

РАЗБИЕНИЕ БОЛЬШИХ СЕТОК

Е.Н. Головченко

АННОТАЦИЯ

При численном решении задач гидрогазодинамики и механики на распределенных вычислительных системах возникает задача сбалансированного распределения сетки по процессорам. Число процессоров, на котором будет считаться вычислительная задача, зачастую заранее неизвестно. Поэтому имеет смысл разбивать сетку на большое число микро-доменов, а потом формировать из них домены. Целью данной работы является получение сбалансированных разбиений треугольных и тетраэдральных сеток, содержащих до 10^9 вершин, на большое число связанных микро-доменов. Создан комплекс программ параллельной декомпозиции сеток. В комплекс вошло два алгоритма: параллельный алгоритм геометрической декомпозиции сеток и параллельный инкрементный алгоритм декомпозиции графов. Проведено сравнение разбиений на микро-домены четырех тетраэдральных сеток (порядка 10^8 вершин, 10^9 тетраэдров, 10^9 ребер), полученных методами созданного комплекса программ, пакета PARMETIS, пакета ZOLTAN и пакета PT-SCOTCH. Проведено сравнение различных разбиений графов микро-доменов на домены, а также разбиений сразу на домены.

ВВЕДЕНИЕ

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 передаваемых 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 [1], отличающиеся видом графа и критериями сбалансированного 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.

Целью данной работы является получение сбалансированных разбиений треугольных и тетраэдральных сеток, содержащих до 10^9 вершин, на большое число связанных микро-доменов.

Число процессоров, на котором будет считаться вычислительная задача, зачастую заранее неизвестно. Поэтому имеет смысл разбивать сетку на большое число микро-доменов, а потом формировать из них домены. Количество микро-доменов на несколько порядков меньше числа вершин, поэтому многократное разбиение микро-доменов на домены быстрее многократного разбиения всей сетки.

Еще одной областью использования разбиения сеток на микро-домены является хранение больших сеток. В микро-доменах функции достаточно гладкие, что позволяет отрезать биты в рамках одного микро-домена. В результате на хранение информации затрачивается меньше памяти. Экономит память также локальная нумерация вершин от нуля в рамках одного микро-домена, что позволяет на хранение одного номера вершины тратить меньше байтов. Существуют и другие способы сжатия информации о сетке в рамках одного микро-домена. При сжатом хранении сеток, разбитых на микро-домены, также важна сбалансированность разбиения.

Областью данного исследования являются нерегулярные (треугольные и тетраэдральные) сетки, содержащие до 10^9 вершин. В настоящее время такие сетки невозможно разместить в памяти одного процессора, следовательно, разбить их можно только параллельным алгоритмом. Методы разбиения графов параллельных пакетов PARMETIS, JOSTLE, PT-SCOTCH и ZOLTAN основываются на иерархических алгоритмах, 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 [2]. 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. Связность доменов также важна при визуализации распределенных сеточных данных, и на связных микро-доменах коэффициент сжатия информации о сетке будет больше.

Пакет JOSTLE разбивает на число доменов, равное числу процессоров, на которых производится разбиение. Поэтому для разбиения на микро-домены он не подходит. Кроме того пакет JOSTLE не выложен в свободном доступе, и получение его проблематично. Пакет PARMETIS является основным пакетом, который используется в данный момент для разбиения сеток. Еще одним недостатком пакета PARMETIS является получение сильно несбалансированных разбиений.

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

Проведено сравнение разбиений на микро-домены четырех тетраэдральных сеток, полученных методами созданного комплекса программ, методами пакета PARMETIS, пакета ZOLTAN, и пакетом PT-SCOTCH. Также проведено сравнение различных разбиений графов микро-доменов на домены и разбиений сразу на домены.

КОМПЛЕКС ПРОГРАММ ПАРАЛЛЕЛЬНОЙ ДЕКОМПОЗИЦИИ СЕТОК

Комплекс программ параллельной декомпозиции сеток включает в себя два алгоритма: параллельный алгоритм геометрической декомпозиции сеточных данных и параллельный инкрементный алгоритм декомпозиции графов. Оба алгоритма реализованы на распределенной памяти с использованием библиотеки MPI.

