

ИЗМЕРЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ И МАСШТАБИРУЕМОСТИ ПРОГРАММНЫХ КОМПЛЕКСОВ

Г.А. БАГДАСАРОВ *, О.Г. ОЛЬХОВСКАЯ *

* Институт прикладной математики им. М.В. Келдыша РАН
Москва, Россия

Ключевые слова: пакеты прикладных программ, высокопроизводительные вычисления, ускорение, эффективность, масштабируемость

Аннотация. Выполнены измерения производительности пакета прикладных программ для трехмерного моделирования задач мультифизики на сетках нерегулярной структуры при решении на высокопроизводительных ЭВМ модельных и производственных задач из предметной области пакета. Вычислены такие характеристики, как ускорение, эффективность и стоимость вычислений. Выявлен показатель, определяющий практическую производительность – количество узлов сетки на расчетном узле, и его оптимальные значения для рассматриваемого программного обеспечения – от 10 до 50 тысяч.

PROGRAM PACKAGES PERFORMANCE AND SCALABILITY MEASUREMENTS

G.A. BAGDASAROV *, O.G. OLKHOVSKAYA *

* Keldysh Institute of Applied Mathematics RAS
Moscow, Russia

e-mail: gennadiy3.14@gmail.com, web page: <http://www.keldysh.ru/persons/genabug.html>

Key words: application software, high performance computations, speedup, efficiency, scalability

Summary. Performance was measured of an application program package for multiphysics simulations at irregular meshes. The measurements were carried out while solving both test and practical problems using high performance computing. Speedup, efficiency and cost of computations were examined. Practical performance criteria was defined which is the number of computational mesh cells per computing node. Its optimal value for the studied software was found to be from 10 to 50 thousand.

1 ВВЕДЕНИЕ

Методы измерения, оценки и анализа производительности и масштабируемости программных комплексов, ориентированных на использование высокопроизводительной и суперкомпьютерной вычислительной техники, рассмотрены на примере параллельного программного комплекса (ППК) для трехмерного моделирования задач мультифизики. ППК создан в ИПМ им. М.В.Келдыша РАН на основе численных методик с применением неструктурированных расчетных сеток. Созданное программное обеспечение (ПО) не только позволяет решать задачи мультифизики, объединяющие разномасштабные конвективные и диссипативные процессы, в трехмерных областях произвольной геометрии, но и включает ряд инноваций, призванных сделать его полнофункциональным исследовательским кодом, использующим как актуальные физические и математические модели и численные методы, так и возможности современной высокопроизводительной вычислительной техники.

ППК представляет собой совокупность вычислительной среды (инфраструктуры) для проведения параллельных расчетов начально-краевых задач на неструктурированных вычислительных сетках и реализованных в рамках этой среды физических солверов.

Основные численные методы:

- ◆ технология неструктурированных и блочно-структурированных сеток, смешанные тетраэдральные, гексаэдральные и призматические элементы;
- ◆ явные TVD/WENO схемы повышенной точности для уравнений гиперболического типа, неявные схемы на основе варианта метода Галеркина с разрывными базисными функциями для диссипативных процессов;
- ◆ схема суммарной аппроксимации для полной системы уравнений;
- ◆ схема "предиктор-корректор" 2-го порядка точности по времени.

Основные принципы архитектуры программного комплекса:

- ◆ использование высокопроизводительных распределенных вычислений;
- ◆ большое число сервисных функций (ввод-вывод данных, операции с сетками, поддержка параллельных вычислений, динамическая работа с вычислительными объектами);
- ◆ динамическое создание, уничтожение и настройка различных вычислительных объектов (солверы, аппроксимации, граничные условия, свойства вещества), автоматизация управления памятью;
- ◆ использование объектно-ориентированного подхода к проектированию ПО, приемов объектно-ориентированного и обобщенного программирования (язык реализации C++), соответствие современным мировым стандартам коммерческой разработки ПО;
- ◆ кросс-платформенность.

Для реализации сложной модели мультифизики на каждом временном шаге необходимо вычислять несколько десятков физических величин в каждой ячейке расчетной сетки. Для моделирования существенно трехмерных процессов с учетом реальной геометрии устройств необходимы расчетные сетки, содержащие от нескольких миллионов до десятков и сотен миллионов ячеек. Решение задач такой размерности возможно только с использованием распределенных вычислений на базе современных высокопроизводительных ЭВМ. Поэтому ППК был сконструирован как программное обеспечение для высокопроизводительных параллельных вычислительных систем. Для выполнения расчетов на массивно-параллельных системах с распределенной памятью применяется известный подход геометрического параллелизма, основанный на декомпозиции расчетной области (разбиении ее на вычислительные домены, каждый из которых обрабатывается отдельным счетным ядром). Поддержка геометрического параллелизма встроена в дискретную модель ППК на уровне структур сеточных данных. Для декомпозиции смешанных сеток нерегулярной структуры использовался один из самых известных и широко применяемых пакетов такого рода – METIS (ParMETIS). Код может запускаться в качестве параллельного приложения MPI на компьютерах с общей или распределенной памятью. METIS обеспечивает в достаточной мере сбалансированное разбиение сетки в том смысле, что все вычислительные домены содержат приблизительно одинаковое число расчетных ячеек, и размеры границ доменов также приблизительно одинаковы. Однако, при решении реальных задач, в которых имеет место существенная пространственная неоднородность моделируемых физических процессов, баланс загрузки вычислительных ядер может нарушаться, и их использование может оказаться неэффективным.

2 ИССЛЕДОВАНИЕ МАСШТАБИРУЕМОСТИ ПАРАЛЛЕЛЬНОГО ПРОГРАММНОГО КОМПЛЕКСА

В нашем случае, когда численная модель включает большое количество различных по ресурсоемкости и организации вычислений солверов, а расчетная сетка имеет нерегулярную структуру, теоретическая оценка эффективности распараллеливания затруднительна. Таким образом, важное значение приобретают практические измерения производительности работы сложных программных комплексов на высокопроизводительных ЭВМ.

Основные параметры, характеризующие эффективность параллельных алгоритмов, методы их оценки и анализа описаны в книге В. П. Гергея¹.

