

Документирование проекта

- Руководство пользователя и разработчика создается с помощью SPHINX

Заключение

Для комплекса НИМФА реализовано программное средство NMC, предназначенное для подготовки концептуальной модели и просмотра результатов расчета в параллельном режиме.

ТЕХНОЛОГИЯ РЕШЕНИЯ ЗАДАЧ ФИЗИКИ ПЛАЗМЫ НА СУПЕР-ЭВМ

А. В. Снытников

Институт вычислительной математики и математической геофизики СО РАН,
г. Новосибирск

Для решения в рамках кинетического подхода актуальных задач физики плазмы, таких как моделирование плазменно-пучкового взаимодействия, моделирование взаимодействия лазерного импульса с плазмой, солнечные вспышки II и III типа и др. была создана суперкомпьютерная технология для решения задач физики плазмы. Основными элементами технологии являются эйлерово-лагранжева декомпозиция расчетной области, реализация наиболее трудоемких расчетных процедур на ускорителях вычислений, параметризованная форма реализации метода частиц в ячейках и методика межархитектурного переноса программ.

Актуальность

Суперкомпьютерное моделирование в последнее время является важным инструментом проведения научных исследований в таких, например, областях науки, как астрофизика, физика плазмы, разработка новых материалов или новых лекарств. СуперЭВМ даже называют «третьей равноправной компонентой научной технологии» наряду с теорией и экспериментом. Приложения математического моделирования на суперЭВМ не ограничиваются собственно научными вопросами, а наоборот, включают в себя множество промышленных и оборонных задач. Это видно по зарубежной статистике использования наиболее крупных вычислительных машин: около половины (46 %) из них работают на промышленных предприятиях.

В последние годы (с 2010 по 2016 год) было опубликовано достаточно много статей (более 300) в различных международных журналах по тематике «эксафлопсные вычисления» (т. е. вычисления на перспективных суперЭВМ производительностью порядка 10^{18} операций с плавающей точкой в секунду), из них около трети посвящены решению различных прикладных задач. Среди множества прикладных задач, решаемых на суперЭВМ, задачи физики плазмы представляют большой интерес как с точки зрения выбора и обоснования модели (равновесная плазма или неравновесная, устойчивая или неустойчивая, магнитогидродинамическое описание или кинетика), так и с точки зрения вычислительных методов (решение уравнения Пуассона или уравнений Максвелла, выбор одного из многих вариантов решения кинетического уравнения или уравнений МГД) и, в особенности, с точки зрения разработки и программной реализации вычислительных алгоритмов (различные варианты декомпозиции расчетной области, организации межпроцессорных обменов, достижения оптимальной производительности).

Все перечисленные вопросы в том или ином виде встречаются и в других областях математического моделирования, тем не менее в физике плазмы есть проблемы, не имеющие аналога по степени сложности (и соответственно, по количеству необходимых для решения вычислительных ресурсов). Эти проблемы возникают при моделировании плазменных неустойчивостей и турбулентностей в установках управляемого термоядерного синтеза (УТС) [1] при высоких температурах (1–5 КэВ, что соответствует 10 млн. градусов), когда плазма не является даже приближенно равновесной в сколь угодно малой окрестности. В этом случае плазма имеет очень большое количество степеней свободы, необходимость учета которых приводит к использованию соответствующих больших объемов памяти. Особенная важность решения задачи о динамике плазмы в установках УТС связана с тем, что управляемый термоядерный синтез является единственным вариантом решения энергетической проблемы. Другая важная область приложения моделирования плазмы – это промышленные плазменные технологии, используемые, в частности, для производства чистого кремния для нужд микроэлектронной промышленности и для производства микропроцессоров (установки молекулярно-лучевой эпитаксии). Также моделирование высокотемпературной плазмы может быть полезно для исследования природы солнечных вспышек и исследования их влияния на магнитосферу Земли и арктическую навигацию.