Параллельный алгоритм геометрической декомпозиции основан на методе рекурсивной координатной бисекции. Параллельные алгоритмы геометрического разбиения реализованы только в пакетах PARMETIS и ZOLTAN. В пакете PARMETIS геометрический алгоритм основывается на разбиении кривой Гильберта. В пакете ZOLTAN реализованы три геометрических алгоритма: алгоритм рекурсивной координатной бисекции, алгоритм рекурсивной инерциальной бисекции и алгоритм, разрезающий кривую Гильберта. Отличие рекурсивной координатной бисекции созданного алгоритма от аналогичного алгоритма в пакете ZOLTAN состоит в том, что разбиение производится одновременно по нескольким координатам. Вершины сортируются сначала по одной координате, потом внутри нее по следующей координате в циклическом порядке и т.д., что позволяет обрабатывать ситуации наличия нескольких узлов с одним значением координаты. В результате медиана проводится точно, и в разбиении на равные домены числа вершин в доменах отличаются не больше, чем на единицу. В пакете ZOLTAN вершины из медианы распределяются по областям произвольным образом, что хотя и сохраняет разницу числа вершин в доменах не превышающей единицу, но увеличивает число разрезанных ребер. Достоинствами параллельного алгоритма геометрической декомпозиции, как и других геометрических методов, являются экономичное использование памяти и относительная быстрота работы. Недостатками является то, что алгоритм не учитывает связи между вершинами, и ему требуется геометрическая информация о сетке. Целью создания данного алгоритма являлось получение сбалансированных разбиений на микро-домены. В частности, при разбиениях на равные части, разница числа вершин в доменах не превышает единицу.

Параллельный инкрементный алгоритм основан на последовательном инкрементном алгоритме декомпозиции графов, созданном Якововским М.В. из ИПМ им. М.В.Келдыша РАН [3]. Достоинством параллельного инкрементного алгоритма является формирование связных доменов. Сбалансированность разбиений, получаемых параллельным инкрементным алгоритмом несколько хуже, чем разбиений, получаемых алгоритмами геометрической декомпозиции, но заметно лучше, чем разбиений, получаемых пакетом PARMETIS, а также пакетом PT-SCOTCH. Наличие геометрической информации для алгоритма необязательно, поскольку предварительное распределение вершин по процессорам можно получить огрублением графа, вместо геометрической декомпозиции.

Подробное описание алгоритмов можно посмотреть в статье [4].

РЕЗУЛЬТАТЫ РАЗБИЕНИЙ НА МИКРО-ДОМЕНЫ

Вычисления производились на кластере MBC-100K (1460 модулей по два четырехядерных процессора Intel Xeon, 140.16 TFlop/s).

Четыре сгущающиеся тетраэдральные сетки, содержащие $10^8 \div 2.7 * 10^8$ вершин, $7 * 10^8 \div 1.6 * 10^9$ тетраэдров, $8 * 10^8 \div 1.9 * 10^9$ ребер, были разбиты на 25600 микро-доменов методами PartKway (на графиках обозначен PK) и PartGeomKway (на графиках PGK) пакета PARMETIS, методами GeomDecomp (параллельный алгоритм геометрической декомпозиции, на графиках G) и IncrDecomp (параллельный инкрементный алгоритм, на графиках I) разработанного комплекса программ параллельной декомпозиции сеток, иерархическим диффузионным алгоритмом пакета PT-SCOTCH и методами RCB, RIB и HSFC 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, HSFC выполняет геометрическое разбиение с использованием кривой Гильберта. Метод PartGeom (аналог HSFC) пакета PARMETIS не подходит для разбиения на микро-домены, так как в нем предполагается равенство числа формируемых доменов числу процессоров, на которых производится разбиение. Метод PHG (разбиение гиперграфа) пакета ZOLTAN не может быть такие большие сетки, а небольшую сетку, содержащую $2 * 10^5$ вершин, разбил на 200 доменов с образованием двух доменов с нулевым числом вершин. Метод REFTREE пакета ZOLTAN разбивает кривую Гильберта с использованием дерева, отображающего процесс адаптивного уточнения, которым была построена сетка. Для рассматриваемых сеток этот метод неприменим.