Исследования масштабируемости параллельного программного комплекса выполнялись на примере двух задач: тестовой задачи о распространении тепловой волны (решается уравнение теплопроводности по неявной схеме, для решения систем линейных алгебраических уравнений (СЛАУ) используется параллельный пакет линейной алгебры Aztec²) и производственной задачи моделирования сжатия плазменного лайнера магнитным полем³ (решается полная система уравнений радиационной магнитной газовой динамики с табличными уравнениями состояния и учетом диссипативных процессов). В обоих случаях для организации параллельных

вычислений использовалась инфраструктура ППК, декомпозиция сеток осуществлялась с помощью пакета METIS. Измерения производились средствами внутренней системы журналирования ППК (фиксировалось астрономическое время выполнения операций на каждом вычислительном узле). Измерения выполнялись на суперкомпьютерах ЛОМОНОСОВ (НИВЦ МГУ) и МВС-100К (МСЦ РАН).

2.1 Исследование масштабируемости при повышении подробности расчетных сеток и увеличении числа вычислительных ядер

Исследование масштабируемости при повышении подробности расчетных сеток (до 10^6 ячеек) и увеличении числа вычислительных ядер (до 128 ядер) выполнялось в НИВЦ МГУ на суперкомпьютере "Ломоносов" с пиковой производительностью 510 Тфлопс.

Измерения выполнялись для тестовой задачи о распространении тепловой волны [4]. Решалось уравнение теплопроводности в форме:

$$\frac{\partial \varepsilon}{\partial t} = \frac{\partial}{\partial s} \left(k \frac{\partial T}{\partial s} \right) + f, \quad \varepsilon(T) = T \quad \left(C_V = \frac{\partial \varepsilon}{\partial T} \Big|_V = 1 \right), \quad k = k_0 T^\alpha.$$

Это уравнение имеет автомодельное решение

$$T(s, t) = \left[\frac{\alpha D}{k_0} (Dt - s + S_1) \right]^{\frac{1}{\alpha}}, \quad S_1 = const.$$

Это решение описывает тепловую волну, распространяющуюся в направлении s со скоростью D . Точка, в которой температура становится нулевой, имеет координату $S_{fr}(t) = S_1 + Dt$. Решение в области $s \geq 0$ существует при следующих начальных и граничных условиях:

$$T(s, 0) = \begin{cases} \left[\frac{\alpha D}{k_0} (S_1 - s) \right]^{\frac{1}{\alpha}}, & 0 < s \leq S_1 \\ 0, & s > S_1 \end{cases}$$

$$T(0, t) = \left[\frac{\alpha D}{k_0} (S_1 + Dt) \right]^{\frac{1}{\alpha}}, \quad t > 0.$$

В качестве автомодельной переменной s выбиралась координата x . Тестовые расчеты выполнялись для следующего набора параметров: $\alpha=2$; $k_0=0,5$; $S_1=0,5$; $D=5$; $0 \leq t \leq 0,4$.

Расчетная область – куб $3 \times 3 \times 3$, шаг по времени $\Delta t = 10^{-4}$ (для решения задачи выполняется 100 шагов).

ИЗМЕРЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ И МАСШТАБИРУЕМОСТИ ...

Использовались 4 расчетные сетки, включающие соответственно 125 тысяч (50x50x50), 250 тысяч (100x50x50), 500 тысяч (200x50x50) и 1 миллион (400x50x50) ячеек. На каждой из этих сеток решалось уравнение теплопроводности с использованием 16, 32, 64 и 128 вычислительных ядер. С помощью встроенных средств журналирования ППК измерялось время решения задачи в целом и отдельно время выполнения первичной подготовки данных. По результатам измерений вычислены ускорение, эффективность использования параллельным алгоритмом процессоров и стоимость вычислений. Результаты измерений и обработки полученных данных приведены в таблице 1 и далее на рисунках 1 – 4. T_p – время счета при использовании p вычислительных ядер.

Особенностью исследуемого алгоритма является наличие этапа предварительной подготовки данных, выполняемого перед началом собственно расчета (решения уравнений). Предварительная подготовка данных заключается в чтении описания расчетной сетки из файла, анализе отношений смежности и инцидентности элементов сетки, выделении сеточных элементов, используемых в межпроцессорных обменах, вычислении геометрических соотношений, необходимых для аппроксимации решаемых уравнений, построении портрета матрицы СЛАУ и др. Такая достаточно сложная предварительная работа с сеточными данными характерна для большинства алгоритмов, использующих сетки нерегулярной структуры. Этот этап заметно отличается от последующего по характеру выполняемых на нем операций с сеткой и данными, и, как следствие, по показателям эффективности распараллеливания. В рассмотренном примере этот этап занимал от 5 до 10 процентов общего времени счета, в других случаях (больше время моделирования) эта доля может быть меньше. Поэтому далее показатели для двух этапов расчета будут обсуждаться отдельно.

Графики показателей эффективности для этапа предварительной подготовки данных подписаны "data preparing", для этапа решения уравнений – "iterations", для расчета в целом – "total".

Измерения проводились для двух разностных аппроксимаций уравнения теплопроводности (обе реализованы и используются в ППК). Обе схемы строятся на основе метода Галеркина с разрывными базисными функциями, аппроксимация 1 использует для потоков базисные функции, равные единице в центрах граней ячеек, аппроксимация 2 – в узлах. Во втором случае шаблон разностных уравнений включает в несколько раз больше ячеек (в 3,85 раза больше для регулярной сетки гексаэдров, использовавшейся в тестовой задаче, в 10 и более раз – для сеток тетраэдров). Этот фактор следует учитывать при оценке вычислительной сложности задачи наряду с размером расчетной сетки. Графики показателей эффективности для аппроксимации 1 приведены в верхней части соответствующих рисунков, для аппроксимации 2 – в нижней.

