

УДК 519.688

ОСОБЕННОСТИ ПРОВЕДЕНИЯ ЭКЗАФЛОПС-РАСЧЕТОВ В ФИЗИКЕ ПЛАЗМЫ**В. А. Вшивков¹, А. В. Снытников¹**

Использование экзафлопс-вычислений для задач физики плазмы принципиально необходимо в тех случаях, когда требуется получить высокоточный численный результат для неравновесной плазмы. Предложен подход для достижения экзафлопс-производительности с помощью суперЭВМ гибридной архитектуры при проведении расчетов на основе метода частиц в ячейках для задач физики плазмы. В настоящий момент производительность составляет 0.8 Teraflops на одну карту Tesla на этапе движения модельных частиц, так что оценка производительности суперкомпьютера Tianhe-1A составляет 5.6 Petaflops, что вдвое превышает производительность на тесте пакета LINPACK. Таким образом, при дальнейшем увеличении мощности гибридных суперЭВМ производительность порядка 1 экзафлопс может быть достигнута в приложениях метода частиц в ячейках. Это означает, что расчеты соответствующей размерности могут быть проведены, но вполне вероятно, что существующие системы хранения и пересылки данных не дадут возможности хранить, обрабатывать и визуализировать результаты экзафлопс-вычислений. Приведены оценки размерности данных в экзафлопс-расчетах в задачах физики плазмы, а также мощностей, необходимых для обработки результатов. Работа выполнена при финансовой поддержке РФФИ (коды проектов 09-02-01131 и 11-01-00249), а также интеграционных проектов СО РАН № 103, № 113 и № 26. Статья рекомендована к печати Программным комитетом Международной суперкомпьютерной конференции “Научный сервис в сети Интернет: экзафлопсное будущее” (<http://agora.guru.ru/abrau2011>).

Ключевые слова: метод частиц в ячейках, параллельные вычисления, физика плазмы.

1. Введение. Актуальность данной работы связана с необходимостью точного расчета характеристик неустойчивостей, возбуждаемых релятивистским электронным пучком в субтермоядерной плазме на установке ГОЛ-3 (ИЯФ СО РАН), а также выяснения причин возникновения аномальной электронной теплопроводности [1]. В силу того что плазма в данном случае является существенно неравновесной, использование упрощенных (например, магнитогидродинамических) методов невозможно, расчет необходимо производить в рамках кинетического подхода, т.е. с помощью метода частиц в ячейках, что требует больших вычислительных ресурсов. Конкретной физической задачей в данном случае является расчет спектра плазменной турбулентности. Для ее решения необходимо проведение трехмерных расчетов с высоким разрешением по всем трем координатам в силу того, что плазменные колебания в данном случае являются существенно анизотропными.

Кроме того, в последнее время ведутся интенсивные дискуссии о проблемах создания программного обеспечения для экзафлопс-компьютеров [2]. Применение метода частиц в ячейках принципиально необходимо в тех случаях, когда физическая ситуация является неясной и нет возможности сделать упрощающие предположения, однако этот метод очень ресурсоемкий, что ограничивает его использование. Поэтому возможность реализации метода частиц в ячейках на экзафлопс-компьютерах, с одной стороны, решит проблему низкой скорости работы метода, а с другой стороны, позволит достичь экзафлопс-производительности быстрее, чем на других физических приложениях, таких как гидро- и газодинамика, теория упругости и др. (такое возможно в силу большей локальности данных в методе частиц по сравнению с решением систем линейных алгебраических уравнений, а также методами конечных элементов или конечных объемов).