Эйлерово-лагранжева декомпозиция

Предложен новый эффективный вариант декомпозиции расчетной области – эйлерово-лагранжева декомпозиция, которая одновременно обеспечивает минимизацию фрагмента вычислительного алгоритма и также соответствует условию линейности алгоритма, необходимому для достижения высокой масштабируемости.

Рассмотрены различные варианты декомпозиции расчетной области на примере параллельной реализации метода частиц в ячейках: лагранжева декомпозиция, эйлерова декомпозиция и смешанная эйлерово-лагранжева декомпозиция. Показаны результаты тестов на суперЭВМ и кластерах различной архитектуры и на основе различных процессорных элементов. Наилучшим достигнутым результатом (в рамках данной работы) является эффективность в слабом смысле 92 %, полученная на суперкомпьютере «Ломоносов», НИВЦ МГУ с использованием 500 графических ускорителей Nvidia Tesla, рис. 1. Важно отметить, что этот результат получен для чисто лагранжевой декомпозиции расчетной области, таким образом он показывает нижний предел эффективности распараллеливания для предложенной в данной работе эйлерово-лагранжевой декомпозиции расчетной области.

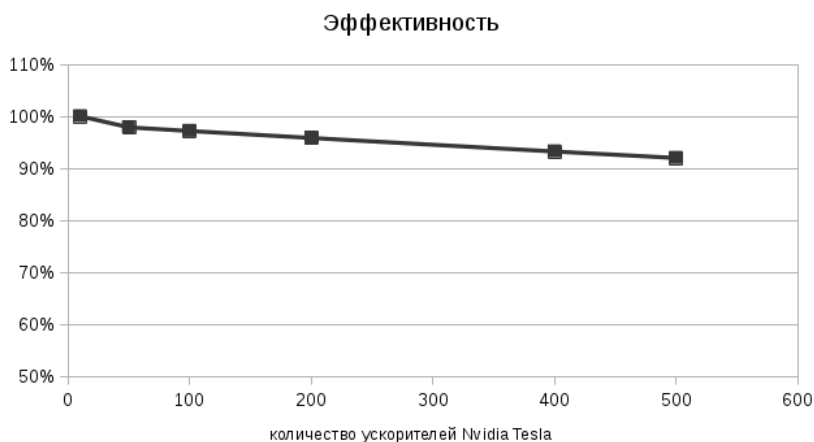


Рис. 1. Эффективность распараллеливания в слабом смысле для **лагранжевой** декомпозиции, расчеты проведены на суперкомпьютере «Ломоносов», НИВЦ МГУ

Для параллельной реализации моделей физики плазмы были использованы следующие виды декомпозиции расчетной области (рис. 2).

- **Эйлерова** декомпозиция: разделение расчетной области (пространства моделирования) между процессорными элементами (ПЭ) на части по координате (ячейки сетки и **находящиеся в них частицы**).

- **Лагранжева** декомпозиция: разделение расчетной области (пространства моделирования) между процессорными элементами (ПЭ) на части по количеству подвижных элементов **частицы распределяются независимо от координаты**.

- Смешанная **эйлерово-лагранжева** декомпозиция [2]: расчетная область разделяется на подобласти для решения уравнений Максвелла, и на каждую подобласть назначается группа из M ПЭ, далее частицы каждой подобласти разделяются дополнительно между этими ПЭ.