аппроксимация 1													
	р	подго- товка данных	решение уравне- ния	весь расчет	подго- товка данных	решение уравне- ния	весь расчет	подго- товка данных	решение уравне- ния	весь расчет	подго- товка данных	решение уравне- ния	весь расчет
сетка (тыс. ячеек)		125	125	125	250	250	250	500	500	500	1000	1000	1000
Tr (с)	16	16	166	182	34	327	361	97	709	806	270	2107	2377
Sp	16	16	16	16	16	16	16	16	16	16	16	16	16
Er	16	1	1	1	1	1	1	1	1	1	1	1	1
Cr (с-ядро)	16	256	2656	2912	544	5232	5776	1552	11344	12896	4320	33712	38032
Tr (с)	32	9	99	108	20	201	221	35	363	398	110	1037	1147
Sp	32	28,44	26,82	26,96	27,2	26,02	26,14	44,34	31,25	32,4	39,27	32,51	33,16
Er	32	0,89	0,84	0,85	0,85	0,81	0,82	1,39	0,98	1,01	1,23	1,02	1,04
Cr (с-ядро)	32	288	3168	3456	640	6432	7072	1120	11616	12736	3520	33184	36704
Tr (с)	64	4	73	77	9	121	130	20	223	243	51	534	585
Sp	64	64	36,38	37,82	60,44	43,24	44,43	77,6	50,87	53,07	84,71	63,13	65,01
Er	64	1,0	0,57	0,59	0,95	0,68	0,7	1,21	0,80	0,83	1,7	0,99	1,02
Cr (с-ядро)	64	256	4672	4928	576	7744	8320	1280	14272	15552	3264	34176	37440
Tr (с)	128	3	69	72	5	98	103	10	149	159	22	327	305
Sp	128	85,33	38,49	40,44	108,8	53,39	56,08	155,2	76,13	81,11	196,4	103,09	124,7
Er	128	0,67	0,3	0,32	0,85	0,42	0,44	1,2	0,6	0,63	1,53	0,86	0,9
Cr (с-ядро)	128	384	8832	9216	640	12544	13184	1280	19072	20352	2816	39040	41856
аппроксимация 2													
	р	подго- товка данных	решение уравне- ния	весь расчет	подго- товка данных	решение уравне- ния	весь расчет	подго- товка данных	решение уравне- ния	весь расчет	подго- товка данных	решение уравне- ния	весь расчет
сетка (тыс. ячеек)		125	125	125	250	250	250	500	500	500	1000	1000	1000
Tr (с)	16	29	409	438	72	933	1005	221	2704	2925	612	8584	9196
Sp	16	16	16	16	16	16	16	16	16	16	16	16	16
Er	16	1	1	1	1	1	1	1	1	1	1	1	1
Cr (с-ядро)	16	464	6544	7008	1152	14928	16080	3536	43264	46800	9792	137344	147136
Tr (с)	32	13	213	226	34	450	484	74	1272	1346	229	3833	4062
Sp	32	35,7	30,7	31,0	33,9	33,2	33,2	47,8	34,0	34,8	42,8	35,8	36,2
Er	32	1,11	0,96	0,97	1,06	1,04	1,04	1,49	1,06	1,09	1,33	1,12	1,13
Cr (с-ядро)	32	416	6816	7232	1088	14400	15488	2368	40704	43072	7328	122656	129984
Tr (с)	64	7	137	144	18	247	265	38	690	728	95	1696	1791
Sp	64	66,3	47,8	48,7	64	60,4	60,7	93,1	62,7	64,3	103,1	81,0	82,2
Er	64	1,04	0,75	0,76	1,00	0,94	0,95	1,43	0,98	1,00	1,61	1,27	1,28
Cr (с-ядро)	64	448	8768	9216	1152	15808	16960	2432	44160	46592	6080	108544	114624
Tr (с)	128	5	131	136	9	170	179	16	336	352	38	850	888
Sp	128	92,8	50,0	51,5	128	87,8	89,8	221	128,8	133,0	257,7	161,6	165,7
Er	128	0,73	0,39	0,40	1,00	0,67	0,70	1,73	1,00	1,04	2,01	1,26	1,29
Cr (с-ядро)	128	640	16768	17408	1152	21760	22912	2048	43008	45056	4864	108800	113664

Таблица 1 : Время расчета при увеличении расчетной сетки и количества вычислительных ядер.

ИЗМЕРЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ И МАСШТАБИРУЕМОСТИ ...

Ускорение S_p рассчитывалось по формуле $S_p = (T_{16} / T_p) \cdot 16$, т.к. минимальное число процессов, для которых проводились измерения, было равно 16.

