



УДК 004.222.3+004.421

## MICROINSTRUCTIONS ARCHITECTURE LEVEL WITH LONG BINARY ARITHMETIC FOR A COMPUTER SYSTEM OF HIGH-PRECISION COMPUTING

### РІВЕНЬ АРХІТЕКТУРИ МІКРОКОМАНД З ДОВГОЮ ДВІЙКОВОЮ АРИФМЕТИКОЮ ДЛЯ КОМП'ЮТЕРНОЇ СИСТЕМИ ВИСОКОТОЧНИХ ОБЧИСЛЕНЬ

Levchenko A.A. / Левченко А.О.

c.t.s., as.prof. / к.т.н., доц.

ORCID: 0000-0001-5550-0027

Golovko O.V. / Головки О.В.

Master of Computer Engineering / магістр комп'ютерної інженерії

Odessa National University, Odessa, Dvoryanska, 2, 65029

Одесский национальный университет, Одесса, Дворянская, 2, 65029

**Анотація.** В роботі розглядаються сукупність програмних кодів виконання наступних операцій, що формують рівень архітектури мікрокоманд з довгою двійковою арифметикою для комп'ютерної системи високоточних обчислень. Наведені часткові результати експериментів перевірки працездатності довгої двійкової арифметики для побутових комп'ютерних систем. Особливості подання двійкових чисел у довгої арифметики дозволяють значно знизити похибку в арифметичних операціях порівняно з IEEE 754 та наблизити похибку до величин похибок властивих десятинному представленню числа. Зазначений час проведення розрахунків для розв'язків систем лінійних алгебраїчних рівнянь, що складає до кількох хвилин для систем 10 ступеню. Час розрахунку залежить від ступеню різниці у кількості значущих розрядів в двох операндах однієї операції.

**Ключеві слова:** інформаційні сервіси, високоточні розрахунки, ієрархічна побудова комп'ютерів, базові арифметичні операції, формування архітектури мікропрограм, подання числових даних.

#### Вступ.

На сьогодні існує цілий ряд галузей, розв'язки задач у яких вимагають підвищеної точності розрахунків. Узагальнення подібних задач проведено в [1]. В оглядовій частині згаданої роботи зазначено, що проблема впливу помилок округлення виникає у великій кількості завдань в різних областях знань, включаючи обчислювальну математику, математичну фізику, біохімію, квантову механіку, математичне програмування. В роботі [2] згадано, що через помилки округлення при використанні стандартної арифметики IEEE 754 обчислений результат підсумовування може виявитися вкрай неточним. Ця робота визначає й цікавий перелік посилань до робіт з проблематики розв'язку завдання забезпечення точності розрахунків в комп'ютерах.

Таким чином розвиток шляхів забезпечення високоточних обчислень для комп'ютерних систем є перспективним напрямком досліджень.

#### Визначення проблеми.

В роботі [1] окреслено два напрямки робіт з забезпечення обчислень при виконанні розрахунків в надвеликих числових діапазонах. До першого напрямку віднесено роботи по модернізації і створення нових технологій гібридних обчислень і обробки даних, другим напрямком визначено розробку спеціалізованих апаратних засобів підтримки операцій над «надвеликими»



операндами з використанням ПЛІС (FPGA) і систем на кристалі.

Сучасні роботи [3-11] природно присвячені розробці різноманітних алгоритмів високоточних обчислень з представленням чисел в форматі з плаваючою комою. Наприклад у згаданих роботах [1, 2] завдання забезпечення точності обчислень пропонується забезпечити шляхом реалізацій високоточних обчислень в базисі модулярно-інтервальної арифметики або з використанням бібліотеки MPRES з використанням представлення двійкових чисел у форматі з плаваючою крапкою.

Взагалі огляд наукових публікацій дозволяє зробити висновок, що більшість досліджень присвячено не усуненню витоків втрати працездатності програмно-апаратних засобів та розвитку способів забезпечення точності комп'ютерних обчислень, а розвитку та пристосуванню алгоритмів обчислень з плаваючою комою для забезпечення точності розрахунків в конкретних задачах. Так [1] в якості способів зменшення помилок комп'ютерних розрахунків вважається використання бібліотек багаторазової точності, що надають структури даних і підпрограми обробки чисел, довжина яких перевищує формати IEEE 754.

