УДК 511.172

РАСПАРАЛЛЕЛИВАНИЕ ПРЕДСТАВЛЕНИЙ МНОГОРАЗРЯДНЫХ ЧИСЕЛ НА МОДУЛЯРНЫХ СТРУКТУРАХ ДАННЫХ

Чернобровкин В.В.

ГБОУ ВПО «Сургутский государственный университет», Сургут, e-mail: kooa@adm.surgu.ru

В работе описывается распараллеливание представлений сверхбольших многоразрядных чисел с помощью системы остаточных классов (СОК), Китайской теоремы об остатках (КТО) и Малой теоремы Ферма. Такое представление многоразрядных чисел дает преимущества по времени для вычислительных операций при значительном большом объеме обрабатываемых данных, так как вся обработка информации проходит параллельно и независимо друг от друга. Параллельное представление допускает отображение миллион и более разрядных чисел и может широко применяться как в кодировании информации, в алгоритмах сверхбыстрых вычислений, так и в тестировании суперкомпьютеров. Несмотря на некоторые недостатки, которые в свою очередь преодолимы, модулярная система или система остаточных классов является уникальной математической средой для параллельных вычислений. Поэтому такую систему можно учитываться при разработке высокопроизводительных систем на базе программируемых интегральных логических схем (ПЛИС) и многоядерных процессоров.

Ключевые слова: распараллеливание, китайская теорема об остатках, малая теорема Ферма, система остаточных классов, сверхбольшие числа, многоразрядность, обработка данных

PARALLELIZATION OF REPRESENTATIONS OF MULTI-DIGIT NUMBERS ON A MODULAR DATA STRUCTURES

Chernobrovkin V.V.

Surgut State University, Surgut, e-mail: kooa@adm.surgu.ru

The paper describes the parallelization of views with the large multi-digit numbers with assistance through the system of residual classes (SRC), the Chinese remainder theorem (CTR) and the Small Fermat's theorem. Such representation of multi-digit numbers gives advantages time for computing operations with significant large volume of processed data, as all the processing information runs parallel to and independently of each other. Parallel presentation allows the presentation of a million or more bit numbers and can be widely used as the encoding of information in the algorithms for high-performance computing and testing of supercomputers. Despite some shortcomings, which in turn can be overcome, modular system, or system of residual classes is unique mathematical environment for parallel computing. Therefore, such a system can be considered in the development of high-performance systems on the basis of PLIC and many-core processors.

Keywords: Parallelization, China's theorem on residues, Fermat's little theorem, system of residual classes, extra-large number bitness, data processing

В современной алгоритмической теории большие и сверхбольшие числа играют важную роль в криптографии для построения шифров различной сложности и для тестирования производительности суперкомпьютеров [4, 8]. Исследование сверхбольших чисел [5] показало, что они имеют величины, которые трудно или почти невозможно поразрядно представить в обычных позиционных системах счисления.

Если число $A = 2^{2147483647}$ возвести в указанную степень, то оно будет иметь более миллиона разрядов. Поэтому такие числа нецелесообразно представлять в обобщенной позиционной системе счисления в натуральном исходном состоянии.

Представим сверхбольшие многоразрядные числа с помощью обычной позиционной системы счисления, в виде информационной системы, показанной на схеме вычисления сверхбольших чисел (рис. 1), и проанализируем их.

На схеме показано, что на вход алгоритма вычисления сверхбольших чисел (стрелки слева) поступают сверхбольшие многоразрядные числа, стрелки «сверху-вниз» обозначают позиционную систему счисления и арифметические операции. На выходе алгоритма сверхбольшое число-результат. Однако если число-результат отобразить в обычной позиционной системе, то возникают проблемы:

1. Число не вмещается в типовой компьютерный диапазон $[0, ..., 2^{31}]$.

2. Временные затраты на представление числа в обобщенной позиционной системе.

3. Нагрузка на объем оперативной памяти, если число состоит из связанных списков.

По схеме абстрактного представления сверхбольшого многоразрядного числа (рис. 2) на выходе получается сверхбольшое число, которое не вмещается в типовой компьютерный диапазон. Наглядно выполнять какие-либо операции с таким числом не представляется возможным, так как оно занимает сотни страниц файлов любого типа из операционной системы Windows.