2. Реализация на GPU метода частиц в ячейках для задачи взаимодействия электронного пучка с плазмой. Численная модель, используемая для решения задачи о релаксации пучка, состоит из уравнений Власова для электронной и ионной компонент плазмы и системы уравнений Максвелла. В данной работе используется алгоритм решения системы уравнений, описанный в работе [3]. Уравнения Власова решаются методом частиц в ячейках (PIC — Particle-In-Cell method). В рамках этого метода

¹ Институт вычислительной математики и математической геофизики СО РАН, просп. акад. М. А. Лаврентьева, 6, 630090, Новосибирск; В. А. Вшивков, зав. лабораторией, e-mail: vsh@ssd.sccc.ru; А. В. Снытников, науч. сотр., e-mail: snytav@ssd.sccc.ru

решаются уравнения движения модельных частиц, которые являются уравнениями характеристик для уравнения Власова. Для решения уравнений движения используется схема с перешагиванием. Для нахождения электрических и магнитных полей используется схема, в которой поля определяются из разностных аналогов законов Фарадея и Ампера. Эта схема имеет второй порядок аппроксимации по пространству и по времени.

Рассмотрим следующую постановку задачи. В начальный момент в трехмерной области решения, которая имеет форму прямоугольного параллелепипеда, находится плазма, состоящая из электронов и ионов. Модельные частицы распределены по области равномерно. Задается плотность плазмы и температура электронов, температура ионов считается нулевой. Дополнительно в области присутствуют электроны пучка, которые также распределены по области равномерно (предполагается, что пучок уже вошел в расчетную область). Электроны пучка отличаются от электронов плазмы тем, что они имеют кинетическую энергию направленного движения 1 МэВ, а их температура равна нулю. Модельные частицы, соответствующие электронам пучка, имеют меньшую массу, нежели модельные частицы, соответствующие электронам плазмы (отношение их масс равно отношению плотности плазмы и плотности пучка). Итак, исходными параметрами задачи являются: плотность и температура электронов плазмы, отношение плотности электронов плазмы к плотности электронов пучка, энергия электронов пучка.

Распараллеливание выполнено методом декомпозиции расчетной области по направлению, перпендикулярному направлению движения электронного пучка. Используется смешанная эйлерово-лагранжева декомпозиция. Сетка, на которой решаются уравнения Максвелла, разделена на одинаковые подобласти по одной из координат. С каждой подобластью связана группа процессоров (в том случае, когда вычисления производятся на многоядерных процессорах, процессором для единообразия будет именоваться отдельное ядро). Далее, модельные частицы каждой из подобластей разделяются между процессорами связанной с этой подобластью группы равномерно, вне зависимости от координаты.

Каждый из процессоров группы решает уравнения Максвелла во всей подобласти. Далее решаются уравнения процессоров группы, производится обмен граничными значениями тока и полей с соседними подобластями, а затем рассылаются полученные граничные значения всем процессорам своей группы.

Актуальность настоящей работы связана с тем, что в экспериментах на многопробочной магнитной ловушке ГОЛ-3 (ИЯФ СО РАН) наблюдается понижение электронной теплопроводности на 2–3 порядка по сравнению с классическим значением. Известно множество работ по моделированию теплопроводности в термоядерных установках. Вычисление температуры в этих работах производится в основном с помощью гидродинамических уравнений. Таким образом, функция распределения электронов по энергиям предполагается максвелловской, что может не соответствовать действительности. Для вычисления неравновесных распределений применяют, в частности, бессеточные модификации метода частиц. Однако для метода частиц в ячейках не известны критерии, позволяющие понять, насколько правильным является полученное в расчетах распределение температуры, а потому корректность моделирования теплопроводности также оказывается под вопросом. Это означает, что удалось на качественном уровне воспроизвести эффект аномального понижения электронной теплопроводности. Однако для более точного моделирования и выяснения причин этого явления необходимо более высокое разрешение и увеличение количества модельных частиц.

3. Оценка потребностей в вычислительных ресурсах. В настоящее время проведены расчеты взаимодействия релятивистского электронного пучка с плазмой, позволившие в квазиодномерном случае точно рассчитать инкремент двухпоточковой неустойчивости: получено 0.081, точное значение 0.077 [4]. Однако для этого пришлось значительно увеличить число модельных частиц, а именно до 1000 в одной ячейке. Число узлов сетки $120 \times 4 \times 4$, размер области 1.2 (в безразмерных единицах), при этом величина дебаевского радиуса равна 0.0089 в тех же единицах. Таким образом, длина области в дебаевских радиусах составляет 134.8.

