

Применение архитектуры multiGPU+CPU для задач прямого численного моделирования ламинарно - турбулентного перехода при рассмотрении задач в качестве нелинейных динамических систем *

Евстигнеев Н.М., Рябков О.И.

Федеральный исследовательский центр "Информатика и управление" Российской Академии наук, Институт Системного Анализа.

В работе обобщаются данные о применении различных параллельных вычислительных архитектур при численном моделировании задач ламинарно - турбулентного перехода (ЛТП). Обычно анализ ЛТП основан на рассмотрении статистических параметров - корреляций пульсаций скорости, энергетических спектров и т.д. Анализ ЛТП как нелинейной динамической системы в дополнение к уже указанному анализу основан на анализе собственных значений якобиана, вида аттракторов систем в фазовом пространстве и собственных значений матрицы монотонности. В результате строятся бифуркационные сценарии и диаграммы. Это дает возможность проследить механизм усложнения для рассматриваемых задач при ЛТП при изменении выбранных параметров (чисел Рейнольдса, Маха, Фруда и т.д.). Рассмотрение процесса ЛТП с точки зрения нелинейных динамических систем накладывает требования точности и быстродействия используемых алгоритмов решения задач. Начиная с 2008 года, в наших работах используются GPU и multiGPU архитектуры совместно с CPU. За это время было рассмотрено восемь постановок задач ЛТП. Для численного моделирования применялись различные методы высокого порядка. В данной работе для каждого класса методов рассматриваются характерные вычислительные операции, приводятся использованные библиотеки и выполняется сравнение эффективности разработанных алгоритмов и примененных библиотек с CPU-версиями кода а также между собой. Показано, что в среднем на один GPU по сравнению с CPU ускорение варьируется от 5 до 35 раз. В связи со сложностью алгоритмов, как при MPI CPU, так и при multiGPU подходе, ускорение редко бывает линейным и пропорционально степенной функции с показателем 0.78-0.81. Для multiGPU анализа алгоритмы тестировались на пять GPU. Показаны результаты при гибридном применении CPU+multiGPU для одной из задач.

Ключевые слова: multiGPU, гибридная архитектура GPU и CPU, Прямое численное моделирование, Ламинарно - турбулентный переход, Динамические системы, численные методы высокого порядка.

1. Введение

Работа является объединением результатов ускорений параллельных вычислений для задач Ламинарно-Турбулентного Перехода (ЛТП). Во многих статьях [6-22] не имелась возможность уделить данному вопросу достаточно внимания. В работе рассматриваются различные методы численного моделирования и особенности, возникающие при этом с точки зрения технологии реализации и программирования параллельных вычислений. Основное внимание уделяется multiGPU подходу.

В общем случае анализ ЛТП основан на построении аттракторов численных аппроксимаций НКЗ и их сечений сечениями Пуанкаре. Пусть M – фазовое пространство, а φ – функция действия фазового потока изучаемой динамической системы на фазовое пространство, порождающая траекторию $\gamma \in M \times \mathbb{R}$. Сечение Пуанкаре S – сечение M в некоторой

*Работа поддержана Российским Фондом Фундаментальных Исследований (грант 14-07-00123).

точке $m \in \mathbb{M}$ такое, что \mathbf{S} трансверсально траектории γ . Таким образом, изучается множество точек пересечения $A = \gamma \cap \mathbf{S}$. Технически для реализации данного метода расчетные данные записываются из некоторой точки физического или функционального пространства за весь период расчета. После чего по имеющимся данным строится сечение Пуанкаре и множество точек пересечения наносится на плоскость. Более детально, с.м. [6].

Работа разделена на семь частей. В первой части в общих терминах ставятся задачи, рассмотренные в различных работах. Далее излагаются архитектуры, на которых производятся исследования. В последующих четырех частях рассматриваются элементы алгоритмов и полные численные решения различными методами на multiGPU архитектурах. В последней части делается общий вывод по применимости multiGPU и гибридных алгоритмов в задачах ЛТП.

2. Исходные начально-краевые задачи

Рассмотрим задачи ЛТП, которые решались с применением (multi)GPU и гибридных архитектур. В общем случае Начально-Краевые Задачи (НКЗ) можно разделить на задачи для несжимаемых и сжимаемых течений. Принципиальное различие заключается в процедуре расчета при численном моделировании. В первом случае требуется либо проекция в пространство соленоидальных вектор-функций, либо разложение решения на потенциальное и соленоидальное. Во втором случае требуется решение задачи распада произвольного разрыва на элементах конечномерного приближения [1]. Для экономии места сформулируем типичную НКЗ для данных случаев.

Пусть задана область $\Omega \subset \mathbb{R}^k, k = 2, 3$, краевые и начальные условия. Для каждого значения бифуркационного параметра (например числа Рейнольдса - R) из заданного отрезка найти $\mathbf{u} : \Omega \times (0, T] \rightarrow \mathbb{R}^k, p : \Omega \times (0, T] \rightarrow \mathbb{R}$ такие, что на Ω выполняется система уравнений Навье-Стокса для вязкой несжимаемой жидкости:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \frac{1}{R} \Delta \mathbf{u} + \mathbf{f} \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (2)$$

Задача для течения вязкого газа записывается как:

Задана $\Omega \subset \mathbb{R}^3$, краевые и начальные условия. Определим скалярные функции f как $f : \Omega \times (0, T] \rightarrow \mathbb{R}$ и вектор-функции \mathbf{f} как $\mathbf{f} : \Omega \times (0, T] \rightarrow \mathbb{R}^3$. Тогда для каждого значения бифуркационного параметра из заданного отрезка найти такие \mathbf{u}, ρ, E , что выполняется система уравнений:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j} [\rho u_j] &= 0; \frac{\partial}{\partial t} (\rho u_i) + \frac{\partial}{\partial x_j} [\rho u_i u_j + p \delta_{ij} - \tau_{ji}] = g_i, \quad i = 1, 2, 3; \\ \frac{\partial}{\partial t} (\rho E) + \frac{\partial}{\partial x_j} [\rho u_j E + u_j p - u_i \tau_{ij}] &= 0; E = \frac{1}{2} \rho \mathbf{u}^2 + \rho e; \\ p &= (\gamma - 1)(E - 1/2 \rho \mathbf{u}^2); \tau_{ij} = 2\nu S_{ij}; \\ S_{ij} &= \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \frac{1}{3} \delta_{ij} \frac{\partial u_k}{\partial x_k}. \end{aligned} \quad (3)$$

Здесь: E - скалярная функция полной энергии газа; e - скалярная функция внутренней энергии газа; γ - показатель адиабаты газа (1.4 для воздуха); p - скалярная функция давления; \mathbf{u} - вектор-функция скорости газа; ρ - скалярная функция плотности газа; \mathbf{g} - внешняя сила; ν - динамическая вязкость газа. Подразумевается суммирование по одинаковым индексам.

