С.А. Инютин

МЕТОДЫ ОРГАНИЗАЦИИ МНОГОРАЗРЯДНЫХ ВЫЧИСЛЕНИЙ

Введены новые структуры данных, виды распараллеливания компьютерных процессов, взаимосвязанные структуры эмулируемых компьютерных арифметик, основных и производных форматов данных, необходимых для анализа проблем, связанных с созданием вычислительных средств поддержки параллельного многоразрядного вычислительного процесса.

Многоразрядный параллельный процесс, сложность вычислений, модулярные компьютерная арифметика и форматы данных.

При математическом моделировании тонких технологических процессов в ряде отраслей техники и производства, а также в теоретической космологии и быстро развивающейся прикладной теории чисел появляются отдельные вычислительные задачи, для решения которых требуется специальная организация вычислительных процессов. В этих процессах значения переменных функций или операндов операций, в частности целочисленных, для ряда вычислительных задач, значительно, на много (сто, тысячу) порядков, превышают максимум типового компьютерного диапазона серийной вычислительной техники. Типовой компьютерный диапазон тесно связан с длиной машинного слова. В машинных словах или их совокупности отображаются данные в различных внутренних форматах. Данные в этих форматах обрабатываются в операционных регистрах процессора. Обработка форматов данных в регистрах поддерживается аппаратными и микропрограммными способами. Машинная арифметика для обработки этих форматов числовых данных, не превышающих максимума типового компьютерного диапазона, является основной в компьютере, поэтому имеет максимальную скорость выполнения [1].

Для корректного анализа проблем организации вычислений с превышением типового компьютерного диапазона серийной вычислительной техники необходимо ввести специальную терминологию.

Назовем диапазонными числовые величины, для отображения которых типовой диапазон достаточен. В зависимости от длины машинного слова вычислительной системы типовой диапазон является 16, или 32, или 64-разрядным (подразумеваются бинарные разряды). Для диапазонных величин, отображаемых в беззнаковом целочисленном формате, для $n = 16 \mid 32 \mid 64$ выполняется неравенство $0 \le A \le 2^n - 1$. Для числовых величин, отображаемых в целочисленном формате со знаком, для $n = 16 \mid 32 \mid 64$ выполняется двухстороннее неравенство $-2^{n-1} \le A \le 2^{n-1} - 1$. Компьютерные форматы данных с двойной точностью и числовые величины, отображаемые в них, удобно условно отнести к диапазонным, так как, как правило, в регистрах процессора обработка форматов данных с двойной точностью также имеет микропрограммную и аппаратную поддержку. Это позволяет считать достаточно быстрой машинную арифметику двойной точности, в которой выполняются операции с компьютерными данными в этих форматах.

Назовем сверхдиапазонными или многоразрядными числовые величины, для отображения которых типовой диапазон недостаточен [2]. В зависимости от длины машинного слова вычислительной системы сверхдиапазонные чи-

словые величины не отображаются в удвоенном типовом 16, или 32, или 64-разрядном диапазоне. Для сверхдиапазонных величин в зависимости от требований вычислительного процесса можно ввести производные беззнаковый целочисленный формат и целочисленный формат со знаком.

Для сверхдиапазонных числовых величин, отображаемых в беззнаковом целочисленном формате, для *n*-разрядного процессора для $n = 16 \mid 32 \mid 64$ выполняется двойное неравенство $2^{2n} - 1 < C \le D$, где *D* — некоторая константа, зависящая от предельных возможностей программно эмулируемой многоразрядной арифметики, операционных ресурсов или объема памяти вычислительной системы, выделяемых для представления и операций с многоразрядными числовыми величинами. Следовательно, множество многоразрядных числовых величин можно назвать многоразрядным диапазоном ***W***, для которого $D = \max W$.

Для сверхдиапазонных числовых величин, отображаемых в целочисленном формате со знаком, для *n* -разрядного процессора для $n = 16 \mid 32 \mid 64$ выполняется двойное неравенство $-D \le C < -2^{2n-1}$ и $2^{2n-1} - 1 < C \le D$.

Задачи тестирования и факторизации больших числовых величин на простоту называют вычислительными проблемами из-за их большой временной и алгоритмической сложности. Проиллюстрируем необходимость многоразрядных вычислений для решения вычислительной проблемы тестирования на простоту чисел Мерсенна при больших показателях степени.

Числами Мерсенна называются числа специального вида $M_q = 2^q - 1$, где $q \in$ **N** — множество натуральных чисел, $q > 2$.