На гистограмме 1 представлено процентное отношение максимального модуля отклонения от среднего арифметического числа вершин в микро-домене в разбиениях четырех тетраэдральных сеток различными методами на 25600 микро-доменов. Как видно из гистограммы 1, наибольшее отклонение у методов PartKway и PartGeomKway пакета PARMETIS (50 - 60 %), у PT-SCOTCH отклонение порядка 10%, у всех остальных методов - в среднем меньше 1 %. Наименьшее отклонение с разницей в одну вершину у методов геометрической декомпозиции GeomDecomp, RCB, RIB и HSFC.

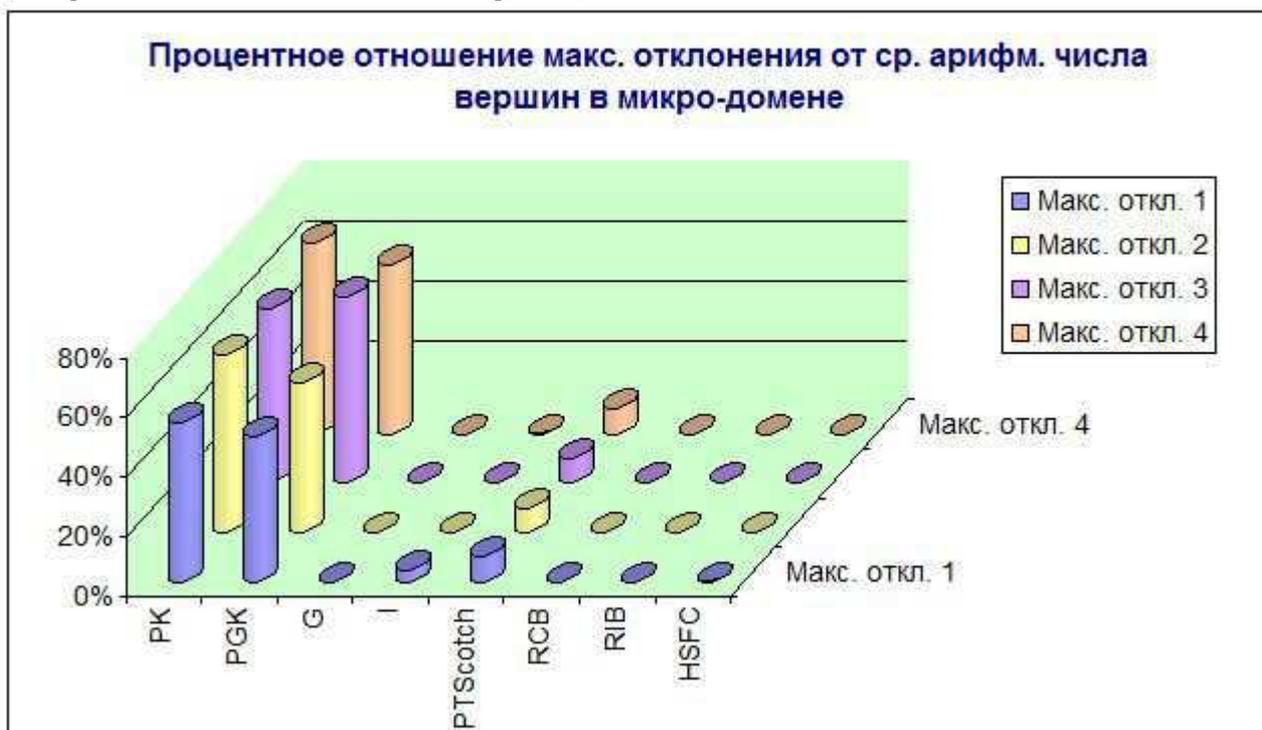


Рис. 1. Процентное отношение макс. модуля отклонения от ср. арифметического числа вершин в микро-домене в разбиениях тетраэдральных сеток на 25600 микро-доменов

