

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ. ТЕНДЕНЦИИ РАЗВИТИЯ АРХИТЕКТУРЫ ГЕТЕРОГЕННЫХ СТРУКТУР

*Институт проблем математических машин и систем НАН Украины, Киев, Украина

**МП «ПАК» Института кибернетики имени В.М. Глушкова АН УССР, Киев, Украина

Анотація. Пропонується будувати архітектуру гетерогенної обчислювальної системи (ГВС), призначеної для вирішення (загальних) завдань типу математичної фізики (МФ), за принципом: центральна інтелектуальна машина-спецпроцесор (ЦИМ – СП). Перехід до надання і обробки інформації в СП ГВС у вигляді складних структур даних (ССД) вимагає наявності ЕОМ з мовою високого рівня (типу АНАЛІТИК). ЦІМ забезпечить перехід від аналітичного опису суті обчислювального експерименту до одного з чисельних методів рішення задачі в СП паралельного типу; перетворення і введення ССД; роботу СП у режимі єдиного обчислювального потоку; запуск системи та індикацію результату; корекцію при необхідності та ін.

Ключові слова: гетерогенна обчислювальна система, спецпроцесор, центральна інтелектуальна машина, мова АНАЛІТИК, складні структури даних, паралельна обробка, процесорний елемент, скалярний помножувач, єдиний обчислювальний потік.

Аннотация. Предлагается строить архитектуру гетерогенной вычислительной системы (ГВС), предназначенной для решения (общих) задач типа математической физики (МФ), по принципу: центральная интеллектуальная машина-спецпроцессор (ЦИМ – СП). Переход к представлению и обработке информации в СП ГВС в виде сложных структур данных (ССД) требует наличия ЭВМ с языком высокого уровня (типа АНАЛИТИК). Она обеспечит переход от аналитического описания сути вычислительного эксперимента к одному из численных методов решения задачи в СП паралельного типа; преобразование и ввод ССД; работу СП в режиме единого вычислительного потока; запуск системы и индикацию результата; коррекцию при необходимости и т.д.

Ключевые слова: гетерогенная вычислительная система, спецпроцессор, центральная интеллектуальная машина, язык АНАЛИТИК, сложные структуры данных, параллельная обработка, процессорный элемент, скалярный множитель, единый вычислительный поток.

Abstract. It is proposed to build the architecture of heterogeneous computing systems (HCS) as a system for the solution (common) tasks of mathematical physics type (MP), according to the principle: central intelligence machine-special processor (CIM – SP). The transition to the submission and processing of information in the SP HCS in the form of complex data structures (CDS) requires a computer with a high-level language (such as ANALYST). CIM will provide: the transition from the analytical description of the essence of computer simulation (when setting) to one of numerical methods for solving the problem in the SP parallel type, conversion and CDS input, SP operation in single mode computational flow, system startup and displays the result; correction if necessary, etc.

Keywords: heterogeneous computing system, special processor, central intelligent machine, ANALYST language, complex data structures, parallel processing, processing element, scalar multiplier, single computational flow.

1. Введение

XX-й век в области вычислительной техники (ВТ) прославился созданием электронных цифровых вычислительных машин. От механического арифмометра перешли к электронным устройствам, обрабатывающим логические операции наравне с числовыми арифметическими с высокой скоростью, с электронными запоминающими устройствами большого объема с большой скоростью считывания информации, а также прикладному и системному программному математическому обеспечению.

К прорывным проектам можно отнести:

– в элементной базе – переход от ламповых к транзисторным и далее к интегральным микросхемам сверхвысокой плотности и быстродействия (нанотехнологии);

– в структурной базе – создание не только вычислительных (счетных) машин, но и интеллектуальных машин для инженерных расчетов (МИР), управляющих, а также коммутационных машин, с помощью которых появилась возможность управлять технологическими процессами и перейти к электронной коммутации сообщений и мобильной телефонной связи;

– на рубеже веков на базе электронных интегральных микросхем были созданы микропроцессоры. Они нашли широчайшее применение в быту: мобильная телефония, спутниковое телевидение, кассовые аппараты, робототехника, автоматическое управление автомобилями и др. транспортными средствами, стиральные машины и др. бытовая техника, что помогло улучшить быт людей, повысить производительность труда, научно-технический и интеллектуальный уровень людей, освободить от рутинной работы специалистов, занятых в сфере быта, производства.

Все это сочетается с повышением уровня интеллектуальности и снижением стоимости производства и стоимости производимой продукции.

Технологи совместно со специалистами по ВТ перешли к созданию параллельных структур для решения научно-технических, экономических, логических задач, к цифровой обработке не только сигналов, но и изображений. Этому способствовали успехи математиков, которые повысили интеллектуальный уровень процессора и всей вычислительной системы, обогатившихся различного рода узлами, блоками, запоминающими устройствами, системами связи узлов и устройств как в объеме одного процессора, так и межпроцессорных связей в пределах одной или нескольких вычислительных систем (ВС) с выходом на интернет.

2. Современное состояние средств ВС

Текущий период характеризуется построением и использованием гетерогенных структур на базе графических (GPU) процессорных устройств с кластерными системами (CPU) на базе классических процессоров [1, 2].

По сути, GPU – это спецпроцессор параллельного типа, предназначенный для выполнения матричных и скалярных операций. Причем матричные операции выполняются в режиме SIMD с тенденцией перехода к выполнению скалярных операций. При решении общих задач кластерная структура CPU решает задачи HOST-системы. Она выполняет функции по загрузке – выгрузке исходных данных решаемой задачи, разбивке на части её между процессорами кластерной системы, хранению в общей памяти программы и результатов полученного решения, включая обязанности администратора ВС.

