

ВЫЧИСЛИТЕЛЬНЫЕ И ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫЕ СИСТЕМЫ

УДК 004.272

СТРАТЕГИИ ДЕКОМПОЗИЦИИ В ЗАДАЧАХ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ АДДИТИВНЫХ ЛАЗЕРНЫХ ТЕХНОЛОГИЙ*

М. Д. Хоменко, А. В. Дубров, Ф. Х. Мирзаде

*Институт проблем лазерных и информационных технологий,
филиал ФНИЦ «Кристаллография и фотоника» РАН,
140700, г. Шатура, Московская обл., ул. Святоозерская, 1
E-mail: hmd@laser.ru*

Развитие аддитивных технологий и их применение в промышленности связано с возможностью прогнозирования финальных свойств закристаллизовавшегося добавленного материала. Исследуется задача, которая характеризуется динамической и пространственно-неоднородной вычислительной сложностью, что при равномерной декомпозиции расчётной области приводит к несбалансированной нагрузке на вычислительные ядра. Использована стратегия разбиения расчётной области, минимизирующая потери процессорного времени при проведении серийных расчётов аддитивного технологического процесса. Выбранная стратегия является оптимальной с точки зрения распределения динамической вычислительной нагрузки неизвестной априори. Определена масштабируемость вычислительной задачи на кластере Института проблем лазерных и информационных технологий РАН, использующем интерконнект InfiniBand. В результате применения параллельного кода с оптимальной декомпозицией удалось значительно сократить время расчёта (до нескольких часов), что важно в контексте реализации проекта разработки пакета программного обеспечения поддержки инженерной деятельности в области аддитивных технологий.

Ключевые слова: декомпозиция области, численное моделирование, аддитивные технологии, OpenFOAM.

DOI: 10.15372/AUT20160614

Введение. Решение современных задач моделирования гидродинамических течений невозможно представить без использования высокопроизводительных параллельных компьютеров. Существуют различные алгоритмы декомпозиции расчётной области для проведения параллельных вычислений. Параметры хорошей декомпозиции заранее не известны, их нахождение не ограничивается лишь одной минимизацией параллельных коммуникаций и зависит от конкретной задачи и вычислительного оборудования [1, 2]. Много работ в этом направлении посвящено динамическому распределению нагрузки, поэтому время декомпозиции является важным параметром [3].

Свободный пакет приложений вычислительной гидродинамики с открытым исходным кодом — OpenFOAM (Open Field Operation and Manipulation) — имеет огромное число пользователей практически в каждой области промышленного производства и науки [4]. Он включает в себя инструменты для моделирования, создания сеток, для пред- и постобработки. Многие решатели из стандартного набора программ по умолчанию могут работать в параллельном режиме, что даёт возможность использовать высокопроизводительные вычислительные системы и расширять круг применимости пакета.

*Работа выполнена при поддержке Президиума РАН (программа I.33П «Фундаментальные проблемы математического моделирования») и Российского фонда фундаментальных исследований (проекты № 16-38-00933 мол-а и № 16-29-11743 офи-м).

В данной работе исследуется производительность численной модели расчёта микроструктуры наплавленного слоя, которая включает решение нелинейного интегрального уравнения Колмогорова — Джонсона — Мейла — Авраами (КЖМА) и дифференциальных уравнений, описывающих тепломассоперенос, с помощью метода конечных объёмов (FVM). Такая задача характеризуется динамической и пространственно-неоднородной вычислительной сложностью, что при равномерной декомпозиции расчётной области приводит к неоднородной нагрузке на вычислительные ядра. Это накладывает дополнительные требования на использование встроенной векторизации расчёта в пакете OpenFOAM, а её реализация осложнена необходимостью проведения множества параметрических исследований с вариацией условий задачи.