В связи с этим сверхбольшие многоразрядные числа $A = 2^{2147483648}$ эффективно представлять в непозиционной системе счисления – системе остаточных классов, где нет переносов из младших разрядов в старшие.

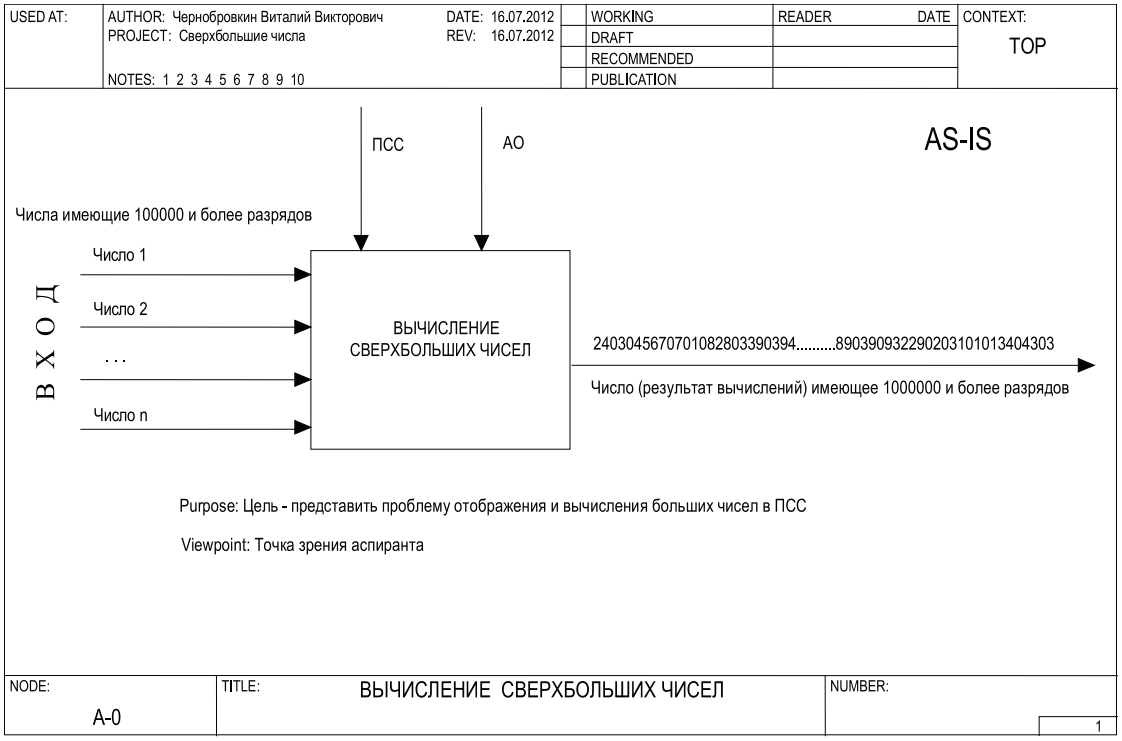

Рис. 1. Схема вычисления сверхбольших чисел

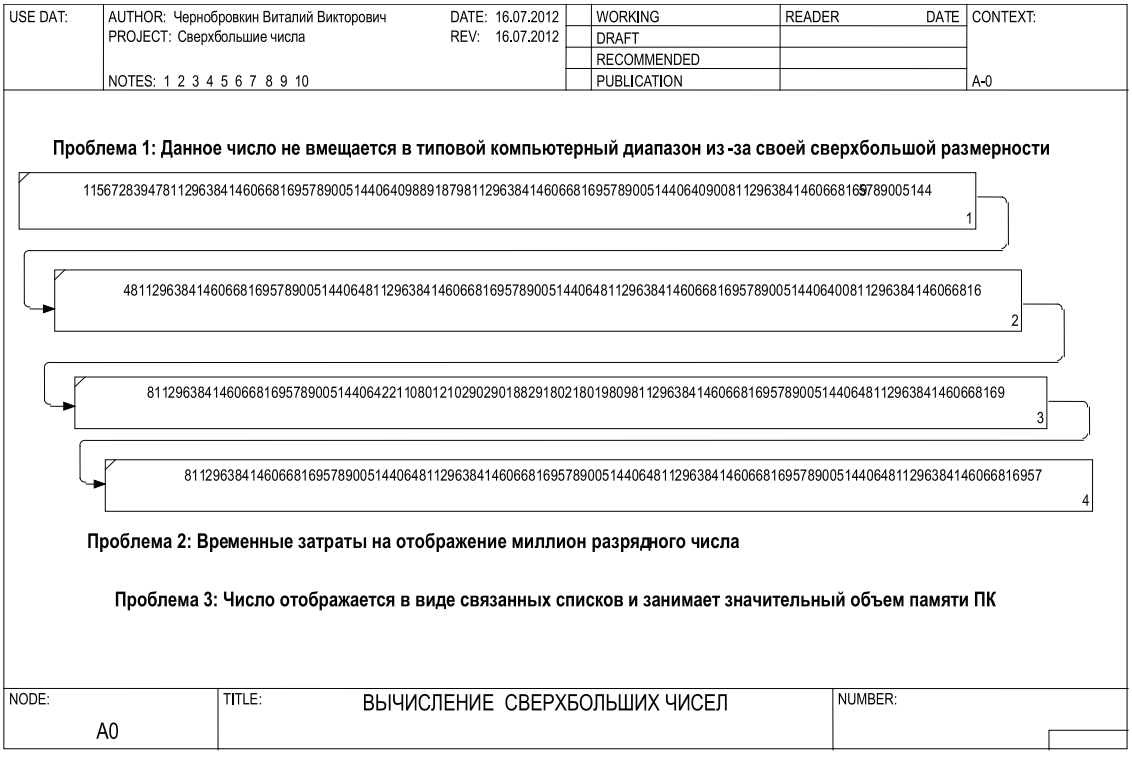

Рис. 2. Схема абстрактного представления сверхбольшого многоразрядного числа в виде связанных списков

Существует проблема, которая состоит в следующем: если сверхбольшое многоразрядное число представить с помощью системы остаточных классов [5, 7], то число остатков, как и выбранная система оснований, будет большим, что в свою очередь влияет на вычислительные процессы.

Введем числовой диапазон в двоичной позиционной системе $[2^{65537}, ..., 2^{2147483648}]$.

Утверждение 1 Число $A = 2^{2147483648}$, имеющее количество разрядов из диапазона $[2^{65537}, ..., 2^{2147483648}]$ и более, можно называть сверхбольшим многоразрядным числом.

Число, имеющее миллион и более разрядов, можно назвать сверхбольшим многоразрядным, так как в реальном исполнении оно имеет большое численное значение.

Утверждение 2 Сверхбольшие многоразрядные числа, представленные в системе остаточных классов, зависят от веса позиционного ранга $|Z|_P$ [1], а значит, могут состоять из большого множества выбранных модулей и соответственно остатков. Вес позиционного ранга $|Z|_P$ зависит от содержания в нем количества простых чисел. При вычислении сверхбольших чисел в системе остаточных классов нужно учитывать следующее равенство:

$$0 \le A \le P, \qquad (1)$$

где A – число, результат вычисления; $P = \prod_{i=1}^{n} p_i$ – основания модулярной системы, взаимно простые числа.

Утверждение 3 Сверхбольшое число можно представить с помощью системы остаточных классов в виде остатков от вычетов [3, 5, 6, 7]

$$a \equiv b(\mathrm{mod}\, p). \qquad (2)$$

Возможность такого представления числа определяется следующими теоремами.

Теорема 1 (Китайская теорема об остатках) [2, 3, 5, 6, 7].

Любое целое положительное число в системе остаточных классов можно представить в виде набора остатков (вычетов) от деления этого числа на выбранные основания (модули) системы.