Для этих чисел известны отдельные результаты, что позволяет упростить алгоритмы тестирования и факторизации [5]:

— необходимое условие, если число M_q — простое, то q — простое. Действительно, при составном $q = q_1 q_2$, $q_1 < q_2$ в разложении числа Мерсенна всегда можно выделить множители: $2^{q_1} - 1$, $2^{q_2} - 1$;

— о делителях чисел Мерсенна, если $q \in$ **N** — простое число, $q \equiv 3 (\mathrm{mod}\ 4)$, то $2 \cdot q + 1$ делит число M_q тогда и только тогда, когда $2 \cdot q + 1$ — простое число.

Один из наиболее быстрых методов проверки на простоту чисел Мерсенна базируется на специальных рекуррентных последовательностях. На основе этих последовательностей разработан кодонезависимый вычислительный метод [4, 5].

Теорема Лукаса — Лемера. Пусть построена рекуррентная последовательность $\{W_n\}$ при следующих условиях:

$$W_0 = 4,\ W_{n+1} = W_n^2 - 2,\ n \ge 0.$$

Число Мерсенна M_q — простое тогда и только тогда, когда M_q делит нацело W_{q-2} член последовательности.

Применение модулярной арифметики и форматов данных позволило разработать вычетный итерационный алгоритм первого рода для решения этой вычислительной проблемы, имеющий меньшую сложность, чем известные ранее [3]. Современные методы ориентированы на тестирование чисел Мерсенна при $q > 300\,000$, а сверхдиапазонные числовые величины для таких вычислительных процессов содержат более 100 тыс. десятичных разрядов позиционного представления.

Диапазон для представления многоразрядных целых числовых величин иногда называют целочисленным большим компьютерным диапазоном [2]. Разработка и программная реализация эффективных методов вычислений в больших диапазонах позволяет создавать методы решения задач, требующих вычислений, как в больших, так и в сверхбольших компьютерных диапазонах, появляющихся при вычислении ряда функций от больших числовых величин, в данном случае аргументов этих функций [3]. В частности, для отдельных задач вычислительной теории чисел максимум сверхбольшого компьютерного диапазона должен достигать значения константы Виноградова — Гольбаха $3^{3^{15}}$. Такого рода задачи возникают, в частности, при вычислении выражений следующего вида:

$$C = \sum_{i=1}^{k} A_i^{f(B)} - [\sum_{i=1}^{k} A_i^{f(B)} / B] \cdot B ,$$

где A_i , *B*, а также область значений функции *f*(*B*) принадлежат множеству больших числовых величин ***W***.

Эффективные вычисления уже в больших компьютерных диапазонах требуют специальной эмулируемой компьютерной арифметики, а также специальной организации вычислительного процесса. Для вычислительных средств, на которых решаются такие задачи с использованием эмулируемой арифметики, критическим параметром является производительность, так как арифметические операции и вычислительные процессы при больших диапазонах изменения переменных реализуются достаточно сложно алгоритмически и программно. Алгоритмы и вычислительные процессы требуют больших затрат машинного времени. Как правило, сложность таких алгоритмов превышает линейную. Все вышеперечисленное обусловливает необходимость в использовании мощного резерва повышения производительности вычислительных систем — различных видов распараллеливания. Для анализа проблем повышения производительности при многоразрядных вычислениях целесообразно рассмотреть распараллеливание на следующих уровнях [3]:

— структуры аппаратных средств вычислительной системы;

— архитектуры программно-аппаратных вычислительных средств;

— потоков вычислительных заданий;

— отдельных вычислительных заданий;

— алгоритмов вычислительных задач, реализующих задания;

— форматов машинных данных и компьютерных арифметик, в которых выполняются вычислительные алгоритмы, обрабатывающие числовые величины, представленные в машинных форматах.

Последний уровень распараллеливания представляет особый интерес. Распараллеливание машинных операций является достаточно универсальным методом увеличения быстродействия вычислительной системы. Для разработки модулярной компьютерной арифметики и форматов представления данных используется математический аппарат модулярных систем счисления, известных в теории чисел. Отметим, что для использования модулярных систем счисления необходим точный учет пределов изменения числовых величин в модулярных компьютерных диапазонах. В модулярных системах счисления для мультипликативных операций над соответствующими операндами достигается линейная временная сложность при определенных ограничениях на пределы изменения операндов в модулярных диапазонах. Линейных оценок сложности для мультипликативных операций нельзя достичь для машин-

ных арифметик и форматов данных, построенных на основе позиционных, полиадических и полулогарифмических систем счисления. Для них выполняется оценка сложности Штрассена [4]. Выигрыш в быстродействии проявляется явно и становится заметным для вычислительных процессов, использующих многоразрядные числовые величины операндов. Таким образом, компьютерные арифметики и соответствующие форматы данных, основанные на традиционно используемых для представления данных в вычислительной технике системах счисления, не позволяют конструировать оптимальные по сложности архитектурные решения для класса распараллеливаемых вычислительных задач, работающих с многоразрядными значениями операндов.

