УДК 004.2 (045)
Д 973.20-02-5-05

Жуков І. А. д-р техн. наук
Гуменюк В. О. канд. техн. наук
Синельников О. О.

НОВИЙ ПІДХІД ДО ПРЕДСТАВЛЕННЯ ДВІЙКОВОЇ ІНФОРМАЦІЇ У ВИСОКОПРОДУКТИВНИХ ПАРАЛЕЛЬНИХ ОБЧИСЛЮВАЛЬНИХ СИСТЕМАХ

Інститут комп'ютерних технологій Національного авіаційного університету

Розглядаються причини виникнення й існування проблеми автоматичного контролю в ЕОМ загалом і у високопродуктивних паралельних обчислювальних системах зокрема. Аналізується можливість застосування у таких системах коду, названого парафазним, з метою підвищення продуктивності та рівня достовірності результатів обчислень. Пропонується новий підхід до організації обчислювальних операцій, з урахуванням особливостей такого представлення двійкової інформації у паралельних обчислювальних системах.

Постановка проблеми. Ріст вимог до продуктивності обчислювальних засобів, що виконують обчислення, пов'язані з проектуванням літаків та ракет, фундаментальними науковими дослідженнями, передбаченням погоди та природних катаклізмів і т. ін. за період найближчих чотирьох років передбачає збільшення продуктивності до 10^{25} флопс [1]. Проблема реалізації масових паралельних обчислювальних процесів, що виникла внаслідок підвищення вимог до продуктивності обчислювальних засобів – ці вимоги перевершили фізичні можливості одного процесора, що працює на принципі фон Неймана [2].

Аналіз останніх досліджень. Практичне рішення проблеми підвищення продуктивності обчислювальних систем звичайно зв'язують, у першу чергу, зі збільшенням тактової частоти роботи елементів і кількості цих елементів, що дозволяє вводити паралелізм обробки й програмованість структури Однак удосконалювання обчислювальних систем завжди супроводжував розрив між швидкодією логічних елементів і елементів пам'яті. Цей розрив при зростанні ступеня інтеграції й швидкодії великих інтегральних схем (ВІС) має тенденцію до збільшення. На кожному рівні розвитку елементної бази в силу обставин, обумовлених необхідністю подолання даного розриву, обмеженням на розмір і кількість виводів у корпусів мікросхем, наявними засобами автоматизації програмування одні архітектури отримували переваги над іншими – наприклад, по показнику «продуктивність/вартість» [1].

Архітектура на базі обміну повідомленнями використовує окремі набори команд читання й запису для роботи з локальною пам'яттю й спеціальні команди типу *send, receive* для керування адаптерами каналів вводу-виводу. Стандартизовані вимоги, пропоновані шиною до адаптерів, дозволяють будувати системи з «великих» блоків – системних плат робочих станцій і ПК, а також мережних плат (*Myrinet*, *Quadrics*, *Dolphin SCI*, *Fast Ethernet* і ін.) і комутаторів комунікаційних середовищ. Для таких систем гостро коштує проблема ефективності паралельних обчислень, тому що вони свідомо мають обмеження пропускної здатності обмінів, обумовлені шиною *PCI*.

Орієнтація розроблювачів на створення систем з розподіленою поділюваною пам'яттю привела до інтеграції в кристал блоку керування когерентністю багаторівневою пам'яттю, доступ до блоків якої виконується через інтегровану в той же кристал комунікаційне середовище. Як приклади цього підходу можна назвати мікропроцесори *Alpha* 21364 і *Power* 4. Інтеграція функцій, з одного боку, дозво-

ляє істотно збільшити пропускну здатність між компонентами кристала в порівнянні із пропускною здатністю між різними кристалами, що реалізують окремо кожну функцію. І, як наслідок, підняти продуктивність систем, підвищити надійність й знизити вартість систем.

Розвиток найпоширенішої з архітектур – кластерної – сприяло створенню серверів-«лез» (*Server Blade*), що дозволяють вирішити завдання не менш складні, чим ті, які прийнято довіряти суперкомп'ютерам. У якості зовнішніх «сполучних» інтерфейсів між «лезами» найчастіше використаються *Fibre Channel* або *Infiniband*, що дозволяє істотно розширити сферу застосування кластерів на базі таких складових, дозволивши їм скласти реальну конкуренцію деяким суперкомп'ютерам.

