

ТЕМА

**ОРГАНІЗАЦІЯ І
ФУНКЦІОНУВАННЯ
ІНТЕРФЕЙСІВ
ПАРАЛЕЛЬНИХ КС**

ПЛАН:

1. Огляд технологій для побудова комунікаційного середовища.
2. Організація і функціонування SCI інтерфейсу.
3. Комунікаційне середовище на основі інтерфейсу SCI.
4. Комерційні мережі міжз'єднань.

1. ОГЛЯД ТЕХНОЛОГІЙ ДЛЯ ПОБУДОВИ КОМУНІКАЦІЙНОГО СЕРЕДОВИЩА.

Огляд технологій для побудови комунікаційних середовищ

Конкретна реалізація з'єднання мультикомп'ютерів між собою, з пам'яттю і з зовнішніми пристроями називається ***комунікаційним середовищем комп'ютера.***

Одним з підходів є поділ пам'яті на незалежні модулі та забезпечення можливості доступу різних процесорів до різних модулів одночасно за допомогою використання різного роду комутаторів.

У комп'ютерах сімейства Cray T3D / T3E всі процесори були об'єднані спеціальними високошвидкісними каналами в тривимірний тор, в якому кожен обчислювальний вузол мав безпосередні зв'язки з шістьма сусідами.

1. Огляд технологій для побудови комунікаційного середовища.

Огляд технологій для побудови сполучних мереж

У даний час існує ряд технологій для побудови з'єднаних мереж мультипроцесорів мультикомп'ютерів:

- Масштабований когерентний інтерфейс SCI.
- Мережева технологія Myrinet.
- Комунікаційне середовище Raceway.
- Комунікаційні середовища на базі трансп'ютероподібних процесорів.

Побудови комунікаційних середовищ на основі масштабованого когерентного інтерфейсу SCI

SCI (Scalable Coherent Interface) прийнятий як стандарт в 1992 (ANSI / IEEE Std 1596-1992).

Призначений для досягнення високих швидкостей передачі з малим часом затримки, при цьому забезпечуючи масштабовану архітектуру, що дозволяє будувати системи, що складаються з безлічі блоків.

Являє собою комбінацію шини та локальної мережі, забезпечує реалізацію когерентності кеш-пам'яті, що розміщується в вузлі SCI, за допомогою механізму розподілених директорій, який покращує продуктивність, приховуючи витрати на доступ до віддалених даних в моделі з розподіленою пам'яттю.

1. Огляд технологій для побудови комунікаційного середовища.

Область застосування SCI

Традиційна область застосування SCI - це комунікаційні середовища багатопроцесорних систем. На основі цієї технології побудовані, зокрема, комп'ютери серії hpcLine від Siemens або модульні сервери NUMA-Q від IBM, раніше відомі як Sequent.

Продуктивність передачі даних зазвичай знаходиться в межах від 200 Мбайт / с до 1000 Мбайт / с на відстанях десятків метрів з використанням електричних кабелів та на відстані кілометрів з використанням оптоволокна. SCI зменшує час міжвузлових комунікацій в порівнянні з традиційними схемами передачі даних в мережах шляхом усунення звернень до програмних рівнів.

1. Огляд технологій для побудови комунікаційного середовища.

SCI (Scalable Coherent Interface)

Виробники обладнання: Dolphin Interconnect Solutions і ін.

Показники продуктивності: Для продуктів Dolphin: пікова пропускна здатність - 667 MB / sec ,
Апаратна латентність - 1.46 мкс, в рамках MPI -
близько 4 мкс.

Зразкові ціни: Мережевий адаптер D331 PCI-64/66 - PCI-SCI (для з'єднання в кільце) - \$ 975.
Мережевий адаптер D334 PCI-64/66 2D 1-Slot - PCI-SCI (для з'єднання в 2-мірний топ) - \$ 1280.
Мережевий адаптер D336 PCI-64/66 3D 1-Slot - PCI-SCI (для з'єднання в 3-мірний топ) - \$ 1 005.
Комутатор D535 MS-8X - 8 (до 8 портів) - \$ 4980. SAN-кабель 1 м - \$ 100.

1. Огляд технологій для побудови комунікаційного середовища.