А представление и обработка арифметической информации в существующих процессорах (и системах) построены по принципу последовательного представления и обработки чисел с фиксированной или плавающей точкой с одинарной или удвоенной точностью (разрядностью).

Доработки в языке СИ++, выполненные фирмой nVIDIA, например (и др.), позволили решать общие (научно-технические, экономические и др.) сложные задачи с высоким быстродействием при низкой стоимости затрат.

Однако гетерогенные системы при существенных преимуществах перед отдельным применением кластеров на базе CPU имеют и свои недостатки.

Кроме традиционной последовательной обработки чисел, от чего для мощных ВС пора отходить, основным недостатком является избыточность системы. Трудно представить набор задач (при суперсовременной разбивке их на параллельные куски), для которых будет обеспечена 100% загрузка процессоров. Или они должны искать себе работу в то время, когда работают потоковые процессоры (ПП) GPU.

Многоуровневость системы GPU-CPU – загрузку (калькирование данных) в память ПП GPU осуществляют из глобальной памяти CPU. Созданы таким образом 2- и 3-уровневые медленные структуры (память всегда медленнее элементной части ЭВМ).

3. Проблема ввода-вывода

Работает системный загрузчик. Он представляет собой сложную структуру, с которой сталкивается исследователь и с которой следует разбираться каждому.

Языки программирования GPU и CPU для прикладных задач различные, но они должны состыковываться через СИ++, и это вносит диссонанс в общую проблему.

Кроме этого, переход к обработке сплайнов, как сложных структур данных (ССД), вызовет дополнительные трудности на уровне задачи, процедуры, операции. Существующие алгоритмы вычисления многочленов, например, типа схемы Горнера, имеют свою специфику, которую следует учитывать уже на стадии создания структуры спецпроцессора (СП). И они предполагают последовательную обработку числа за числом, выполняя операции типа скалярного умножения. А матричные операции выполняются в режиме SIMD. В силу этого обработку матричного массива проще выполнять в режиме скалярного вычисления отдельных пар сомножителей.

В кластерных системах на базе CPU реализуют алгоритмы, которые распараллеливаются благодаря параллелизму, заложенному в методы (системы линейных алгебраических уравнений (СЛАУ)). Но аналитические описания исследуемых объектов зачастую плохо стыкуются с цифровыми методами, заложенными в CPU и с помощью которых их аппроксимируют. Это создает известные трудности совместного использования их при решении задачи исследователем.

Требования к точности вычисления, необходимость решения объемных краевых, нестационарных и др. задач МФ также предполагают разработку спецпроцессоров параллельного типа.

Архитектура с применением графических устройств (GPU) в сочетании с (уже) традиционными кластерными многоядерными процессорами (CPU) (фирм nVidia CUDA TESLA, AMD) и др. изначально ориентирована на обработку графических задач, а потом уже и информационных фрагментов, свойственных общим задачам. Применение спецпроцессоров GPU является подспорьем для решения общих задач. Оно расширяет возможности спецпроцессора GPU, но возникают вопросы в обоснованности (с точки зрения построения ВС такого назначения) для решения общих задач со свойственной им спецификой.

4. О параллельной обработке задач МФ в рабочих станциях и кластерных системах на базе CPU

Задачи математической физики (МФ) описываются с помощью дифференциальных уравнений в частных производных. Для их решения путем аппроксимации аналитических уравнений дискретными функциями приходят к системам линейных алгебраических уравнений (СЛАУ), решение которых и позволяет пользователю с некоторой погрешностью получить псевдорешение исходной задачи.

Задачам МФ свойственен параллелизм. В узлах сетки, покрывающей область решаемой задачи после ее триангуляции, вычисляются значения узловой функции по однотипному алгоритму. Количество узлов зависит от размеров области для данной задачи и точности, с которой хотят получить псевдорешение.

Для метода конечных разностей (МКР) с его линейной аппроксимацией число точек сетки составляет величину порядка 10^3 и более по каждой пространственной переменной. В спецпроцессоре (СП) узлы связаны в соответствии с принятым для данной задачи вы-

числительным шаблоном. В универсальных ЭВМ эта связь определена индексами коэффициентов решаемой матрицы СЛАУ и для МКР имеет вид разреженной матрицы ленточной структуры с диагональным преобладанием.

Метод конечных элементов (МКЭ) позволяет применить нелинейную аппроксимацию кусочно-гладкими полиномиальными функциями (сплайнами). Размерность сетки (количество узлов в сетке, которой покрывают область решаемой задачи) резко сокращается, а точность псевдорешения повышается. Но вычисления с помощью сплайнов на машинах с последовательной обработкой числа за число весьма трудоемкий и времязатратный процесс. А вычисления их в виде ССД практически невозможны на применяемой структуре процессорного блока.

Применение методов аппроксимации уравнений типа МКЭ [3, 4] позволяет переходить к обработке полиномов.

Использование сплайнов, с помощью которых описывается исследуемый объект, разбитый сеткой на набор конечных элементов, обеспечивает высокую точность аппроксимации при резком сокращении числа точек триангулируемого объекта (исследуемой области) при сокращении времени и стоимости решения в целом.

Однако вычисление полиномов на ЭВМ классической структуры, обрабатывающих их на уровне процедуры, представляет собой длительный процесс с многократным (как правило) обращением к памяти (ЗУ) машины.

Запись сплайнов в виде регулярного матричного представления (РМП) [4–9] позволяет перейти к обработке их (по времени выполнения) к времени выполнения одной арифметической операции. Такой подход поднимает уровень обработки информации в процессоре на более высокий.