На гистограмме 2 представлено число несвязных микро-доменов из 25600 в разбиениях четырех тетраэдральных сеток различными методами. На гистограмме 3 представлен фрагмент данной гистограммы, не учитывающий разбиения методами RIB и HSFC пакета ZOLTAN. Для наглядности гистограммы 3 порядок сеток изменен, при этом цвета, относящиеся к сеткам, сохранены. Как видно из гистограмм 2 и 3, наибольшее число несвязных микро-доменов в разбиениях, полученных методом HSFC пакета ZOLTAN (4 %). Методом RIB формируется 1 - 2 % несвязных микро-доменов, остальными методами - меньше 1 % (порядка 30 - 40 микро-доменов из 25600). Наименьшее число несвязных микро-доменов в разбиениях, полученных методом IncrDecomp (не больше 1 несвязного микро-домена) и PT-SCOTCH (7 несвязных микро-доменов).

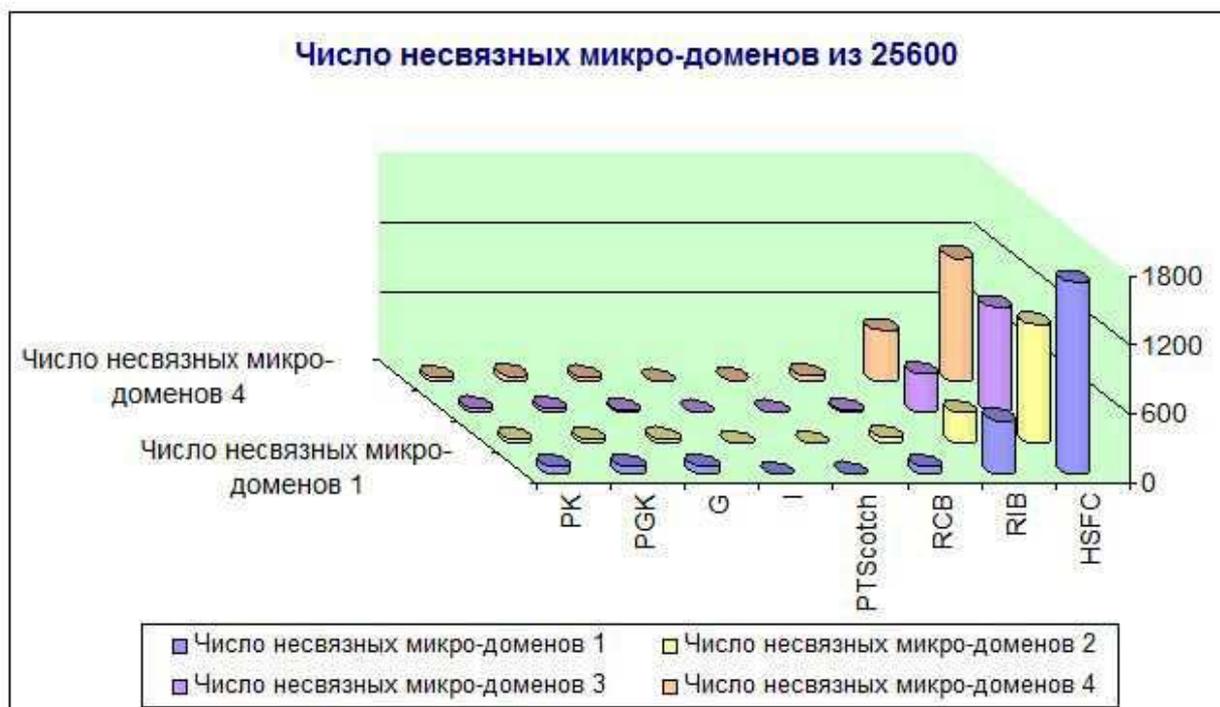


Рис. 2. Число несвязных микро-доменов в разбиениях тетраэдральных сеток на 25600 микро-доменов.



Рис. 3. Число несвязных микро-доменов в разбиениях тетраэдральных сеток на 25600 микро-доменов (фрагмент).

