

В.П. Іващенко, Г.Г. Швачич, М.О. Ткач

ОСОБЛИВОСТІ ПОВУДОВИ РЕЖИМІВ РОБОТИ КОМУНІКАЦІЙНОЇ МЕРЕЖІ БАГАТОПРОЦЕСОРНОЇ СИСТЕМИ З РОЗПОДІЛЕНОЮ ОБЛАСТЮ ОБЧИСЛЕНЬ

Анотація. Статтю присвячено висвітленню питань дослідження особливостей побудови мережевого інтерфейсу в багатопроцесорних системах з розподіленою областю обчислень. Система орієнтована для розв'язування широкого кола прикладних задач. Для цього запропоновано чотири типи топології комунікаційної мережі багатопроцесорної системи: "лінійку", "кільце", "зірку", "решітку". Висвітлюються особливості налаштування мережевого інтерфейсу багатопроцесорної системи для реалізації вказаних режимів роботи комунікаційної мережі.

Ключові слова: багатопроцесорна обчислювальна система, обчислювальні вузли, мережевий інтерфейс, топологія.

Вступ

У сучасних умовах особливої ваги набуває створення багатопроцесорних обчислювальних систем на базі стандартних загальнодоступних технологій та компонентів [1]. Завдяки високому попиту й пропозиції на лезові конфігурації в науковій практиці запропоновано саме "блейд"-кластерний обчислювальний комплекс для розв'язування задач з розподіленою областю обчислень [2]. Під час проектування кластера велике значення має його конструкція. Саме на етапі конструювання багатопроцесорної системи необхідно передбачити можливості її розширення або модифікації в майбутньому. Відзначимо, що найбільш вдалим рішенням вважається розміщення багатопроцесорної системи в стійці. Таке облаштування виявилось доречним навіть для невеликої обчислювальної системи. У середині стійки містяться вузли, апаратура для ефективного з'єднання компонентів, засоби керування внутрішньою мережею системи та под. Кожне лезо працює під керуванням своєї копії стандартної операційної системи. Склад і потужність вузлів багатопроцесорної

системи може бути різним. В даній роботі розглядається однорідна система. Взаємодія між вузлами обчислювальної системи встановлюється за допомогою спеціалізованих бібліотек.

Проте при проектуванні та ефективному використанні багатопроцесорної системи основна увага приділяється комутаційній мережі системи та її топології [3 – 7]. Топологія кластера та його швидкодія при розв'язуванні обчислювальних задач, звичайно, речі пов'язані. У класичному розумінні існує три базові топології мережі багатопроцесорної системи – “шина”, “кільце” та “зірка”. Але всі сучасні мережі, як правило, є гібридними варіантами, тобто складаються із сегментів, що використовують різні базові топології.

Приймаючи до уваги відзначене, в даній роботі розглядаються особливості побудови та режими роботи комунікаційної мережі багатопроцесорної системи з розподіленою областю обчислень.

Аналіз останніх досліджень

Аналіз перспектив розвитку багатопроцесорних систем показує, що на сьогодні відомі два основні методи підвищення продуктивності й швидкодії обчислювальних систем:

- використання усе більш досконалої елементної бази;
- паралельне виконання обчислювальних операцій.

Перший метод потребує вельми значних капіталовкладень. Досвід фірми CRAY, яка створила суперкомп'ютер на базі арсеніду галію, показав, що розробка принципово нової елементної бази для високопродуктивних обчислювальних систем – занадто складне завдання навіть для іменитих корпорацій. Другий метод почав домінувати після оголошення в США урядової програми, відомої як “Прискорена стратегічна комп'ютерна ініціатива” (ASCI).

З огляду на подані вище міркування, відзначимо, що останнім часом процес створення високопродуктивних систем обчислення, здебільшого, розвивався, одним шляхом – об'єднанням багатьох процесорів для паралельного розв'язку однієї великої та складної задачі. На цій підставі сьогодні ототожнюють поняття суперкомп'ютера і паралельної (багатопроцесорної) обчислювальної системи.