Применение РМП с обработкой полиномов с помощью СП в виде ССД на одном или нескольких этапах всей цепи подготовки, вычисления коэффициентов матрицы, решения СЛАУ и вывода полученного решения обеспечивает большой эффект. Но не полный. Повысить его можно, если обеспечить комплексное решение проблемы обработки информации в проблемно-ориентированном спецпроцессоре (ПОС) от аналитического описания исследуемого объекта, начиная с его постановки, до получения решения с его визуализацией и представлением в виде твердой копии. Следует автоматизировать все этапы от аналитического описания проблемы до записи в виде сплайнов с последующим решением на СП.

Часто такую работу выполняют постановщики задач или программисты (если исследователи её не выполняют).

Впервые задачу повышения интеллектуального уровня обработки информации и сокращения времени решения решили разработчики языка АНАЛИТИК, реализовав его в машине для инженерных расчетов (ЭВМ МИР). Однако ранее проблема обработки информации в виде ССД не стояла так остро, как в настоящее время. Использовались методы, которые традиционно обрабатывали число за число. И такая форма обработки остается в кластерных структурах, построенных с использованием классических одно- и многоядерных процессоров.

5. Возможности построения параллельных ПОС с обработкой ССД

Многие проблемы гетерогенных ВС снимаются при разработке и создании спецпроцессора, например, предложенного для решения задач математической физики (МФ) [4] в режиме единого технологического (вычислительного) потока (ЕТП) с обработкой ССД на параллельной структуре. Каждый процессорный элемент (ПЭ) такого СП строится в виде транспьютероподобного элемента (ТПЭ) со скалярным умножителем (СУ) в составе каждого процессорного элемента (ПЭ).

Преимущество такого СП заключается в следующем: исходные данные задачи, вычисляемые коэффициенты в узлах каждого КЭ, коэффициенты глобальной матрицы жесткости и вектора нагрузок (коэффициенты системы линейных алгебраических элементов (СЛАУ)), решение СЛАУ (получение псевдорешения) и вывод его на индикацию поступают, обрабатываются и хранятся в каждом ПЭ и его памяти (без обращения к глобальной памяти системы и HOST, как в кластерной системе на базе CPU). Практически задача решается от ввода краевых условий и до индикации её решения с помощью параллельной структуры из ПЭ с СУ и памятью, встроенной в каждый ПЭ.

В такой архитектуре HOST-машина необходима для задания границ решаемой области (внешних узлов области – сетки КЭ) и вида краевых условий. В архитектуре ВС исследователь берет на себя функции по преобразованию аналитического уравнения к виду, удобному для записи в виде РМП для обработки в виде ССД, то есть дифференцирование или взятие интеграла (в аналитическом виде) от сложных исходных функций, раскрытие скобок в выражениях, приведение подобных членов и проч. Функции, которые свойственны языку типа АНАЛИТИК.

В архитектуре, предложенной в [4], заменив название (и применение) ТПЭ, например, на одноядерный процессор CPU со скалярным умножителем в его составе, можно переходить к решению задач МФ с помощью МКЭ в параллельном режиме единого вычислительного (технологического) потока с обработкой сплайнов в виде ССД. По сути, ПЭ представляет процессор, в котором арифметическое устройство выполнено в виде СУ. Этим он (в значительной степени) и отличается от стандартных сумматоров. А создание ПЭ с СУ в своём составе при современном развитии нанотехнологии не является проблемой.

Известно, что применение МКЭ с его плотно заполненной матрицей коэффициентов СЛАУ (взамен разреженной матрицы МКР) по порядку снижает размерность решаемых систем (при одинаковой точности псевдорешения).

Обработка ССД в ПЭ с СУ в своем составе (выполнение процедуры перемножения, численного дифференцирования и интегрирования полиномов и др. с распараллеливанием вычислительного процесса на уровне микроопераций) осуществляется за время, соизмеримое со временем выполнения одной арифметической операции типа умножения. Это означает и сокращение времени выполнения всех вычислений.

Вычислительный процесс от ввода краевых условий до получения псевдорешения выполняется с помощью однотипных математических выражений или ЕТП на однотипном оборудовании, покрывающем математический алгоритм.

Это дает основание считать целесообразным и оправданным применение СУ в качестве обрабатывающего устройства ПЭ с использованием основ элементов арифметики, заложенных в него.

6. Что предлагается

Идея создания ЕТП явилась продолжением изложенной в [4] (и ранее) идеологии архитектуры двухуровневой системы решения задач МФ в СП, когда на первом уровне с помощью процессора вычислялись коэффициенты системы СЛАУ и загружались в СП, который решал СЛАУ, на втором уровне.

Следует заметить, что та архитектура, как и последующее ее развитие в виде ЕТП, предвосхищали идею использования специализированных (в данное время графических) устройств в сочетании с многоядерными кластерными системами. Отличие в том, что функции многоядерных процессоров возлагали на ПЭ, включавший транспьютер и СУ под общим названием транспьютероподобный элемент (ТПЭ).

Транспьютер имел ограниченный, но достаточный для решения задач, набор арифметико-логических операций. Достижения технологов, которые могут размещать на одном

кристалле миллиарды вентилях, позволили расширить функциональные возможности транспьютера. Заменяя (в CPU) транспьютеры на выпускаемые в текущее время стандартные одно- и многоядерные процессоры из набора CPU, состыковав в единой системе кластеры GPU и CPU, расширив язык СИ++, разработчики гетерогенной системы смогли решать общие задачи.

Обилие математических функций, алгебраических и арифметических операций над ними требует наличия ЭВМ, которая возьмет на себя функции помощника исследователю (и вычислителю), а также СП, который требует задания исходных данных в виде (конкретных) сложных структур. Эти обязанности можно возложить на исследователя или программиста. Но прогресс требует автоматизации рутинного процесса. Тем более, что основа уже заложена создателями ЭВМ МИР в языке АНАЛИТИК.

