

## Методы повышения эффективности широкомасштабных распределенных вычислительных экспериментов на неструктурированных сетках\*

С.А. Суков

Институт прикладной математики им. М.В. Келдыша РАН

В работе представлено описание методов повышения эффективности вычислительных экспериментов на неструктурированных сетках большого размера. Приводятся результаты тестирования предложенных алгоритмов на примере расчетов с использованием тетраэдральных сеток, содержащих до 1.5 миллиарда элементов.

Вычислительная мощность современных суперкомпьютеров позволяет решать актуальные задачи математической физики, связанные с обработкой сверхбольших объемов данных на неструктурированных сетках. Разработанные на основе метода геометрического параллелизма распределенные алгоритмы по определению характеризуются высокой масштабируемостью. Число параллельных процессов приложения здесь, теоретически, ограничивается только числом сеточных элементов. В свою очередь число сеточных элементов среднестатистической расчетной сетки, как минимум, на несколько порядков превосходит число процессорных ядер лидирующих по показателю быстродействия многопроцессорных систем. А, следовательно, есть основания утверждать, что на любом из существующих суперкомпьютеров можно провести расчет, например, задачи газодинамического обтекания на неструктурированной сетке с таким числом элементов, что лимит параллелизма не будет исчерпан.

Фактическая производительность распределенных вычислений на неструктурированных сетках определяется эффективностью методов решения следующих прикладных задач:

- балансировка загрузки процессов;
- стабилизация производительности вычислений на неструктурированных сетках;
- распределенная запись данных контрольных точек;
- распределенное хранение и сжатие больших объемов сеточных данных.

Первые два пункта в списке влияют на быстродействие вычислительного ядра. Последующие два определяют эффективность решения не менее важных с точки зрения затрат времени проблем распределенного хранения и обработки данных широкомасштабных вычислительных экспериментов.

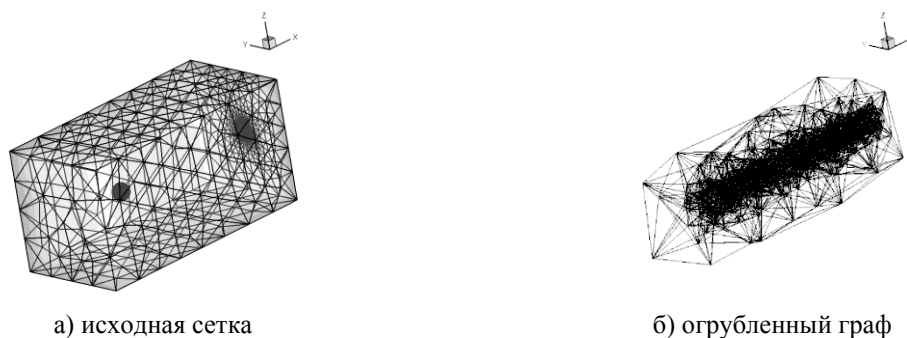


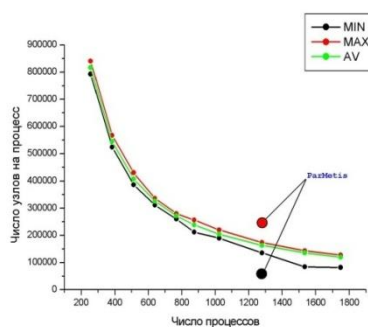
Рис 1. Двухуровневое представление сетки

Одним из методов эффективной обработки неструктурированных сеток, содержащих до миллиарда элементов, является метод двухуровневого распределенного представления, хранения и обработки топологии неструктурированных сеток [1]. Идея метода заключается в предварительном преобразовании топологии сетки путем ее декомпозиции на множество подобластей (доменов) с соответствующим разбиению изменением индексации сеточных узлов и элементов.

\* Работа выполнена при поддержке Российского фонда фундаментальных исследований, проект № 15-07-04213-а.

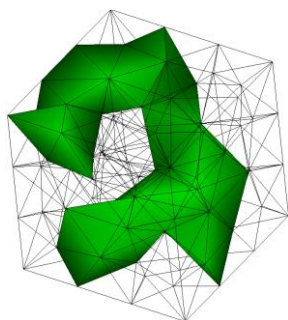
Далее топология преобразованной сетки задается ее огрубленным графом - графом доменов и их связей (**рис. 1**). Огрубленный граф используется на этапе балансировки загрузки. Число вершин графа примерно на два порядка больше максимального числа параллельных процессов приложения, но существенно меньше числа сеточных элементов. Поэтому декомпозиция графа на части, обрабатываемые параллельными процессами приложения, может быть выполнена в последовательном режиме.