Размерность конечномерного аппроксимирующего пространства выбирается из оценки на размерность аттрактора для каждой задачи при верхнем значении бифуркационного параметра, соответствующему наибольшей оценке размерности, с.м. [6, 22]. В каждом подразделе приведен максимальный масштаб дискретизации. В большинстве случаев расчеты

проводились на более грубых аппроксимациях, когда предварительные результаты совпадали. Это также дало возможность провести сопоставление эффективности вычислительных методов при разных размерах задач.

2.1. Задача течения вязкой несжимаемой жидкости с периодической возмущающей силой (задача А.Н.Колмогорова) на \mathbb{T}^k , $k = 2, 3$.

Данные задачи рассматривались в работах [20, 21] а также на некоторых конференциях и в докладах. Особенности НКЗ: область $\Omega = \mathbb{T}^k(\alpha)$, $0 < \alpha \leq 1$, α - параметр удлинения тора, для $k = 3$: $\Omega = [0, 2\pi/\alpha] \times [0, 2\pi] \times [0, 2\pi]$, $R = \frac{\sqrt{A}}{\nu} \left(\frac{L_y L_z}{4\pi^2} \right)^{\frac{3}{4}}$, $\mathbf{f} = (\sin(ny) \sin(mz); 0; 0)^T$; для $k = 2$: $\Omega = [0, 2\pi/\alpha] \times [0, 2\pi]$, $R = \frac{\sqrt{A}}{\nu} \left(\frac{L_y}{2\pi} \right)^{\frac{3}{2}}$, $\mathbf{f} = (\sin(ny); 0)^T$; m, n - параметры силового воздействия; A - амплитуда силового воздействия; L_j - размеры физического пространства; ν - кинематическая вязкость жидкости.

В связи с периодичностью задачи, для решения (1) применяются два метода - метод Галеркина и спектральный метод. Для двухмерной задачи решение ищется в терминах $\omega - \Psi$ ("завихренность - функция тока", более детально с.м. [20]) и приводится к конечномерной системе ОДУ методом Бубнова-Галеркина. В качестве конечномерного пространства выбрана линейная оболочка, натянутая на пространство специальных сплайн-функций (S-сплайны, с.м. [4]). В результате получается система ОДУ, для решения которой применяется n - стадийный метод Рунге-Кутты. Для трехмерной задачи применяется метод Петрова-Галеркина и спектральный метод с применением Дискретного Преобразования Фурье (ДПФ, более детально с.м. [21]). Для метода Петрова-Галеркина используются пробные функции из пространства S-сплайн-функций, а поверочные функции - полиномы Якоби.

Во всех системах тензор валентности три (свертка за счет нелинейности) заменяется на операцию произведения в физическом пространстве для ускорения расчета. Данная операция выполняется с применением либо квадратур наивысшей алгебраической точности (для методов Галеркина), либо с применением быстрого преобразования Фурье. Размерность пространств - для спектрального метода: $512 \times 128 = 65536$ для 2D задачи и $512 \times 128 \times 128 = 8388608$ для 3D задачи (с учетом действительности функций в спектральном методе) и для метода Галеркина: $240 \times 60 = 14400$ для 2D задачи и $40 \times 40 \times 81 = 129600$ для 3D задачи.

Основными операциями для данных задач являются применение матрицы к вектору и дискретное преобразование Фурье.

2.2. Задача течения вязкой жидкости в канале с расширением.

Здесь рассматриваются две задачи - 3D задача течения с уступа и 2D задача течения проводящей жидкости в канале с расширением. Область $\Omega = \Omega_1 \setminus (\Omega_2 \cap \Omega_3)$, $\Omega_1 = L \times H(\times W)^*$, $\Omega_2 = l \times h(\times W)^*$, $\Omega_3 = (L - l) \times h$, где $L > l$, $H > h$. Выражение в скобках - (*) присутствует только для 3D задачи (т.е. глубина), а также ставится несимметричное расширение, т.е. $\Omega_3 = \emptyset$. На границах Ω ставятся граничные условия - стенки(условия непротекания и прилипания), периодические, входные и выходные. Более детально постановки задач описаны в работах [8, 16]. В задаче течения проводящей жидкости в канале, НКЗ ставится аналогично, дополняясь системой уравнений Максвелла и соответствующими граничными условиями.

Для задачи течения с уступа применяется метод расщепления с проектором Гельмгольца-Хаджа. Это приводит к необходимости решения уравнения Пуассона на каждом шаге метода интегрирования системы ОДУ (например, для метода Рунге-Кутты k -ого порядка, требуется решать уравнение Пуассона k раз для одного полного шага по времени).

В связи со сложностью геометрии для данных задач использовались гибридные конечно-

объемные и конечно-элементные методы. Для аппроксимации нелинейных частей уравнений применялись схемы WENO 9 и 11 порядков (еще не опубликованы), а для аппроксимации линейных членов - метод конечных элементов с применением полиномов Лагранжа шестого порядка. Размерность конечномерного пространства для 3D задачи течения с уступа составляет $512 \times 256 \times 128 = 16777216$ для периодического по направлению z граничного условия и $512 \times 256 \times 256 = 33554432$ для граничного условия стенки с одной стороны. Для 2D задачи течения проводящей жидкости - $1024 \times 256 = 262144$. *Основные операции - расчет конвективных членов уравнения WENO - схемой, решение уравнения Пуассона и применение матрицы к вектору для МКЭ.*

2.3. Задача конвекции Релея-Бенара.