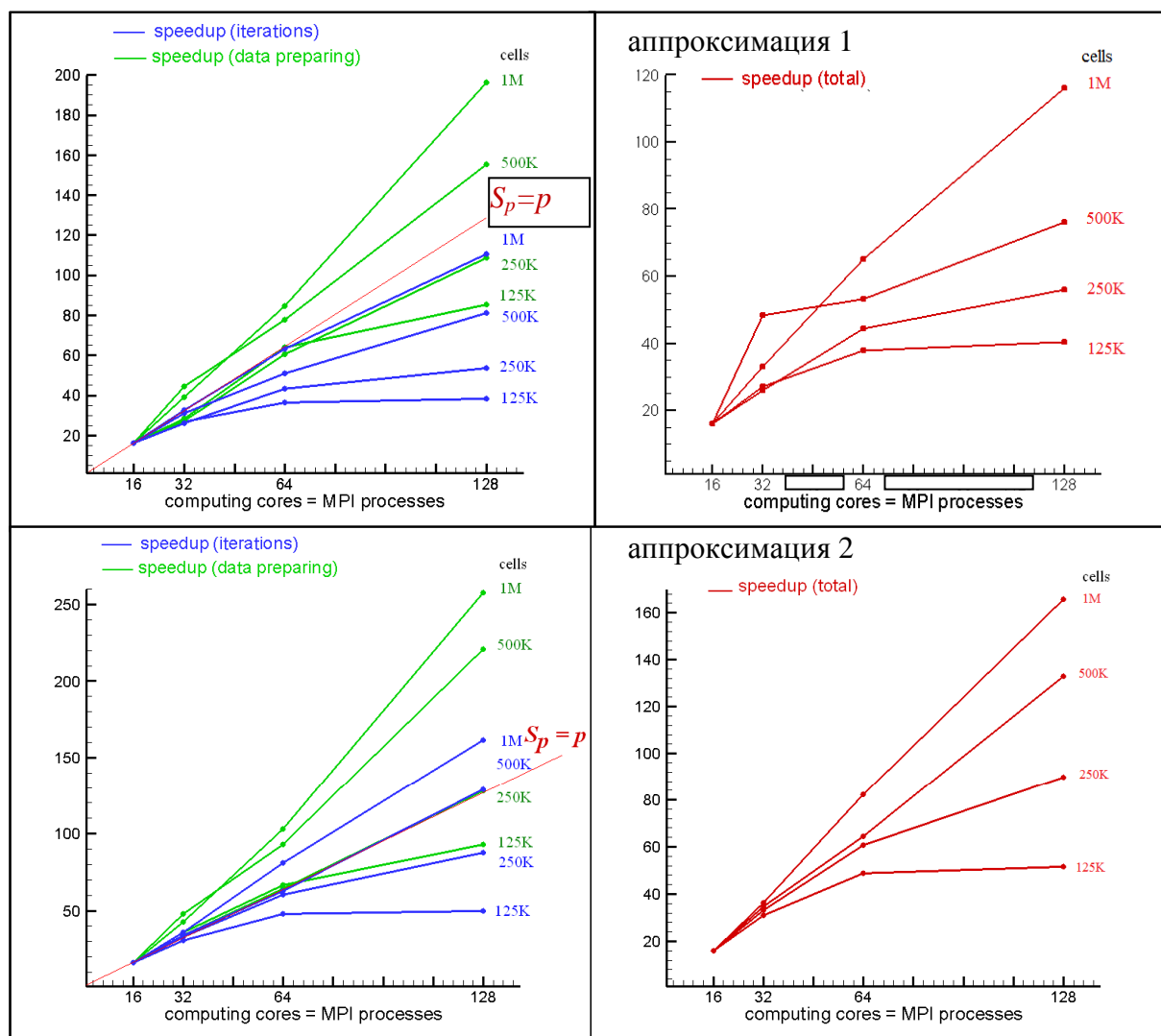


Рис. 1. Ускорение

Синие графики – ускорение при расчете шага по времени, зеленые – ускорение первичной подготовки данных, красные – расчета в целом. Справа от графиков указан размер расчетной сетки (количество ячеек), для которой проводились измерения.

Для первичной подготовки данных на больших сетках (500 тысяч и 1 миллион ячеек) имеет место сверхлинейное ускорение (ускорение в 9,7 и 12,3 раза при увеличении числа процессов в 8 раз), это связано с особенностями алгоритмов работы со структурами данных.

Из рисунка 1 видно, что масштабируемость алгоритма улучшается при увеличении размера сетки. На сетке в 1 миллион ячеек для расчета шага по времени практически достигается ускорение, близкое к максимально возможному (красная линия $S_p = p$ на графике) - ускорение в 6,5 раз при увеличении числа процессов в 8 раз. Для меньших

сеток ускорение значительно меньше. Таким образом, наблюдается эффект Амдаля¹ – ускорение является возрастающей функцией от вычислительной сложности задачи (размера сетки). Эффект усиливается при увеличении числа вычислительных ядер. По этой же причине ускорение выше для аппроксимации 2 (вычислительная сложность выше из-за расширения шаблона разностных уравнений).

Эффективность E_p использования параллельным алгоритмом процессоров при решении задачи определялась соотношением $E_p = (T_{16} \cdot 16) / (pT_p) = S_p/p$, т.е. за 1 принималась эффективность использования 16 процессоров (минимальное число, для которого проводились измерения). С этим связано незначительное превышение единичного уровня на графиках эффективности на рисунке 2.

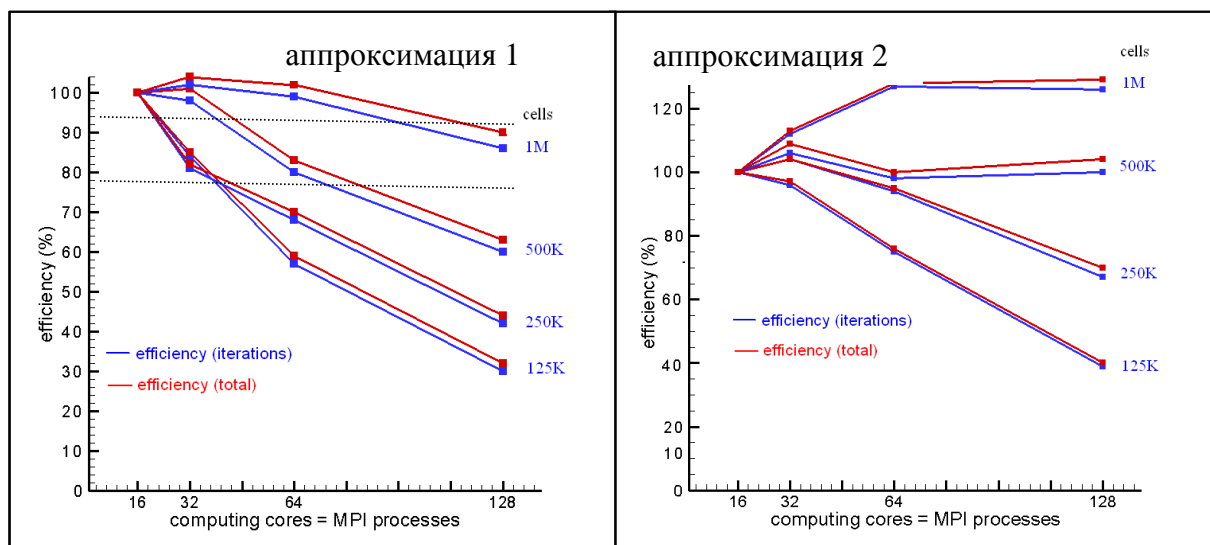


Рис. 2. Эффективность

Синие графики – эффективность при расчете шага по времени, красные – эффективность всего расчета в целом. Справа от графиков указан размер расчетной сетки (количество ячеек), для которой проводились измерения. Видно, что эффективность, близкая к максимально возможной, достигается при расчете на сетке, содержащей 1 миллион ячеек. Кроме того, в этом случае эффективность мало снижается с увеличением числа процессов (уменьшается от 100% до 90% при увеличении числа процессов в 8 раз для аппроксимации 1 и остается практически постоянной для аппроксимации 2), тогда как при использовании сеток меньшего размера эффективность падает в 2-3 раза.

Данных измерений недостаточно для полного анализа функции изоэффективности¹, однако можно утверждать, что для высоких значений эффективности $80\% < E < 90\%$ (пунктирные линии на рисунке 2) эта функция практически линейная $n = Kp$. Здесь мерой сложности вычислений n служит число ячеек расчетной сетки, p – количество вычислительных ядер. Это означает, что при увеличении сетки в q раз для сохранения эффективности распараллеливания следует использовать в q раз больше вычислительных ядер. Размер сетки, обрабатываемой одним ядром, остается при этом постоянным.

ИЗМЕРЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ И МАСШТАБИРУЕМОСТИ ...

Эффективность аппроксимации 2 во всех случаях выше, чем аппроксимации 1, хотя время решения на каждом этапе в 2-4 раза больше. Линейная функция изоэффективности имеет место только для сеток меньших размеров (125 и 250 тысяч). Таким образом, можно сделать вывод, что распараллеливание решения задачи более эффективно при использовании больших (миллион и более ячеек) расчетных сеток. В этом случае параллельный алгоритм можно считать масштабируемым, т.к. при росте числа процессоров он обеспечивает увеличение ускорения при сохранении постоянного уровня эффективности использования процессоров. Эффективность также повышается при увеличении сложности вычислений.