Результаты балансировки загрузки для преобразованной к двухуровневому представлению тетраэдральной сетки (209 028 730 узлов и 1 244 316 672 тетраэдров) показаны на **рис. 2**. Значения расчетных параметров определяются в узлах сетки. На графиках отображены зависимости минимального и максимального числа обрабатываемых процессами сеточных узлов на фоне равномерного распределения. Дисбаланс (отношение максимального размера подобласти к осредненной величине) не превосходит 15%. В то же время прямое разбиение сеточного графа с использованием процедур библиотеки ParMETIS [2] при запуске 1280 параллельных процессов дает дисбаланс распределения узлов порядка 52%. Таким образом, применение двухуровневой схемы не только сокращает затраты времени на балансировку загрузки, но и качественно улучшает результат решения задачи при использовании идентичных подходов к декомпозиции.

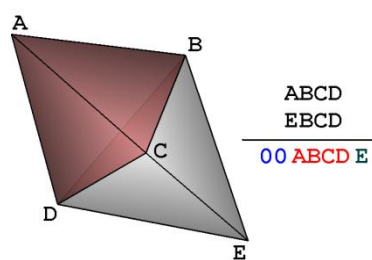


**Рис. 2.** Результаты балансировки загрузки

Геометрия неструктурированных сеток задается координатами сеточных узлов и индексами вершин сеточных элементов. Распределенное хранение этих данных может быть организовано в рамках той же двухуровневой схемы. Геометрия сетки записывается на диск как множество файлов, каждый из которых содержит описание одного или нескольких доменов. Аналогичный принцип структурирования данных можно использовать и при записи расчетных параметров.



а) цепочка по сеточным элементам



б) преобразование двух элементов последовательно  
сти

**Рис. 3.** Циклический метод сжатия

Для сокращения объема записываемых на диск данных целесообразно использовать комбинацию стандартного [2] и специализированных методов сжатия. Специализированный метод сжатия ориентирован на обработку сеток конкретного типа. Например, для сжатия описания топологии элементов тетраэдральных сеток (четверок вершин тетраэдров) были предложены блочный и циклический методы сжатия. Идея циклического метода сжатия состоит в построении цепочек по вершинам дуального графа (графа тетраэдров и их связей по общим граням) сетки. Стартовый тетраэдр задается четверкой индексов своих вершин, а каждый последующий

элемент - номером общей с предыдущим элементом последовательности грани (два бита) и индексом четвертой вершины (рис. 3).

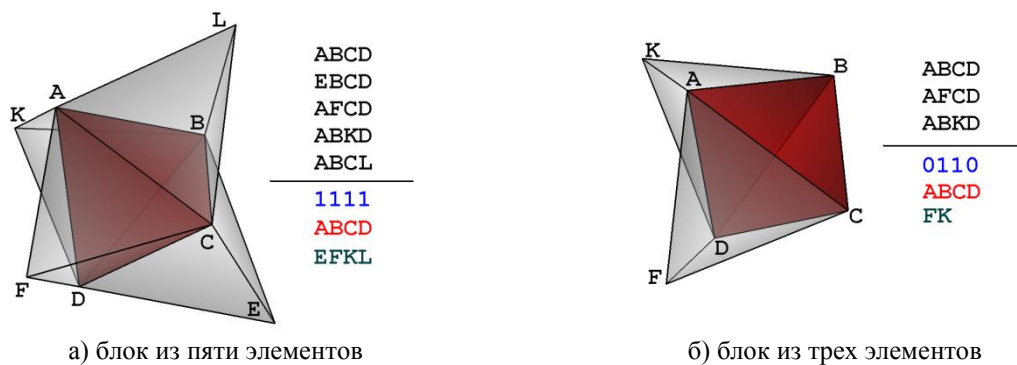


Рис. 4. Блочный метод сжатия

При использовании блочного метода сжатия сеточные элементы разбиваются на группы размером от одного до пяти тетраэдров. Каждая группа описывается четверкой индексов вершин центрального элемента, признаками присутствия связей по каждой из граней (четыре бита) и индексами четвертых вершин смежных центральному по соответствующей грани тетраэдров (рис. 4). Сформированные в результате сжатия циклическим или блочным методом структуры данных далее повторно упаковываются с использованием процедур библиотеки GZIP.

Результаты упаковки геометрии трех тетраэдральных сеток циклическим (С) методом, блочным (В) методом, методом GZIP и их комбинацией представлены в таблице 1. Эффективность упаковки данных специализированными методами оказывается выше по сравнению с алгоритмом GZIP, а применение комбинации методов сжатия улучшает коэффициент упаковки более чем в два раза.