Классическая задача свободной конвекции рассматривалась в работах [9, 11] для системы приближения Обербека-Буссинеска. Рассмотрены две постановки задачи: $\Omega_1 = h \times \mathbb{T}^2$ в цилиндрической геометрии и $\Omega_2 = L \times H \times W$ - кубическая область, взятая, поскольку в книге [3] имеется экспериментальное исследование начальных бифуркаций. Уравнения (1), (2) дополняются уравнением переноса температуры $T_t + (\mathbf{u}, \nabla)T = \Delta T$, и силовым воздействием $\mathbf{f} = RaPr(T - T_0)(0; 0; -1)^T$. Бифуркационными параметрами являются числа Прандтля - $Pr = \frac{\nu}{\chi}$ и Рэлея - $Ra = \frac{\beta gh^3 \Delta T}{\nu \chi}$, построенные по временному масштабу, связанному с масштабом диффузии. Здесь χ - температуропроводность жидкости; β - коэффициент температурного расширения; h - высота слоя; ΔT - перепад температуры на слое; g - ускорение свободного падения. Для Ω_1 применялся спектральный метод на тригонометрических полиномах с введением фиктивного условия на верхних границах, с.м. [9]. Для избежания свертки применялся переход в физическое пространство с применением ДПФ. Для Ω_2 с граничными условиями на стенке применялся комбинированный конечно-объемный + конечно-элементный метод.

Размерность конечномерного аппроксимирующего пространства для Ω_1 составляет $256 \times 256 \times 256 = 16777216$ элементов, а для Ω_2 - $292 \times 292 \times 292 = 24897088$.

Основные операции - ДПФ, расчет конвективных членов уравнения WENO - схемой, решение уравнения Пуассона и применение матрицы к вектору для МКЭ.

2.4. Конвективная сверхзвуковая неустойчивость течения газа.

Задача ставится как обтекание твердого тела потоком вязкого газа при $M_\infty = 2$ для системы уравнений (3). Особенностью задачи является образование сверхзвуковых и дозвуковых зон как в самой области расчета, так и на границе. Это приводит к необходимости рассмотрения граничных условий неотражающего типа. Бифуркационные параметры не выбираются а рассматривается поведение фазового пространства в зависимости от точки в физической области расчета, т.е. параметром выступает координата конфигурационного пространства. Для аппроксимации вязких членов применяется разрывной метод Галеркина на полиномах Лежандра, для аппроксимации невязких членов при решении задачи Римана - метод WENO 9 и 11 порядков для реконструкции потоков [17]. Размерность задачи $800 \times 265 \times 256 = 52428800$. *Основные операции - решение задачи распада произвольного разрыва, расчет потоков схемами WENO, применение матрицы к вектору для разрывного метода Галеркина.*

2.5. Неустойчивость Кельвина - Гельмгольца течения вязкого газа.

Задача ставится в классическом виде слоя смешения двух разноплотностных течений, двигающихся с различной скоростью. Критерием устойчивости течения, является число Ричардсона, определяющееся как $Ri = \frac{\Delta \rho g h}{\rho_A \Delta U^2}$, где g - ускорение свободного падения,

h - высота слоя смешения, ΔU - разница скоростей сдвигового слоя, ρ_A - средняя плотность жидкости (газа), $\Delta \rho$ - разница плотностей сдвиговых слоев. На основе линейного анализа устойчивости отмечается, что турбулентное перемешивание наступает при $Ri < \frac{1}{4}$. Область Ω - прямоугольная, размером $X = 2.8$, $Y = 0.3$, $Z = 0.7$, сила гравитации направлена по направлению оси $-Y$. Граничные условия: условия симметрии по оси Y при $y = 0$ и $y = 0.3$, условия периодичности по оси Z , условия свободного истечения из области по оси X при $x = 2.8$. При $x = 0.0$ задаются начальное возмущение потока с амплитудой $\delta \leq 1 \cdot 10^{-5}$. В данной задаче также требуется применение неотражающих граничных условий, с.м. [19]. Сила гравитации подбирается так, чтобы обеспечить неустойчивость по критерию Ричардсона. Давление задается так, чтобы максимальное число Маха составляло 0.8. Расчет числа Ричардсона выполняется по указанному выражению. Здесь $h = 0.15$, $\Delta \rho = 0.4$, $\rho_A = 0.8$. Для учета вязкости вводится единственное число Рейнольдса, определяемое как $R = \frac{\Delta U h \rho_A}{\nu}$. Получается двухпараметрическая задача, где бифуркационными параметрами являются Ri и R .

Как и в задаче выше, для аппроксимации вязких членов применяется разрывной метод Галеркина на полиномах Лежандра, для аппроксимации невязких членов при решении задачи Римана - метод WENO 9 и 11 порядков для реконструкции потоков [17]. Размерность конечномерного пространства - $800 \times 350 \times 350 = 98000000$. Аналогично предыдущей задаче, основные операции - решение задачи распада произвольного разрыва, расчет потоков схемами WENO, применение матрицы к вектору для разрывного метода Галеркина.