Стоимость вычислений C_p определялась как произведение времени параллельного решения задачи и числа используемых процессоров $C_p = pT_p$.

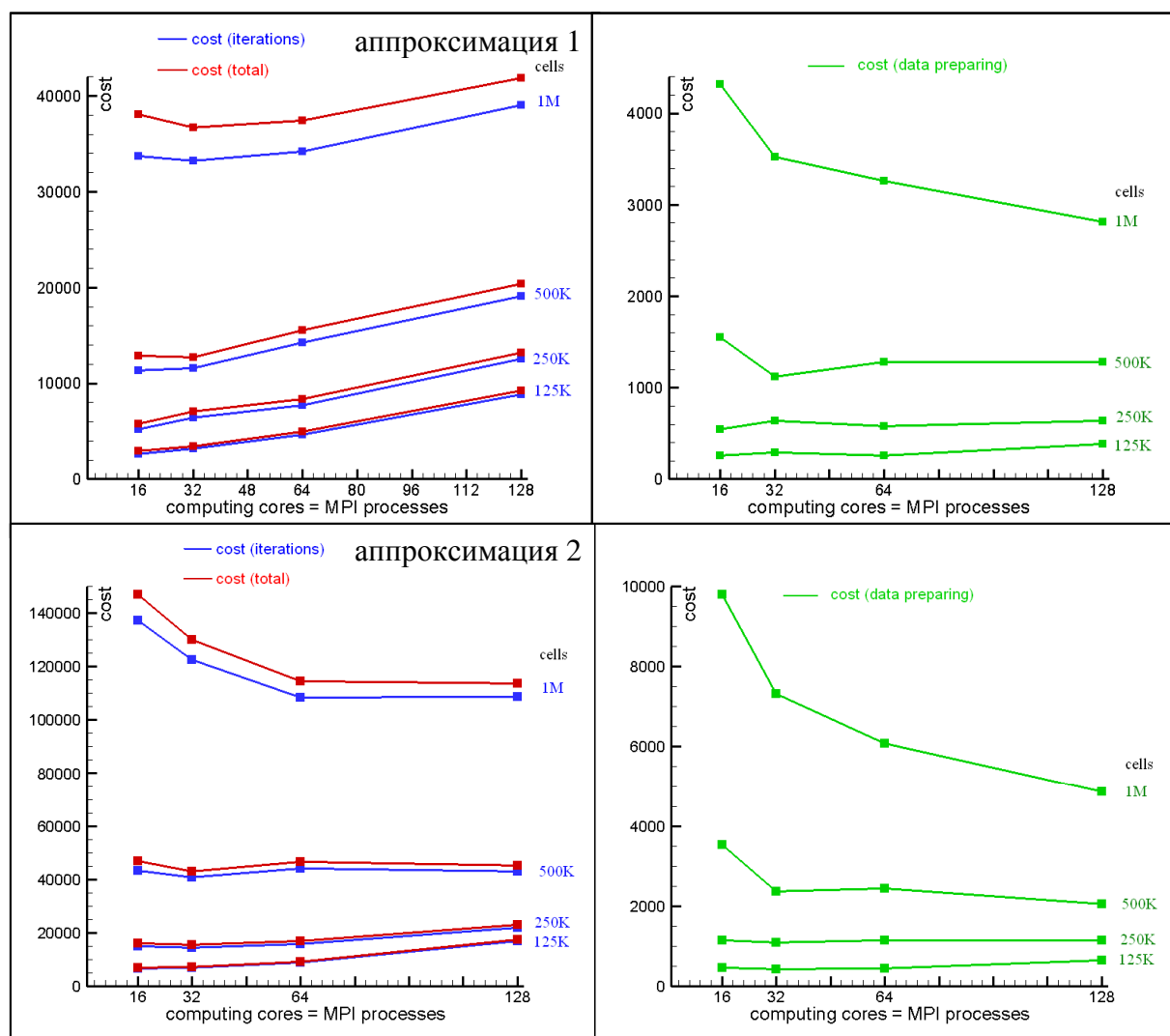


Рис. 3. Стоимость вычислений (сек·ядро)

На рисунке 3 слева представлена полная стоимость вычислений (красные графики) и стоимость вычислений при расчете шагов по времени (синие графики). Справа от графиков указан размер расчетной сетки (количество ячеек). Справа представлена стоимость первичной подготовки данных (зеленые графики). Поведение этих графиков заметно различается. Для аппроксимации 1 полная стоимость вычислений остается практически постоянной при расчете на сетке в 1 миллион ячеек и существенно возрастает с увеличением числа процессов (в 2 и даже 3 раза) для сеток меньших размеров. Для аппроксимации 2 полная стоимость вычислений неприемлемо высока при расчете на сетке в 1 миллион ячеек с малым числом процессов (16 и 32), падая почти в 1,5 раза при переходе к 64 и 128 процессам, и остается практически постоянной в остальных случаях. Стоимость подготовки данных в обоих случаях на меньших сетках практически постоянна, а на сетке в 1 миллион ячеек уменьшается в 1,5-2 раза при увеличении числа процессов в 8 раз. Стоимость подготовки данных для аппроксимации 2 во всех случаях приблизительно в 2 раза выше, чем для аппроксимации 1, что, очевидно, связано с большими вычислительными затратами на построение портрета матрицы СЛАУ и вычисление геометрических факторов матричных коэффициентов.

На рисунке 4 представлена зависимость стоимости вычислений (сек·ядро) от размера расчетной сетки (тысяч ячеек): слева – полная стоимость вычислений (красные графики) и стоимость вычислений при расчете шагов по времени (синие графики), справа – стоимость первичной подготовки данных (зеленые графики). Числа рядом с графиками показывают число вычислительных процессов, для которых проводились измерения.

На левом графике видно, что для аппроксимации 1 стоимость вычислений при 128 процессах заметно выше для всех проверенных сеток. При малом числе параллельных процессов (16, 32) рост стоимости быстрее линейного – при увеличении сетки в 8 раз стоимость возрастает в 13 и 11 раз соответственно. Практически линейный рост стоимости наблюдается в случае 64 процессов (рост стоимости в 7,6 раза при увеличении сетки в 8 раз). Для аппроксимации 2 это справедливо только для меньших сеток (125 и 250 тысяч ячеек). Для больших сеток (500 тысяч и 1 миллион ячеек) наоборот, заметно выше стоимость вычислений при малом числе процессов (16 и 32).