До недоліків таких обчислювальних систем, що стримують ріст їхньої продуктивності, варто віднести збільшене тепловиділення потужного процесора у відносно малому обсязі і його підвищені вимоги до електроживлення.

Додавання процесорів до системи, як відомо, далеко не завжди приводить до лінійного росту її продуктивності. Втрати продуктивності можуть виникати, наприклад, при недостатній пропускній здатності шин через зростання трафіку між процесорами й основною пам'яттю, а також між пам'яттю й пристроями уведення/виводу.

У самому загальному змісті архітектуру комп'ютера можна визначити як спосіб з'єднання комп'ютерів між собою, з пам'яттю й із зовнішніми пристроями. Реалізація цього з'єднання може йти різними шляхами. Конкретна реалізація з'єднань такого роду називається комунікаційним середовищем комп'ютера. Одна з найпростіших реалізацій – це використання загальної шини, до якої підключаються як процесори, так і пам'ять. Сама шина складається з певного числа ліній зв'язку, необхідних для передачі адрес, даних і керуючих сигналів між процесором і пам'яттю. Цей спосіб реалізований в *SMP* системах. Основним недоліком таких систем є погана масштабованість. Збільшення, навіть незначне числа пристроїв на шині викликає помітні затримки при обміні з пам'яттю й катастрофічне падіння продуктивності системи в цілому. Необхідні інші підходи для побудови комунікаційного середовища, і одним з них є поділ пам'яті на незалежні модулі й забезпечення можливості доступу різних процесорів до різних модулів одночасно за допомогою використання різного роду комутаторів. При цьому можливі різні конфігурації систем зв'язку. Оригінальні унікальні рішення значно збільшують ціну комп'ютерів.

Істотно більше простим і більше дешевим виявилося використання зв'язків на базі мереж *Ethernet*. Спочатку використовувалася звичайна 10-мегабітна мережа, потім стали застосовувати *Fast Ethernet*, а останнім часом іноді й *Gigabit Ethernet*. Але для *Fast Ethernet* характерна більша латентність (затримка в передачі даних), оцінювана в 160-180 мікросекунд, а *Gigabit Ethernet* відрізняється високою вартістю [3].

Ключовою особливістю системи *SGI Altix* 3000, заснованої на архітектурі глобальної поділюваної пам'яті *SGI Numaflex*, є використання каскадуємих комутаторів у маршрутизуючих елементах. Каскадуємі комутатори забезпечують системі відносно невеликі часові затримки, або збільшення часу доступу до пам'яті, незважаючи на модульну конструкцію. Це критично для машин, що використають архітектуру неоднорідного доступу до пам'яті (*NUMA*). Затримки є однією з проблем в архітектурі *NUMA*, тому що пам'ять розподіляється між вузлами, а не зосереджена в одному місці.

Між окремими провідниками шини для рівнобіжної передачі даних існує електрична ємність, тому при зміні сигналу, переданого по одному з провідників, виникає перешкода (короткий викид напруги) на інших провідниках. Зі збільшенням довжини шини (збільшенням ємності провідників) перешкоди зростають і можуть сприйматися приймачем як сигнали.

Тому робоча відстань для шини рівнобіжної передачі даних обмежується довжиною 1-2 м, і тільки за рахунок істотного подорожчання шини або зниження швидкості передачі довжину шини можна збільшити до 10-20 м.

Таким чином, огляд архітектур найбільш продуктивних паралельних обчислювальних систем й порівняльний аналіз їх характеристик дозволяють зробити висновок, що окрім переваг, кожна з них має суттєві недоліки. Результатом цього є існуюча проблема реалізації масових паралельних обчислювальних процесів.

Основною метою статті є пошук додаткових можливостей підвищення загальної продуктивності паралельних обчислювальних систем у значній мірі незалежно від їх архітектури. Звернемося до особливостей подання та зберігання інформації в ЕОМ та мікропроцесорних системах.