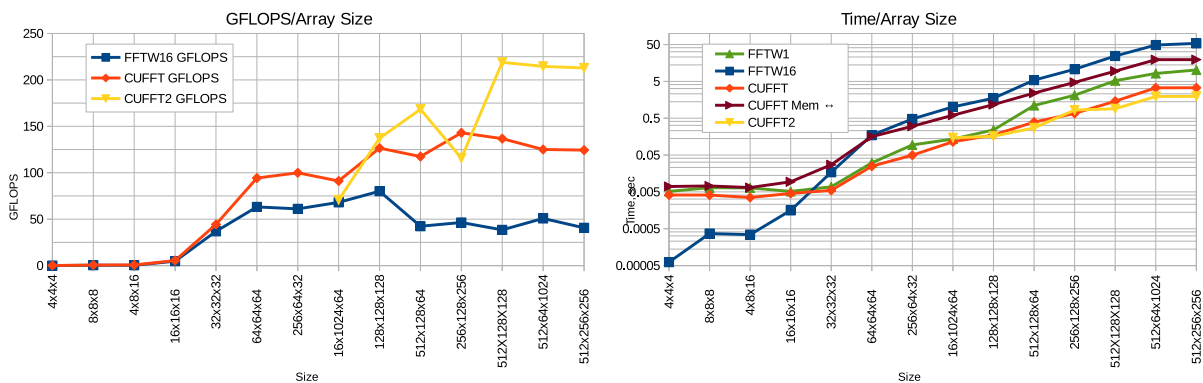
3. Мини multi-GPU кластер

За долгое время проведения работ сменилось много поколений графических сопроцессоров и центральных процессоров. Исторически, все работы выполнены на GPU компании NVIDIA, центральных процессорах компании INTEL. Расчеты в работах с 2007 до 2009 года проводились с использованием двух GTX 8800, с 2009 по 2010 - Tesla C1060, с 2010 по 2013г. - четырех k20, с 2013г по настоящее время - пяти GK110-B1 на базе GeForce. Центральные процессы - Intel Xeon, начиная с L7445, 4 ядра X4, до E5 Ivy Bridge, 16 ядер X2. Для современного сравнения все методы были повторно переметрированы на конфигурацию собранного multi-GPU кластера: Intel Xeon E5 Ivy Bridge 16 ядер, пять GPU NVIDIA GTX TITAN BLACK (GK110-B1), конфигурация собрана с общей памятью (64Gb). Такая конфигурация является относительно дешевой и обладает достаточной вычислительной мощностью для проведения прямого численного моделирования задач ЛТП на начальных стадиях. Все метрики выполнены на данной машине, если не оговорено обратное. Отдельно приводятся данные по производительности в количествах операций с плавающей точкой в секунду (FLOPS). Однако по опыту расчетов, авторы считают оценку FLOPS недостаточной и приводится сопоставление различных методов по времени выполнения. Все вычисления приводятся в арифметике с плавающей точкой с двойной точностью.

4. Дискретное Преобразование Фурье

Для расчета ДПФ используются две реализации алгоритма быстрого преобразования Фурье (БПФ): CUFFT (CUDA v.6.5) и FFTW(3.3.4). Для ускорения расчета CUFFT использовался в режиме двух GPU, FFTW в однопоточном и многопоточном (16 потоков) режиме.

На рис. 1 показано сопоставление времени расчета одного полного преобразования $\mathbb{R}^3 \rightarrow \mathbb{C}^3$ и $\mathbb{C}^3 \rightarrow \mathbb{R}^3$ на реальных данных соответствующих задач при применении одной и двух GPU при размерах области вида $2^p \times 2^q \times 2^n$, $\{p, q, n\} \in \mathbb{N}$. Если сопоставить результаты более раннего теста при версиях CUFFT (CUDA 1.1/2.0) и FFTW(3.2) [5] видно, что библиотека CUFFT улучшена по сравнению с предыдущими версиями. Максимальное



(a) Производительность относительно размера задачи. (b) Время выполнения относительно размера задачи.

Рис. 1. Сопоставление производительности библиотек FFTW и CUFFT на 3D задаче А.Н.Колмогорова.

ускорение составило 17 раз, по сравнению с однопоточным FFTW и в 3.6 раза по сравнению с 16 поточным FFTW. При этом, для задач малой размерности (2D задачи, тестовые 3D задачи), использование CUFFT не является оправданным. Работа при двух GPU для CUFFT оправдана, начиная с размерности задачи 128^3 и выше. При этом для рассмотренных выше задач все результаты были получены на одном GPU ввиду уменьшения конечномерного пространства при сравнении результатов. Если линейные размеры области не являются степенью двойки, то эффективность CUFFT падает, и ускорение по сравнению с 16 поточной FFTW составляет 2.4 раза.

В целом, для спектральных (псевдоспектральных) методов применение CUFFT и CUDA для остальной части кода приводит к ускорению в пять-шесть раз, по сравнению с 16 поточной OpenMP версией кода.

5. Применение матрицы к вектору

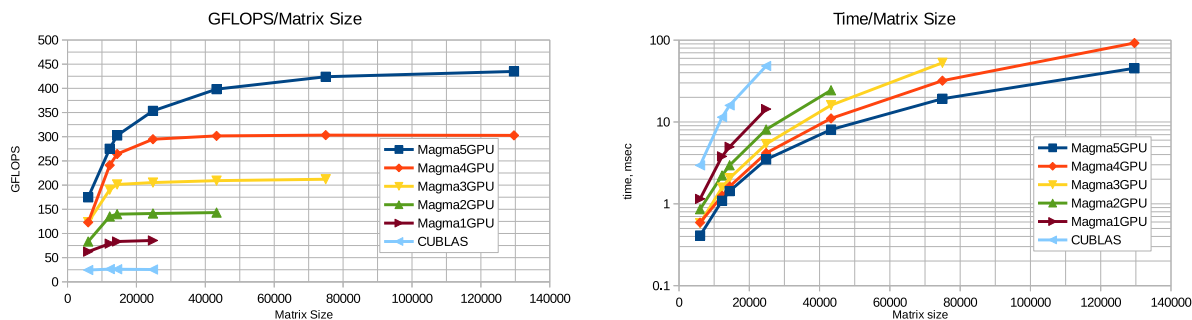
Данная операция применяется, в основном, в методах Галеркина и МКЭ, при этом матрицы могут быть как плотные, так и разреженные (МКЭ). Для плотных матриц применяются библиотеки CUBLAS и MAGMA. Схожие по структуре, они имеют принципиальную разницу, заключающуюся в том, что MAGMA может быть скомпилирована под конкретную архитектуру с автоматическим учетом количества GPU, установленных на одной ноды кластера. Поэтому указанная выше конфигурация мини кластера является достаточно успешной.

На рис. 2 показано сравнение производительности и чистого времени для операций `magmablas_dsymv_mgpu` и `cublasDsymv`.

Таким образом, достигается производительность 85 GFLOPS на одной GPU и 435 GFLOPS на пяти GPU. Как видно, применение MAGMA оправдывается при расчете задач ЛТП и позволяет значительно сократить трудозатраты при написании кода.

6. Схема WENO и обмен данных на multiGPU

Вопрос о коммуникациях для multiGPU не поднимался при использовании библиотек. При написании собственных частей кода данный вопрос возникает. Во всех реализациях используется MPI обмен данными, в котором выделяется буфер в оперативной памяти, относящейся к центральному процессору. Альтернативой такому подходу является DirectGPU, поддерживаемый CUDA начиная с версии 5.5. Авторы остановились на применении перво-



(а) Производительность относительно размера задачи. (б) Время одного шага по времени относительно размера задачи.

Рис. 2. Производительность $y = \alpha Ax + \beta u$ для MAGMA в режиме Multi-GPU и CUBLAS.

го подхода, поскольку он может быть напрямую применен при переносе кода с одной ноды на несколько. Минусами такого подхода является загрузка центрального процессора при передаче данных и большее время, затрачиваемое на перенос данных. На практике в вышеуказанных задачах затраты на передачу данных составляли 0.3-0.5% от всех затрат (по purgof). При применении неявных методов затраты на пересылку возрастают и составляют от 1.5-3%. Снижения затрат можно добиться асинхронным вызовом пересылки данных.