На гистограмме 4 представлено число разрезанных ребер в разбиениях четырех тетраэдральных сеток различными методами. Для наглядности гистограммы порядок сеток изменен, при этом цвета, относящиеся к сеткам, сохранены. Как видно из гистограммы 4, наибольшее число разрезанных ребер у метода HSFC пакета ZOLTAN, наименьшее – у методов PartKway и PartGeomKway пакета PARMETIS. Методы геометрической декомпозиции GeomDecomp, RCB и RIB, не учитывающие связи между вершинами, дают разбиения с большим числом разрезанных ребер, чем методы разбиения графов. В разбиениях, полученных IncrDecomp и PT-SCOTCH, число разрезанных ребер немного больше, чем у методов PartKway и PartGeomKway пакета

PARMETIS, PT-SCOTCH несколько выигрывает у IncrDecomp. Такое число разрезанных ребер в разбиениях, получаемых методом IncrDecomp, объясняется достаточно большой степенью локальности формирования микро-доменов на процессорах.

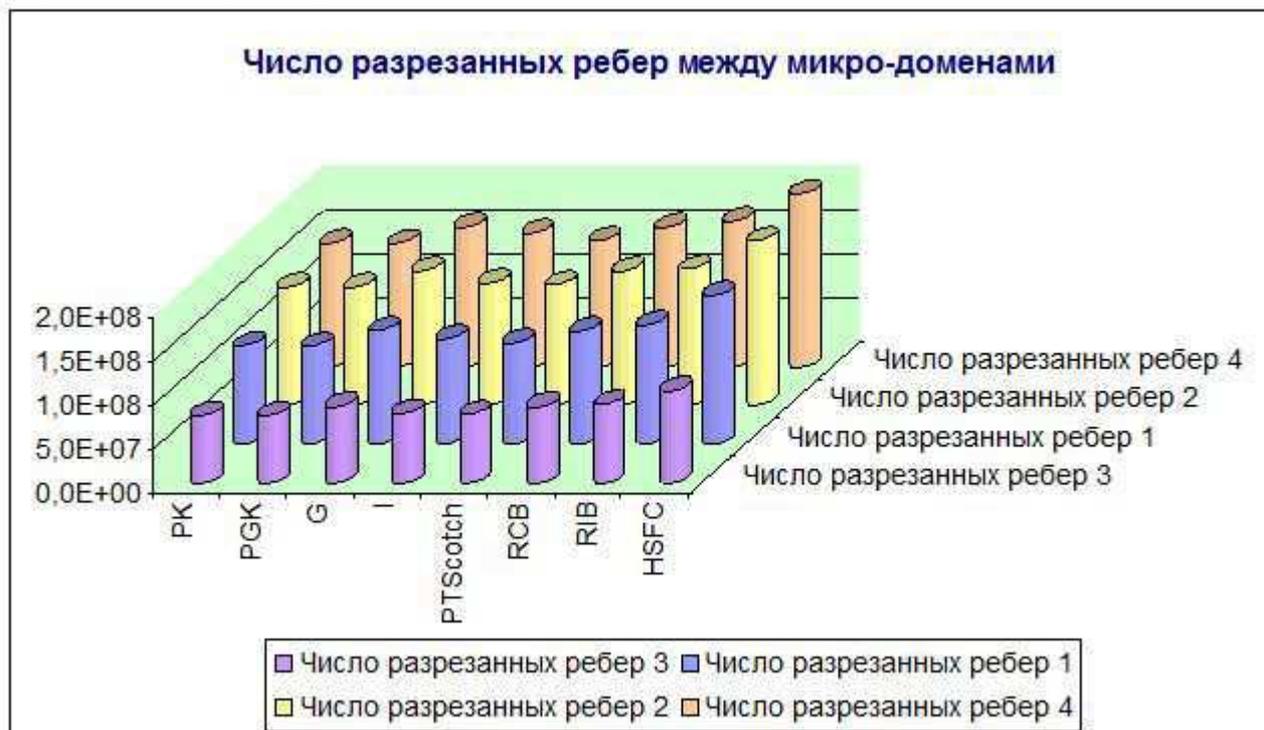


Рис. 4. Число разрезанных ребер в разбиениях тетраэдральных сеток на 25600 микро-доменов.