Далее для оценки полного размера задачи могут быть высказаны следующие соображения.

1. Для адекватного моделирования плазменных неустойчивостей необходимо иметь не менее 16 узлов сетки на длине дебаевского радиуса.

2. Число модельных частиц не может быть уменьшено, иначе полученные результаты будут носить лишь приближенный характер в известных физических случаях, в новых же физических случаях вовсе ничего невозможно будет получить.

3. Задача является существенно трехмерной. В силу того что возбуждаемые пучком плазменные волны распространяются под углом к направлению движения пучка, задача не может быть сведена к двумерной или одномерной без потери существенной части физического смысла. Это означает, что число узлов сетки в направлениях, ортогональных направлению движения пучка (Y и Z), должно быть

сравнимым с числом узлов по направлению X .

Таким образом, получаем следующую оценку размера сетки: $2156 \times 2156 \times 2156$ при 1000 модельных частиц каждого типа в ячейке. Это означает, что требуемый объем памяти составляет 1.4 Петабайт, а вычислительная нагрузка становится порядка 1.5 Petaflops (около 50 операций на каждую частицу) в течение одного временного шага. Следовательно, можно сказать, что для решения рассматриваемой задачи во всей полноте, а именно для высокоточного расчета спектра плазменной турбулентности необходимо применение машин с производительностью более 1 Petaflops, т.е. перспективных экзафлопс-компьютеров. Может показаться, что применения современных суперЭВМ петафлопного класса будет достаточно, но, во-первых, приведенные выше цифры оценочные, во-вторых, они показывают минимально необходимый объем вычислений (134 дебаевских радиуса, 16 ячеек на дебаевский радиус, 1000 частиц), в-третьих, 1.5 Petaflops — это на один временной шаг, которых требуется десятки тысяч [5, 6].

В настоящий момент время пересылок не превышает 4–5 % от общего времени работы программы на сетках размерностью до 500 по каждому координатному направлению и до 1000 ядер. Измерения времени проведены с помощью Intel Trace Analyzer & Collector на кластере НКС-30Т, ИВМиМГ СО РАН.

4. Перспективы достижения экзафлопс-производительности для метода частиц в ячейках на GPU. Используемая в настоящий момент одномерная декомпозиция области не может обеспечить достаточную масштабируемость для расчетов на десятках тысяч и более ядер. Невозможно это в силу того, что в такой декомпозиции минимальным фрагментом является слой сетки и принадлежащие ему частицы. Это достаточно много (для приведенной выше оценки размера задачи — 660 Гб). Такой фрагмент очевидно не поместится в память одного процессорного элемента, а при обработке его группой процессоров будут сильно перегружены коммуникации. Расчет динамики частиц, выполненный на графической карте GeForce 9400 на 16 ядрах, показал значительное повышение производительности по сравнению с расчетом на процессорах Xeon или Nehalem. Необходимо отметить, что время счета включает в себя также время обращения к памяти и, таким образом, сильно различается на разных системах (аналогично, преимущество графических ускорителей достигается за счет быстрой локальной памяти).

Выполнены расчеты движения модельных частиц на графическом ускорителе Tesla C2050 (супер-ЭВМ “ГраФИТ!”, НИВЦ МГУ). Время счета движения частиц на одном временном шаге составило 0.3 миллисекунды для одного миллиона частиц с двойной точностью. Так как для каждой частицы выполняется приблизительно 250 операций, то производительность одной карты Tesla может быть оценена как 833 Гигафлопс (0.8 Терафлопс), притом что для перспективной карты Tesla C2090 планируется производительность порядка 600 Гигафлопс на тесте пакета LINPACK (большая производительность объясняется большей локальностью данных: модельные частицы считаются независимо друг от друга). Для оптимизации доступа к массиву атрибутов частиц (координат и импульсов) использованы текстуры CUDA. Такая производительность означает возможность расчета движения 1 млрд модельных частиц на 25 картах Tesla за 0.12 секунды (исходя из того, что оперативная память Tesla составляет 2–3 Гб в зависимости от модели и для размещения 1 млн частиц нужно 48 Мб; таким образом, на одной карте Tesla можно разместить 40 млн модельных частиц).