Пути решения проблемы могут быть различны. Можно создать одну центральную интеллектуальную машину (ЦИМ), то есть сверхмощную ЭВМ, которая будет преобразовывать (и вычислять) все функции, встречающиеся в практике применения, и представлять их в требуемом исследователю (СП) виде ССД. К такой ЭВМ каждый пользователь может обращаться через интернет (fine reader или с помощью др. вида электронной связи) со своего СП (ВС).

Создание мощной ЦИМ потребует значительных временных затрат и средств, но они окупятся при четкой организации процесса обслуживания пользователей.

Другой путь – создание (или доработка существующих) персональных средств под конкретные классы задач и требуемое представление информации в виде соответствующих ССД.

Удобными и экономически целесообразными могут оказаться спецприставки к СП, выполненные на ПЛИС. С их помощью требуемое аналитическое уравнение подвергнут, например, необходимой обработке в ЦИМ и с помощью системы GRID передадут пользователю в виде ССД, необходимой данному СП.

Создание глобальной ЦИМ кажется предпочтительным. Обилие задач, их аналитических описаний, методов аппроксимации, а также множество разных пользователей конкретного класса уравнений, создают предпосылки к такому решению. Создание ЦИМ для одного пользователя может быть слишком затратным, но сочетание сверхмощной глобальной ЦИМ с персональными средствами может быть оптимальным.

7. Зачем нужна ЦИМ в СП с параллельной обработкой ССД при решении задач МФ в режиме ЕТП

Класс задач МФ необычайно широк. Он охватывает наиболее сложные научно-технические задачи. Описанию их посвящено множество книг, статей, программ. Следуя А.А. Самарскому и А.В. Гулину [12], схема вычислительного эксперимента состоит в:

- формулировке основных законов, управляющих объектом исследования;
- построении математической модели в форме системы уравнений (дифференциальных, интегральных, алгебраических и пр., как правило, нелинейных (с последующей линеаризацией их);
- математической модели ставят в соответствие дискретную модель, которая с помощью численных методов, вычислительных программ на вычислительных средствах может обеспечить решение поставленной задачи.

Проведение каждого этапа ВЭ требует временных затрат и дает свои погрешности (модели, методы аппроксимации, округление при выполнении арифметических операций на вычислительных устройствах с конечной разрядной сеткой).

Для аппроксимации аналитических уравнений задач МФ часто применяют методы конечных разностей (МКР), конечных элементов (МКЭ), граничных элементов (МГЭ). Очень перспективным для широкого класса задач МФ, совмещающим относительно про-

стоту в описании в сочетании с высокой точностью получаемого решения за счет аппроксимации нелинейными (и линейными в том числе) кусочно-гладкими полиномами, является МКЭ. Но относительная простота в описании конечных элементов (КЭ) сопровождается усложнением вычислительных алгоритмов и именно за счет нелинейного вида сплайнов.

Потребность в создании и использовании ЦИМ типа ЭВМ серии МИР с последними версиями языка АНАЛИТИК для решения общих научно-технических задач возникла давно. В нашем случае с помощью ЭВМ, реализующей язык типа АНАЛИТИК, можно задавать требуемую интерпретацию уравнений, описывающих исследуемую математическую модель. Она должна подготовить исходную информацию в виде ССД, необходимых для покрытия дискретной модели (задачи МФ). После этого ВС параллельного типа в режиме ЕТП обработки ССД решает поставленную экспериментатором задачу МФ.

В такой архитектуре ВС ЦИМ может взять на себя функции подготовителя информации; загрузчика исходных данных; планировщика выполняемой задачи. В части программирования она должна задать ВС необходимую структуру данных, которые СП с высокой производительностью обработают в параллельном режиме с помощью ПЭ с СУ в составе каждого из них.

Если, например, дискретная модель описывается дифференциальным уравнением и для аппроксимации его применяется МКЭ, то результатом должен быть набор уравнений сплайнов, с помощью которых (в результате вычислений) в ВС получают псевдорешение поставленной задачи.

ЦИМ возьмёт на себя подготовку всего разнообразия сплайнов, необходимых для них в виде РМП, с учетом вида областей; типа конечных элементов (линейные, треугольные, квадратные, тетраэдральные и проч.); порядка (и вида) интерполяционного полинома (одно-, двух-, трехмерные); дискретной модели, которая строится на множестве кусочно-непрерывных функций, определенных на множестве КЭ; группы элементов (симплекс, комплекс, мультиплекс); вида краевых задач Дирихле, Неймана и др. смешанные задачи; используемого вычислительного метода вариационного или взвешенных невязок (метода Галеркина, моментов, коллокаций, наименьших квадратов), с помощью которых развит МКЭ с множеством его модификаций (спектральный метод или МГЭ) и проч.

В вычислительной триаде [модель – метод (алгоритм) – программа] ЦИМ должна обеспечить стыковку трех этапов ВЭ, направленных на решение поставленной задачи с помощью ВС. Переход от математической модели к численному методу связан с возникновением погрешностей дискретизации (ПД) и округления (ПО).

ПД в основном зависит от вида КЭ и порядка интерполяционного полинома. Это определяет экспериментатор. Задача ЦИМ должна состоять в приведении исходных данных к виду, удобному для РМП в конкретном СП. Для выполнения этой задачи ЦИМ в своей базе должна иметь необходимые сведения или получить их с помощью интернета (и системы GRID) через сеть глобальной паутины и преобразовать к соответствующему виду. От выполнения ею поставленной задачи зависит и проблема сведения к нулю погрешности получаемого псевдорешения, и обеспечивается производительность работы.