Следует также отметить, что рост стоимости вычислений в зависимости от вычислительной сложности (выбора аппроксимации) сильно зависит от размера сетки в целом. Если для сетки 125 тысяч ячеек стоимость расчета с аппроксимацией 2 выше стоимости аналогичного расчета с аппроксимацией 1 в 2-2,5 раза, то для сетки миллион ячеек – уже в 3-4 раза. Это различие проявляется сильнее при малом числе используемых процессоров. Такая тенденция связана с особенностями решения СЛАУ в распределенном режиме. На этапе подготовке данных различия значительно меньше, рост стоимости во всех случаях лежит в пределах от 1,7 до 2,2.

ИЗМЕРЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ И МАСШТАБИРУЕМОСТИ ...

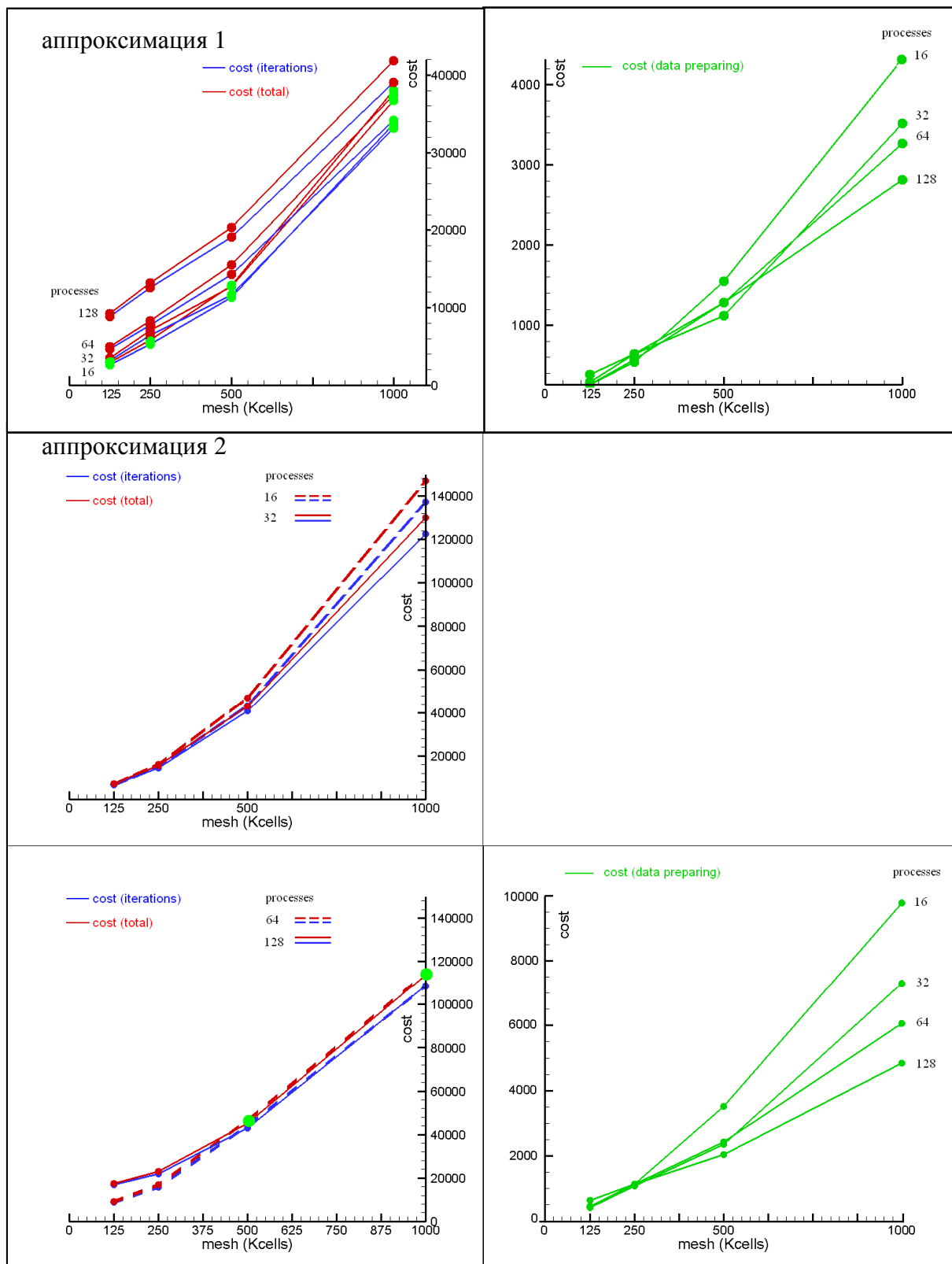


Рис. 4. Стоимость вычислений (сек·ядро) в зависимости от размера сетки

В таблице 2 приведены приблизительные размеры сетки (тысяч ячеек), приходящейся на каждый вычислительный процесс в рассмотренных случаях.

Размер исходной сетки количество вычислительных процессов (тыс. ячеек)	125	250	500	1000
16	7,8	15,6	31,25	62,5
32	3,9	7,8	15,6	31,25
64	1,95	3,9	7,8	15,6
128	0,97	1,95	3,9	7,8

Таблица 2 : Размер сетки в вычислительном домене.

Слева на рисунке 4 отмечены зеленым цветом точки, соответствующие верхнему правому треугольнику таблицы (ячейки с серым фоном). Точки, соответствующие нижнему левому треугольнику таблицы (ячейки с белым фоном), отмечены красным. Зеленым точкам соответствуют стоимостно-оптимальные вычисления. Таким образом, можно утверждать, что существует оптимальный размер сетки в вычислительном домене, обрабатываемом одним процессом. Для данного алгоритма он составляет приблизительно от 10 до 50 тысяч ячеек. При меньшем числе ячеек резко возрастает стоимость вычислений из-за роста доли накладных расходов на междоменные обмены. При увеличении размера сетки в домене снижается эффективность и уменьшается ускорение.

2.2 Исследование масштабируемости решения задачи плазменной мультифизики

Исследование масштабируемости ППК при увеличении числа вычислительных ядер (до 1024 ядер) выполнялось в МСЦ РАН с использованием суперкомпьютера МВС-100К с пиковой производительностью 140,16 ТФлопс.