Парафазний спосіб записування та зчитування інформації. Як відомо, у регістрах на *RS*- або *JK*- тригерах можливий однофазний або парафазний спосіб записування інформації. При однофазному записуванні [9] частота обміну інформацією відносно зменшується, оскільки процеси введення і скидання чергуються. При парафазному записуванні інформації значення кожного розряду слова A передається по двох лініях зв'язку. При цьому пряме значення A_i надходить на вхід S (або D) відповідних тригерів, а інверсне значення A_i – на вхід R (або K). У цьому випадку не потрібне попереднє скидання регістра в стан "0", тому що таку функцію виконує сигнал $\overline{A}_i$ (рис. 1).

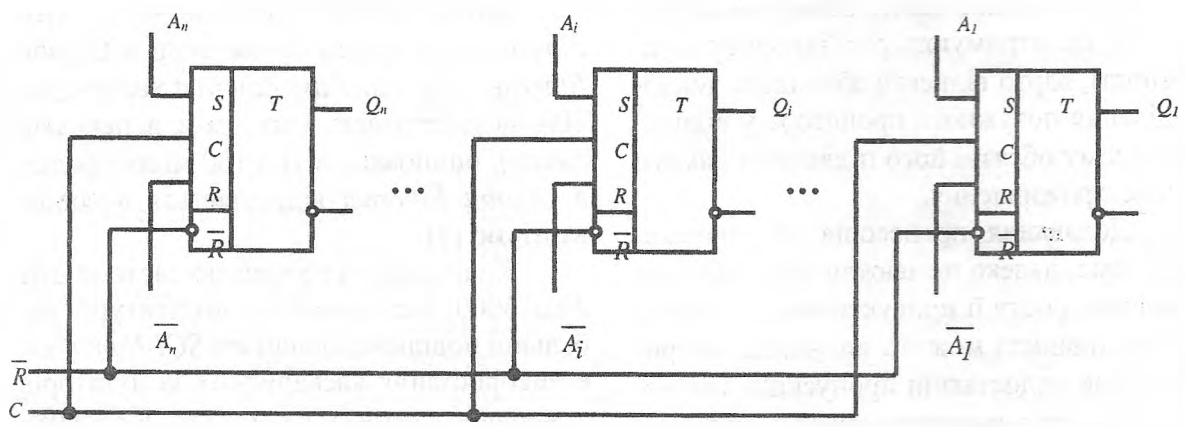

Рис. 1. Схема регістра з парафазним записом даних

Для записування інформації від декількох джерел (напрямків) на вході кожного тригера ставлять додаткові комбінаційні схеми, які створюють вхідну логіку регістра. Для записування в регістр на *JK*-тригерах парафазним кодом слів A і B потрібно реалізувати такі порозрядні функції збудження входів J_i і K_i :

$$J_i = Y_1 A_i \vee Y_2 B_i,\ K_i = Y_1 \overline{A}_i \vee Y_2 \overline{B}_i \quad (1)$$

Схема вхідної логіки i-го розряду регістра на основі рівнянь (1) показана на рис. 2.

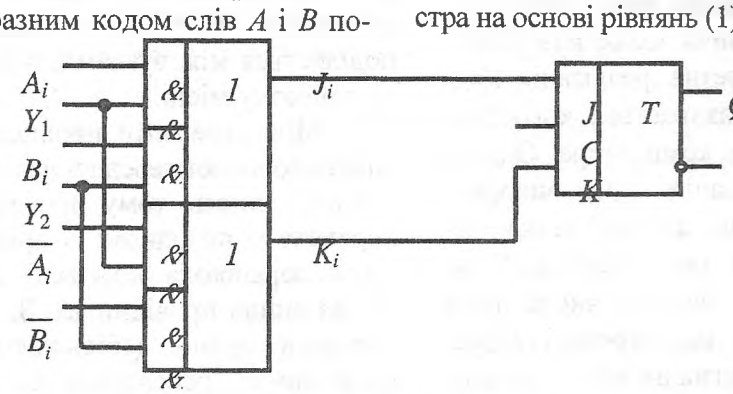

Рис. 2. Схема розряду регістра із записом слів від двох джерел

Інформація, яка зберігається в регістрах, може передаватися у зовнішні схеми парафазним способом у прямому або оберненому коді. Для реалізації мікрооперацій зчитування до виходів кожного тригера підключаються комбінаційні схеми, які створюють вихідну логіку регістра.