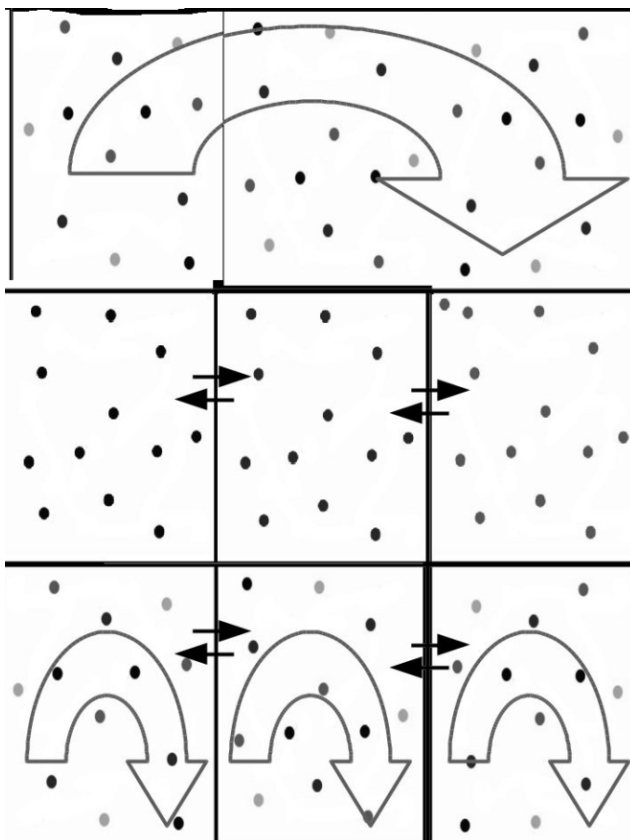


Рис. 2. Схема различных видов декомпозиции. В верхней трети рисунка – лагранжева декомпозиция, в средней части рисунка – эйлерова декомпозиция, и в нижней части – смешанная эйлерово-лагранжева декомпозиция. Коллективные коммуникации (MPI_Allreduce) показаны круговой контурной стрелкой, коммуникации типа «точка-точка» (MPI_Send/MPI_Recv) показаны маленькими черными стрелками.

Частицы, принадлежащие разным процессорам показаны разными оттенками серого

Если использовать графический ускоритель как замену группы процессоров, выполняющих расчет движения модельных частиц в рамках одной подобласти для эйлерово-лагранжевой декомпозиции, то можно, во-первых, значительно ускорить расчет (на 1–2 порядка), во вторых, полностью исключить коммуникации между потоками, связанными с одной подобластью (внутри одного графического ускорителя), что же касается пересылок между подобластями (связанными с разными графическими ускорителями), то они не могут быть более медленными, чем если данные пересылаются между группами процессоров, из-за меньшего количества физических связей в суперкомпьютере.

Адаптация вычислительных методов к архитектуре супер-ЭВМ

Создана параметризованная реализация метода частиц в ячейках. Этот инструмент позволяет с помощью средств языка C++ осуществлять замену численных методов в работающей, отлаженной параллельной программе, в том числе ориентированной на GPU. Это исключительно важно для обеспечения возможности адаптации вычислительных методов к архитектуре суперЭВМ путем замены реализации этих методов на те варианты реализации, которые лучше соответствуют архитектуре. Кроме того, это означает возможность инкорпорирования в код наилучших последних достижений в области метода частиц и возможность смены решаемой физической задачи без переделки кода в целом, путем замены только лишь частей программы, зависящих от задачи, например, граничных условий, что и было продемонстрировано в рамках диссертационной работы, когда программа для моделирования динамики плазмы была перестроена в программу для моделирования встречных пучков путем замены двух процедур. Следует отметить, что переход на гибридные супер-ЭВМ на основе GPU также был выполнен с помощью параметризованной реализации метода частиц. Был создан параметризованный класс «область моделирования» и классы «ячейка» и «частица». Путем наследования были получены классы GPU-ячейка и GPU-область. В класс GPU-ячейка были добавлены процедура перемещения содержимого ячейки на GPU, сравнение данных внутри ячейки на CPU и GPU. Аналогично, в класс GPU-область была добавлена процедура перемещения содержимого области на GPU, сравнение данных на CPU и GPU и вызов процедур (ядер) для расчета поля на GPU.