Все дискретные модели в итоге сводятся к решению СЛАУ. Конечное решение большого количества алгебраических уравнений зависит от спецпроцессора и архитектуры всей ВС. Это приводит к погрешности округления, которую должна предусмотреть ЦИМ. Большое значение имеет точность представления исходных данных [5] и алгоритмов их обработки в ВС, поскольку возможность переполнения разрядной сетки арифметического устройства (АУ) машины может переводить решение задачи в разряд неустойчивых решений. Кроме того, применение методов взвешенных невязок для решения дифференциальных уравнений с использованием РМП, например, предполагает ситуацию, когда ЦИМ должна «грамотно» выполнить разбивку уравнения на части. Это необходимо, чтобы размерность вектора [7] сочеталась с размерностью (количеством) СУ и размерностью раз-

рядной сетки их в структуре ПЭ, заложенных в СП. Необходимо также поставить в соответствие точность исходных данных с точностью коэффициентов, получаемых в конечных элементах при их вычислении, а затем и в результате решения СЛАУ всей ВС [9].

ЦИМ должна быть помощником ВС в части организации вычислений, интерпретации полученного псевдорешения и, возможно, советчика исследователю по модели в целом, должна обеспечивать реализацию конкретного численного метода с помощью данной ВС. Это в первую очередь связано с конечным числом ПЭ и желаемым количеством КЭ, на которое экспериментатор хочет разбить исследуемую область. Она должна подсказать [6] возможно необходимый переход от линейной (или квадратичной) схемы к квадратичной (или кубичной) соответственно схемам интерполяции конечного элемента. Такой переход обеспечит соответствие количества узлов сетки исследуемой области и количества КЭ в используемом СП. Но при этом ЦИМ должна учесть возможные ограничения по ёмкости памяти в каждом ПЭ и др., поскольку предлагаемая ей дискретная модель ставит в соответствие и количество уравнений, и количество коэффициентов [5] в каждом уравнении СЛАУ, чтобы в целом обеспечить корректность численного метода.

Реализованная в СП однотипность вычислительного процесса (ЕТП) имеет преимущества. Она обеспечивает унификацию логических элементов, узлов и блоков устройства и проч., но создаёт некоторые дополнительные требования. К преимуществам следует отнести общую память исходных данных на всех этапах решения. На этапе задания граничных условий – в памяти ПЭ, которые выполняют функции граничных (внешних узлов решаемой области), хранят коэффициенты полиномов, являющихся исходными данными для вычисления коэффициентов в триангулируемой области решаемой задачи МФ.

Вычисленные в каждом ПЭ_{ij} коэффициенты, которые определяют сплайны в КЭ_{ij} решаемой области, являются исходными данными для глобальной матрицы жесткости и вектора нагрузок для СЛАУ. Коэффициенты СЛАУ являются исходными данными при решении системы (факторизация матрицы и вычисление неизвестных [9]), которые и являются значениями сеточных функций. Следует помнить, что при аппроксимации с помощью МКР вычисленные неизвестные являются лишь численными значениями соответствующих сеточных функций в узлах триангулированной области. А применительно к МКЭ вычисленные значения задают уравнение линии между узлами, то есть, вычисленные значения узловых функций в соседних узлах определяют кусочно-гладкий полином этого участка КЭ_{ij} в пределах области решения.

Рывок, сделанный разработчиками и технологами в создании CPU и GPU, а затем и гетерогенных структур, показал возможность разработки и создания СП для решения класса задач, показав на практике их целесообразность в силу высокой производительности в сочетании с низкой стоимостью получаемого решения. Используя опыт создания GPU, CPU-гетерогенных систем, следует строить новые СП, предназначенные для широкого круга общих (научно-технических) задач, разрабатывать на базе новых знаний и элементной базы новые архитектуры высокопроизводительных гетерогенных систем. Необходимость видна, пути решения – тоже. Следует развивать архитектуры с использованием новых математических технологий: параллелизм как на уровне задачи, так и на уровне процедур в сочетании с обработкой ССД в режиме ЕТП.

Наличие архитектур гетерогенных систем на базе серийно выпускаемых спецпроцессоров GPU nVIDIA TESLA и др. совместно с CPU на базе кластерных многоядерных систем классических процессоров показал и на несоответствие параллельных методов вычисления традиционным последовательным алгоритмам.

ЦИМ как системный процессор организует работу системы по переводу аналитических выражений в выражения, с помощью которых возможна численная реализация их в параллельном устройстве СП виде ССД в режиме ЕТП с помощью РМП с использованием

своих ресурсов и ресурсов системы GRID; система должна готовить СП и в целом ВС к параллельному выполнению всего вычислительного процесса; подготовить СП к выполнению не только арифметических операций в виде ССД к выполнению им алгебраических операций типа перемножения векторов, умножения матрицы на вектор и на матрицу, матричного интегрирования и дифференцирования полиномов, записанных с помощью РМП в виде векторов и матриц, решения СЛАУ и др. (что повышает суммарный эффект производительности ВС).

С помощью локальной сети можно в пределах исследовательского подразделения обеспечить использование мощной ЦИМ для работы каждому постановщику ВЭ одной крупной задачи.

Но если ЦИМ выполнена в виде персональной машины, то она может обеспечивать и высокую мобильность экспериментатору. С ее помощью можно проводить ВЭ на расстоянии, используя интернет и СП стандартного типа. Наличие ЦИМ в ВС позволит замкнуть ВЭ. Наличие в ПЭ исходных данных (уже в виде псевдорешения, полученного после решения СЛАУ) обеспечит исследователю обратную связь с постановкой задачи. Все исходные данные первой итерации ВЭ могут использоваться ЦИМ для коррекции исходных данных во второй, ..., i -й итерации в её постановке.