Характеристики SCI

Для HPC-кластерів компанії Dolphin і SCALI пропонують комплекти "Wulfkit" - \$ 1595 за комплект на вузол (при об'єднанні в 3-мірний топ) і \$1095 (при об'єднанні в 2-мерий топ).

Коментарі: SCI (ANSI / IEEE 1596-1992) - добре стандартизована технологія. Крім стандартного мережевого середовища, SCI підтримує побудову систем з пам'яттю та з когерентністю кешей.

Порядок обміну даними

Зазвичай звернення до даних, фізично розташованих у пам'яті іншого обчислювального вузла, які не знаходяться в кеші, призводить до формування запиту на віддалений вузол для отримання необхідних даних, які протягом декількох мікросекунд доставляються в локальний кеш, і виконання програми продовжується.

Особливості SCI

SCI має можливість переносити пакети інших протоколів.

SCI можуть використовувати для з'єднання комутаторів або ж з'єднуватися в кільце.

Протокол передачі даних забезпечує гарантовану доставку даних.

Дана технологія оптимізована для роботи з динамічним трафіком, проте може бути менш ефективна при роботі з великими блоками даних.

Застосування SCI

- Протокол SCI досить складний, він містить великі можливості по управленню трафіком, але вимагає наявності розвиненого програмного забезпечення.
- На комунікаційній технології SCI заснована система зв'язку гіпервузлів CTI (Convex Torroidal Interconnect) в системах HP / Convex Exemplar X-class, крім того на ній побудовані кластерні системи SCALI Computer, системи сімейства hpcLine компанії Siemens, а також cc-NUMA сервера Data General і Sequent
- Традиційна **область застосування SCI** - це комунікаційні середовища багатопроцесорних систем.
- На основі цієї технології побудовані, зокрема, комп'ютери серії hpcLine від Siemens або модульні сервери NUMA-Q від IBM, раніше відомі як Sequent.

Можливості SCI

Модульні SCI комутатори Dolphin дозволяють споживачам будувати масштабовані, кластерні рішення класу підприємства на платформах Windows NT / 2000 / XP, Linux, Solaris, VxWorks, LynuxWorks і NetWare з використанням стандартизованого обладнання та програмного забезпечення.

Програмна підтримка

- ✓ Драйвери для Linux, Windows NT, Solaris. ScaMPI - реалізація MPI компанії Scali Computer для систем на базі SCI. SISCO API - інтерфейс програмування нижнього рівня.
- ✓ Добре стандартизована технологія.
- ✓ Крім стандартного мережевого середовища, SCI підтримує побудову систем з пам'яттю та з когерентністю кешей.

1. Огляд технологій для побудови комунікаційного середовища.

2. ОРГАНІЗАЦІЯ І ФУНКЦІОНУВАННЯ SCI ІНТЕРФЕЙСУ.

Організація та функціонування SCI

- Функціональна організація вузла SCI
- Архітектура SCI
- Протоколи когерентності
- Реалізація GigaRing Channel
- Реалізація SCI фірми Dolphin

Механізм когерентності

SCI передбачає реалізацію когерентності за допомогою стандартно організованою кеш-пам'яттю, що розміщується в вузлі SCI.

Ця кеш-пам'ять розташовується в інтерфейсі між VM та вузлом і вписується в механізм реалізації когерентності VM наступним чином.

Якщо в попередніх рівнях ієрархії пам'яті VM виявляється відсутність необхідних даних, то проводиться пошук цих даних в кеш-пам'яті вузла.

Реалізація комунікаційного середовища SCI

На структурному рівні комунікаційне середовище складається з трьох компонентів:

- адаптерів, які здійснюють інтерфейс між VM і мережею передачі пакетів;
- комутаторів мережі передачі пакетів;
- кабелів, службовців для під'єднання вхідних і вихідних каналів (лінків) адаптерів до портів комутатора і з'єднання комутаторів один з одним для обладнання необхідного числа портів і заданої конфігурації мережі.

Комунікаційне середовище MYRINET

Технологія Myrinet заснована на використанні багатопортових комутаторів при обмежених декількома метрами довжинах зв'язків вузлів з портами комутатора.

Вузли в Myrinet з'єднуються один з одним через комутатор (до 16 портів).