Таким образом, если использовать графический ускоритель как замену группы процессоров, используемых для расчета движения модельных частиц в рамках одной подобласти, то можно, во-первых, значительно ускорить расчет, во-вторых, тем самым полностью исключаются 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 сейчас, когда данные пересылаются между группами процессоров, из-за меньшего количества физических связей в суперкомпьютере. Рассмотрим пересылку частиц между двумя соседними подобластями. Если сейчас с одной подобластью связано, например, 8 процессов (каждый занимает одно процессорное ядро), а пересылки частиц из одной подобласти в другую выполняются только между двумя выделенными процессами (условно “головными” процессами), то необходимо выстраивать маршрутизацию для сообщений в системе, физически состоящей из 16 ядер (или, с точки зрения MPI, из 16 потоков). В том случае, когда расчет выполняется на графическом ускорителе (в 8, 100, 200 потоков CUDA), пересылки будут выполняться только между двумя MPI-потоками. Вероятно, такая пересылка займет не больше времени, чем в первом случае.

Для рассмотрения того, как можно достичь экзафлопс-производительности в расчетах на основе метода частиц в ячейках, возьмем за основу характеристики компьютера Tianhe-1A (№ 1 в списке Top-500 на октябрь 2010 г.). Именно этот компьютер рассматривается, во-первых, из-за того, что, как показано выше, графические ускорители дают значительное преимущество в расчетах по методу частиц, во-вторых, в силу того, что в недавнее время компании NVIDIA и CRAY [7] представили проекты экзафлопс-компьютеров

именно на основе гибридной архитектуры, сходной с архитектурой Tianhe-1A [8]. Исходя из полученной выше производительности одной карты Tesla и количества таких устройств в Tianhe-1A (7000) можно получить оценку производительности этого компьютера для приложений метода частиц в ячейках: 5.6 Petaflops. Превышение над производительностью на тесте пакета LINPACK (2.566 Petalops) может объясняться большей локальностью данных в методе частиц в ячейках по сравнению с задачами линейной алгебры.

Ясно, что в таком случае до производительности в 1 Exaflops остается большое расстояние; тем не менее, с помощью использования графических ускорителей такая производительность достигается легче и быстрее, чем при использовании только лишь процессоров общего назначения. Для рассмотренной задачи производительность в 0.8 Teraflops достигается при использовании 150 процессоров Xeon, а 5.6 Petaflops (предположительно) — при использовании 1050 млн процессоров Xeon.

Следует отметить, что высокая производительность на карте Tesla для метода частиц в ячейках достигается в основном вследствие свойств метода частиц (локальность данных); из особенностей физической задачи (невозможность резких изменений плотности на масштабах 10–100 дебаевских радиусов) следует невозможность пересылки большого количества частиц на одном шаге. Поэтому можно ожидать, что эти оценки останутся верными даже и при других характеристиках используемого оборудования.

5. Оценки ресурсов для проведения экзафлопс-расчетов. Результат расчета в задачах физики плазмы (не только в рассмотренной выше) — это, прежде всего, трехмерные распределения плотности частиц, токов, распределения электромагнитного поля. При этом они должны выдаваться в несокращенном варианте. Это означает, что если расчет проводится на сетке $2156 \times 2156 \times 2156$, то и выдача должна быть такого же размера. Это связано с тем, что для сравнения численного результата с известными физическими закономерностями необходимо вычислить фурье-образ рассматриваемой величины (теоретический анализ в физике плазмы, как правило, выполняется в фурье-пространстве). Если выдача выполняется с сокращением размера, то и погрешность анализа, соответственно, будет больше. Минимально необходимо выдавать плотность тока электронов и трех компонент электрического поля. Если они выдаются в двоичном формате, то размер одной такой выдачи составит 60 Гб. Однако это относится только к одному моменту времени, тогда как требуется от 100 до 300 моментов времени с выдачей в течение одного расчета, т.е. около 18 Петабайт. Таким образом, возникает вопрос о хранении результатов даже одного экзафлопс-расчета. Во всяком случае, в данный момент авторам статьи не известны суперЭВМ с такой или близкой емкостью жесткого диска. Более того, для решения какого-то отдельного вопроса в рамках задачи о релаксации мощного релятивистского пучка в высокотемпературной плазме необходимо несколько (5–10) расчетов, т.е. всего получается около 200 Петабайт. Так как передача такого объема данных из Москвы, например, в Новосибирск не представляется возможной, необходимо обрабатывать результаты на месте, но для этого они должны храниться хотя бы в течение нескольких месяцев.