Были разработаны методы программной оптимизации алгоритмов, пригодные для использования на всех типах суперкомпьютерных архитектур, от процессоров общего назначения до графических ускорителей. Эти методы включают в себя активизацию использования кэш-памяти, внедрение более эффективных методов доступа к памяти, таких как связанные списки или наоборот, в зависимости от архитектуры, статические массивы и др. Предложен ряд оригинальных подходов для повышения скорости расчетов динамики плазмы с применением метода частиц за счет использования графических ускорителей. Подробно рассмотрены различные варианты реализации метода частиц, с использованием текстурной и разделяемой памяти GPU, оптимизация процедуры вычисления токов, а также процедуры переупорядочивания (сортировки) модельных частиц. Движение модельных частиц является наиболее времяемкой частью расчета (до 90 %). В то же время именно эта часть алгоритма наиболее заметно ускоряется при переходе на GPU, как показано в таблице. В перспективе это дает возможность провести трехмерное моделирование динамики высокотемпературной плазмы.

Ускорение для различных GPU по сравнению с расчетом на 4 ядрах процессора Intel Xeon (данные на 2013 год)

Этап вычислительного алгоритма	Ускорение для Nvidia Tesla	Ускорение для Nvidia Kepler
Движение частиц	8,4	41,8
Расчет поля	4,4	25,3

Методика межархитектурного переноса программ

Созданная технология позволяет переносить программу, написанную с помощью технологии CUDA для графических ускорителей фирмы Nvidia, за время, соответствующее времени перекомпиляции на ускорители вычислений Intel Xeon Phi. Это делается с целью обеспечения возможности проведения расчетов на любых доступных архитектурах суперЭВМ, в частности, на наиболее мощных компьютерах из Top10. Для этого программа должна быть написана с использованием специальной разметки, также разработанной в рамках диссертационной работы, а именно: архитектурно-зависимые фрагменты (в частности, ядра CUDA) должны быть оформлены в виде функций со специальной сигнатурой, а коммуникационные процедуры (функции MPI или PVM) должны вызываться через специальную библиотеку оболочек. Рассматривается перенос программы с CUDA на MIC

(не наоборот!!!) и не рассматривается на данный момент вопрос оптимизации под ту или иную архитектуру.

Основные проблемы переноса с архитектуры CUDA на архитектуру MIC заключаются в следующем. Необходимо обеспечить компиляцию ядер CUDA и в особенности вызовов ядер CUDA без компилятора Nvidia. Операции копирования между различными видами памяти в CUDA должны быть пропущены при компиляции на MIC. Необходимо определить типы данных и ключевые слова, входящих в расширение языка C, используемое в CUDA. Для этого вычислительные функции, такие как обработка узлов сетки, модельных частиц, границ расчетной области и пр.

Эти функции оформлены в виде ядер CUDA. Таким образом они не могут быть скомпилированы с помощью компилятора Intel и пр. Основной принцип предлагаемой методики переноса программ: сделать такой участок кода (запуск ядра) по крайней мере **единственным**. Для этого оформляется специальная процедура запуска вычислительных функций. Вычислительная функция передается этой процедуре в качестве параметра, также передается размерность сетки потоковых блоков и потокового блока технологии CUDA. Далее, в случае компиляции компилятором Nvidia, происходит запуск ядра, которому опять же в качестве параметра передается вычислительная функция. Таким образом в программе есть только одно ядро CUDA. Если же компиляция проводится с помощью другого компилятора (напр. Intel), то переданная в качестве параметра процедурная переменная (вычислительная функция) вызывается внутри 6-мерного цикла, количество итераций которого определяется размерностью сетки потоковых блоков и потокового блока, как показано на рис. 3.

```

#ifdef __CUDACC__
    dim3 blocks(grid_size_x,grid_size_y,grid_size_z),
    threads(block_size_x,block_size_y,block_size_z);
    GPU_Universal_Kernel<<<blocks,threads,shmem_size>>>
    (cells,params,h_snf);
#else
#pragma omp parallel for
    for(int i = 0;i < grid_size_x;i++)
    {
        // ....
        h_snf(cells,params,i,j,k,i1,j1,k1);
    }
// ....
}

```

Рис. 3. Универсальная процедура запуска

Примеры решаемых физических задач

Математическая модель высокотемпературной бесстолкновительной плазмы представляется кинетическим уравнением Власова и системой уравнений Максвелла [3]. Для решения этих уравнений использована реализация метода FDTD, описанная в статье [4]. Проведенное моделирование показало количественное соответствие с теоретическими результатами вычисления инкрементов двухпоточковой неустойчивости, как описано в статье [5].