Myrinet може одночасно передавати кілька пакетів, кожен з яких йде зі швидкістю, близькою до 2 Гбіт / с.

Myrinet найчастіше використовують як локальну мережу (LAN) порівняно невеликого фізичного розміру, пов'язуючи разом комп'ютери усередині кімнати або будівлі.

Myrinet особливо популярний для об'єднання комп'ютерів в кластери.

Технологія Myrinet

З'єднання обчислювальних вузлів з адаптерами **Myrinet** в мережу відбувається за допомогою комутаторів, які мають зараз 4, 8, 12 або 16 портів.

Комутатори як і мережеві адаптери побудовані на спеціалізованих мікропроцесорах LANai фірми Myricom.

Середовище Myrinet сильно відрізняється від багатьох інших середовищ передачі, зокрема SCI, простотою концепції і апаратною реалізацією протоколів.

Технологія Myrinet дає великі можливості масштабування мережі і в даний час широко використовується при побудові високопродуктивних обчислювальних

2. Організація і функціонування SCI інтерфейсу,

Виробники обладнання: Myricom

Показники продуктивності: Пікова пропускна здатність - 2 Gbit / sec, повний дуплекс.

Латентність - близько 4 мкс. На MPI-додатках латентність становить близько 10 мксек, швидкість передачі даних - до 200 MB / sec (до 400 MB / sec на дуплексних операціях).

Програмна підтримка: Драйвера для Linux (Alpha, x86, PowerPC, UltraSPARC), Windows NT (x86)

Технологія Myrinet дає високі можливості масштабування мережі і в даний час дуже широко використовується при побудові високопродуктивних кластерів

пропускну здатність на рівні 1 Гбайт/с; середовище передачі створюється за допомогою комутатора фірми Cypress і відповідних мережевих адаптерів.

Комутатор має 6 портів, пропускна здатність кожного порту становить 160 Мбайт/с. Порт складається з 32 сигнальних ліній даних і 5 керуючих ліній. При транзакції середовищем Raceway попередньо встановлює з'єднання, затримка в комутаторі при встановленні з'єднання складає приблизно 125 нс.

Структури обчислювальних систем, що створюються за допомогою Raceway, аналогічні тим, які застосовуються в разі використання мережі Myrinet або комутаторів і адаптерів SCI. Різниця полягає в кількості портів комутаторів, форматах переданих пакетів і в протоколах.