По сумме критериев наилучшими оказались методы GeomDecomp, RCB, IncrDecomp и диффузионный алгоритм пакета PT-SCOTCH. Методы геометрического разбиения GeomDecomp и RCB за наименьшее время на небольшом числе процессоров получают разбиения с разницей числа вершин в микро-доменах, не превышающей единицу. Метод GeomDecomp несколько выигрывает у RCB по числу разрезанных ребер и числу несвязных микро-доменов, но проигрывает по времени. Метод IncrDecomp и PT-SCOTCH получают разбиения с меньшими числами несвязных микро-доменов и разрезанных ребер. При сравнении двух методов, в разбиениях IncrDecomp меньший дисбаланс числа вершин в микро-доменах и меньшее число несвязных микро-доменов, чем в разбиениях PT-SCOTCH, но несколько большее число разрезанных ребер, и время получения разбиения.

РЕЗУЛЬТАТЫ РАЗБИЕНИЙ НА ДОМЕНЫ

По разбиениям всех сеток на микро-домены методами пакета PARAMETIS и созданного комплекса программ были составлены графы связей между микро-доменами с весами вершин, соответствующими количеству вершин в микро-доменах. Графы связей были разбиты на 512 доменов на одном процессоре методами PartGraphRecursive (PGR) и PartGraphKway (PGrK) пакета METIS, методом PartKway (PK) пакета PARAMETIS и методом IncrDecomp (I) созданного комплекса программ, запущенными на одном процессоре. Проведено сравнение различных вариантов разбиений микро-доменов на домены между собой и с разбиениями сразу на домены методами PartKway (PK), PartGeomKway (PGK) и PartGeom (PG) пакета PARAMETIS, GeomDecomp (G) созданного комплекса программ, диффузионным алгоритмом пакета PT-SCOTCH и методами RCB, RIB и HSFC пакета ZOLTAN. Метод IncrDecomp не позволяет разбивать на такое маленькое число доменов, так как в нем предполагается, что при выполнении разбиения на процессоре формируется порядка 10 - 50 микро-доменов.

На гистограмме 5 представлено процентное отношение максимального модуля отклонения от среднего арифметического числа вершин в домене для различных методов на четырех сетках. Слева направо сначала отображены результаты разбиений сразу на 512 доменов, а затем для каждого разбиения на микро-домены представлены различные разбиения графов микро-доменов. Разбиения графов микро-доменов обозначены, как 'метод разбиения на микро-домены + метод разбиения графа микро-доменов'. Как видно из гистограммы 5, разбиения методами PartKway и PartGeomKway пакета PARAMETIS сразу на домены заметно проигрывают всем остальным разбиениям по дисбалансу числа вершин в доменах. Дисбаланс числа вершин в доменах, сформированных из микро-доменов, не зависит от дисбаланса числа вершин в микро-доменах. Наименьший дисбаланс числа вершин в доменах при разбиениях графов микро-доменов получен методами PartGraphRecursive и PartGraphKway пакета METIS. Среди всех разбиений наименьший дисбаланс числа вершин в доменах, равный единице, получен методами PartGeom пакета PARAMETIS, GeomDecomp созданного комплекса программ и методами RCB, RIB и HSFC пакета ZOLTAN.