Доказательство: Пусть число $A = 2^{2147483648}$ представлено в системе остаточных классов как

$$A(a_1, a_2, ..., a_n), \qquad (3)$$

где a_1, a_2,..., a_n – наименьшие неотрицательные остатков (вычетов), образованные путем целочисленного деления числа A на выбранные основания

$$P_i = p_{12} p \cdot ... \cdot p_n = \prod_{i=1}^{n} p_i, \qquad (4)$$

где p_i – взаимно простые числа.

Пример:

$$a_i = A(\mathrm{mod}\, p_i) = A - \left[\frac{A}{p_i}\right] - p_i, \ (\forall A_i \in [1, n]), \qquad (5)$$

И если $\forall A_i \ne j$, $(p_i, p_j) = 1$, то представление числа (3) является единственным при условии

$$0 \le A \le P_i. \qquad (6)$$

Тогда число A будет выглядеть следующим образом:

$$\begin{aligned} A &\equiv a_1(\mathrm{mod}\, p_1); \\ A &\equiv a_2(\mathrm{mod}\, p_2); \\ &\cdots\cdots\cdots \\ A &\equiv a_n(\mathrm{mod}\, p_n). \end{aligned} \qquad (7)$$

Теорема доказана.

Докажем п. 1.6.

Теорема 2 Если числа в $p_i = p_1, p_2, ..., p_n$ попарно взаимно просты, то для любых остатков $a_1, a_2, ..., a_n$, таких, что $0 \le a_n \le p_n$ при всех $i = 1, 2, ..., n$ найдется число A, которое при делении на p_i дает остаток a_i при всех $i = 1, 2, ..., n$

Доказательство: Применим индукцию по. При $n = 1$ утверждение теоремы очевидно. Пусть теорема справедлива при $n = k - 1$, т.е. существует число M, дающее остаток r_i при делении на p_i при $i \in \{1, 2, ..., k - 1\}$. Обозначим

$$d = a_1, a_2, ..., a_{k-1}. \qquad (8)$$

Рассмотрим числа M, $M + d$, $M + 2d$, ..., $M(a_k - 1)d$. Покажем, что хотя бы одно из этих чисел даёт остаток r_k при делении на a_k. Предположим, что это не так. Поскольку количество чисел равно a_k, а возможных остатков при делении этих чисел на a_k может быть не более чем $a_k - 1$ (т.к. ни одно число не даёт остаток r_k), то среди них найдутся два числа, имеющих равные остатки. Пусть это числа

$$M + sd \text{и } M + td \text{М при } 0 \le s \le a_k$$

$$\text{и } 0 \le t \le a_k \text{ и } s \ne t. \qquad (9)$$

Тогда их разность

$$(M + sd) - (M + td) = (s - t)d$$

делится на a_k, что невозможно, т.к. $0 < |s - t| < a_k$ и $d = a_1, a_2, ..., a_{k-1}$ взаимно просто с a_k, ибо числа $a_1, a_2, ..., a_{k-1}$ попарно взаимно просты (по условию). Противоречие.

Теорема доказана.

Так как речь идет о миллион разрядных числах, то опишем реальные проблемы их построения в системе остаточных классов:

- Выбор количества оснований
- Разрядность выбранных оснований
- Многоразрядность остатков

Это означает, что, если мы строим сверхбольшие числа, у которых миллион и более разрядов, то количество оснований системы будет зависеть от разрядности этих чисел. Так, например, если разрядность каждого из выбранных оснований маленькая, то соответственно количество оснований нужно увеличивать. И, наоборот, если разрядность выбранных оснований большая, то их нужно уменьшить до необходимого количества. То есть речь идет о расширении или сужении оснований системы остаточных классов [7]. Еще одна проблема состоит в том, что при отображении сверхбольшого числа $A = 2^{2147483648}$ в системе остаточных классов остатки могут быть тоже многоразрядными.

В общем случае решить вышеперечисленные проблемы можно с помощью Малой теоремы Ферма, доказательство которой приведено в [2, 6] и Китайской теоремы об остатках [2, 3, 5, 6, 7].

Основной смысл Малой теоремы Ферма состоит в том, что если p – положительное простое число, a – целое, тогда

$$a^{p-1} \equiv 1(\mathrm{mod}\, p), \text{ если } p - \text{простое число}, \quad (10)$$

значит, $a^k \equiv a^r(\mathrm{mod}\ p)$ и достаточно делать вычисления только для экспонент, показатель которых меньше $p - 1$.