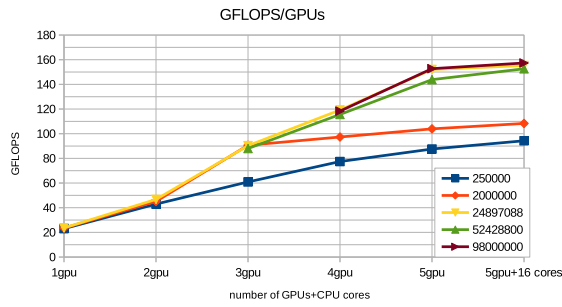
Отдельно кратко следует упомянуть достаточно сложную процедуру вычисления схемы WENO. Поскольку шаблон схемы является достаточно громоздким (17×3 точек шаблона для WENO-9 и 21×3 точек шаблона для WENO-11) вычисление весов полиномов Лагранжа выполнялось в четырех (шести для WENO-11) отдельных ядрах. Сетка красилась в "цвета" и каждое ядро работало на своем цвете, таким образом минимизируется coalescing. После сборки весов, результаты объединялись в промежуточных ядрах, после чего выполнялась окончательная сборка TVB схемы с использованием весов шаблонов. Несмотря на сложность, такой метод оказался в 4.5 раза быстрее, чем наивный метод.

Реализация multiGPU+CPU была выполнена на задаче о неустойчивости Кельвина - Гельмгольца. Суть в том, что рассмотрение характеристических граничных условий вместе с буферным слоем приводит к большому использованию регистров для GPU, что значительно замедляет выполнение соответствующих ядер. Поэтому расчет характеристических граничных условий выполняется на CPU на основе MPI. Результаты, представленные на рис. 3, для данной конфигурации получены на машине с двумя CPU, 16 ядер на каждый. Один центральный процессор отвечал за коммуникацию между GPU, другой - за расчет граничных условий. В конфигурации с одним CPU прироста производительности не было обнаружено.

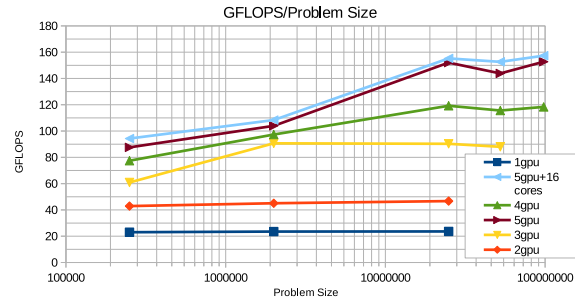
Характерный локальный минимум на рис. 3b как раз связан с выходными граничными условиями, на которых возникает трансзвуковой режим течения. При применении CPU+multiGPU данный минимум уменьшается.

7. Уравнение Пуассона, решение СЛАУ

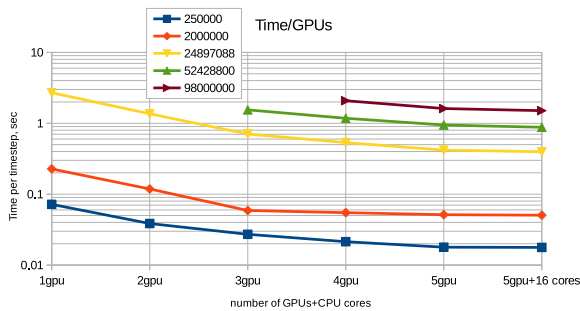
Для решения уравнения Пуассона применяется многосеточный геометрический метод (МГМ), с.м. [10]. Он же применяется при необходимости решения СЛАУ. Такая необходимость возникает при использовании неявных методов цель которых - выйти на квази-стационарное решение перед применением высокоточных явных методов (ограничение по устойчивости при прямом численном моделировании задач ЛТП менее строгое, чем ограничение по точности). В отличие от работы [10], на верхнем уровне МГМ применяются различные методы Крыловского типа. Из практики использования наиболее удачным ока-



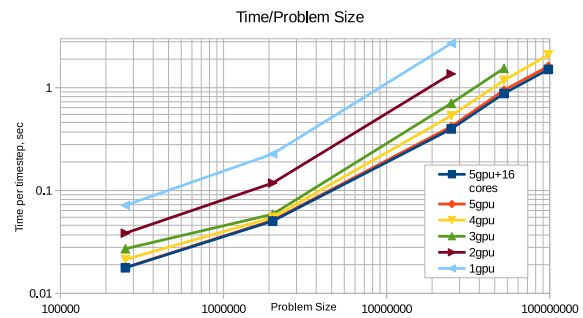
(а) Производительность всего кода относительно количества GPU (для различных размеров задач).



(б) Производительность всего кода относительно размера задачи.



(с) Время одного шага по времени относительно количества GPU (для различных размеров задач).



(д) Время одного шага по времени относительно размера задачи.

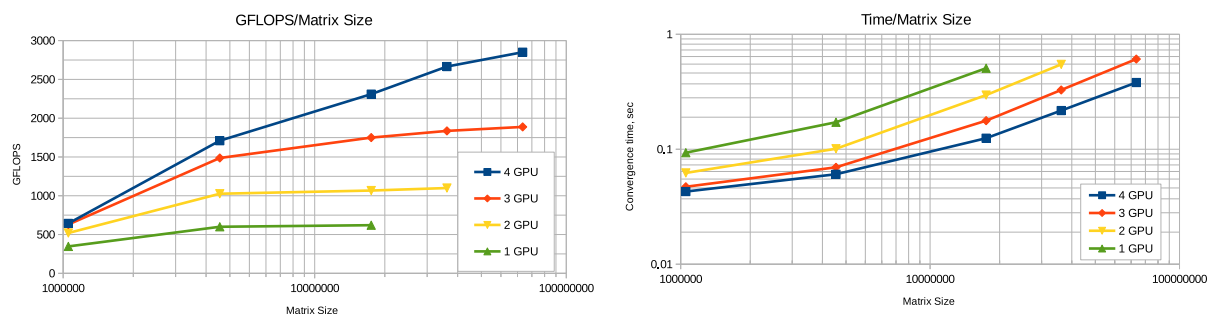
Рис. 3. Производительность всего кода для моделирования течения вязкого газа.