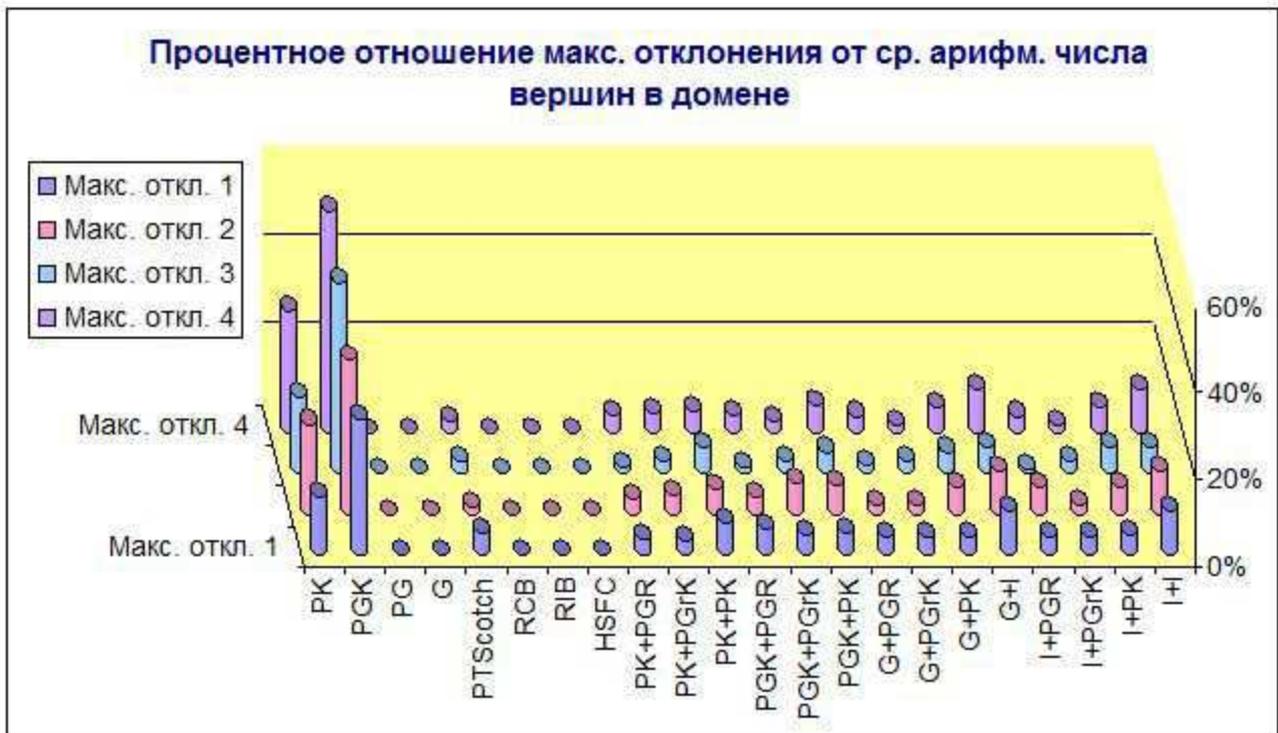


Рис. 5. Процентное отношение максимального модуля отклонения от среднего арифметического числа вершин в домене в разбиениях тетраэдральных сеток на 512 доменов.

На гистограмме 6 представлено число несвязных доменов в различных разбиениях тетраэдральных сеток на 512 доменов. Как видно из гистограммы, метод PartGeom пакета PARMETIS формирует практически все домены несвязными. На гистограмме 7 представлен фрагмент данной гистограммы. Для наглядности порядок сеток изменен, при этом цвета, относящиеся к сеткам, сохранены. Как видно из гистограмм 6 и 7, наибольшее число несвязных доменов (не считая, метод PartGeom пакета PARMETIS) сформировано методами RIB и HSFC пакета ZOLTAN, также при разбиениях графа микро-доменов, полученного методом геометрической декомпозиции GeomDecomp. Остальные разбиения содержат примерно одинаковое число несвязных доменов. Среди разбиений сразу на домены наименьшее число несвязных доменов в разбиениях, полученных пакетом PT-SCOTCH.



Рис. 6. Число несвязных доменов в разбиениях тетраэдральных сеток на 512 доменов.



Рис. 7. Число несвязных доменов в разбиениях тетраэдральных сеток на 512 доменов (фрагмент).