Особенности физико-математической модели. Типичный лазерный аддитивный процесс схематически показан на рис. 1. В начальный момент времени на твёрдую поверхность мишени падает лазерный луч, движущийся от некоторой точки x_0 в положительном направлении оси x со скоростью V_{scan} . Энергия, непосредственно поглощённая поверхностью, и энергия, перенесённая на поверхность частицами, расходуется на нагревание поверхностного слоя подложки и плавление [5]. Образовавшаяся «ванна» расплава движется вслед за лазерным лучом. Гидродинамические течения локализованы в ванне расплава, в её задней части происходит процесс кристаллизации. При расчёте микроструктуры применяется кинетическая модель [6–8], а при расчёте гидродинамических процессов — алгоритм PIMPLE (комбинация методов PISO и SIMPLE) на основе естественных физических переменных (скорости, давления). Для описания движения жидкости в переходной области, когда изменяется фазовое состояние вещества, используются закон Дарси и уравнение Козени — Кармана [9]. Задача характеризуется различными пространственными масштабами вдоль и поперёк движения луча. Расчётная область обычно вытянута вдоль движения луча, а представляющие интерес физические процессы, требующие основных вычислительных ресурсов, происходят в некоторой локальной области, зависящей от положения луча.

Для расчёта были взяты две трёхмерные ортогональные сетки: малая с количеством ячеек равным $5 \cdot 10^5$ ($210 \times 45 \times 55$) и большая — 10^6 ($256 \times 64 \times 64$), созданные с помощью встроенной утилиты blockMesh. В полученной неравномерной сетке основная часть ячеек сконцентрирована в зоне, где предполагается действие лазерного луча. Недостатком blockMesh для нашего случая является отсутствие возможности объединения ячеек для разрежения краевых областей. Для решения систем линейных уравнений используется метод предобусловленных бисопряжённых градиентов. Выход на стационарный режим параметров ванны расплава происходит за время порядка 0,7 с от начала моделирования. Расчёт производительности проводится на интервале 0,04 с после выхода на стационарный режим. Время расчёта всего процесса составляет порядка 8 ч на 128 ядрах.

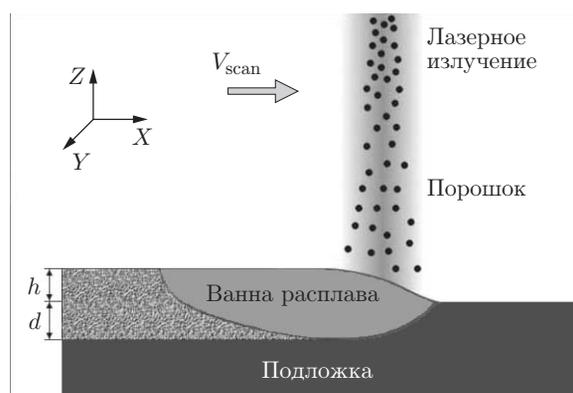


Рис. 1. Схема лазерной наплавки

Схемы декомпозиции. Сетка и поля с помощью утилиты `decomposePar` [10] разбиваются на подобласти в соответствии с числом используемых ядер. У пользователя OpenFOAM есть возможность выбрать «ручной» или один из трёх автоматических методов декомпозиции. Методы `Simple` и `Hierarchical` — простые геометрические разбиения, когда расчётная область разделяется на части по направлениям, например две части по координате x , одна часть по y ; в `Hierarchical` есть также возможность выбора очередности направления разбиения; `Scotch` — автоматическое разбиение, где минимизируются межпроцессорные границы.

При выборе стратегии разбиения следует учитывать как особенности конкретной вычислительной задачи, так и общий принцип, связанный с уменьшением межпроцессорного обмена.

Рассматриваемая задача трёхмерного моделирования лазерных аддитивных технологий характеризуется динамической локализацией вычислительной нагрузки, т. е. физические процессы, моделирование которых требует большего количества вычислений, меняют своё пространственное расположение в зависимости от времени. Это связано с перемещением области ванны расплава вдоль поверхности обрабатываемой детали. Для увеличения производительности необходимо максимизировать количество вычислительных ядер в течение всего расчёта, на которые распределяется нагрузка в таких сложных процессах.