Для повышения эффективности вычислительных процессов, использующих многоразрядные числовые величины, необходимо учесть другой важный момент. Вычислительные алгоритмы для расчетов на основе математических соотношений в моделях, как правило, разрабатываются кодонезависимыми, т.е. не зависящими от арифметики, в которой они реализуются, и от способа представления в некоторой системе счисления переменных величин, обрабатываемых в алгоритме. При реализации нетривиальных вычислительных процессов, оперирующих с целочисленными величинами, принимающими значения из больших и сверхбольших диапазонов, используются предельные возможности вычислительной техники. При этом из-за технико-экономических ограничений критическими параметрами становятся особенности разрабатываемой для поддержки вычислительного процесса архитектуры программно-аппаратных средств, на которых алгоритмы реализуются. Компьютерные арифметики являются теоретической базой для разработки таких вычислительных процессов.

Применительно к проблеме организации многоразрядного вычислительного процесса на серийном компьютере или вычислительном кластере введем следующую иерархическую структуру, описывающую взаимосвязь алгоритмов:

— первичная математическая постановка задачи, решаемой вычислительными методами, или математическая модель с соотношениями, соответствующими содержательному наполнению проблемы; используются кодонезависимые переменные, инвариантные к типу системы счисления и способу представления переменных;

— проблемный вычислительный алгоритм, решающий поставленную математическую задачу; учитываются сложность, сходимость алгоритма, погрешности дискретизации переменных и т.п.;

— расчетный алгоритм как модификация вычислительного алгоритма; учитывается тип системы счисления, способ и диапазоны представления переменных;

— эмулируемая многоразрядная компьютерная арифметика; учитываются тип системы счисления, способ и диапазоны представления операндов операций этой арифметики;

— базовая машинная арифметика, являющаяся основой эмулируемой многоразрядной арифметики; учитываются базовые машинные форматы данных.

Для оптимального применения модулярных систем в качестве арифметико-логической базы вычислительных устройств сами проблемные алгоритмы нуждаются в преобразовании их в «вычетную форму» с максимальным количеством операций вычисления по модулю [2]. В этом случае после преобразования проблемные алгоритмы оптимально, с минимальным количеством немодульных операций, отображаются на модулярные аппаратные или про-

граммные арифметико-логические компоненты вычислительных средств: устройства или программного обеспечения, ориентированного на сетевое многомашинное применение [3].

Для решения проблем, возникающих при разработке вычислительных средств, ориентированных на работу с многоразрядными числовыми величинами, актуально исследование вопросов:

— в каких диапазонах с использованием эмулируемой арифметики еще возможны вычисления, т.е. нахождение грани, до которой можно дойти при соответствующей вычислительной архитектуре и программно-аппаратной базе при приемлемых на практике затратах машинного времени (например, решение отдельных вычислительных проблем требует месяцев работы одиночного компьютера, недель вычислительного кластера или многопроцессорного суперкомпьютера);

— оптимизации структуры программно-аппаратных вычислительных средств, обеспечивающих вычисления при многократном превышении машинного диапазона;

— разработки формальных и эвристических методов целенаправленного преобразования вычислительных алгоритмов в расчетные алгоритмы, учитывающие ограничения и особенности вычислительных средств, ориентированных на вычисления в больших диапазонах;

— оптимальной организации вычислительного процесса с минимальной временной и алгоритмической сложностью по преобразованным в расчетные вычислительным алгоритмам на таких средствах;

— разработки методов верификации промежуточных и итоговых результатов многоразрядного вычислительного процесса и методов графической визуализации и интерпретации результата.

ЛИТЕРАТУРА

1. *Инютин С.А.* Основы многоразрядной алгоритмики. Сургут: Изд-во РИЦ, 2002. 138 с.

2. *Инютин С.А.* Модулярные вычисления в сверхбольших компьютерных диапазонах // Электроника. 2003. № 6. С. 54–61.

3. *Инютин С.А.* Основы модулярной алгоритмики. Ханты-Мансийск: Полиграфист, 2009. 297 с.

4. *Ноден П., Китте К.* Алгебраическая алгоритмика. М.: Мир, 2003. 720 с.

5. *Riesel H.* Prime Numbers and Computer Methods for Factorization. Stuttgart; Boston: Birhhauser, 2004. 452 p.

S.A. Inyutin

Organization methods for multi-digit calculations

The paper introduces new data structures, types of parallelizing computer processes, interrelated structures of emulated computer arithmetics, primary and derivative data formats, necessary for analyzing problems solved under the creation of a computational tool support for parallel multi-digit computation process.

Multi-digit parallel process, computational complexity, modular computer arithmetic and data formats.