зался метод *BiCGstab*. При условии использования пролонгаторов и рестрикторов высокого порядка (с 4 до 8) [10] удается добиться асимптотики $\mathcal{O}(N_x N_y N_z)$ для времени сходимости при решении уравнения Пуассона и $\mathcal{O}((N_x N_y N_z)^{1.15-1.2})$ при применении неявных методов и решении неявной части для разрывных методов Галеркина (где необходимо). На рис. 4 показана производительность решения уравнения Пуассона для задачи течения с уступа. Оптимальным для данной задачи оказался W -цикл, 5 сеток с изменением размера в 2 раза, пролонгаторы/рестрикторы шестого порядка. При этом, на последних ветках пролонгаторов применяются *BiCGstab* методы в тестовом режиме, для оценки скорости сходимости и следующая пролонгирующая ветвь подправляется с учетом этих результатов. Критерий сходимости на верхнем уровне - по норме C , $\varepsilon \leq 5.0 \cdot 10^{-15}$.

Как видно из рис. 4, на данной задаче удалось добиться практически идеального масштабирования, характерного для многосеточных геометрических методов. Кроме того, производительность одного GPU составила 620.33 GFLOPS, что составляет 32.5% от максимальной теоретической производительности на операциях с плавающей точкой двойной точности.

8. Общий анализ результатов

За длительное время написания кодов для прямого численного моделирования ЛТП, авторы пришли к выводу, что относительно небольшие машины, построенные на архитектуре multiGPU+CPU являются оптимальными при требованиях к сеткам не больше 100 млн. элементов. Причем более низкая производительность, по сравнению с НРС кластерами, может быть компенсирована монопольным доступом к машине. В дальнейшем, после отладки всех узлов на такой машине, задачи с большим количеством элементов могут быть лег-



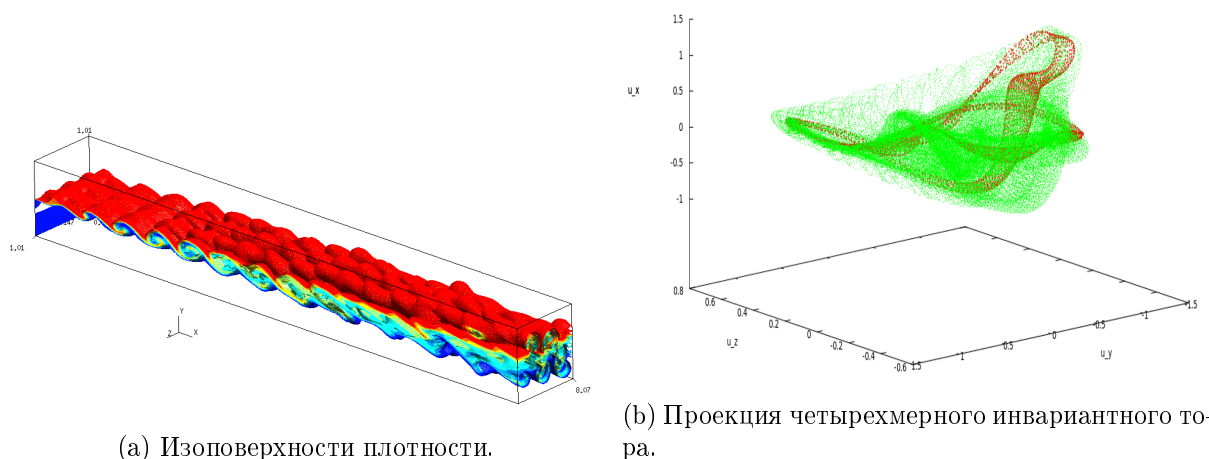
(а) Производительность относительно размера задачи. (б) Время одного шага по времени относительно размера задачи.

Рис. 4. Производительность при решении уравнения Пуассона геометрическим многосеточным методом.

ко перенесены на промышленный кластер. Наиболее затратными для данной архитектуры до сих пор остаются эллиптические задачи, неявные методы и ограничение на количество регистров на ядро.

В целом, использование нами численных методов на данной архитектуре позволило в 5-7 раз ускорить расчеты по сравнению с более ранними конфигурациями, упомянутыми выше. Ускорение, по сравнению с CPU версиями кодов составляет от 5-6 раз при использовании библиотек, до 35, а в отдельных задачах и до 50 раз.

Несмотря на все еще слабую поддержку двойной точности в GPU, производительности уже достаточно чтобы получать очень детальное разрешение течения. Например, режим течения и проекция фазового портрета в \mathbb{R}^3 (и его многомерные сечения) для задачи Кельвина—Гельмгольца показаны на рис. 5. Видно, что точности достаточно чтобы разрешить течение, в котором был обнаружен устойчивый четырехмерный инвариантный тор, при том что в данной задаче наблюдается масштабный эффект мультиустойчивости.



(а) Изоповерхности плотности.

(б) Проекция четырехмерного инвариантного тора.

Рис. 5. Один из режимов при ЛТП в задаче Кельвина—Гельмгольца.