Поскольку требуется определить универсальную стратегию декомпозиции, оптимальную для проведения серий вычислительных экспериментов, использующих комбинации значений констант и параметров технологического процесса, необходимо учитывать возможность изменения размеров ванны расплава в различных экспериментах. Выбираются такие размеры расчётной области и структура сетки, чтобы обеспечить расчёт области фазовых переходов и зоны термического влияния на сетке с максимальной плотностью вычислительных ячеек. Для этого следует ориентироваться на возможно максимальные размеры ванны расплава во всех сериях экспериментов. Это приводит к тому, что в части экспериментов, где наблюдаются меньшие размеры ванны расплава, по краям расчётной области не будут протекать процессы, требующие ресурсоёмких вычислений, и, значит, соответствующие таким областям вычислительные ядра будут работать в «холостом» режиме. Вариантом решения данной проблемы является ориентация подобластей декомпозиции поперёк направления движения ванны расплава.

Цель декомпозиции — максимально распределить сконцентрированную нагрузку по процессорам в каждый момент времени. Простым правилом декомпозиции является равномерное распределение числа ячеек между процессорами. Также для снижения количества переданных данных следует уменьшать «площадь» (количество граничных ячеек) межпроцессорных границ. В предлагаемой работе использовалось `Hierarchical`-разбиение и исследовались 16 стратегий декомпозиции (табл. 1) для двух сеток. Следует отметить, что такой метод требует минимального времени на декомпозицию, не зависящего от числа процессоров [3], поэтому он подходит для динамического распределения нагрузки.

Разбиения типа `Slice N-norm` получены рассечением расчётной области множеством плоскостей, перпендикулярных оси X , Y или Z . Количество сечений вдоль главной оси варьировалось в экспериментах. Вдоль остальных осей расчётная область всегда делилась на две части.

Разбиения типа `Column-N` состоят из множества квадратных в сечении подобластей, вытянутых вдоль одной из осей X , Y или Z . Площадь поперечного сечения подобластей и соответственно их количество варьировалось в экспериментах.

Вытянутая форма подобласти может привести к увеличению накладных расходов, связанных с межпроцессорным обменом. Чтобы уменьшить их влияние на производительность, исследовались также стратегии `Block2` и `Block4`, которые отличаются от `Column-Z` тем, что расчётная область сначала делится на две или четыре части вдоль направления

Х. Для сравнения используется встроенная декомпозиция Scotch, минимизирующая число межпроцессорных границ.

Расчёты проводились на суперкомпьютере Института проблем лазерных и информационных технологий РАН производительностью 3,3 Тфлопс. Кластер на платформе Scientific Linux 6 состоит из семи серверных четырёхпроцессорных материнских плат, соединённых через коммутатор Infiniband. Четыре материнские платы работают с процессором AMD Opteron 6378 (2,4 ГГц), который имеет 16 расчётных ядер. Три платы содержат 12-ядерные процессоры AMD Opteron 6168 (1,9 ГГц). В [11] показано, что интерконнект Infiniband выигрывает в производительности у интерконнекта Gigabit Ethernet уже на 16 ядрах на подобной по объёму задаче. В данной работе минимальное число ядер, использованное для расчётов, равно 24, поэтому интерконнект Gigabit Ethernet не исследуется.

Результаты расчётов производительности. Для осуществления вычислений в параллельном режиме требуется регулярное согласование граничных условий у смежных подобластей в процессе счёта. Объём передаваемой при этом информации пропорционален количеству граничных ячеек. Утилита decomposePar выводит подробную информацию о сетке, например количество внутренних и граничных ячеек у каждого ядра, и рассчитывает средние N_{aver} и максимальные N_{max} значения этих параметров по процессорам. Среднее количество граничных ячеек также можно оценить, зная пропорции декомпозиции и размер расчётной области (см. табл. 1).

На рис. 2, *a, c* приведены зависимости количества граничных ячеек от числа подобластей для различных стратегий разбиения исследованных сеток. Оптимальной стратегией с этой точки зрения является стратегия Column-Y. Также близкие результаты имеют стратегии Block4, Column-Z и Slice X-norm. По рис. 2, *c* видно, что Scotch показывает минимальное среднее значение площади межпроцессорных границ.