Комунікаційне середовище Raceway прийняте в якості стандарту (ANSI/

2. Організація і функціонування SCI інтерфейсу,

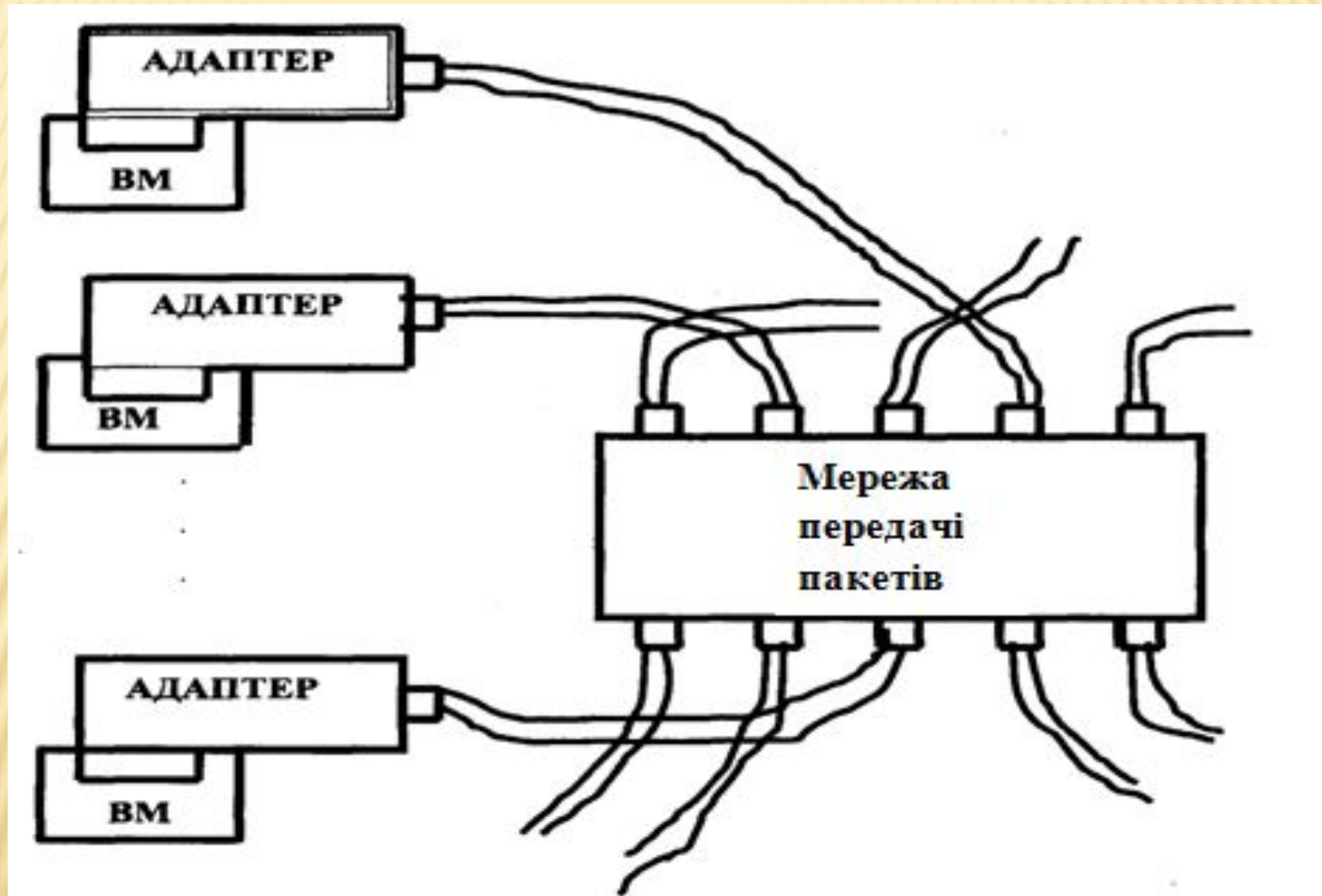
Призначення SCI

Стандарт SCI забезпечує побудову легкого в реалізації, що масштабується, ефективного у вартісному аспекті комунікаційного середовища для об'єднання процесорів і пам'ятей.

Для створення розподіленої мережі робочих станцій.

Для організації введення/виведення суперЕОМ, високопродуктивних серверів і робочих станцій на базі сучасних мікропроцесорів.