Схема вихідної логіки парафазним прямим або оберненим кодом будується на основі таких порозрядних логічних рівнянь:

$$Ш_i^* = Y_{пр} Q_i \vee Y_{пр} \overline{Q_i};\ \overline{Ш_i}^* = Y_{об} Q_i \vee Y_{об} \overline{Q_i}, \quad (2)$$

де $Y_{пр}$ і $Y_{об}$ – керуючі сигнали видачі відповідно прямого або оберненого коду; Q_i і $\overline{Q_i}$ – пряме та інверсне значення виходу i-го розряду регістра; $Ш_i$ – розряд однофазної шини даних; $Ш_i^*$ і $\overline{Ш_i}^*$ – розряди парафазної шини даних.

Очевидно, що керуючі сигнали $Y_{пр}$ і $Y_{об}$ не повинні збігатися в часі. При зчитуванні інформації парафазним оберненим кодом отримаємо:

$$Y_{пр} = 0;\ Y_{об} = 1;\ Ш_i^* = \overline{Q_i};\ \overline{Ш_i}^* = Q_i.$$

Схеми вихідної логіки для i-го розряду на основі рівнянь (2) показані на рис. 3.

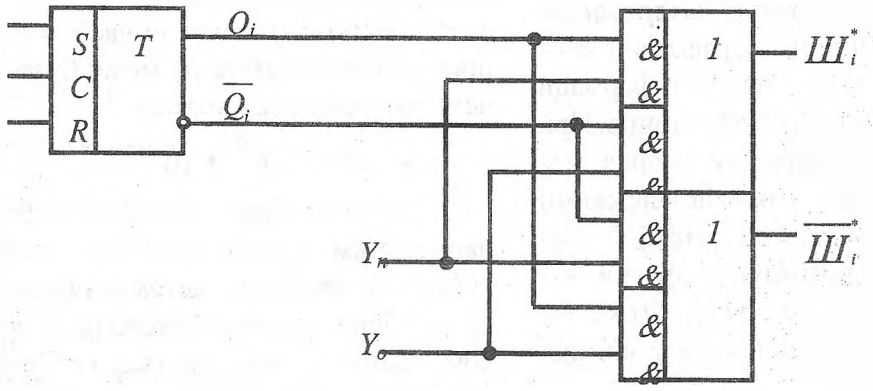

Рис. 3. Схема вихідної логіки i-го розряду регістра для зчитування інформації парафазним кодом

Отже, як відзначається в [9], парафазний спосіб представлення інформації має перевагу у продуктивності над однофазним способом за рахунок виключення попереднього скидання регістра в стан "0". Розглянемо далі інші можливості, пов'язані із застосуванням парафазного коду.

В [7] установлено, що традиційний спосіб подання двійкової інформації в ЕОМ, при якому обидві цифри двійкового розряду представлені одним тригером (взаємозалежне подання) привів до втрати природної контролездатності позиційних числень, що у свою чергу обумовило використання в ЕОМ різних надлишкових кодових побудов, що дозволяють виявляти або виправляти помилки. При такому принципі побудови апаратного контролю не може бути забезпечене безвідмовне функціонування органів контролю за наступними причинами.

Система контролю будується на елементах, по паспортній інтенсивності відмов однакових з елементами контрольованих вузлів. Енергетичні режими роботи елементів системи контролю по суті не відрізняються від режимів роботи контрольованих елементів, оскільки процедура контролю по модулю заснована на виконанні обчислювальних операцій за правилами розрахувань одночасно з виконанням контрольованих операцій. Резервування елементів контролю приводить до погіршення ряду параметрів обчислювальної системи.

Імовірність D_l одержання у обчислювальній системі безпомилкового (достовірного) результату визначається в [7] як

$$D_l = 1 - (1 - P_{kl} * P_{nl} * P_{ml}) * R_l, \quad (3)$$

де P_{kl} – імовірність безвідмовної роботи органів контролю; P_{nl} – імовірність охоплення контролем устаткування обчислювальної системи; P_{ml} – методична ймовірність виявлення помилок категорії l; R_l – імовірність виникнення помилок категорії l ($l = 1, 2, \ldots, \Psi$).