Отже, нині набули поширення два основні напрями у застосуванні паралельної обробки даних. Перший з них – це паралельні комп'ютери із спільною пам'яттю (SMP). Уся їх оперативна пам'ять розділяється декількома однаковими процесорами. Це найбільш зручний для програміста спосіб об'єднання багатьох процесорів у єдиний. На жаль, він є, одночасно, і найдорожчим в апаратній реалізації. Таким чином, сфера застосування SMP-підходу – це досить коштовні, не дуже великі за кількістю вузлів системи, як зазначалось, досить зручні для програмування. При цьому доводиться визнати, що тут існує проблема масштабування. Другий напрям – масивно-паралельні комп'ютери з розподіленою пам'яттю (MPP). Об'єднання окремих самостійних ЕОМ спеціалізованими каналами зв'язку виявилось менш зручним для програміста, але на порядок дешевим, до того ж цей підхід допускає потенційно нескінченне масштабування пам'яті. Архітектура такого типу має багато переваг: коли треба, можна додавати процесори, збільшуючи продуктивність кластера; якщо фінансові можливості обмежені або заздалегідь відома необхідна обчислювальна потужність, то легко підбрати потрібну конфігурацію системи. Назва таких систем підкреслює теоретично необмежену масштабованість пристроїв подібного класу.

Наведений аналіз напрямів розвитку високопродуктивних установок показує, що реального перелому в оволодінні технологіями паралельних обчислень можна досягти введенням додаткового (фактично базового) рівня в ієрархії потужностей апаратних засобів багатопроцесорних обчислювальних систем MPP-архітектури – персональних обчислювальних кластерів (ПОК). Таким чином, пропонується створити фундамент піраміди апаратних засобів технології паралельних обчислень у вигляді персональних обчислювальних кластерів, що є аналогічним наявному фундаменту піраміди апаратури традиційних технологій послідовних обчислень у вигляді ПЕОМ. Так само як ЕОМ свого часу перестали бути екзотикою після масового поширення персональних комп'ютерів, так і оволодіння

технологіями паралельних обчислень можливе тільки внаслідок широкого застосування ПОК.

Отже, реальний перелом в оволодінні технологіями паралельного моделювання можливий у разі запровадження багатопроцесорних обчислювальних систем модульного типу, сконструйованих на основі MPP-архітектури. Крім того, аналіз недоліків сучасних багатопроцесорних систем показує, що на сьогоднішньому етапі їх розвитку персональні обчислювальні кластери повинні конструюватися для розв'язування певного класу задач. Ці питання і знайшли висвітлення в даній роботі.

Постановка проблеми досліджень

Нині проблема моделювання архітектури високопродуктивних багатопроцесорних систем обробки даних, які використовуються при розв'язуванні задач із розподіленою областю обчислень не знайшла свого розвитку. Зауважимо, що практичне застосування таких систем орієнтовано на розв'язування наступного типу задач. Існує різницєва сітка розмірності M ; час обчислення задачі при використанні однопроцесорної системи являє собою параметр t . Але він не є визначальним. Принциповим виступає збільшення розміру сітки, причому понад такий, який може оброблятися в пам'яті одного процесора. Така процедура є визначальною для більш детального розрахунку або отримання деяких нових ефектів досліджуваних процесів.

З іншого боку, щоб отримати допустиму точність числового розв'язку зазначених задач, їх область визначення необхідно покрити сіткою, яка містить не менше $100 \times 100 \times 100$ вузлів. У кожній точці сітки потрібно обчислити 5 – 20 функцій, а стан усього ансамблю функцій потрібно визначити в $10^2 - 10^4$ моментах часу. Тому потрібно отримати порядку $10^9 - 10^{11}$ результатів проміжних обчислень. Крім того, необхідно взяти до уваги, що для обчислення проміжних результатів, як показує практика, потрібно в середньому виконати $10^2 - 10^3$ арифметичних операцій. За таких умов для проведення тільки одного варіанта числового експерименту кількість операцій

порядку $10^{11} - 10^{14}$ є для зазначеного типу задач цілком рядовим. А якщо врахувати необхідну кількість варіантів, компіляцію та роботу операційної системи, то стає очевидним, що швидкодія обчислювальної техніки повинна вимірюватися мільярдами операцій за секунду.

Отже, стає актуальною не тільки проблема моделювання архітектури високопродуктивних багатопроекторних систем із розширюваною областю, а й особливостей побудови режимів роботи комунікаційної мережі таких систем для розв'язування заявленого типу задач.