Але існує і інший шлях створення комп'ютерних систем високоточних розрахунків [12], що ґрунтується на ієрархічній побудові комп'ютерних систем зі зміною способу подання двійкових даних на мікроархітектурному рівні та рівні архітектури мікрокоманд процесору. В роботах [13-15] запропоновано спосіб подання двійкових чисел у вигляді масивів, що усуває витки помилок подання чисел з плаваючою комою. З огляду на ієрархічну побудову комп'ютерних систем напрямком подальших робіт є формування саме мікроархітектурного рівня подання двійкових даних та розроблення архітектури базових мікрокоманд процесору. **Метою роботи** автори визначили обґрунтування структури базових мікрокоманд процесору з довгою двійковою арифметикою для комп'ютерної системи високоточних обчислень.

#### **Розв'язок основного завдання роботи.**

Останнім часом технологічні рішення дозволяють створити малогабаритні комп'ютери потужність яких перевищує потужність супер ЄОМ 60-70 років ХХ сторіччя. Але реалізація систем розрахунків з підвищеною точністю стикається с низкою штучних обмежень обробки даних, що є наслідком комерційного створення програмних продуктів високого рівня для побутових задач користувачів.

Лімітуючим фактором обмеження точності в ПЄОМ є використання типів числових даних, що представленні широко розповсюдженим стандартом представлення чисел та арифметики – IEEE 754, впровадженим ще у 1985 році, з якого використовується тип з максимальним розміром у 64 біт – довжина регістру сучасного процесору. Це дозволяє зберігати і опрацьовувати дійсно дуже великі або маленькі числа, що дозволяє використовувати ПЄОМ в дослідженнях, але з точністю лише у середньому 16.3 знаки.

З іншого боку широке розповсюдження потужних мікрокомп'ютерів привело до їх поступового використання в наукових дослідженнях та інженерних задачах, але обмеження в особливостях представлення чисел не



дозволяють проводити високоточні розрахунки.

Істотною перевагою для програмної реалізації базових арифметичних операцій системи розрахунків підвищеної точності на рівні архітектури базових мікрокоманд процесору є наявність алгоритмів виконання операцій з обмеженою розрядною сіткою.

Заналізу недоліків довгої арифметики [13-15] можна зробити висновок, що зараз при найпопулярнішому об'ємі пам'яті у 16 Гб і довгому числі з кількістю значущих розрядів розміром 1 Мб проблемою залишиться обмежена можливість використання довгої арифметики у системах реального часу.

43.24% протестованих сучасних комп'ютерів що мають вбудовану оперативну пам'ять 16 Гб з застосуванням довгої арифметики забезпечують проведення розрахунків з величиною похибки властивою десятинній арифметиці. Для експериментів використовувались розрядність подання двійкових даних, що при використанні десятинної основи буде мати 1048576 значущих цифри. Виявленим недоліком є істотне збільшення часу розрахунку, в порівнянні з стандартними для користувачів, пакетами прикладних програм. Так для розв'язку системи лінійних рівнянь 10 ступеню потребується від декількох десятків секунд до навіть хвилин.

Але наявність готових алгоритмів базових операцій було вирішальним фактором вибору подання чисел для практичної реалізації програмно-апаратної системи високоточних розрахунків [13]. Для реалізації програмного рішення обрана мова C++ стандарту ISO/IEC 14882:2017. Рішення було реалізовано у вигляді класу C++ з трьома полями:

- 32-біт цілочисельного поля для знаку;
- 64-біт цілочисельного поля для положення коми;
- вектор з 64-біт цілочисельними елементами для десятичних розрядів.

Рівень архітектури мікрокоманд з довгою двійковою арифметикою для комп'ютерної системи високоточних обчислень сформовано сукупністю програмних кодів виконання наступних операцій:

- вивод довгого числа з пам'яті;
- запис довгого числа у пам'ять;
- порівняння двох довгих чисел з врахуванням знаку числа;
- порівняння двох довгих чисел без врахування знаку числа;
- нормалізація довгих чисел;
- додавання довгих чисел без врахування знаку числа;
- віднімання довгих чисел без врахування знаку числа;
- додавання довгих чисел з врахуванням знаку числа;
- віднімання довгих чисел з врахуванням знаку числа;
- множення довгих чисел з врахуванням знаку числа;
- множення довгих чисел без врахування знаку числа;
- ділення довгих чисел з врахуванням знаку числа;
- ділення довгих чисел без врахування знаку числа.