Измерения проводились для типичной задачи плазменной мультифизики. Постановка задачи и результаты расчетов подробно описаны в ³. Моделировалось сжатие проволочной сборки магнитным полем мощного токового импульса. Под действием токового импульса вещество проволочек нагревается, испаряется и переходит в плазменное состояние. Возникающее азимутальное магнитное поле сжимает плазму к оси сборки, где в момент максимального сжатия происходит быстрый переход кинетической энергии в энергию излучения. Рассматривалась одиночная проволочная сборка цилиндрической формы, состоящая из 240 алюминиевых проволок, имеющих диаметр 10,4 мкм и высоту 50 мм. Диаметр сборки 140 мм, диаметр разрядной камеры 150 мм. В качестве граничного условия задавался ток генератора (амплитуда 6 МА, время нарастания тока 700 нс), и вычислялась напряженность магнитного поля на внешней стенке камеры. Начальная конфигурация проволочной сборки задавалась в форме плазменной оболочки с плотностью $1,5 \cdot 10^{-4}$ г/см³ и температурой 2 эВ. Плотность фоновой плазмы 10^{-7} г/см³. Уравнения состояния, транспортные коэффициенты и оптические свойства использованы в форме таблиц. Постановка задачи соответствует эксперименту № 637 на установке SPHINX Исследовательского Центра CEA Gramat (Франция).

ИЗМЕРЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ И МАСШТАБИРУЕМОСТИ ...

Использовалась расчетная сетка блочной структуры со смешанными элементами (гексаэдры и треугольные призмы), состоящая из 4 миллионов ячеек. Такой размер сетки обусловлен требованиями физической точности – на менее подробной сетке невозможно с необходимой точностью вычислить параметры центрального пинча, формирующегося на оси сборки. На этой сетке решалась полная система уравнений РМГД, включающая диссипативные процессы, с использованием 16, 32, 64, 128, 256, 512 и 1024 вычислительных ядер. Результаты измерений приведены в таблице 3.

количество вычислительных ядер	тыс. ячеек/ядро	подготовка данных (мин:сек)	шаг по времени без расчета переноса излучения (мин:сек)	шаг по времени с расчетом переноса излучения (мин:сек)
16	250	498:56	2:00	4:00
32	125	119:41	1:09	2:24
64	60	30:43	0:41	1:17
128	30	8:48	0:20	0:55
256	15	2:18	0:12	0:26
512	7	1:15	0:07	0:17
1024	3.6	0:31	0:04	0:13

Таблица 3 : Время расчета при увеличении числа вычислительных ядер.

По результатам проведенных измерений в алгоритме выявлено три части, которые масштабируются принципиально по-разному: начальная подготовка данных, расчет временного шага без расчета переноса излучения и расчет шага с расчетом переноса излучения. Для расчета переноса излучения используется многогрупповое (по спектру) диффузионное приближение, что связано с многократным (по числу спектральных групп) решением уравнения диффузии по неявной схеме, а также с интенсивной работой с таблицами оптических свойств. Поэтому расчет переноса излучения не только занимает от 50% до 70% времени расчета временного шага, но и показатели эффективности его распараллеливания отличаются от показателей других частей алгоритма. Первичная подготовка данных занимает в производственном расчете значительно больше времени, чем в тестовой задаче. Это связано с вычислением геометрических соотношений, необходимых для аппроксимации решаемых уравнений, в нескольких различных солверах. Тем самым становятся практически неприемлемыми расчеты на малом количестве вычислительных ядер. Например, расчет "первого шага" на 16 ядрах занял более 8,5 часов.

Вычисленные по результатам измерений показатели эффективности распараллеливания представлены в таблице 4 и на рисунках 5 и 6, n/p - средний размер сетки в домене, который обрабатывается на одном вычислительном ядре.

p	n/p (тыс.)	подготовка данных			шаг по времени без расчета переноса излучения			шаг по времени с расчетом переноса излучения		
		S_p	E_p	C_p	S_p	E_p	C_p	S_p	E_p	C_p
16	250	16	1	478976	16	1	1920	16	1	3840
32	125	66,7	2,1	229792	27,8	0,87	2208	26,7	0,83	4608
64	60	259,8	4,1	117952	46,8	0,73	2624	49,9	0,78	4928
128	30	907,2	7,1	67584	96	0,75	2560	69,8	0,55	7040
256	15	3470,4	13,6	35328	160	0,63	3072	147,7	0,58	6656
512	7	6386,2	12,5	38400	274,3	0,54	3584	225,9	0,44	8704
1024	3.6	15451,2	15,1	31744	480	0,47	4096	295,4	0,29	13312

Таблица 4 : Показатели эффективности распараллеливания при увеличении числа вычислительных ядер

Ускорение S_p рассчитывалось по формуле $S_p=(T_{16}/T_p)\cdot 16$, т.к. минимальное число процессов, для которых проводились измерение, было равно 16.

На рисунке 5 слева – ускорение при расчете шага по времени, справа – ускорение первичной подготовки данных.

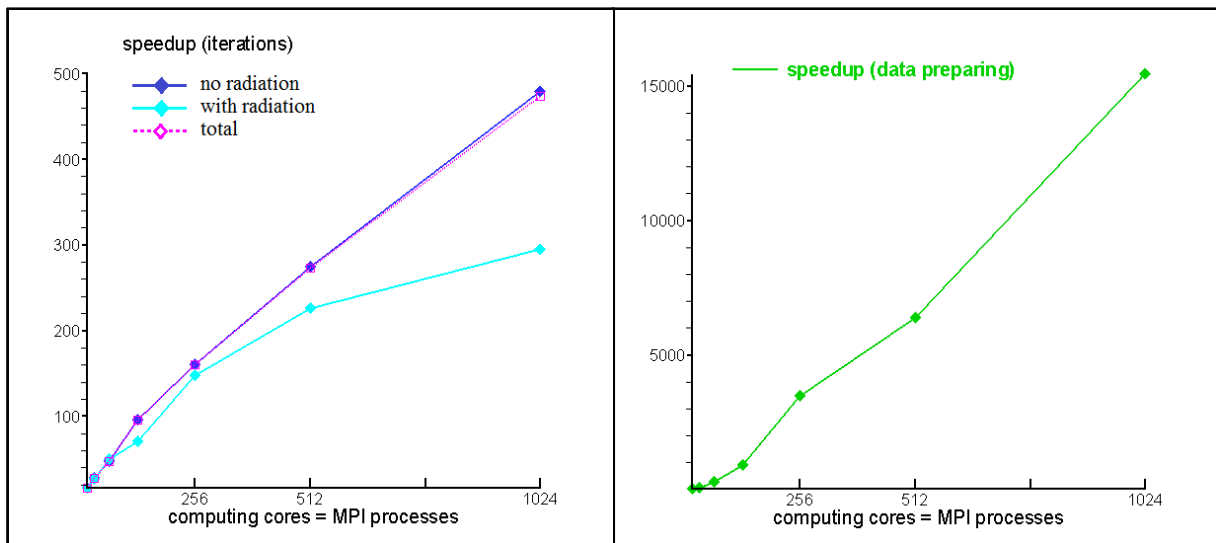


Рис. 5. Ускорение

Для первичной подготовки данных имеет место сверхлинейное ускорение (ускорение в 966 раз при увеличении числа процессов в 64 раза). Зависимость ускорения от числа ядер имеет почти квадратичный характер $S_p \propto p^2$. Этому соответствует квадратичная зависимость времени обработки сеточных данных от размера сетки на вычислительном узле $T_p \propto (n/p)^2$.