Мета та задачі дослідження

Мета даної роботи полягає у висвітленні питань дослідження особливостей побудови мережевого інтерфейсу в багатопроекторних системах з розподіленою областю обчислень. При цьому система повинна застосовуватися для розв'язування широкого кола прикладних задач. Для цього необхідно розв'язати наступні задачі:

- запропонувати чотири типи топології комунікаційної мережі багатопроекторної системи: “лінійку”, “кільце”, “зірку”, “решітку”;
- висвітлити особливості налаштування мережевого інтерфейсу багатопроекторної системи для реалізації вказаних режимів роботи комунікаційної мережі системи.

Виклад основного матеріалу досліджень

Для досягнення поставленої мети було розроблено модуль багатопроекторної обчислювальної системи з розподіленою областю обчислень. Структурна схема модуля високопродуктивної багатопроекторної системи з розширюваною областю обчислень подана на рис 1.

Модуль містить один майстер-вузол (PM001) і N обчислювальних slave-вузлів (PN001, PN002, ..., PN00n), два керовані комутатори (SW1, IB1), реконфігуровану мережу для обміну даними між обчислювальними вузлами, систему локального збереження основних результатів та етапів проміжних обчислень (TCA Controller Storage System), механізм резервування ключових компонентів, а також передбачає мережеве завантаження вузлів у мережі GI (Gigabit Ethernet) за допомогою комутатора SW1. У майстер-вузлі та в slave-вузлах застосовуються одні й ті самі комплектувальні елементи (материнські плати, процесори, мережеві плати Gigabit Ethernet, зовнішні двопортові мережеві плати InfiniBand ConnectX-3 HCA

(MCX354A-FCBT). Зокрема, майстер-вузол додатково обладнано накопичувачем жорсткого диска (SSD) та DVD.

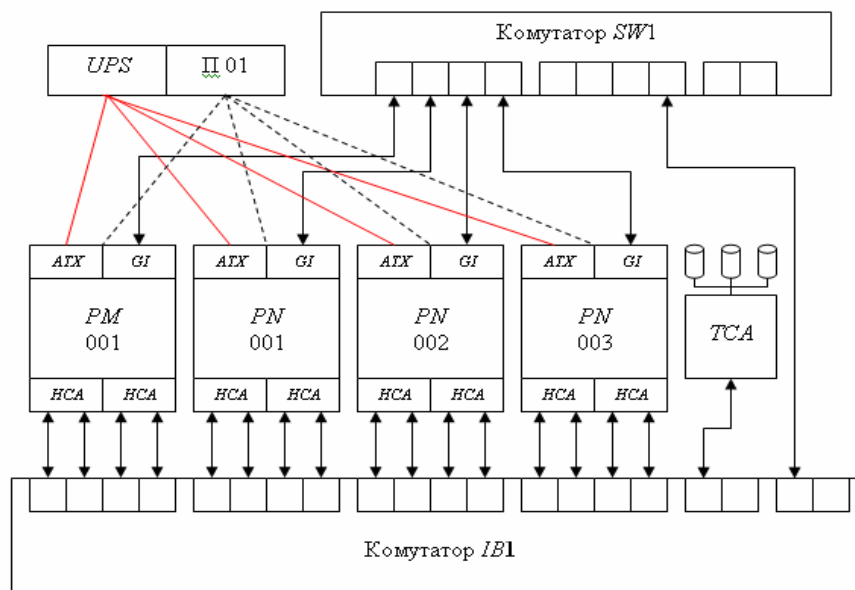


Рисунок 1 – Структурна схема модуля високоефективної багатопроцесорної системи з розподіленою областю обчислень

Комутаційна мережа багатопроцесорної обчислювальної системи підтримує чотири режими її конфігурації. Їх було орієнтовано на реалізацію граничного обміну даними, що відображають особливості задач, які розв’язуються за допомогою такої багатопроцесорної системи.

Розглянемо особливості побудови та налаштування мережевого інтерфейсу багатопроцесорної системи для реалізації вказаних режимів.

Перший режим. Частина прикладних задач передбачає, що граничний обмін даними відбувається між сусідніми обчислювальними вузлами, крім першого та останнього. У такому випадку застосовується топологія комунікаційної мережі типу «лінійка».

Такий режим роботи мережевого інтерфейсу, з одного боку, просто реалізується, з іншого боку, відповідає структурі передачі даних при розв’язуванні багатьох обчислювальних задач (наприклад, при організації конвеєрних обчислень) [8]. Структурна схема режиму роботи реконфігурованої мережі за топологією типу “лінійка” подана на рис. 2.