Применяя Китайскую теорему об остатках, можно упростить отображение сверхбольших чисел. Показав их в виде вычетов степеней по модулю n, если известно его разложение на простые множители. Предположим, что каждый простой множитель входит в это разложение с кратностью 1, потому что именно в таком случае метод наиболее эффективен.

Допустим, что разложение имеет вид

$$n = p_1 \ldots p_k, \quad (11)$$

где $0 < p_1 < \ldots p_k$ – простые числа.

Для целых чисел a и m сначала находим вычет a^m по каждому модулю p_i. Если простые множители не слишком велики, то вычисления будут очень быстрыми даже для больших m и a, поскольку нам помогает это делать Малая теорема Ферма (п. 10).

Предположим, что вычисления сделаны, причем

$$\begin{gathered} a^m \equiv r_1(\mathrm{mod}\, p_1) \text{ и } 0 \le r_1 \le p_1; \\ a^m \equiv r_2(\mathrm{mod}\, p_2) \text{ и } 0 \le r_2 \le p_2; \quad (12) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a^m \equiv r_k(\mathrm{mod}\, p_k) \text{ и } 0 \le r_k \le p_k. \end{gathered}$$

Поэтому для определения вычета a^m по модулю n нужно решить систему сравнений:

$$\begin{cases} x^m = r_1 \left(\mathrm{mod}\, p_1\right), \\ x^m = r_2 \left(\mathrm{mod}\, p_2\right), \\ \ldots\ldots\ldots\ldots\ldots \\ x^m = r_k \left(\mathrm{mod}\, p_k\right). \end{cases} \quad (13)$$

Напомним, что модули системы – различные попарно взаимно простые числа. Значит, по Китайской теореме об остатках система всегда имеет решение, к примеру r, $0 \le r_i \le p_i - 1$. Более того, любые два таких решения сравнимы по модулю $p_1 \ldots p_k = n \ldots$. Так как a^m тоже решение системы, имеем $a^m \equiv r(\mathrm{mod}\, n)$. Следовательно, r – вычет a^mпо модулю n.

Пример. Сюръективно отобразим число $A = 2^{2147483648}$ в виде вычетов по выбранным основаниям $p_1 = 11$, $p_2 = 13$, $p_3 = 17$, $p_4 = 19$. Изменив при этом степень указанного числа на единицу 2147483648 – 1 = 2147483647, применим Малую теорему Ферма, для чего найдем остатки от деления степени 2147483647 (на $p - 1$) каждого выбранного основания p_1, p_2, p_3, p_4. Тогда

$$\begin{gathered} p_{i-1} = 11 - 1 = 10;\ p_{i-2} = 13 - 1 = 12; \\ p_{i-3} = 17 - 1 = 16;\ p_{i-4} = 19 - 1 = 18; \quad (14) \end{gathered}$$

$$\begin{gathered} a'_1 = 2147483647 \bmod 10 = 7; \\ a'_2 = 2147483647 \bmod 12 = 7; \quad (15) \\ a'_3 = 2147483647 \bmod 16 = 15; \\ a'_4 = 2147483647 \bmod 18 = 1. \end{gathered}$$

Следовательно

$$\begin{gathered} a_1 = 2^{2147483647} \equiv 2^7(\mathrm{mod}\, 11); \\ a_2 = 2^{2147483647} \equiv 2^7(\mathrm{mod}\, 13); \quad (16) \\ a_3 = 2^{2147483647} \equiv 2^{15}(\mathrm{mod}\, 17); \\ a_4 = 2^{2147483647} \equiv 2^1(\mathrm{mod}\, 19). \end{gathered}$$

Окончательно число $A = 2^{2147483647}$ будет выглядеть:

$$\begin{gathered} 2^7 \equiv 7(\mathrm{mod}\, 11); \\ 2^7 \equiv 11(\mathrm{mod}\, 13); \quad (17) \\ 2^{15} \equiv 9(\mathrm{mod}\, 17); \\ 2 \equiv 2(\mathrm{mod}\, 19). \end{gathered}$$

Выводы

1. Распараллеливание представлений сверхбольших многоразрядных чисел с помощью системы остаточных классов дает преимущество по времени для вычисле-

ний при большом объеме обрабатываемых данных, так как вся обработка информации происходит параллельно и независимо друг от друга.