Операції додавання довгих чисел з врахуванням знаку числа та віднімання довгих чисел з врахуванням знаку числа є похідними від додавання та віднімання довгих чисел без врахування знаку числа. Передбачено й виконання операцій множення та ділення як похідних від операцій додавання та віднімання довгих чисел з врахуванням знаку числа.

Під час виконання операції додавання – знак двох чисел перемножується, вектори доповнюються нулями для того щоб вирівняти розряди чисел, кожен елемент вектора (розряд числа) додається до відповідного елементу другого числа. Потім переповнення переносяться у старші розряди.

Під час виконання операції віднімання – менше відіймається від більшого і в залежності від порядку операндів визначається знак різниці. Вектори доповнюються нулями для того щоб вирівняти розряди чисел, кожен елемент вектора (розряд числа) відіймається від відповідного елементу другого числа і розряди займаються при необхідності. Потім з вектора прибирається зайві нулі.

Під час виконання операції порівняння – по-перше порівнюються знаки чисел, потім змінна для визначення положення коми і тільки потім порозрядно числа. Реалізовані алгоритмічно порівняння чи більше та чи дорівнює, інші порівняння реалізовані через цих два алгоритми.

Під час виконання операції множення – знаки перемножуються, положення ком додаються, а розряди перемножуються кожен на кожний. Потім виконується перенос переповнень у старші розряди.

Під час виконання операції ділення - знаходження оберненого дільника та помноження на це число. Для знаходження оберненого числа реалізовано алгоритм – визначається скільки разів можна відняти число від 1 з подальшим зсувом коми при наявності залишку. Оскільки існують ірраціональні числа, то точність такої операції була обмежена 100 знаками – результат буде повернено коли буде досягнута максимальна точність або коли не буде залишку від ділення.

Таким чином виконана сукупність програмного коду базових операцій що складають систему обчислення різнорозмірних даних без втрати точності. Особливості подання двійкових чисел у довгої арифметики дозволяють значно знизити похибку в арифметичних операціях порівняно з IEEE 754 та наблизити похибку до величин похибок властивих десятинному представленню числа, що забезпечує проведення високоточних розрахунків на сучасних побутових комп'ютерах.

Сукупність зазначених операцій дозволяє виконувати програмування на мові C++ будь-яких математичних розв'язків сучасних задач з абсолютною величиною похибки наближеною до теоретичною похибки методу. Зазначена сукупність операцій складає рівень архітектури мікрокоманд з довгою двійковою арифметикою для комп'ютерної системи високоточних обчислень.

#### **Заключення та висновки.**

В роботі було розглянуто сукупність операцій команд з довгою двійковою арифметикою для комп'ютерної системи високоточних обчислень. Протестовані сучасні побутові комп'ютери з вбудованою оперативною пам'яттю 16 Гб з застосуванням мікрокоманд з довгої арифметики забезпечують



проведення розрахунків з величиною похибки властивою десятиною арифметиці. Тестові приклади розв'язку системи лінійних рівнянь 10 ступеню найбільш ресурсним з кількості базових операцій методом Гауса показали виключення похибки округлення та подання даних але для розв'язку потребується від декількох десятків секунд до кількох хвилин. Час розрахунку залежав від ступеню різномірності даних. Під різномірністю мається на увазі істотно різна кількість значущих розрядів в двох операндів однієї операції.

#### Литература:

1. А. С. Коржавина, В. С. Князьков Реализация высокоточных вычислений в базе модулярно-интервальной арифметики [Online]. – Access mode: <https://cyberleninka.ru/article/n/realizatsiya-vysokotochnyh-vychisleniy-v-bazise-modulyarno-intervalnoy-arifmetiki>.

2. К.С. Исупов В.С. Князьков А.С. Куваев Суммирование многократной точности на центральных и графических процессорах с использованием библиотеки MPRES Известия ЮФУ. Технические науки. Раздел III. Математическое и программное обеспечение № 8 (2018). pp. 191-203. DOI 10.23683/2311-3103-2018-8-191-203.

3. Kornerup P., Lefèvre V., Louvet N., Muller J. On the Computation of Correctly Rounded Sums // IEEE Transactions on Computers. – 2012. – Vol. 61, No. 3. – P. 289 - 298.