На рис. 2, *b, d* дано время расчёта для различных методов декомпозиции. Низкие значения площади межпроцессорных границ приводят к лучшей производительности расчёта, однако это не обязательное условие. Декомпозиции с большим значением площади границ также могут показывать удовлетворительную производительность. Стоит отметить ухудшение производительности в расчётах, использующих количество подобластей, не кратное числу ядер процессора, который задействован не полностью (см. табл. 1).

Если число расчётных ячеек не делится нацело на количество ядер, объём холостой работы увеличивается, снижая эффективность расчёта. Можно взять отношение внутренних ячеек $\Delta N = (N_{\text{max}} - N_{\text{aver}})/N_{\text{aver}}$ для оценки неравномерности распределения данных по процессорам. Например, увеличение неравномерности распределения данных с $\Delta N = 0,1\%$ до $\Delta N = 0,4\%$ приводит к ухудшению производительности на 14% (с 1010 до 1152 с) для декомпозиции Column-Z на 192 ядрах. Идеальным является полностью равномерное распределение числа ячеек по подобластям. В алгоритме декомпозиции Scotch допускается различие порядка 1%, при котором возможно ухудшение производительности.

В табл. 2 представлено распределение нагрузки по ядрам в сериях вычислений для трёх декомпозиций. Сравняются две декомпозиции, полностью идентичные по площади границ и числу ячеек на подобласть, а также встроенная декомпозиция Scotch, которая имеет минимальные межпроцессорные границы (рис. 3). На практике поперечный размер ванны расплава в горизонтальном направлении (*Y*) чаще всего больше, чем в вертикальном (*Z*). Это приводит к тому, что метод декомпозиции Column-Z распределяет её на большее число ядер, что даёт выигрыш в производительности такой декомпозиции (см. рис. 3). Встроенная декомпозиция Scotch не учитывает специфику задачи, делит ванну расплава на меньшее число подобластей и показывает несколько худшую производительность.

Очерёдность нумерации подобластей по направлениям *X, Y, Z* также влияет на производительность. Особенно это заметно при использовании большого числа вычислитель-