Обработка результатов заключается в следующем: расчет инкрементов плазменных неустойчивостей (в простейшем варианте — приближение экспонентой временной зависимости для каждой из фурье-гармоник электрического поля, где общее число гармоник равно размерности сетки), визуализация плотностей для избранных моментов времени, статистический анализ (например, по методу главных компонент) для выявления зависимостей между входными и выходными параметрами. В данном случае возникает вопрос об автоматизации отбора данных для просмотра. Несмотря на то что существуют программы, способные отрисовать трехмерное распределение описанного выше размера или, возможно, даже для всего расчета за разумное время, очевидно, что для человека невозможно просмотреть такое количество картинок с высоким разрешением и заметить разницу.

Итак, можно указать следующие требования к жесткому диску экзафлопс-вычислителя:

- объем диска — 200 Петабайт,
- скорость диска — 270 Гбайт/сек. (для обработки указанного массива данных в течение часа).

Сейчас для SSD-дисков скорость чтения — порядка 0.7 Гб/сек. Скорость сетевого соединения составляет 11 Гб/сек. (для передачи этого массива данных по сети в течение суток), тогда как сейчас время передачи 1 Гб данных по внутренней сети ННЦ СО РАН занимает около получаса, т.е. 0.0005 Гб/сек. Видно, что недостаток мощности систем хранения и передачи данных между текущим состоянием и перспективными экзафлопс-компьютерами ненамного меньше, чем по вычислительным мощностям (притом что системам хранения данных традиционно уделяется меньше внимания). Таким образом, даже с учетом перспективных средств хранения и передачи данных проведение экзафлопс-расчетов будет представлять собой серьезную техническую проблему. Однако о ее решении следует задуматься уже сейчас, чтобы мощности экзафлопс-систем действительно могли эффективно использоваться и не возникло проблемы, которая существует сейчас на многих кластерах, когда можно провести большой трехмерный расчет, но

нельзя ни сохранить его результат, ни передать по сети для обработки.

СПИСОК ЛИТЕРАТУРЫ

1. *Astrelin V.T., Burdakov A.V., Postupaev V.V.* Generation of ion-acoustic waves and suppression of heat transport during plasma heating by an electron beam // *Plasma Physics Reports*. 1998. **24**, N 5. 450–462.
2. ExaScale Software Study: Software Challenges in Extreme Scale Systems. DARPA report. September 14, 2009 (<http://www.er.doe.gov/ascr/Research/CS/DARPA>).
3. *Вишневков В.А., Вишневков К.В., Дудникова Г.И.* Алгоритмы решения задачи взаимодействия лазерного импульса с плазмой // *Вычислительные технологии*. 2001. **6**, № 2. 47–63.
4. *Лотов К.В., Терехов А.В., Тимофеев И.В.* О насыщении двухпотоковой неустойчивости электронного пучка в плазме // *Физика плазмы*. 2009. **35**, № 6. 567–574.
5. *Григорьев Ю.Н., Вишневков В.А.* Численные методы “частицы-в-ячейках”. Новосибирск: Наука, 2000.
6. *Бедсел Ч., Лэнгдон Б.* Физика плазмы и математическое моделирование. М: Мир, 1989.
7. *Степаненко А.С.* Оценки ускорения вычислений гибридными системами // Тр. Пятой Международной конференции “Параллельные вычисления и задачи управления”. Москва. 26–28 октября 2010. 29–38.
8. <http://www.top500.org>

Поступила в редакцию
25.11.2011