ЦИМ может использоваться для создания электронной и твердой копий в качестве корректирующего звена ВЭ, когда постановщик задачи вносит коррекцию в эксперимент, не используя математический аппарат. Известно, что аналитическое уравнение не всегда является наглядным, если исследуется сложная область. А с помощью «графического пера» можно изменять контуры изображения и проч.

В глобальной сети может быть задействовано множество ВС. Каждая из них, имея свою ЦИМ, через GRID-систему может общаться с другой ВС, подкачивать новые данные или передавать результаты своей работы другим. Причем совсем не обязательно, чтобы все ЦИМ приводили аналитическое описание задач своего класса к одному информационному типу (вида РМП или в другой форме ССД).

В предлагаемой архитектуре структура СП представляет собой арифметический процессор, обрабатывающий ССД, представленных в регулярной форме в виде матриц, векторов. Обрабатываются они на регулярных ПЭ, не требующих дополнительной подкачки исходных данных из главного ЗУ или из памяти ЦИМ, выполняющей роль HOST-машины. СП выполняет в параллельном режиме скалярные произведения [17], матричные и векторные операции над ССД в виде полиномов (сплайнов), записанных в виде РМП в памяти каждого ПЭ (не требует подкачки). При этом ПЭ работают в синхронном режиме (в пределах данного блока). Поскольку уровень машинного языка повышается до высокого уровня (типа АНАЛИТИК [10, 11, 14, 15]), производительность системы повышается, сокращается количество команд управления за счет их аппаратной реализации.

ЦИМ – машина для решения системных вопросов проблемных процессоров. Цель её – обеспечить функционирование ПЭ на базе СУ в режиме обработки ССД. Кроме преобразования аналитических функций в числовые с реализацией РМП, она должна с помощью системы GRID, например, обеспечивать переход к обработке процедур, как операций. Она является частью вычислительного процесса проблемного СП. После подготовки исходных данных на этапе задания краевых условий следует калькирование их в соответствующие узлы СП. После расчета коэффициентов КЭ следует формирование глобальной матрицы жесткости и вектора нагрузок, далее – решение СЛАУ, а полученное псевдорешение следует отобразить с помощью устройства индикации. И каждую смену этапов в описанном выше вычислительном процессе ЦИМ должна отобразить на связях ПЭ в структуре СП, а при необходимости обеспечить и коррекцию ВЭ.

Следует предположить, что системное МО не будет расти в объеме. Наоборот, оно будет опираться на МО базовой ЦИМ и, следовательно, упрощаться. Рост непрофессио-

нальных пользователей (исследователей, постановщиков эксперимента) для существующих мощных кластерных вычислительных систем приведет к увеличению вспомогательных подсистем. Но в архитектуре вычислительной системы предлагаемого типа ЦИМ с СП облегчит труд исследователя, обеспечит резкий прирост производительности и снижение стоимости получаемого решения.

Как показал опыт, описание с помощью дифференциальных уравнений (в частных производных) разнообразных математических моделей различных процессов не является проблемой. Это лишь дополнительный труд математикам, которые занимаются разработкой структур СП и численных методов, который для данной предметной области может быть разовым (хотя и трудоемким). И в конечном итоге все, как правило, сводится к решению СЛАУ. А решение СЛАУ с помощью СУ на достаточно долгий период можно считать решенным [9] (в части методов и структур для решения задач МФ, например), по крайней мере, в части структуры данных и возможных вычислительных методов. Уровень обработки информации от последовательной (числа за числом) можно переводить к уровню обработки ССД, выполняя (по времени исполнения и количеству команд управления) процедуру перемножения полиномов как обычную операцию умножения 2-х чисел. А если учесть, что выполнение процесса осуществляется параллельно в каждом ПЭ и на множестве ПЭ в режиме ЕТП с понижением количества уровней обрабатываемой части ВС, это даст резкий прирост производительности ВС.

Эффективность проблемных вычислений повысится и за счет управления: исключается набор команд, микропрограмм, объем памяти для их хранения и время выборки, количество регистров КЭШ-памяти и проч., так как задача решается в режиме SIMD для всего СП. А набор устройств, блоков и разновидностей ПЭ в структуре СП сокращается до одного с памятью в каждом ПЭ. По сути, ЦИМ является первым уровнем в СП из ПЭ (как HOST-машина).

8. Структура ЦИМ и возможная реализация ПЭ

Обычно «жесткие» структуры ведут к специализации системы или к потере итоговой скорости (производительности) вычислений и неэффективному использованию оборудования при потоке разнообразных заданий. Часто выход видят в объединении в одной системе подсистем; в приспособлении структуры ЭВМ к структурам решаемых задач.

Последняя – более перспективная идеология, но приспособление – через методы, описывающие исходные задачи, ещё перспективней. Следует разрабатывать новые и приспособлять известные методы, которые на конкретной (ориентированной на реализацию этих методов) параллельной структуре имеют возможность решать различный набор задач с обработкой ССД в режиме ЕТП. Это повысит эффективность всей ВС. Структура может быть перестраиваемой, но она должна обладать и модульностью, и параллелизмом и обрабатывать ССД в ЕТП для набора задач определенного класса и др., где встречаются отдельные их фрагменты.

Структура решаемых задач должна естественным образом отражаться и на ЦИМ, которая с помощью методов, ориентированных на аппаратную реализацию, должна хорошо накладываться на структуру ПЭ, отдельных модулей СП и ВС в целом.

Следует искать единство в организации вычислительных процессов (описаний аналитических и численных) со структурой ВС (СП), их реализующих, чтобы устранить существующий разрыв.

А единству в организации вычислительных процессов должна способствовать и структура ПЭ. Он является частью вычислительной среды постоянной структуры, но с переменными функциями. СУ (в пределах одного или всех ПЭ блока) должен быть с перестраиваемой конфигурацией в рамках проблемной ориентации СП на реализацию известных и предлагаемых методов решения задач (МФ).

