы и легче сгущаются.

Задача сбалансированного разбиения сетки на домены сводится к более общей задаче разбиения графа на домены. В этом случае бьется граф, соответствующий коммуникационным нагрузкам сетки. Существует несколько моделей декомпозиции графов, отличающиеся видом графа и критериями сбалансированного разбиения: стандартная модель графа, модель двудольного графа, модель гиперграфа, модель декомпозиции с несколькими ограничениями, модель декомпозиции с несколькими целевыми функциями, модель асимметричного разбиения. В случае разбиения сеток хорошо себя зарекомендовал наиболее распространенный подход, использующий стандартную модель графа. В нем сетка аппроксимируется неориентированным графом  $G = (V, E)$ , где  $V$  – множество вершин,  $E$  – множество ребер. И вершины, и ребра, имеют вес. Оптимальным считается разбиение на домены, при котором выровнен суммарный вес вершин в доменах и минимизирован суммарный вес разрезанных ребер между доменами (разрезанное ребро – ребро, соединяющее вершины из разных доменов). В данной модели суммарный вес вершин в доменах отвечает за равномерность разбиения сетки на домены (а в дальнейшем равномерность распределения по процессорам, которые будут обрабатывать эти домены), а суммарный вес разрезанных ребер – за коммуникационную нагрузку между процессорами. Как известно, данная задача декомпозиции графа является NP-полной, поэтому для ее решения используются различные эвристические методы: алгоритмы рекурсивных координатной бисекции, инерциальной бисекции, спектральной бисекции, алгоритм Kernighan-Lin (KL) и Fiduccia-Mattheyses (FM), жадные алгоритмы (greedy methods), иерархические алгоритмы. Эти алгоритмы реализованы в следующих последовательных пакетах декомпозиции графов: METIS, JOSTLE, SCOTCH, CHACO и PARTY. К параллельным пакетам относятся PARMETIS (параллельная версия пакета METIS), JOSTLE, PT-SCOTCH (параллельная версия пакета SCOTCH) и ZOLTAN.

Областью данного исследования являются нерегулярные сетки, содержащие до  $10^9$  вершин. В настоящее время такие сетки невозможно разместить в памяти одного процессора, следовательно, разбить их можно только параллельными алгоритмами. Методы разбиения графов параллельных пакетов PARMETIS, JOSTLE, PT-SCOTCH и ZOLTAN основываются на иерархических алгоритмах [1, 2, 3], состоящих из следующих частей: поэтапное огрубление графа, декомпозиция самого маленького из полученных графов и отображение разбиения на предыдущие графы с периодическим локальным уточнением границ доменов. Недостатком таких алгоритмов является образование доменов, границы которых состоят из неоптимальных наборов сегментов. В частности, домены могут оказаться несвязными. Такое ухудшение качества доменов для некоторых задач является критичным. Например, на доменах с длинными границами, или сложной конфигурацией, алгоритмы решения систем линейных уравнений сходятся за большее число итераций [4].

Еще одним недостатком существующих пакетов является получение сильно несбалансированных разбиений. В частности, в разбиениях, получаемых пакетом PARMETIS, числа вершин в доменах могут отличаться в два раза.

В рамках данной работы был создан комплекс программ GRIDSPIDERPAR декомпозиции больших сеток (до  $10^9$  вершин). В него вошли два алгоритма: параллельный алгоритм геометрической декомпозиции сеточных данных и параллельный инкрементный алгоритм декомпозиции графов.

Параллельный алгоритм геометрической декомпозиции комплекса программ GRIDSPIDERPAR основан на методе рекурсивной координатной бисекции. На каждом этапе рекурсивной координатной бисекции параллелепипед, заключающий в себе сетку, разбивается на две части. Выбирается координатная ось, вдоль которой параллелепипед имеет наибольшую протяженность. Параллелепипед разрезается перпендикулярно полученной оси. Достоинством данного метода является то, что в разбиениях на равные домены числа вершин в доменах отличаются не больше, чем на единицу. Подобный алгоритм реализован в пакете ZOLTAN [3].

Отличие рекурсивной координатной бисекции созданного алгоритма от аналогичного алгоритма в пакете ZOLTAN состоит в том, что в нем секущая плоскость (медиана) разрезается также по остальным координатам, что позволяет обрабатывать ситуации наличия нескольких узлов с одним значением координаты. В пакете ZOLTAN вершины из медианы распределяются по областям произвольным образом, что увеличивает число разрезанных ребер.