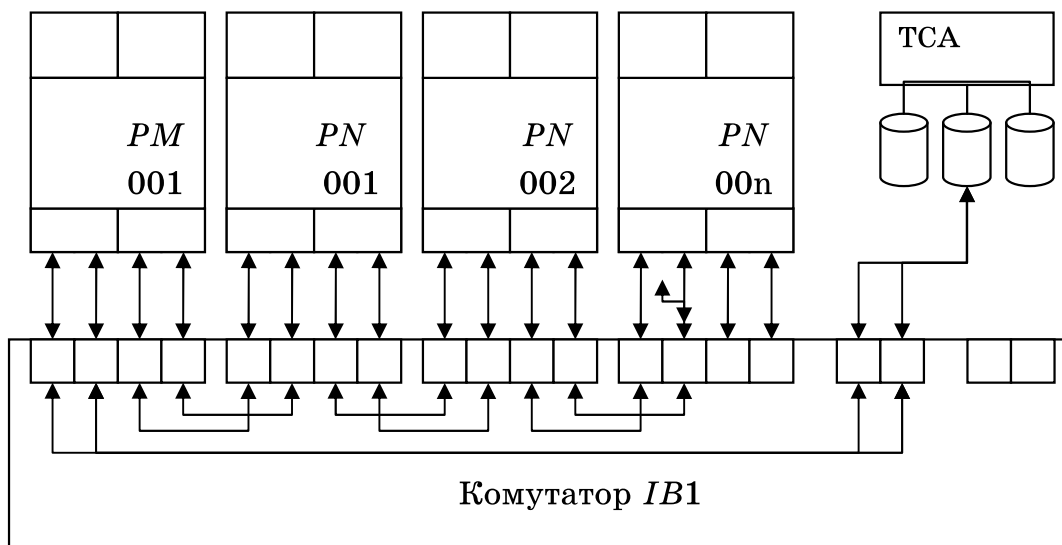


Рисунок 2 – Структурна схема режиму роботи мережевого інтерфейсу за топологією типу “лінійка”

Для формування цього режиму роботи між портами керованого комутатора IB1 формуються “розподілені VLAN” мережі: VS01a між портами 03 і 05, VS01b між портами 04 і 06, VS12a між портами 07 і 09, VS12b між портами 08 і 10, VS23a між портами 11 і 13, VS23b між портами 12 і 14, мережі VS03ab між портами 01, 02, 17 і 18.

Налаштування комутатора IB1 та його конфігурування виконується майстер-вузлом за допомогою двох портів такого комутатора стандарту Gigabit Ethernet (IB1GI.i1-керування, IB1GI.i2-масштабування).

Систему локального збереження результатів та проміжних обчислень TCA Controller Storage System під'єднано до портів 17, 18 керованого комутатора IB1.

Другий режим. Частина прикладних задач передбачає, що граничний обмін даними відбувається тільки між сусідніми обчислювальними вузлами. У такому разі зв'язок між обчислювальними slave-вузлами організовується за топологією типу “кільце”.

Структурна схема режиму роботи реконфігурованої мережі зображена на рис. 3.