4. Higham N.J. Accuracy and stability of numerical algorithms. – Philadelphia, USA: SIAM, 2002. – 680 p.

5. Priest D.M. Algorithms for arbitrary precision floating point arithmetic // Proceedings of the 10th IEEE Symposium on Computer Arithmetic. Grenoble, France: IEEE. – 1991. – P. 132 - 143.

6. Higham N.J. The Accuracy of Floating Point Summation // SIAM Journal on Scientific Computing. – 1993. – Vol. 14, No. 4. – P. 783 - 799.

7. Anderson I.J. A Distillation Algorithm for Floating - Point Summation // SIAM Journal on Scientific Computing. – 1999. – Vol. 20, No. 5. – P. 1797 - 1806.

8. Zhu Y. - K., Yong J. - H., Zheng G. - Q. A New Distillation Algorithm for Floating - Point Summation // SIAM Journal on Scientific Computing. – 2005. – Vol. 26, No. 6. – P. 2066 - 2078.

9. Rump S.M. Ultimately Fast Accurate Summation // SIAM Journal on Scientific Computing. – 2009. – Vol. 31, No. 5. – P. 3466 - 3502.

10. Goodrich M.T., Eldawy A. Parallel Algorithms for Summing Floating - Point Numbers // Proceedings of the 28th ACM Symposium on Parallelism in Algorithms and Architectures. – Pacific Grove, California, USA: ACM, 2016. – P. 13 - 22.

11. Demmel J., Nguyen H.D. Fast reproducible floating - point summation // Proceedings of the 21st IEEE Symposium on Computer Arithmetic (ARITH). – Austin, Texas: IEEE, 2013. – P. 163 - 172.





12. Kadric E. , Gurniak P. and DeHon A. Accurate Parallel Floating - Point Accumulation // IEEE Transactions on Computers. – 2016. – Vol. 65, No. 11. – P. 3224 – 3238.

12. А.О.Левченко, О.В. Головка Порівняльний аналіз можливостей сервісів високоточних розрахунків// Modern engineering and innovative technologies // International periodic scientific journal – Karlsruhe, germany: – 2020. – №14, part 1

13. Левченко А.О., Arithmetic operation for binari numbers repressetated as arrays // Modern engineering and innovative technologies / International periodic scientific journal – Karlsruhe, Germany: – 2019. – №9, part 1. – p. 51-59. DOI: 10.30890/2567-5273.

14. Левченко А.О., Берков Ю.М., Головка О.В.Помилки арифметичних операцій з двійковими числами представлених як масиви для системи прогнозу// «Modern scientific researches» / международный научный периодический рецензируемый журнал – ЧП Ёлнать, Минск, совместно с институтом «Ise&E», Беларусь: – 2020. – №13, part 1. – p. 81-88. DOI: 10.30889/2523-4692.

15. Левченко А.О., Р.М. Войтенков Граничні точності обчислень в інформаційних системах з представленням чисел із плаваючою комою Збірник наукових праць ВА (м. Одеса), Випуск № 2 (2), 2014 р., С.157-161.

**Abstract.** *The paper considers a set of program codes for the following operations, which form the level of architecture of microinstructions with long binary arithmetic for a computer system of high-precision calculations. Partial results of long binary arithmetic validation experiments for household computer systems are given. As operations that make up the level of architecture of microinstructions with long binary arithmetic for a computer system of high-precision calculations, the content of codes is considered to perform the following operations: call a long number from memory; recording a long number in memory; comparison of two long numbers taking into account the sign of the number; comparison of two long numbers without taking into account the sign of the number; normalization of long numbers; adding long numbers without taking into account the sign of the number; subtraction of long numbers without taking into account the sign of the number; adding long numbers taking into account the sign of the number; subtraction of long numbers taking into account the sign of the number; multiplication of long numbers taking into account the sign of the number; multiplication of long numbers without taking into account the sign of the number; division of long numbers taking into account the sign of the number; division of long numbers without taking into account the sign of the number. Features of the representation of binary numbers in long arithmetic can significantly reduce the error in arithmetic operations compared to IEEE 754 and bring the error closer to the values of errors inherent in the decimal representation of the number. The calculation time for solutions of systems of linear algebraic equations is specified, which is up to several minutes for systems of 10 degrees. The calculation time depends on the degree of difference in the number of significant bits in the two operands of one operation.*

**Key words:** *information services, high-precision calculations, hierarchical construction of computers, basic arithmetic operations, formation of firmware architecture, representation of numerical data.*

Статья отправлена: 30.11.2020 г.

© Левченко А.О.