Одна из возможных параллельных архитектур создания ВС на ПЭ транспьютерного типа, в основе которых лежит скалярный СУ векторов, была предложена в [4].

Имеются методы умножения [13, 14], которые начинаются с умножения младших разрядов двух n -разрядных чисел. Эти методы дают результат для чисел, представленных в прямом коде на n -разрядной сетке сумматоров за $2n$ – микротакта. А для чисел, представленных в дополнительном коде, но на $2n$ – разрядном регистре сумматоров, дают результат за те же $2n$ – микротактов. Эти методы гарантируют выполнение операции за время, соизмеримое с временем получения одного частичного произведения (ЧП) на сумматорах параллельного типа в традиционных способах умножения чисел с последовательным прогоном переносов из младшего разряда в старший. При $n \geq 64$, когда при выполнении операции умножения требуется сложить порядка 63 ЧП, $i = 1, \dots, n$, это существенно.

В серийно выпускаемых процессорах (в части их АЛУ), которые являются основой кластерных систем CPU, следует предусмотреть возможность работы как со сквозным переносом из младшего разряда в старший на каждом такте суммирования частичных произведений, так и с запоминанием значений результата и переноса на каждом микротакте текущего такта выполняемой арифметической операции. Незначительные функциональные логические элементы, которые потребуются в качестве аппаратурной «добавки» (практически регистр для хранения переносов, который можно использовать в других целях), расширят область применения таких АЛУ. При этом появится возможность использовать стандартный суммирующий блок в качестве АЛУ как в CPU, так и в СУ ПЭ проектируемых ВС. При современных технологических возможностях это реализуемо. Такое решение позволит поднять производительность в одних случаях и расширит область применения в других.

9. Заключение

Создание спецпроцессоров параллельного типа – важный шаг на пути построения цифровых аналогов физической среды.

Потребность в высокопроизводительных гетерогенных ВС, ориентированных на решение общих задач (и, в частности, задач МФ), есть и строить их надо под конкретные классы задач. Но следует идти не путем перестройки спецпроцессоров, разработанных для графических задач (GPU), а с учетом специфики опыта их создания.

Назрела потребность в переходе к обработке ССД, к параллелизму как на уровне решения всей задачи, так и на этапе обработки ССД.

Переход к обработке ССД связан с подготовкой структуры исходных данных, записанных экспериментатором в виде аналитических выражений, к виду, на который ориентирована их численная обработка в СП. Для этого требуется включение ЦИМ в состав ВС.

Центральная интеллектуальная машина должна обеспечить подготовку СП к работе; определить порядок интерполяционного полинома; вид используемого вычислительного метода; задать вид уравнений, область решения, тип конечных элементов и краевых условий; обеспечить ввод-вывод исходных данных; выполнить необходимые преобразования исходных данных ВЭ с ССД (необходимыми в данном СП); обеспечить коррекцию связей и сигналов управления в структуре СП, связанную с однотипностью вычислительных узлов в ПЭ СП (при разнообразии выполняемых ими функций на различных этапах ВЭ).

Многие функции, которые следует выполнять ЦИМ, заложены в языке АНАЛИТИК последних модификаций. Идеология языка АНАЛИТИК, заложенная в ЭВМ серии МИР, в целом отвечает требованиям к ЦИМ гетерогенных ВС. Их следует доработать, дополнив функциями, вытекающими из обработки ССД в СП, реализующем ЕТП в параллельном режиме. Задача ЦИМ – освободить экспериментатора (или программиста) от рутинного

процесса подготовки ССД для конкретной решаемой экспериментатором задачи и обеспечить комплексное выполнение ВЭ.

Структура предполагаемого ПЭ с СУ в своем составе включает узлы и блоки, обработанные на этапах проектирования и производства. И если в серийно выпускаемые АЛУ процессора добавить некоторые функциональные логические элементы (свойственные операции умножения чисел в параллельном режиме младшими разрядами вперед с запоминанием значений результата и переноса), то вопрос с элементной базой может быть закрыт в части и серийности, и стоимости.

Отдельно следует заметить, что предлагаемая структура ССД на базе РМП, будучи пионерской, не является, по-видимому, единственной. Но любая другая структура ССД потребует включения ЦИМ в архитектуру ВС. Её наличие обеспечит комфорт экспериментатору при постановке ВЭ, а всей ВС – стыковку аналитических и численных методов при обработке ССД в СП в параллельном режиме и едином вычислительном потоке решения научно-технических и экономических задач общего назначения.

Разработка и создание гетерогенных ВС, ориентированных на решение научно-технических, экономических и др. общих задач – перспективное направление, но оно должно развиваться, исходя из требований конкретных задач (с учетом опыта разработки и эксплуатации GPU-CPU-систем).

При проектировании и разработке гетерогенных систем от идеологии представления в ВС исходных данных (с обработкой их последовательно числа за числом) следует уходить, переходя к представлению и обработке данных в виде сложных структур (ССД).

Кроме представления данных в виде ССД, в проектируемые и создаваемые гетерогенные ВС следует закладывать параллелизм: при обработке ССД; на уровне решения всей задачи и выполнять весь вычислительный процесс в режиме ЕТП от ввода исходных данных до получения псевдорешения с индикацией его с помощью средств отображения.

Наличие ЦИМ в ВС позволит замкнуть ВЭ.

Необходимо выполнять в ЦИМ дифференцирование и интегрирование сложных функций в аналитическом виде, переходя затем в СП к выполнению этих процедур в виде операций над простыми функциями (кусочно-гладкими полиномами) в цифровом виде. Это обеспечивает переход к решению сложных нелинейных задач.