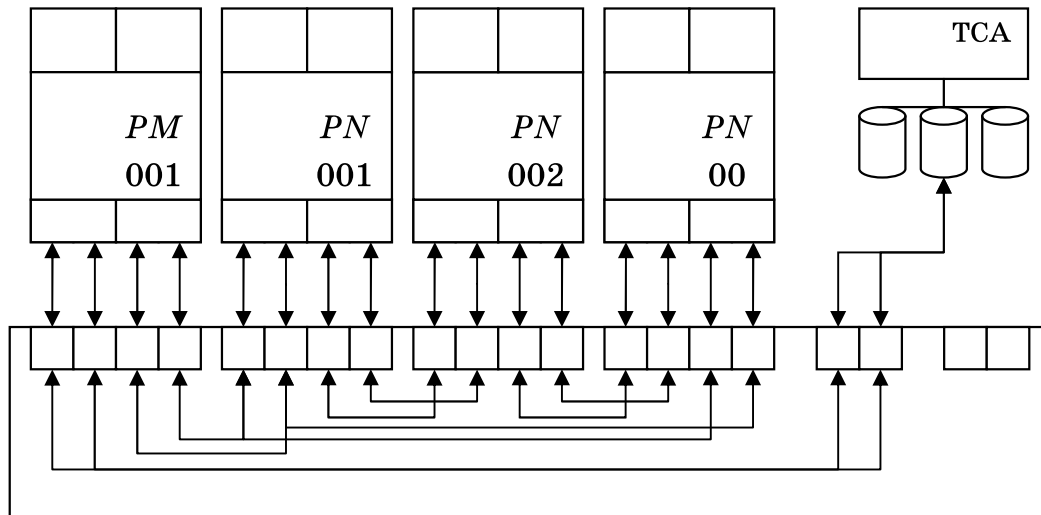


Рисунок 3 – Структурна схема режиму роботи мережевого інтерфейсу за топологією типу “кільце”

Для формування цього режиму роботи між портами керованого комутатора IB1 формуються типу “розподілені VLAN” мережі: VS01a між портами 03 і 06, VS01b між портами 04 і 05, мережі VS12a між портами 07 і 09, VS12b між портами 08 і 10, мережі VS23a між портами 11 і 13, VS23b між портами 12 і 14, мережі VS13a між портами 05 і 15, VS13b між портами 06 і 16, мережі VS00ab між портами 01, 02 і 18, 17. Налаштування комутатора IB1 та його конфігурування виконується майстер-вузлом за допомогою двох портів такого комутатора стандарту Gigabit Ethernet (IB1GI.i1-керування, IB1GI.i2-масштабування). Систему локального збереження результатів та проміжних обчислень TCA Controller Storage System під'єднано до портів 17, 18 керованого комутатора IB1.

Третій режим. Частина прикладних задач передбачає, що граничний обмін даними відбувається між усіма обчислювальними вузлами. Така схема зв'язку між обчислювальними slave-вузлами носить назву типу “зірка”. Дана топологія є найбільш продуктивною, наприклад, при організації централізованих схем паралельних обчислень [8]. Особливість цієї топології полягає в тому, що всі slave-вузли системи мають зв'язок з керуючим master-вузлом. Структурна схема такого режиму роботи реконфігурованої мережі подана на рис. 4.

Для формування цього режиму роботи між портами керованого комутатора IB1 формуються “розподілені VLAN” мережі: VS123a між портами 01, 03, 05, 07, 09, 11, 13, 15, 17 та VS123b між портами 02, 04, 06, 08, 10, 12, 14, 16, 18.

Налаштування комутатора IB1 та його конфігурування виконується майстер-вузлом за допомогою двох портів такого комутатора стандарту Gigabit Ethernet (IB1GI.i1-керування, IB1GI.i2-масштабування). Систему локального збереження результатів та проміжних обчислень TCA Controller Storage System під'єднано до портів 17, 18 керованого комутатора IB1.