2. Любое многоразрядное число можно представить в виде остатков от степенных вычетов.

3. Над представленными параллельным образом числами можно проводить ускоренные параллельные вычисления.

4. Если образование остатков a_i, производится независимо друг от друга, то каждое вычисление будет также происходить не зависимо друг от друга.

Список литературы

1. Амербаев В.М. Теоретические основы машинной арифметики. – Алма-Ата: Наука, 1976. – 320 с.

2. Бухштаб А.А. Теория чисел. – М.: Наука, 1975.

3. Василенко О.Н. Современные способы проверки простоты чисел // Кибернетический сборник. Новая серия. – 1988. – Вып. 25. – С. 162–187.

4. Горбунов В.С., Эйсымонт Л.К., Речинский А.В., Заборовский В.С., Забеднов П.В. Суперкомьютеры для промышленности – вопросы тестирования, анализа и разработки // Суперкомпьютерные технологии: материалы 2-й Всесоюзной конференции (СКТ-2012), С. 360–364.

5. Инютин С.А. Теория и методы моделирования вычислительных структур с параллелизмом машинных операций: дис. ... д-ра техн. наук. – М., 2001. – 5 – 264 с.

6. Коутинхо С. Введение в теорию чисел алгоритм RSA. – М.: Постмаркет, 2001. – 328 с.

7. Червяков Н.И., Сахнюк П.А., Шапошников А.В., Ряднов С.А. Модулярные параллельные вычислительные структуры нейропроцессорных систем. – М.: Физматлит 2003. – 43с.

8. Agarwal R., Alpern B., Carter L. High-performance parallel implementation of the NAS kernel benchmarks on IBM SP2 // IBM System Journal, Edit. 34, February, 1995.

References

1. Amerbaev V.M. Theoretical foundations of computer arithmetic. Alma-Ata: Nauka, 1976. -320 p.

2. Bukhshtab A.A. Theory of numbers M: Nauka, 1975.

3. Vasilenko O.N. Modern means of verification simplicity numbers // Cyber collection. New series. 1988. Vol. 25. pp. 162–187.

4. Gorbunov V.S., Eisymont L.K, Rechinsky A.V., Zaborovsky V.S., Zabednov P.V., Super computers for industry testing, analysis and development. Materials of the 2nd all-Union conference «Supercomputer technologies» (SKT-2012), pp. 360–364.

5. Inyutin S.A. Theory and methods of modeling of computing structures with a connection for use with computers-abuses machine operations: диссерт. on competitions for the degree of doctor of technical Sciences. M., 2001. pp. 5–264.

6. Coutinho C.K. Introduction Numbers Theory and RSA Cryptography. Moscow: Postmarket, 2001. 328 p.

7. Chervyakov N.I., Sakhnyuк P.A., Shaposhnikov A.V., Ryadnov S.A. Modular parallel nye computing structures neuroprocessor systems M.: Fizmatlit 2003. 43 p.

8. Agarwal R., Alpern Century, L. Carter. High-performance parallel implementation of the NAS kernel benchmarks on the IBM SP2, IBM System Journal, Vol. 34, February, 1995.

Рецензенты:

Инютин С.А., д.т.н., профессор кафедры «Проектирование вычислительных комплексов», ФГБОУ ВПО «Российский государственный технологический университет им. К.Э. Циолковского (МАТИ)», г. Москва;

Бахарев М.С., д.т.н., профессор кафедры «Нефтегазовое дело», ФГБОУ ВПО «Сургутский институт нефти и газа, филиал Тюменского государственного нефтегазового университета», г. Сургут;

Криштоп В.В., д.ф.-м.н., профессор, заведующий кафедрой «Физика», Дальневосточный государственный университет путей сообщения, г. Хабаровск, профессор Университета Kwangwoon University, Korea.

Работа поступила в редакцию 19.12.2013.