Параллельный инкрементный алгоритм декомпозиции графов комплекса программ GRIDSPIDERPAR основан на последовательном инкрементном алгоритме декомпозиции графов, созданном Якобовским М.В. из ИПМ им. М.В.Келдыша РАН [5]. Достоинством инкрементного алгоритма является формирование связанных доменов. Инкрементный алгоритм не основывается на иерархическом подходе. Наиболее близкими к нему являются алгоритм пузырькового роста [6] и диффузионные алгоритмы. Однако алгоритм пузырькового роста, в отличие от инкрементного алгоритма, не гарантирует получение сбалансированного разбиения. А одним из отличий инкрементного алгоритма от диффузионных алгоритмов является то, что в инкрементном алгоритме происходит освобождение части вершин из плохих доменов и последующий рост доменов. Другим отличием инкрементного алгоритма от существующих алгоритмов является новый критерий проверки качества доменов, в соответствии с которым проверяется связность оболочек доменов (подробное описание см. в статье [5]). Параллельный инкрементный алгоритм комплекса программ GRIDSPIDERPAR является параллельной версией последовательного инкрементного алгоритма декомпозиции графов. Однако в нем есть отличия. В параллельный алгоритм декомпозиции графов добавлены выделение групп плохих доменов и отдельная работа с ними. Ещё одним отличием является то, что в параллельном инкрементном алгоритме рост доменов происходит не просто поиском в ширину, но с учетом минимизации суммарного веса разрезанных ребер. Несколько изменен критерий оценки качества доменов. Учитывается не только связность оболочек, но также количество плохих доменов и суммарный вес разрезанных ребер. Нововведения позволили ускорить нахождение алгоритмом оптимальных разбиений. Подробное описание алгоритмов пакета GRIDSPIDERPAR опубликовано в статье [7].

В данной работе был поставлен ряд численных экспериментов. Рассматривались следующие физические задачи: моделирование газоплазменных потоков в диверторе токамака, моделирование распространения ударной волны от взрыва химического взрывчатого вещества в протяженном сооружении с нетривиальной геометрией и моделирование распространения ударной волны от приземного источника энергии взрывного типа. На физических задачах проведено тестирование разбиений, полученных методами созданного комплекса программ GRIDSPIDERPAR, пакета PARMETIS, пакета ZOLTAN и пакета PT-SCOTCH. Сравнивалась эффективность параллельного счета физических задач пакетом MARPLE3D [8] при распределении сеток по ядрам в соответствии с различными разбиениями. Параллельный программный комплекс MARPLE3D создан в ИПМ им. М.В.Келдыша РАН, и его предметной областью являются задачи двухтемпературной радиационной магнитной гидродинамики.

Вычисления производились на кластерах MBC-100K (227,94 TFlop/s), "Ломоносов" (1700 TFlops) и «Helios» (1524.1 TFlop/s).

В расчетных областях были построены тетраэдральные и гексаэдральная сетки. Для всех сеток построены дуальные графы, учитывающие связи между ячейками через ребро. Число вершин в графах  $2.8 \cdot 10^6$  —  $6.1 \cdot 10^7$ , число ребер  $2.3 \cdot 10^7$  —  $5.4 \cdot 10^8$ . Графы разбивались на 256 и 4096 доменов.

Разбиения графов на домены были получены следующими методами: PartKway и PartGeomKway пакета PARMETIS, методами GeomDecomp (параллельный алгоритм геометрической декомпозиции) и IncrDecomp (параллельный инкрементный алгоритм) разработанного комплекса программ параллельной декомпозиции сеток GRIDSPIDERPAR, иерархическим диффузионным алгоритмом пакета PT-SCOTCH и методами RCB, RIB, HSFC и PHG пакета ZOLTAN. Методы PartKway и PartGeomKway пакета PARMETIS основаны на иерархическом алгоритме разбиения графов. В PartGeomKway выполняется предварительное геометрическое разбиение. Метод RCB пакета ZOLTAN основан на алгоритме рекурсивной координатной бисекции, метод RIB — на алгоритме рекурсивной инерциальной бисекции. Метод HSFC выполняет геометрическое разбиение с использованием кривой Гильберта. Метод PHG разбивает гиперграф.

При сравнении разбиений наибольший дисбаланс числа вершин в доменах обнаружен в разбиениях, полученных методами PartKway и PartGeomKway пакета PARMETIS (до 80 %), наименьший (с разницей в одну вершину) — в разбиениях геометрическими методами GeomDecomp, RCB, RIB и HSFC. Пакетом PT-SCOTCH получены разбиения с отклонением от среднего арифметического порядка 5 %, методом IncrDecomp программного комплекса GRIDSPIDERPAR — до 0.1 %. В разбиениях, полученных методом PHG пакета ZOLTAN, присутствуют домены с нулевым числом вершин, поэтому вычисление физических задач на данных разбиениях не проводилось.

Наименьшее число разрезанных ребер получено методами PartKway и PartGeomKway пакета PARMETIS (для двух из трех графов). В разбиениях, полученных пакетом PT-SCOTCH, число разрезанных ребер немного больше, чем у методов PartKway и PartGeomKway пакета PARMETIS. Далее идут разбиения, полученные методом IncrDecomp комплекса программ GRIDSPIDERPAR. Методы геометрической декомпозиции GeomDecomp, RCB, RIB и HSFC, не учитывающие связи между вершинами, дают разбиения с большим числом разрезанных ребер, чем методы разбиения графов. При этом методом GeomDecomp комплекса

программ GRIDSPIDERPAR получены разбиения с меньшим числом разрезанных ребер, чем аналогичным методом RCB пакета ZOLTAN. Наибольшее и сильно выделяющееся число разрезанных ребер получено методом HSFC пакета ZOLTAN.