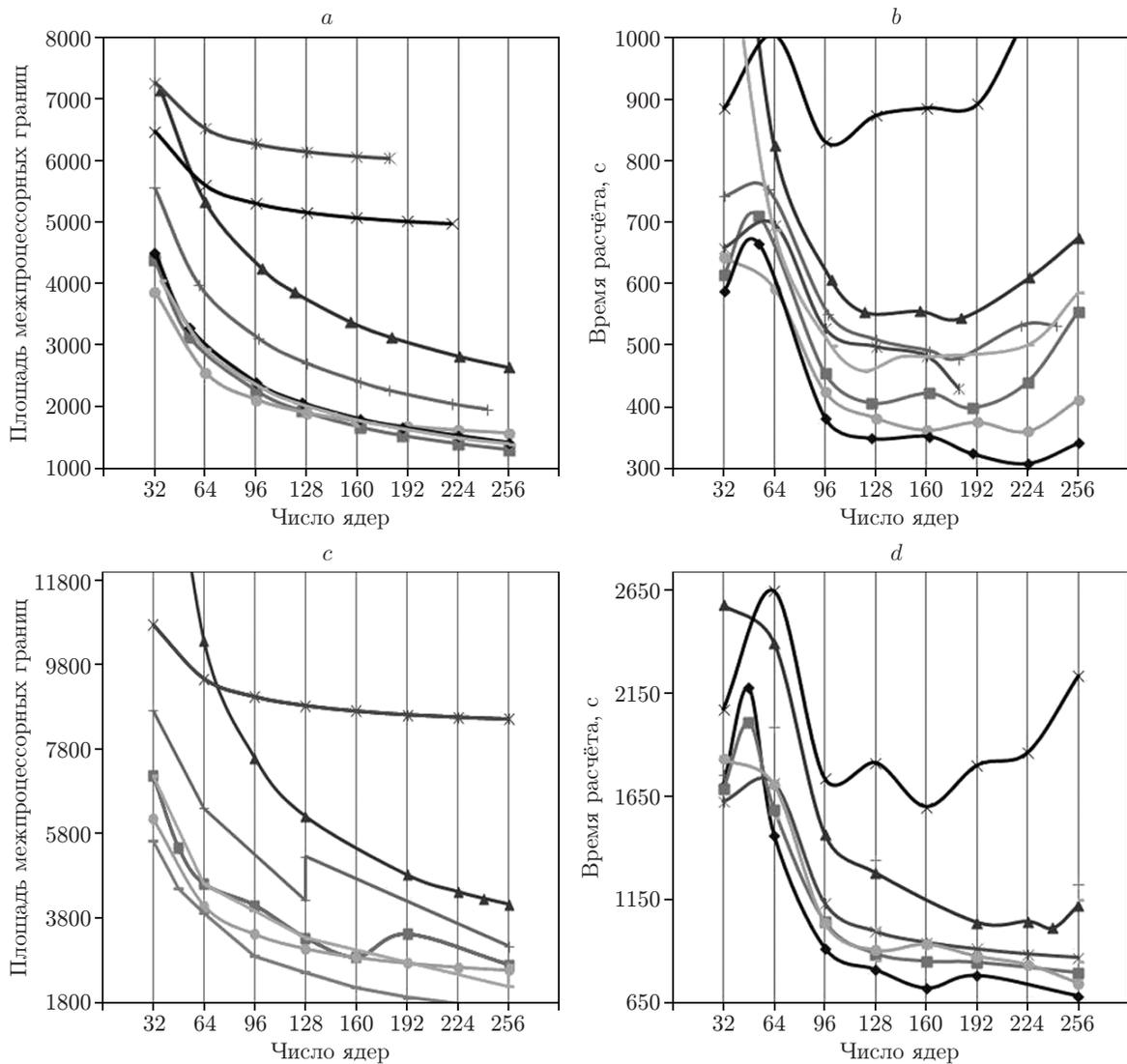


Рис. 2. Число граничных ячеек и время расчёта для различных декомпозиций: *a, b* — малая, *c, d* — большая сетки

Таблица 2

Сравнение распределения сконцентрированной нагрузки для лучших декомпозиций

Число ядер	Column-Z (<i>xy1</i>)		Column-Y (<i>x1z</i>)		Scotch	
	Число ядер с ванной расплава	Процент ядер с ванной расплава	Число ядер с ванной расплава	Процент ядер с ванной расплава	Число ядер с ванной расплава	Процент ядер с ванной расплава
32	7	21,87	6	18,75	7	21,87
48	7	14,58	7	14,58	8	16,66
64	11	17,18	9	14,06	9	14,06
96	13	13,54	13	13,54	10	10,41
128	18	14,06	15	11,71	14	10,93
160	20	12,4	18	11,25	15	9,37
192	23	11,97	23	11,97	17	8,85
256	33	12,89	27	10,54	22	8,59

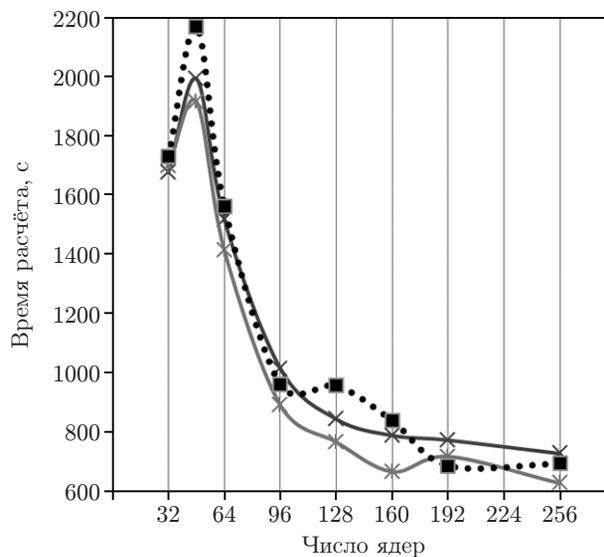


Рис. 3. Время расчёта для лучших декомпозиций на большой сетке (× — Column-Y, * — Column-Z, ■ — Scotch)