На гистограмме 8 отображены числа разрезанных ребер в разбиениях тетраэдральных сеток на 512 доменов. Как видно из гистограммы, методы PartGeom пакета PARMETIS и HSFC пакета ZOLTAN также сильно проигрывают по числу разрезанных ребер, что неудивительно при таком большом числе несвязных доменов. Остальные разбиения сразу на домены имеют меньшее число разрезанных ребер, чем разбиения графов микро-доменов. Среди разбиений графов микро-доменов наименьшее число разрезанных ребер получено методами PartGraphKway пакета METIS и PartKway пакета PARMETIS. Среди всех разбиений наименьшее число разрезанных ребер в разбиениях, полученных методами PartKway и PartGeomKway пакета PARMETIS и пакетом PT-SCOTCH.



Рис. 8. Число разрезанных ребер в разбиениях тетраэдральных сеток на 512 доменов.

Основные результаты по разбиениям на домены следующие. Вне зависимости от того, насколько различным был дисбаланс числа вершин в микро-доменах, дисбаланс числа вершин в образуемых из микро-доменов доменах оказался на одном уровне. Видимо, это связано с недостаточной чувствительностью алгоритмов разбиения графов к весам вершин. Наилучшие разбиения графов микро-доменов получены методом PartGraphKway пакета METIS. Среди всех разбиений наиболее качественными оказались разбиения, полученные методами GeomDecomp созданного комплекса программ, RCB пакета ZOLTAN и пакетом PT-SCOTCH. Методы GeomDecomp и RCB формируют домены с меньшим дисбалансом числа вершин в доменах. В разбиениях, полученных пакетом PT-SCOTCH, дисбаланс числа вершин в доменах сравним с разбиениями графов микро-доменов. Однако пакетом PT-SCOTCH получают разбиения с меньшим числом разрезанных ребер и несколько меньшим числом несвязных доменов.

ВЫВОДЫ

Создан комплекс программ параллельной декомпозиции больших сеток (до 10^9 вершин) на большое число микро-доменов. В комплекс вошло два алгоритма: параллельный алгоритм геометрической декомпозиции сеток и параллельный инкрементный алгоритм декомпозиции графов.

Проведено сравнение разбиений на микро-домены четырех тетраэдральных сеток ($10^8 \div 2.7 * 10^8$ вершин, $7 * 10^8 \div 1.6 * 10^9$ тетраэдров, $8 * 10^8 \div 1.9 * 10^9$ ребер), полученных методами созданного комплекса программ, пакета PARMETIS, пакета ZOLTAN и пакета PT-SCOTCH. Наиболее качественные разбиения получены методами GeomDecomp (параллельный алгоритм геометрической декомпозиции) и IncrDecomp (параллельный инкрементный алгоритм) созданного комплекса программ, RCB пакета ZOLTAN и пакетом PT-SCOTCH. Качество разбиений проверялось по дисбалансу числа вершин в микро-доменах, числу несвязных микро-доменов и числу разрезанных ребер.

Проведено сравнение различных разбиений графов микро-доменов на домены, а также разбиений сразу на домены. Результаты показали, что дисбаланс числа вершин в доменах, сформированных из микро-доменов, не зависит от дисбаланса числа вершин в микро-доменах. Лучшие разбиения графов микро-доменов на домены получены пакетом METIS. Наиболее качественные разбиения на домены получены методами GeomDecomp (параллельный инкрементный алгоритм) созданного комплекса программ, RCB пакета ZOLTAN и пакетом PT-SCOTCH.

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

ЛИТЕРАТУРА:

1. Bruce Hendrickson, Tamara G. Kolda. Graph partitioning models for parallel computing. *Parallel Computing*, 26, 2000, 1519-1534.
2. Francois Pelegrini. A parallelizable multi-level banded diffusion scheme for computing balanced partitions with smooth boundaries // ENSEIRB, LaBRI and INRIA Futurs, Universite Bordeaux I.

3. М. В. Якобовский. Инкрементный алгоритм декомпозиции графов. Вестник Нижегородского университета им. Н.И.Лобачевского. Серия «Математическое моделирование и оптимальное управление», Вып. 1(28). Нижний Новгород: Издательство ННГУ, 2005, с. 243-250.
4. Е.Н. Головченко. Параллельный пакет декомпозиции больших сеток // Математическое моделирование. 2011. Т. 23. № 10. 3-18.