Литература

1. Toro E.F. Riemann Solvers and Numerical Methods for Fluid Dynamics. Springer-Verlag, 1999.

2. Sarrisy I. E., Kassinosz S., Knaepen B., Carati D. Large-eddy simulations of the turbulent Hartmann flow close to the transitional regime. // Center for Turbulence Research: Proceedings of the Summer Program. 2006. P. 387 – 397.
3. Фрик П.Г. Турбулентность: подходы и модели. Москва-Ижевск: Институт компьютерных исследований. 2003. 292 с.
4. Силаев Д.А., Коротаев Д.О. Решение краевых задач с помощью S-сплайна. // Компьютерные исследования и моделирование. 2009. Т.1. №2. С. 161 – 171
5. CUFFT 1.1 / 2.0 vs FFTW 3.1.2 (x86_64) vs FFTW 3.2 (Cell) comparison. URL:
http://www.sharcnet.ca/~merz/CUDA_benchFFT/
6. Evstigneev N.M. and Magnitskii N.A. FSM Scenarios of Laminar-Turbulent Transition in Incompressible Fluids. Chapter 10 in Nonlinearity, Bifurcation and Chaos – Theory and Applications. INTECH. 2013. P. 250 – 280.
7. Евстигнеев Н.М., Магницкий Н.А., Сидоров С.В. Новый подход к объяснению природы турбулентности вязкой несжимаемой жидкости. // Труды ИСА РАН. 2008. Т.33. С. 49 – 65.
8. Евстигнеев Н.М., Магницкий Н.А., Сидоров С.В. О природе турбулентности в задаче движения жидкости за уступом. // Дифференциальные уравнения. 2009. Т.45. №1. С. 69 – 73.
9. Евстигнеев Н.М., Магницкий Н.А., Сидоров С.В. О природе турбулентности в конвекции Релея-Бенара. // Дифференциальные уравнения. 2009. Т.45. №6. С. 909 – 912.
10. Евстигнеев Н.М. Интегрирование уравнения Пуассона с использованием графического процессора технологии CUDA. // Вычислительные методы и программирование. 2009. Т10, С. 268 – 274.
11. Евстигнеев Н.М., Магницкий Н.А. О возможных сценариях перехода к турбулентности в конвекции Рэлея-Бенара. // Доклады Академии Наук. 2010. Т.433. №3. С. 1 – 5
12. Евстигнеев Н.М., Магницкий Н.А. Нелинейная Динамика В Начально-Краевой Задаче Течения Жидкости С Уступа Для Гидродинамического Приближения Уравнений Больцмана. // Дифференциальные Уравнения. 2010. Т. 46. №12. С. 1794 – 1798
13. Evstigneev N.M., Magnitskii N.A., Sidorov S.V. Nonlinear dynamics of laminar-turbulent transition in three dimensional Rayleigh–Benard convection. // Communications in Nonlinear Science and Numerical Simulation. Volume 15. Issue 10. October 2010. P. 2851 – 2859
14. Евстигнеев Н.М. Применение графического процессора для ускорения численного сеточного метода Больцмана с энтропийной стабилизацией // Труды ИСА РАН. 2010. Т.53. №14. С. 111 – 123.
15. Evstigneev N.M. and Magnitskii N.A. Nonlinear Dynamics of Laminar-Turbulent Transition in Back Facing Step Problem for Boltzmann Equations in Hydrodynamic Limit. // AIP. 2010. Volume 1281. P. 896 – 900.
16. Евстигнеев Н. М., Магницкий Н. А., Рябков О. И. Численное исследование перехода к турбулентности в задаче о двумерном течении вязкой сжимаемой проводящей жидкости в канале с симметричным расширением. // Труды ИСА РАН. 2012. Т.62. №1. С. 55 – 62.

17. Евстигнеев Н.М., Магницкий Н.А. Особенности фазового пространства уравнений динамики газа для трансзвуковой начально-краевой задачи // Труды ИСА РАН. 2012. Т.62, №4. С. 85 – 102.
 18. Евстигнеев Н.М., Магницкий Н.А. Нелинейная динамика начальной стадии ламинарно-турбулентного перехода в задаче развития неустойчивости Кельвина—Гельмгольца. // Труды ИСА РАН. 2013. Т.63 №3. с. 45 – 52.
 19. Евстигнеев Н.М., Магницкий Н.А. О развитии неустойчивости Кельвина—Гельмгольца на начальной стадии ламинарно-турбулентного перехода в вязком газе. // Труды ИСА РАН. 2014. Т.64. №3. С. 41 – 52.
 20. Евстигнеев Н.М., Магницкий Н.А., Силаев Д.А. Качественный анализ динамики в задаче А.Н. Колмогорова течения вязкой несжимаемой жидкости. // Дифференциальные Уравнения. 2015. Т.51. №10. С. 1302 – 1314.
 21. Евстигнеев Н.М., Магницкий Н.А. Численное исследование нелинейной динамики в обобщенной задаче течения А.Н.Колмогорова. // Системный анализ и информационные технологии (САИТ'2015) Труды шестой международной конференции (Светлогорск, 10-20 июня, 2015г.) Светлогорск: Издательский центр БФУ им.И.Канта, 2015. С. 49 – 55.
 22. Евстигнеев Н.М. Об одном способе осреднения уравнений динамики сжимаемой и несжимаемой жидкости. // Вестник МГОУ, серия ФМ. 2010. №2. С. 47 – 52.
-

Application of multiGPU+CPU architecture for DNS of laminar - turbulent transition problems which are considered as nonlinear dynamical systems*

Evstigneev N.M., Ryabkov O.I.

Federal Research Center "Computer Science and Control" of Russian Academy of Science, Institute for System Analysis.

In the paper we summarize the use of parallel computational architectures for the direct numerical simulation of laminar - turbulent transition problems (LTTPs). Usually, DNS results are analyzed by a set of statistical parameters, namely, pulsation velocities correlations, energy spectra etc. When a dynamical system analysis approach is applied to DNS results, it is necessary to evaluate additional parameters: phase space attractors and Poincare sections, eigenvalues of Monodromy matrix etc. This allows one to present results as bifurcation scenarios and diagrams and bring up more details concerning LTTP scenario as functions of bifurcation parameters (e.g. Reynolds, Mach, Froude numbers). This task is computationally expensive and algorithms are complex. This brings up more demands on hardware efficiency and software algorithm optimization. We've being using GPU and multiGPU together with CPU architectures since 2008 in this kind of DNS. We've considered eight separate LTTPs since then. Different high order methods were applied. In the paper we show common computational problems for each problem class. We illustrate the application of libraries and stand alone algorithms, perform efficiency benchmarking across GPUs and with CPU versions. It is shown that in general one GPU is 5 to 35 times faster than CPU. The acceleration is worse than linear and scales up as a power function with exponent in 0.78-0.81. This is, presumably, due to MPI communication bandwidth and complexity of algorithms in multiCPU and multiGPU approach. We use five GPUs for multiGPU and show CPU+multiGPU efficiency for one of the considered problems.

Keywords: multiGPU, hybrid GPU and CPU architecture, Direct Numerical Simulation, Laminar - Turbulent Transition, Dynamical Systems, High Order Numerical methods.

References

1. Toro E.F. Riemann Solvers and Numerical Methods for Fluid Dynamics. Springer-Verlag, 1999.
2. Sarrisy I. E., Kassinosz S., Knaepen B. , Carati D. Large-eddy simulations of the turbulent Hartmann flow close to the transitional regime.// Center for Turbulence Research: Proceedings of the Summer Program 2006. P. 387 – 397.
3. Frik P.G. Turbulentnost': podkhody i modeli. [Turbulence: approach and models]. - Moskva-Izhevsk: Institut komp'yuternykh issledovaniy [Moscow-Izhevsk: Institute for Computer Research]. 2003. 292 P.
4. Silaev D.A., Korotaev D.O. Reshenie kraevykh zadach s pomoshch'yu S-splayna [Numerical solution of boundary problems with S-splines].// Komp'yuternye issledovaniya i modelirovanie [Computer Research and Modeling]. 2009. Vol.1. No.2. P. 161 – 171
5. CUFFT 1.1 / 2.0 vs FFTW 3.1.2 (x86_64) vs FFTW 3.2 (Cell) comparison. URL:

http://www.sharcnet.ca/~merz/CUDA_benchFFT/

*Paper is supported by Russian Foundation for Basic Research (RFBR) grant 14-07-00123.