ных ядер. При очередности XYZ общую границу имеют подобласти с большей разницей номеров, чем при очередности YXZ , т. е. соответствующие им ядра обмениваются информацией. Различие порядкового номера >16 в нашем случае свидетельствует о межпроцессорном обмене, а >64 — об обмене данными между материнскими платами. В табл. 3 приведена длительность расчётов с различной очередностью нумерации для декомпозиции Column-Z. Нумерация, начинающаяся с X , показывает худшее время, так как в этом случае имеется больше смежных подобластей, рассчитываемых ядрами разных материнских плат. Обмен данными между такими ядрами осуществляется дольше, что ухудшает производительность. Можно сделать вывод, что нумерацию следует начинать с направления, которое распределено на меньшее число подобластей.

На рис. 4 показано время решения отдельных уравнений. Считается, что соответствующие пропорции, полученные для всего расчёта и исследуемого временного интервала, совпадают. Вычисление интеграла для доли новой фазы, итерационное согласование её с температурой (TempTime), а также расчёт поля давления (PreqTime) занимают значительную часть расчёта. Уравнения для скорости (UeqnTime) и концентрации (CeqnTime) не требуют много времени.

Для разбиений с плохим ($\gg 1$) аспектным соотношением у подобластей (например, Column-X или Slice Z-norm) итерационный расчёт давления занимает большую часть

Таблица 3

Сравнение влияния направления нумерации различного числа ядер для декомпозиции Column-Z на время расчёта

Число ядер (число частей по направлениям)	Время расчёта, с	
	XYZ	YXZ
64 ($16 \times 4 \times 1$)	1424	1404
128 ($32 \times 4 \times 1$)	847	833
256 ($64 \times 4 \times 1$)	689	624
512 ($64 \times 8 \times 1$)	1049	833

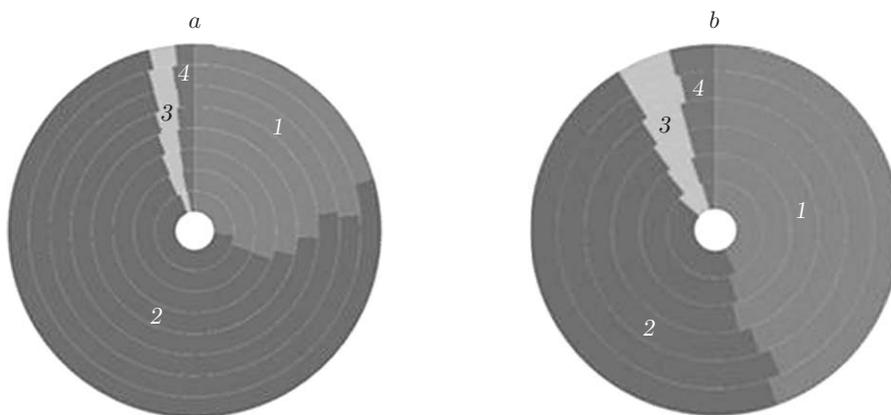


Рис. 4. Распределение времени решения уравнений для двух методов декомпозиции для малой сетки: *a* — Column-*X*, *b* — Slice *Y*-norm (1 — TeqnTime, 2 — PeqnTime, 3 — SeqnTime, 4 — UeqnTime)

времени (~75 %). Его длительность также зависит от количества ядер при использовании декомпозиции Column-*X*. Для разбиения Slice *Y*-norm такой зависимости не наблюдается. Несмотря на максимальное значение количества граничных ячеек (рис. 4, *a*) декомпозиция Slice *Y*-norm показывает удовлетворительные результаты по времени расчёта. Процент времени вычисления поля давления остаётся на одном уровне и не замедляет расчёт при увеличении числа ядер.