В соответствии с разбиениями дуальных графов были получены разбиения вершин сеток. Каждая физическая задача пускалась на одинаковое время на всех разбиениях. Были получены числа итераций и физические времена, до которых досчитали задачи. Распределения разбиений по числам итераций и физическим временам оказались аналогичными, поэтому далее рассматриваются только числа итераций. На Рис. 1 представлены числа итераций для моделирования газоплазменных потоков в диверторе токамака.

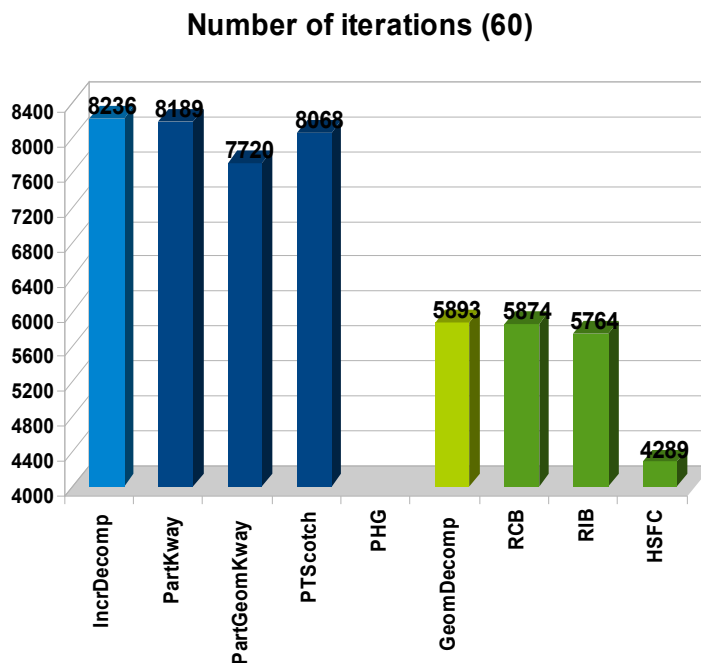


Рис. 1. Числа итераций для моделирования газоплазменных потоков в диверторе токамака.

По числу итераций метод IncrDecomp пакета GRIDSPIDERPAR опережает остальные методы декомпозиции графов, а метод GeomDecomp пакета GRIDSPIDERPAR опережает похожий метод RCB пакета ZOLTAN. На разбиениях, полученных методами декомпозиции графов, числа итераций больше, чем на разбиениях, полученных геометрическими методами, не учитывающими связи между вершинами. На остальных физических задачах картина сохраняется с некоторыми вариациями.

Из описанных результатов можно сделать выводы, что на разбиениях, полученных алгоритмами созданного комплекса программ GRIDSPIDERPAR, параллельный счет рассмотренных физических задач идет быстрее, чем на разбиениях, полученных другими пакетами.

Работа выполнена при поддержке РФФИ (гранты 11-01-12086-офи-м-2011, 11-07-00779-а, 11-07-93939-ДЖИ8-а, 13-01-12073-офи\_м) и Гранта Президента РФ НШ-4580.2012.1.

#### ЛИТЕРАТУРА:

1. G. Karypis, K. Schloegel. ParMETIS – Parallel Graph Partitioning and Sparse Matrix Ordering Library, Version 4.0 // University of Minnesota, Minneapolis, August 4, 2011.
2. G. Karypis, V. Kumar. Parallel Multilevel k-Way Partitioning Scheme for Irregular Graphs // SIAM REVIEW. 1999. Vol. 41. No 2. 278-300.
3. E. Boman, K. Devine, U. Catalyurek, D. Bozdog, B. Hendrickson, W.F. Mitchell, J. Teresco. Zoltan: Parallel Partitioning, Load Balancing and Data-Management Services. Developer's Guide, Version 3.3. Sandia National Laboratories, Copyright©2000-2010.
4. F. Pelegrini. A parallelizable multi-level banded diffusion scheme for computing balanced partitions with smooth boundaries // ENSEIRB, LaBRI and INRIA Futurs, Universite Bordeaux I.
5. М.В. Якововский. Параллельные алгоритмы сортировки больших объемов данных // Фундаментальные физико-математические проблемы и моделирование технико-технологических систем: Сб. науч. тр., Выпуск 7 / Под ред. Л.А. Уваровой. – М.: Издательство "Янус-К", 2004. – С. 235-249.
6. R. Preis, R. Diekmann. PARTY – A Software Library for Graph Partitioning. Advances in Computational Mechanics with Parallel and Distributed Processing, CIVIL-COMP PRESS, 1997, 63 – 71.
7. Е.Н. Головченко. Параллельный пакет декомпозиции больших сеток // Математическое моделирование. 2011. Т. 23. № 10. 3-18.

8. В.А. Гасилов и др. Пакет прикладных программ MARPLE3D для моделирования на высокопроизводительных ЭВМ импульсной магнитоускоренной плазмы // Матем. моделирование, 2012, Т.24, №1, С. 55–87.