Применение в СП ПЭ (с включением в него СУ), имеющих собственное запоминающее устройство, при наличии ЦИМ обеспечивает сокращение количества уровней в ВС (исключая дополнительное обращение к глобальной памяти).

Учитывая, что практически всевозможные методы аппроксимации уравнений сводятся к решению СЛАУ, возможность факторизации матрицы СЛАУ (без выполнения операции деления на прямом ходе [9] на структуре ПЭ с СУ в его составе) обеспечит решение систем $Ax = b$ (для плотно заполненных матриц); выполнение процедуры обращения матриц; итерационное уточнение решений СЛАУ; вычисление невязок, исправление элементов приближенной матрицы и др. в режиме ЕТП.

Разработке новых архитектур гетерогенных ВС должна сопутствовать разработка новых, модернизация и адаптация известных численных и арифметических методов, с представлением данных в виде ССД, ориентированных на параллельное выполнение операций и аппаратную реализацию алгоритма. Это позволит внести новую струю в развитие архитектур мощных сверхвысокопроизводительных ВС.

В серийно выпускаемый универсальный сумматор следует заложить некоторые функциональные логические элементы, которые позволят программно в процессе работы процессора переключать АЛУ в режим СУ и обратно; осуществлять операции типа суммирования с запоминанием значений результата и переноса, возникших в каждом из разрядов параллельного АЛУ. Такое решение понизит стоимость разработки и изготовления АЛУ

для СУ, будет способствовать серийному выпуску СП на базе ПЭ с СУ, а в архитектуре ВС – повышению эффективности и производительности СП.

СПИСОК ЛИТЕРАТУРЫ

1. Адинец А. Графический вызов суперкомпьютерам / А. Адинец, В. Воеводин // Открытые системы. – 2008. – № 4. – С. 35 – 41.
2. Губайдуллин Д.А. Об особенностях использования архитектуры гетерогенного кластера для решения задач механики сплошных сред / Д.А. Губайдуллин, А.И. Никифоров, Р.В. Садовников // Вычислительные методы и программирование. – 2011. – № 12. – С. 450 – 460.
3. Зенкевич О., Морган К. Конечные элементы и аппроксимация / О. Зенкевич, К. Морган; пер. с англ. – М.: Мир, 1986. – 318 с.
4. Ледянкин Ю.Я. Единый технологический поток в организации вычислений – способ повышения производительности параллельных структур на процессорных элементах транспьютерного типа. – Киев, 1989. – 20 с. – (Препр. /АН УССР. Ин-т кибернетики им. В.М. Глушкова; 89-57).
5. Ледянкин Ю.Я. К вопросу преобразования и параллельного ввода граничных условий при решении краевых задач в едином вычислительном потоке / Ю.Я. Ледянкин // Математичні машини і системи. – 2012. – № 1. – С. 28 – 25.
6. Ледянкин Ю.Я. Методы взвешенных невязок, коллокаций, моментов. Способ параллельной реализации в едином вычислительном потоке решения задач математической физики / Ю.Я. Ледянкин // Математичні машини і системи. – 2012. – № 2. – С. 17 – 28.
7. Ледянкин Ю.Я. Метод Галеркина. Способ параллельной реализации задач математической физики в едином вычислительном потоке / Ю.Я. Ледянкин // Математичні машини і системи. – 2012. – № 3. – С. 69 – 80.
8. Ледянкин Ю.Я. Метод наименьших квадратов. Способ параллельной реализации в едином вычислительном потоке решения задач математической физики / Ю.Я. Ледянкин // Математичні машини і системи. – 2012. – № 4. – С. 59 – 69.
9. Ледянкин Ю.Я. Способы параллельного решения систем $Ax = b$ в едином технологическом потоке решения задач математической физики / Ю.Я. Ледянкин // Математичні машини і системи. – 2013. – № 1. – С. 63 – 74.
10. АНАЛИТИК (Аналитический язык для описания вычислительных процессов с использованием аналитических преобразований) / В.М. Глушков, В.Г. Боднарчук, Т.А. Гринченко [и др.] // Кибернетика. – 1971. – № 3. – С. 102 – 134.
11. Клименко В.П. Аналитические преобразования на вычислительных машинах / В.П. Клименко, Ю.С. Фишман // Энциклопедия кибернетики. – Киев: Гл. ред. Укр.сов.энцикл., 1974. – Т. 1. – С. 116 – 117.
12. Самарский А.А. Численные методы: учеб. пособие для вузов / А.А. Самарский, А.В. Гулин. – М.: Наука. Гл. ред. Физ-мат. лит., 1989. – 432 с.
13. Ледянкин Ю.Я. Ускоренный метод умножения чисел, представленных в последовательном дополнительном коде / Ю.Я. Ледянкин // Автоматика. – 1989. – № 3. – С. 76 – 82.
14. Ledyankin Yu.Ya. Accelerated Method of Multiplying Numbers Represented in Sequential Code / Yu.Ya. Ledyankin // Soviet Journal of Automation and Information Sciens. – 1989. – N 3. – P. 88 – 94.
15. Клименко В.П. Основные принципы построения систем интерпретации языков, проблемно-ориентированных на научные и инженерные задачи / В.П. Клименко // Системы компьютерной алгебры семейства АНАЛИТИК. Теория. Реализация. – Киев: Применение. Изд.ООО «НПП Интерсервис», 2010. – С. 8 – 24.
16. Клименко В.П. Развитие и реализация идей МИР / В.П. Клименко // Системы компьютерной алгебры семейства АНАЛИТИК. Теория. Реализация. – Киев: Применение. Изд. ООО «НПП Интерсервис», 2010. – С. 186 – 190.

Стаття надійшла до редакції 18.03.2013