Рис. 5 иллюстрирует эффективность распараллеливания: $M = t_1 / (Nt_N)$, где t_i — время расчёта на i ядрах, N — число используемых ядер. Видно, что в нашем случае удастся получить удовлетворительные результаты. Значения больше единицы, по-видимому, связаны с неверной оценкой времени последовательного расчёта.

В соответствии с полученными результатами оптимальной стратегией является Column-*Z*. Она позволяет добиться максимальной эффективности и наилучшего ускорения при серийных расчётах. Вертикальная ориентированность подобластей даёт возможность наиболее эффективно расходовать процессорное время при проведении расчётов с пере-

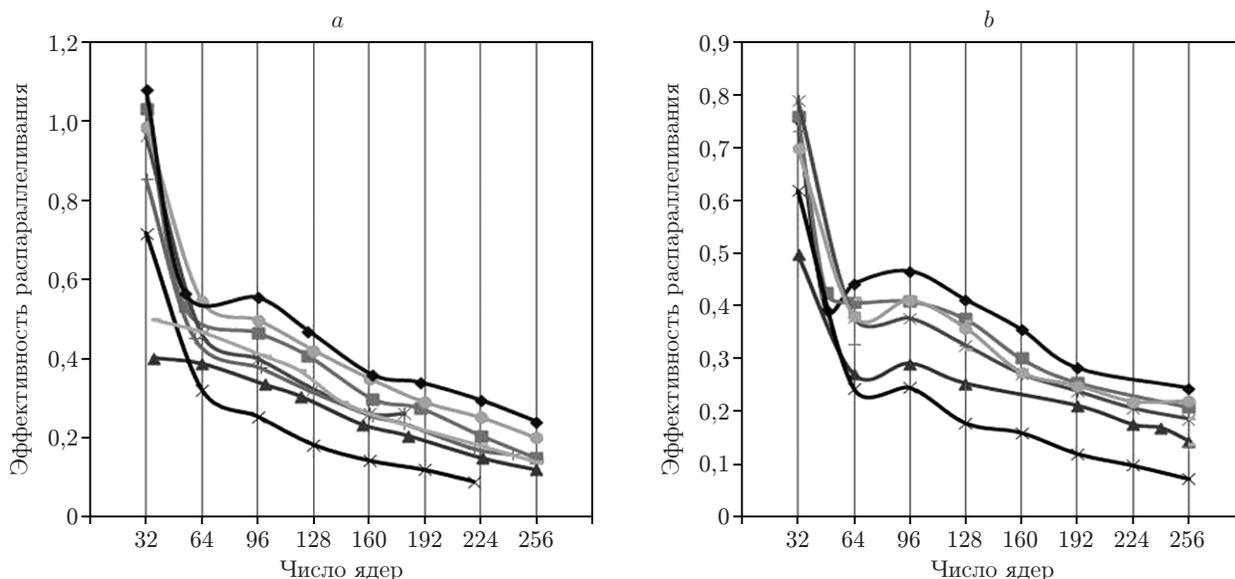


Рис. 5. Эффективность распараллеливания для различных декомпозиций: *a* — малая, *b* — большая сетки

менными комбинациями значений констант и параметров технологического процесса, что влечёт за собой, в частности, изменение глубины ванны расплава.

Таким образом, при исследовании режимов лазерных аддитивных технологий с низкими значениями мощности лазерного излучения или высокими скоростями сканирования, при которых глубина ванны расплава оказывается невелика, потеря эффективности из-за холостого режима работы части вычислительных ядер минимальна среди использованных стратегий декомпозиции. Масштабируемость расчёта наблюдалась вплоть до 224 ядер (28 подобластей вдоль X и 8 — вдоль Y).

Удлиненная вдоль направления X форма расчётной области приводит к существенному увеличению площади межпроцессорных границ при применении формы подобластей декомпозиции вытянутой вдоль траектории движения ванны расплава (Slice Y -norm, Slice Z -norm, Column- X). Такие стратегии декомпозиции оказываются неэффективными в задаче проведения серийных расчётов лазерных аддитивных технологий.