З виразу (3) випливає, що проблема не може бути вирішена на основі тради-

чного" і "нульового" виходів тригера, що представляє цифру "1" в Тому розряді операнда ***X;*** х,ох„0 - сигнали високого рівня з "одиничного" і "нульового" виходів тригера, що представляє цифру "0" в г'-ому розряді операнда ***X***\ значення сигналів ***уц уа*** і ***у i,o Уіо*** Щ° представляють операнд У - аналогічні.

Формування значень переносів "1" у тетраді двійкових розрядів суматора відбувається відповідно до виразів:

ПІ,- = 2,i,oi 3.2,01 2l/-3.01 П1/-4 V 2,i,01 2 , 2,01 2 А-3,10 V 2;.їді 2, 2Д0 (2,ідо 2,ю);

ПІМ = 2.2,01 у(-3,01 П I, 4 V 2 г2,01 2,здо у (*Xi-2,io* 2, ідо);

П1г-2 = 2 , 3,01 П I, 4 V (2г_здо 3-гдо);

пі,-з = 2/-ЗД0 ПІ/.4.,

Вираз для г'+I-го розряду, з урахуванням необхідності, при складанні повнорозрядних операндів, узгодження даної тетради двійкових розрядів з наступною, більш старшою:

П1(+i = 2,oi 2,4.01 2,-2,oi 2,-з,oi II 1,-4 V 2,.oi 2,_i.oi 2 , 2,01 2,здо у —у 2;,oi 2 , i,oi3-2.ю у у 2,oi 2,ідо у (2,ю 2,+i.ю),

де ПІ,- - перенос "1" в і-й розряд суми з множини молодших розрядів.

Висновки

Аналіз властивостей парафазного коду і структур відповідних обчислювальних пристроїв показує, функції контролю виконуються одночасно з процесом обчислення, не додаючи непродуктивних часових витрат. Можливість контролю кожного розряду операнду окремо дозволяє планувати процес обчислення поразрядно, або частками операндів, що сприяє скороченню апаратних витрат. Використання парафазного коду, як і кодів «М з А» [5] надає можливості для практичного вирішення проблеми автоматичного контролю ЕОМ.

Список літератури

1. *Бурцев В.С.* Система массового параллелизма с автоматическим распределением аппаратных средств суперЭВМ в процессе решения задачи // Юбилейный сб. трудов институтов Отделения информатики, вычислительной техники и автоматизации РАН. - М„ 1993. - Т. 2. - С. 5-27.

2. *Бурцев В.С.* Новые подходы к оценке качества вычислительных средств // Параллелизм вычислительных процессов и развитие архитектуры суперЭВМ. - М.: Нефть и газ, 1997. - С. 28-40.

3. *Гуменюк В.О.* Технології кластерних архітектур// Проблеми інформатизації та управління: зб. наук, праць: Вип. 11. -К .: НАУ, 2004. - С.151-156.

4. *Воеводин В.В., Воеводин Вл.В.* Параллельные вычисления. - С.Пб.: БХВ-Санкт-Петербург, 2002. - 357 с.

5. *Корнеев В.В.* Параллельные вычислительные системы. - М.: Нолидж, 1999. - 320 с.

6. *Корнеев В.В.* Современные микропроцессоры. - С.Пб.: БХВ-Санкт-Петербург, 2003. - 440 с.

7. *Брюховин Е.И.* О проблеме автоматического контроля в ЭВМ и контролеспособности позиционных счислений// УСиМ. - 1977. - №4. - С. 71-75.

8. *Брюховин Е.И., Гуменюк В.А.* Исследование достоверности результатов вычислений при представлении чисел в ЭВМ *«М* из *N»* кодом / Техническая диагностика электронных систем:, сб. науч. трудов - К.: Наукова думка. - 1982. - 172 с.

9. *Бабич Н.П., Жуков И.А.* Компьютерная схемотехника. Методы построения и проектирования: Учебное пособие. - К: МК-Пресс, 2004. - 576 с.

10. *Гуменюк В.А., Жуков И.А., Гуменюк А.В.* Применение неразделимых кодов *«М* из *N»* в высокопроизводительных параллельных вычислительных системах // Проблеми інформатизації та управління: зб. наук, праць: Вип. И. - К.: НАУ, 2004. - С. 256-263.