При расчете временных шагов без переноса излучения (синий график слева на

рисунке 6) ускорение на всем исследованном диапазоне остается линейным, но приблизительно в 2 раза меньше максимально возможного. Ускорение расчета переноса излучения (голубой график слева на рисунке 6) еще меньше и его рост заметно замедляется с увеличением числа MPI процессов. Для 1024 процессов ускорение почти в 4 раза меньше максимально возможного. В рассматриваемой задаче оказалось возможным использовать схему дробных шагов, в которой временной шаг расчета диффузии излучения приблизительно на 2 порядка больше временного шага для остальных процессов. Такая схема позволила существенно увеличить скорость и масштабируемость расчета в целом. Ускорение данного алгоритма показано слева на рисунке 6 сиреневой пунктирной линией, практически совпадающей с синим графиком "без излучения".

Эффективность использования параллельным алгоритмом процессоров при решении задачи и **стоимость вычислений** приведены на рисунке 6. Эффективность E_p определялась соотношением $E_p = (T_{16} \cdot 16) / (pT_p) = S_p/p$, т.е. за 1 принималась эффективность использования 16 процессоров (минимальное число, для которого проводились измерения). Относительная стоимость вычислений C_p определялась как произведение времени параллельного решения задачи и числа используемых процессоров, отнесенная к стоимости вычислений на 16 ядрах $C_p = (pT_p)/C_{16}$. В расчете стоимости и эффективности учитывалось только время решения уравнений (без начальной подготовки данных). При использовании 1024 вычислительных ядер стоимость вычислений возросла более чем вдвое, соответственно эффективность упала до 50%.

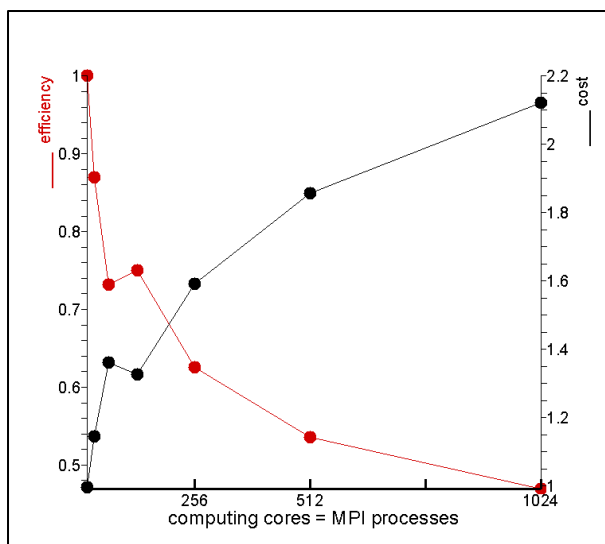


Рис. 6. Эффективность и стоимость вычислений

Как и в тестовой задаче, практически приемлемая эффективность вычислений оказалась зависящей от выявленного выше параметра – размера сетки на вычислительном ядре. В данном случае значение этого параметра также лежит в пределах от 50 до 10 тысяч ячеек. Т.е. эффективным было решение задачи на 64 – 256 ядрах. При меньшем числе ядер неприемлемо долгое время потребовалось для предварительной подготовки данных, при большем – заметно возросла стоимость временных шагов.

3 ЗАКЛЮЧЕНИЕ

Для параллельного программного комплекса, предназначенного для моделирования задач мультифизики на высокопроизводительных ЭВМ, были проведены серии измерений производительности на суперкомпьютерах Ломоносов (НИВЦ МГУ) и МВС-100К (МСЦ РАН). Изучение таких показателей, как ускорение, эффективность использования процессоров и стоимость вычислений, позволило проанализировать масштабируемость параллельных алгоритмов ППК и выявить следующие их особенности:

- ◆ Имеется ресурсоемкий, но хорошо масштабируемый этап предварительной подготовки данных, характерный для большинства алгоритмов, использующих сетки нерегулярной структуры. Наличие этого этапа в начале расчета делает практически неприемлемым решение сложных задач на больших сетках при малом числе вычислительных ядер.
- ◆ Имеются различия в масштабируемости разных физических солверов, что необходимо учитывать при построении хорошо масштабируемых алгоритмов.
- ◆ Определен показатель практически приемлемой эффективности – размер сетки в вычислительном домене, обрабатываемом одним MPI процессом. Значение этого показателя составило от 50 до 10 тысяч ячеек на вычислительное ядро.

Эти результаты справедливы как для тестовой задачи теплопроводности, так и для типичной производственной задачи из предметной области ППК, что позволяет предположить применимость сделанных выводов и для других задач, решаемых средствами параллельного программного комплекса. Таким образом, становится возможным априорный анализ эффективности их решения на высокопроизводительных многоядерных системах. Изложенная методика измерения и анализа производительности и масштабируемости может быть применена для широкого круга научных программных кодов, ориентированных на высокопроизводительные вычисления.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации по государственному контракту от 17.05.2012 г. № 07.524.12.4019 в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы».

ЛИТЕРАТУРА

- [1] В. П. Гергель. Теория и практика параллельных вычислений. Интернет-университет информационных технологий - ИНТУИТ.ру, БИНОМ. Лаборатория знаний, 2007 г., 424 стр.
- [2] R.S.Tuminaro, M.Heroux et al Official Aztec User's Guide Version 2.1, Sandia National Laboratories, Albuquerque, 1999.
- [3] В. А. Гасилов, А. С. Болдарев, С. В. Дьяченко, и др. Пакет прикладных программ MARPLE3D для моделирования на высоко-производительных ЭВМ импульсной магнитоускоренной плазмы. - Математическое моделирование, 2012, том 24, номер 1, стр. 55–87.
- [4] Я. Б. Зельдович, Ю. П. Райзер. Физика ударных волн и высокотемпературных гидродинамических явлений. М., ФИЗМАТЛИТ, 2008.

Received November 11, 2012