Заключение. В данной работе исследовалась производительность свободного пакета OpenFOAM для задач моделирования кристаллизации и тепломассопереноса в лазерных аддитивных технологиях. Для повышения производительности необходимо придерживаться следующих рекомендаций:

- 1) максимально распределять сконцентрированную нагрузку по вычислительным ядрам в каждый момент времени;
- 2) снижать площадь межпроцессорных границ;
- 3) начинать нумерацию подобластей по направлениям, которые распределены на меньшее число частей;
- 4) использовать количество подобластей разбиения, кратное числу ядер процессора;
- 5) равномерно распределять число ячеек между вычислительными ядрами;
- 6) избегать больших аспектных соотношений сторон для подобластей.

Максимальную эффективность распараллеливания при проведении серийных расчётов для задач аддитивных технологий демонстрирует разбиение Column- Z . Вертикальная ориентированность подобластей позволяет наиболее эффективно задействовать процессорное время при проведении расчётов с переменными комбинациями значений констант и параметров технологического процесса, что может изменить глубину ванны расплава. Масштабируемость расчёта наблюдается при использовании до 256 вычислительных ядер и зависит от размера сетки.

СПИСОК ЛИТЕРАТУРЫ

1. **Hendrickson B., Kolda T. G.** Graph partitioning models for parallel computing // *Parallel Computing*. 2000. **26**, Is. 12. P. 1519–1534.
2. **Павский К. В., Курносков М. Г., Поляков А. Ю.** Инструментарий оптимизации параллельного моделирования наноструктур с квантовыми точками // *Автометрия*. 2014. **50**, № 3. С. 56–61.
3. **Wang M., Tang Y., Guo X., Ren X.** Performance analysis of the graph-partitioning algorithms used in OpenFOAM // *Proc. of the IEEE 5th Intern. Conf. on Advanced Computational Intelligence (ICACI)*. Nanjing, China. 2012. P. 99–104.
4. **OpenFOAM.** URL: www.openfoam.org (дата обращения: 22.09.2016).
5. **Гришаев Р. В., Мирзаде Ф. Х., Низьев В. Г. и др.** Моделирование плавления и кристаллизации при селективном лазерном спекании с инъекцией металлических порошков // *Физ. и химия обработки материалов*. 2013. № 1. С. 12–23.
6. **Mirzade F. Kh., Niziev V. G., Panchenko V. Ya. et al.** Kinetic approach in numerical modeling of melting and crystallization at laser cladding with powder injection // *Phys. B: Condensed Matter*. 2013. **423**. P. 69–76.

7. **Хоменко М. Д., Панченко В. Я., Низьев В. Г. и др.** Исследование микроструктуры наплавленного слоя при лазерном спекании с коаксиальной подачей порошков // Изв. РАН. Сер. Физическая. 2016. **80**, № 4. С. 422–427.
8. **Mirzade F. Kh., Khomenko M. D., Niziev V. G.** Numerical simulation of solute evolution during laser cladding with nickel superalloy powder injection // Opt. and Quant. Electron. 2016. **48**, Is. 11. P. 513-1–513-12.
9. **Rösler F., Brüggemann D.** Shell-and-tube type latent heat thermal energy storage: Numerical analysis and comparison with experiments // Heat and Mass Transfer. 2011. **47**, Is. 8. P. 1027–1033.
10. **OpenFOAM Foundation.** The Open Source CFD Toolbox User Guide. 2014. URL: <http://cfd.direct/openfoam/user-guide/> (дата обращения: 31.10.2016).
11. **Nilsson H.** Some experiences on the accuracy and parallel performance of OpenFOAM for CFD in water turbines // Proc. of the 8th Intern. Conf. on Applied Parallel Computing: State of the Art in Scientific Computing. Ser. Lecture Notes in Comput. Sci. Berlin — Heidelberg: Springer-Verlag, 2007. Vol. 4699. P. 168–176.

Поступила в редакцию 5 мая 2016 г.