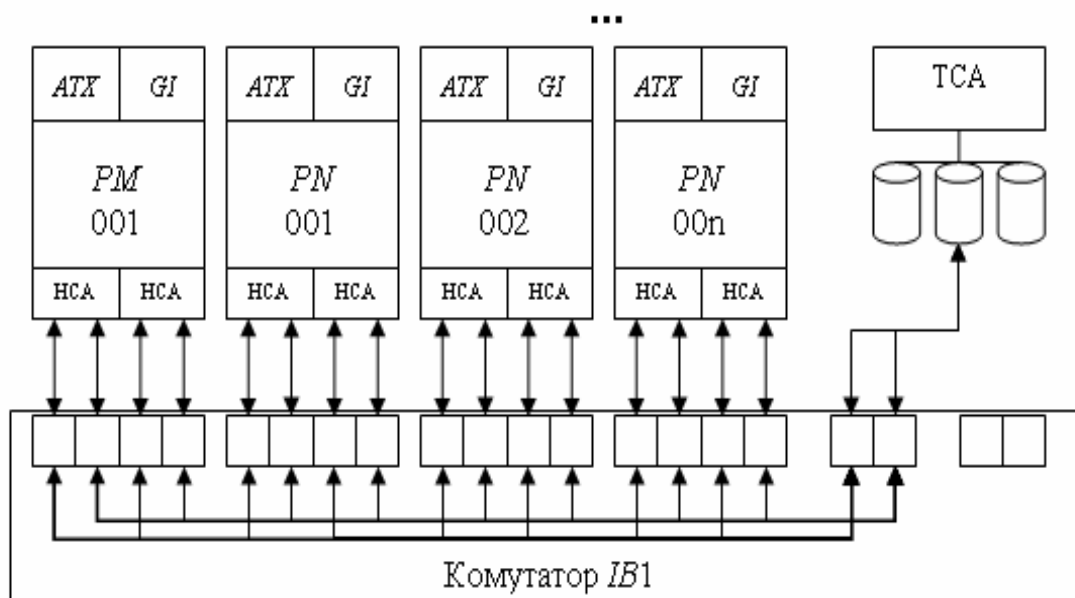


Рисунок 4 – Структурна схема режиму роботи мережевого інтерфейсу за топологією типу "зірка"

Четвертий режим. Частина прикладних задач передбачає, що граничний обмін даними відбувається за прямокутною сіткою між суміжними обчислювальними вузлами за топологією типу "решітка" (Grid network, mesh, 3D-mesh). Вона може бути досить просто реалізована та, крім того, використана більш продуктивна при паралельному виконанні багатьох числових алгоритмів (наприклад, при реалізації методів аналізу математичних моделей, що описуються диференціальними рівняннями в частинних похідних) [8].

Це така топологія, де вузли утворюють регулярну багатовимірну решітку. Крім того, кожне ребро решітки паралельно її осі з'єднує два суміжних вузла. Структурна схема режиму роботи такої реконфігурованої мережі зображена на рис. 5.

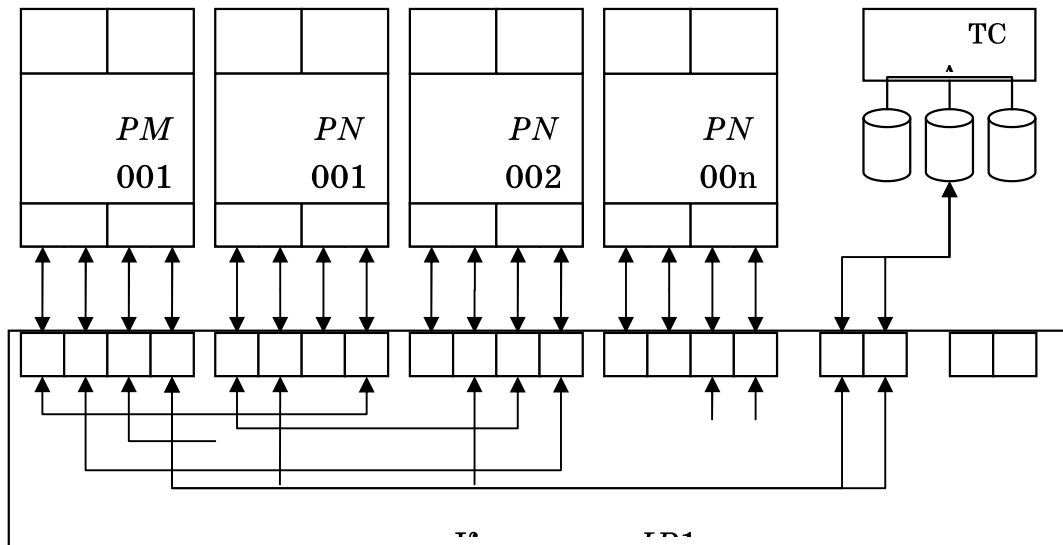


Рисунок 5 – Структурна схема режиму роботи мережевого інтерфейсу за топологією типу “решітка”

Для формування цього режиму роботи між портами керованого комутатора IB1 формуються “розподілені VLAN” мережі: VS01 між портами 01 і 08, мережі VS02 між портами 02 і 12, VS03 між портами 03 і n-3, мережі VS12 між портами 05 і 11, мережі VS23 між портами 09 і n-2, мережі VS3n між портами 16 і n-1.

Налаштування комутатора IB1 та його конфігурування виконується майстер-вузлом за допомогою двох портів такого комутатора стандарту Gigabit Ethernet (IB1GI.i1-керування, IB1GI.i2-масштабування). Систему локального збереження результатів та проміжних обчислень TCA Controller Storage System під’єднано до портів 17, 18 керованого комутатора IB1.