В качестве важного примера использования плазмы в технологических процессах была рассмотрена численная модель тлеющего ВЧ-разряда в силановой плазме на основе метода частиц в ячейках. Она состоит из кинетического уравнения Больцмана, дополненного уравнением Пуассона для потенциала в цилиндрической системе координат. Столкновения моделируются с помощью метода, именуемого в зарубежной литературе null collision technique [6]. Уравнение Пуассона решается с помощью комбинированного метода, включающего в себя преобразование Фурье по угловой координате и по z , а также прогонку по радиальному направлению. Проведено сравнение результатов численных расчетов с экспериментальными данными, показано качественное соответствие.

Работа выполнена при поддержке гранта РФФ 14-11-00485. Разработка программы была поддержана грантами РФФИ 14-07-00241, 15-01-00508 и 16-07-00434.

Литература

1. Бэдсел Ч., Лэнгдон Б. Физика плазмы и математическое моделирование. М: Мир, 1989.
2. Снытников А. В. Моделирование на суперЭВМ аномальной теплопроводности в плазме термоядерной ловушки // Всероссийская суперкомпьютерная конференция «Научный сервис в сети Интернет: масштабируемость, параллельность, эффективность»: сб. трудов. Новороссийск, 21–26 сентября, 2009. С. 87–90.
3. Вшивков В. А., Григорьев Ю. Н., Федорук М. П. Численные методы «частицы-в-ячейках». М.: Наука, 2000.
4. Вшивков В. А., Вшивков К. В., Дудникова Г. И. Алгоритмы решения задачи взаимодействия лазерного импульса с плазмой // Вычислительные технологии. 2001. Т. 6, № 2. С. 47–63.
5. Месяц Е. А., Снытников А. В., Лотов К. В. О выборе числа частиц в методе частиц-в-ячейках для моделирования задач физики плазмы // Вычислительные технологии. 2013. Т. 18, № 6. С. 83–96.
6. Birdsall C. K., Particle-In-Cell Charged-Particle Simulations, Plus Monte Carlo Collisions with Neutral Atoms, PIC-MCC // IEEE Transactions on Plasma Science. 1991. Vol. 19, N 2. P. 65–85.

СШИВАНИЕ МНОГОГРАННЫХ НЕСТРУКТУРИРОВАННЫХ СЕТОК

И. В. Соболев, А. В. Шурыгин

Российский Федеральный Ядерный Центр –
Всероссийский НИИ экспериментальной физики, г. Саров

В мировой практике, при расчете динамики различных конструкций, достаточно часто возникает общеизвестная проблема построения начальной сетки в геометриях сложной формы. Один из возможных подходов к решению данной проблемы заключается в том, чтобы сначала разбить сложную геометрию на более простые фрагменты, построить в них сетку, а затем сшить полученные сетки в единую счетную сетку. Такой подход позволяет строить сетку достаточно хорошего качества в областях сложной формы.

В настоящее время подобный инструмент разработан и реализован в некоторых коммерческих пакетах программ, например, таких как: CD-Adapco Star-CCM+, Ansys Icem CFD и другие, но нам не удалось найти публикаций на эту тему. В то же время решение данной проблемы является достаточно актуальной и практически значимой задачей и для методики ТИМ-3D [1], предназначенной для решения задач механики сплошной среды на неструктурированных многогранных сетках.

В данной работе представлен метод сшивания, основанный на быстром установлении соответствия между двумя многоугольными поверхностями в пространстве и дроблении граней одной из поверхностей проекциями ребер другой поверхности. Этот метод обобщается на произвольное количество тел, кроме того, границы сшиваемых поверхностей этих тел могут полностью или частично не совпадать. Программная реализация и верификация разработанных алгоритмов проведена в рамках методики ТИМ-3D.

Данный метод работает как в ручном режиме, когда пользователь задает необходимые параметры для сшивки, так и в автоматическом режиме.