6. Evstigneev N.M. and Magnitskii N.A. FSM Scenarios of Laminar-Turbulent Transition in Incompressible Fluids. - Chapter 10 in Nonlinearity, Bifurcation and Chaos – Theory and Applications. INTECH. 2013. P. 250 – 280.
7. Evstigneev N.M., Magnitskii N.A. and Sidorov S.V. Novyy podkhod k ob"yasneniyu prirody turbulentnosti vyazkoy neszhimaemoy zhidkosti [New approach for turbulence in a problem for incompressible viscous flow]. // Trudy ISA RAN [Proc. ISA RAS]. 2008. Vol.33. P. 49 – 65.
8. Evstigneev N.M., Magnitskii N.A. and Sidorov S.V. On the nature of turbulence in a problem on the motion of a fluid behind a ledge.// Differential Equations. January 2009. Vol.45. Iss.1. P. 68 – 72.
9. Evstigneev N.M., Magnitskii N.A. and Sidorov S.V. On the nature of turbulence in Rayley-Benard convection.// Differential Equations. June 2009. Vol.45. Iss.6. P. 909 – 912.
10. Evstigneev N.M. Numerical integration of Poisson's equation using a graphics processing unit with CUDA-technology. //Vychisl. Metody Programm. 2009. Vol.10. P. 268 – 274.
11. Evstigneev N.M. and Magnitskii N.A. On possible scenarios of the transition to turbulence in Rayleigh-Benard convection.// Doklady Mathematics. August 2010. Vol.82. Iss.1. P. 659 – 662.
12. Evstigneev N.M. and Magnitskii N.A. Nonlinear dynamics in the initial-boundary value problem on the fluid flow from a ledge for the hydrodynamic approximation to the boltzmann equations. // Differential Equations. December 2010. Vol.46. Iss.12. P. 1794 – 1798.
13. Evstigneev N.M., Magnitskii N.A. and Sidorov S.V. Nonlinear dynamics of laminar-turbulent transition in three dimensional Rayleigh–Benard convection. //Communications in Nonlinear Science and Numerical Simulation. October 2010. Vol.15. Iss.10. P. 2851 – 2859.
14. Evstigneev N.M. Primenenie graficheskogo protsessora dlya uskoreniya chislennogo setochnogo metoda Bol'tsmana s entropiynoy stabilizatsiey [Application of GPU for acceleration of Lattice - Boltzmann method with entropy fix].// Trudy ISA RAN [Proc ISA RAS]. 2010. Vol.53, No.14. P. 111 – 123.
15. Evstigneev N.M. and Magnitskii N.A. Nonlinear Dynamics of Laminar-Turbulent Transition in Back Facing Step Problem for Boltzmann Equations in Hydrodynamic Limit.//AIP. 2010. Vol.1281. P. 896 – 900.
16. Evstigneev N.M., Magnitskii N.A. and Ryabkov O.I. Chislennoe issledovanie perekhoda k turbulentnosti v zadache o dvumernom techenii vyazkoy szhimaemoy provodyashchey zhidkosti v kanale s simmetrichnym rasshireniem [Simulation of laminar - turbulent transfer in 2D Hartmann flow problem with symmetrically expanding channel]. // Trudy ISA RAN [Proc ISA RAS]. 2012. Vol.62, No.1. P. 55 – 62.
17. Evstigneev N.M. and Magnitskii N.A. Osobennosti fazovogo prostranstva uravneniy dinamiki gaza dlya tranzvukovoy nachal'no-kraevoy zadachi [On phase space peculiarities of gas dynamics equations for a supersonic initial-boundary value problem]. // Trudy ISA RAN [Proc.ISA RAS]. 2012. Vol.62. No.4, P. 85 – 102.
18. Evstigneev N.M. and Magnitskii N.A. Nelineynaya dinamika nachal'noy stadii laminarno-turbulentnogo perekhoda v zadache razvitiya neustoychivosti Kel'vina—Gel'mgol'tsa [Nonlinear dynamics of initial stage in Kelvin-Helmholtz

- laminar-turbulent instability problem]. // Trudy ISA RAN [Proc ISA RAS]. 2013. Vol.63. No.3. P. 45 – 52.
19. Evstigneev N.M. and Magnitskii N.A. O razvitii neustoychivosti Kel'vina—Gel'mgol'tsa na nachal'noy stadii laminarno-turbulentnogo perekhoda v vyazkom gaze [On the initial stage of Kelvin—Helmholtz instability for laminar-turbulent transition in viscous gas flow]. // Trudy ISA RAN [Proc. ISA RAS]. 2014. Vol.64. No.3. P. 41 – 52.
 20. Evstigneev N.M., Magnitskii N.A. and Silaev D.A. Qualitative Analysis of Dynamics in Kolmogorov's Problem on a Flow of a Viscous Incompressible Fluid. // Differential Equations. October 2015. Vol.51. Iss.10. P. 1292 – 1305.
 21. Evstigneev N.M. and Magnitskii N.A. Chislennoe issledovanie nelineynoy dinamiki v obobshchennoy zadache techeniya A.N.Kolmogorova [Numerical simulation of nonlinear dynamics in generalized Kolmogorov problem for 3D viscous incompressible fluid]. // Sistemnyy analiz i informatsionnye tekhnologii (SAIT'2015) Trudy shestoy mezhdunarodnoy konferentsii (Svetlogorsk, 10-20 iyunya, 2015g.) Svetlogorsk: Izdatel'skiy tsentr BFU im.I.Kanta, 2015. S. 49 – 55. [System Analysis and Informational Technologies. (SAIT'2015) Proceedings of the International Scientific Conference], Svetlogorsk, Published by Immanuel Kant Baltic Federal University. 2015. P. 49 – 55.
 22. Evstigneev N.M. Ob odnom sposobe osredeniya uravneniy dinamiki szhimaemoy i neszhimaemoy zhidkosti [About one averaging method of the equations of a compressible and incompressible fluid]. // Vestnik MGOU, seriya FM [Bulletin MSRU, PhM series]. 2010. No.2, P. 47 – 52.