Висновки і перспективи подальших досліджень

Запропоновано нову структуру багатопроесорної обчислювальної системи з розподіленою областю обчислень, яка порівняно з відомими системами враховує специфіку досліджуваного класу задач. При цьому система спрямована для розв’язування широкого кола прикладних задач. Досліджено теоретичні й практичні аспекти перспектив застосування новітніх комунікаційних технологій у багатопроесорних кластерних системах з розподілюваною областю обчислень. При цьому:

1. Запропоновано чотири типи топології комунікаційної мережі багатопроесорної системи: “лінійку”, “кільце”, “зірку”, “решітку”.

2. Висвітлено особливості налаштування мережевого інтерфейсу багатопроцесорної системи для реалізації вказаних режимів роботи комунікаційної мережі системи.

3. Для розв'язування багатьох сильнозв'язаних задач із застосуванням модульної багатопроцесорної системи в режимі її найбільшої продуктивності може бути недостатньо запропонованих лез. З огляду на це автори рекомендують розвивати новий якісний напрям конструювання багатопроцесорних систем – спряження декількох модулів в єдиний обчислювальний комплекс. У своїх подальших дослідженнях автори мають намір розкрити особливості спряження модулів багатопроцесорних систем з метою розширення їх обчислювальних можливостей. При цьому планується розробити варіант, за яким спряжено два модулі багатопроцесорної системи – головний модуль та модуль-розширення.

ЛИТЕРАТУРА

1. Башков Є.О. Високопродуктивна багатопроцесорна система на базі персонального обчислювального кластера / Є.О. Башков, В.П. Иващенко, Г.Г. Швачич // Наукові праці Донецького національного технічного університету. Проблеми моделювання та автоматизації проектування/ – Вип. 9 (179). – Донецьк: ДонНТУ, 2011. – С.312 – 324.
2. Some aspects of design of multiprocessor systems when solving problems with the expandable area calculations / Ivaschenko V.P., Alishov N.I., Shvschych G.G., Tkach M.A. / Mathematics and Computer Science: Journal of Qafqaz University. Baku, Azerbaijan. – Vol. 3. – Numb. 1, 2015 – P. 129 – 136.
3. Перспективы применения современных коммуникационных технологий и исследование их влияния на эффективность многопроцессорных кластерных систем / Е.А. Башков, В.П. Иващенко, Г.Г. Швачич, М.А. Ткач // Наукові праці Донецького національного технічного університету. Інформатика, кібернетика та обчислювальна техніка.– Вип. 14 (188). – Донецьк: ДонНТУ, 2011. – С. 100 – 111.
4. Иващенко В.П. Про проблему узгодження компонентів модульної багатопроцесорної обчислювальної системи в задачі створення нових технологічних процесів / В.П. Иващенко, М.А. Ткач // Информационные технологии в металлургии и машиностроении:

- материалы науч.-техн. конф., Днепропетровск, 2013 г. – Дніпропетровськ, 2013. – С. 18 – 19.
5. Иващенко В.П. Дослідження впливу мережевого інтерфейсу на ефективність багатопроцесорної кластерної системи / В.П. Иващенко, Г.Г. Швачич, М.О. Ткач // Strategy of Quality in Industry and Education : VII International Conference, June, 3 – 10, 2011, Varna; Bulgaria. – Varna, 2008. – V. 2. Proceedings. – P. 289 – 294.
 6. Иващенко В.П. Некоторые аспекты применения сетевой технологии InfiniBand в многопроцессорных моделирующих средах / В.П. Иващенко, Г.Г. Швачич, М.А. Ткач // Информационные технологии в металлургии и машиностроении: междунар. конф.: Днепропетровск, 2014 г.: тез. Докладов.– Днепропетровск: 2014. – С. 95 – 96.
 7. Исследование влияния сетевого интерфейса на эффективность модульной многопроцессорной системы / Е.А. Башков, В.П. Иващенко, Г.Г. Швачич, М.А. Ткач // Наукові праці Донецького національного технічного університету. – Вип. 14 (188). – Донецьк: ДонНТУ, 2011. – С. 89 – 99. – (Інформатика, кібернетика та обчислювальна техніка).
 8. Гергель В.П. Основы параллельных вычислений для многопроцессорных вычислительных систем: учеб. пособие / В.П. Гергель, Р.Г. Стронгин. – Нижний Новгород: ННГУ, 2003. – 184 с.