Таблица 1. Коэффициент упаковки данных

Сетка (число узлов/число тетраэдров)	Коэффициент сжатия данных		
	В / В + GZIP	С / С + GZIP	GZIP
7320 / 37964	2.65 / 5.99	3.09 / 6.65	2.09
58926 / 363124	2.69 / 5.04	3.19 / 6.07	1.74
455160 / 2647040	2.72 / 4.48	3.28 / 5.68	2.59

Алгоритмы моделирования задач газовой динамики на неструктурированных сетках характеризуются низким соотношением числа арифметических операций на единицу данных. Поэтому быстродействие ядра сильно зависит от метода индексации сеточных элементов, определяющего порядок доступа к данным в оперативной памяти. Исходная индексация сеточных элементов соответствует методу генерации сетки и не всегда оптимальна с точки зрения особенностей программной реализации вычислительного алгоритма. Поэтому производительность вычислений на неструктурированных сетках без оптимизации индексации может быть низкой или нестабильной с ростом числа параллельных процессов.

Для оптимизации индексации сеточных элементов неструктурированных сеток предлагается использовать метод, основанный на идее присвоения индексов в порядке обхода вершин дуального графа. Первый элемент последовательности обхода выбирается произвольным образом. Далее в конец последовательности в произвольном порядке добавляются индексы элементов, имеющих с первым элементом общую грань. Затем в последовательность добавляются элементы, имеющие общую грань со вторым элементом и ранее не вошедшие в последовательность, и так далее до момента обхода всех сеточных элементов. Сравнительные результаты быстродействия модуля по вычислению конвективных потоков на основе метода конечного объема с полиномиальной реконструкцией переменных в ячейках тетраэдральной сетки для различных вариантов индексации элементов приведены в таблице 2. Расчет задачи невязкого обтекания сферы проводился на локально сгущающейся сетке, содержащей 119286 узлов и 679339 тетраэдров. Позитивная индексация соответствует описанному выше методу. Негативная индексация формируется таким образом, чтобы исключить кэширование данных. Представленные

результаты демонстрируют существенное влияние метода индексации сеточных элементов на время работы модуля.

**Таблица 2.** Время работы модуля вычисления конвективных потоков

Число процессорных ядер	Индексация элементов	Время (секунд)	Относительная производительность
1	Позитивная	0.2608	1.73
6	Позитивная	0.0590	7.64
1	Негативная	0.4506	1
6	Негативная	0.1069	4.22

Обобщенное тестирование эффективности методов оптимизации вычислений на неструктурированных сетках большого размера проводилось на примере задачи невязкого обтекания тела сложной формы на тетраэдральной сетке, содержащей 260 517 739 узлов и 1 555 767 296 тетраэдров, с определением значений в узлах. Расчет проводился с использованием ресурсов МВС «Ломоносов». Замерялись время выполнения 500 вычислительных шагов и производительность вычислений при запуске от 1000 до 29600 параллельных процессов (таблица 3).

**Таблица 3.** Параметры вычислительного эксперимента

Число процессов	Время (секунд)	Ускорение	Эффективность	Производительность (TFlops)
1000	209.69	1	1	1.43
2000	110.03	1.91	0.95	2.72
4000	59.14	3.55	0.89	5.05
8000	32.36	6.48	0.81	9.23
12000	22.63	9.27	0.77	13.20
16000	17.70	11.85	0.74	16.88
24000	12.63	16.60	0.69	23.66
28000	11.37	18.44	0.66	26.29
29600	10.72	19.56	0.66	27.88

Представленные выше результаты свидетельствуют о высокой эффективности использования программных реализаций разработанных методов распределенной обработки на примере тетраэдральных сеток. В дальнейшем предполагается адаптация методов на случай обработки гибридных сеток с элементами типа шестигранник, тетраэдр, четырехугольная пирамида и треугольная призма.

## Литература

1. Суков С.А., Якобовский М.В., Обработка трехмерных неструктурированных сеток на многопроцессорных системах с распределенной памятью. В сб. "Фундаментальные физико-математические проблемы и моделирование технико-технологических систем", вып. 6, под ред. Л.А. Уваровой. М., Изд-во "Janus-K", 2003, с. 233-239.
2. G. Karypis and V. Kumar. Multilevel algorithms for multi-containt graph partitioning. Technical Report TR 98-019, Department of Computer Science, University of Minnesota, 1998.
3. Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео, Изд-во Диалог-МИФИ, 2002.