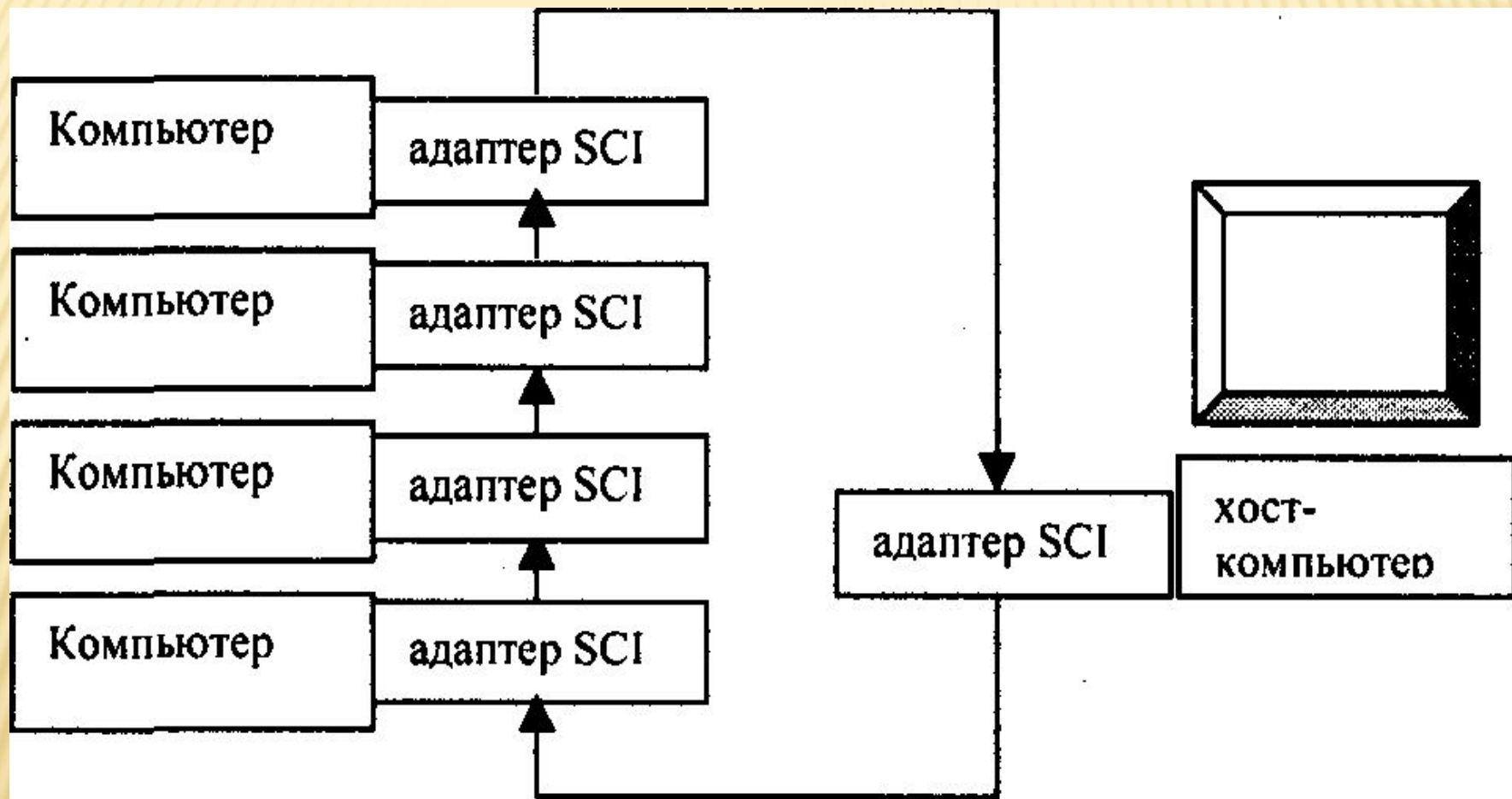
Склад комунікаційного середовища SCI



2. Організація і функціонування SCI інтерфейсу.

3. КОМУНІКАЦІЙНЕ СЕРЕДОВИЩЕ НА ОСНОВІ ІНТЕРФЕЙСУ SCI.

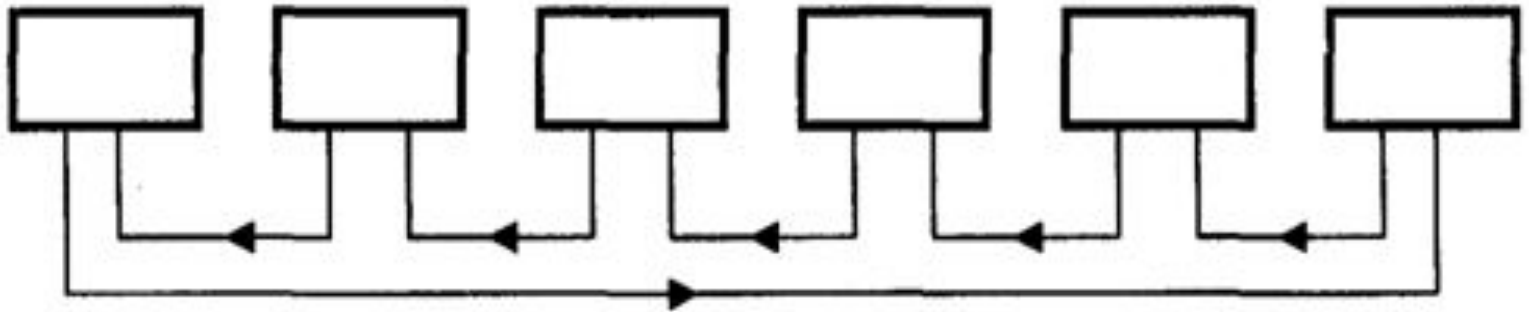
Комунакацiйне середовище на основi iнтерфейсу SCI



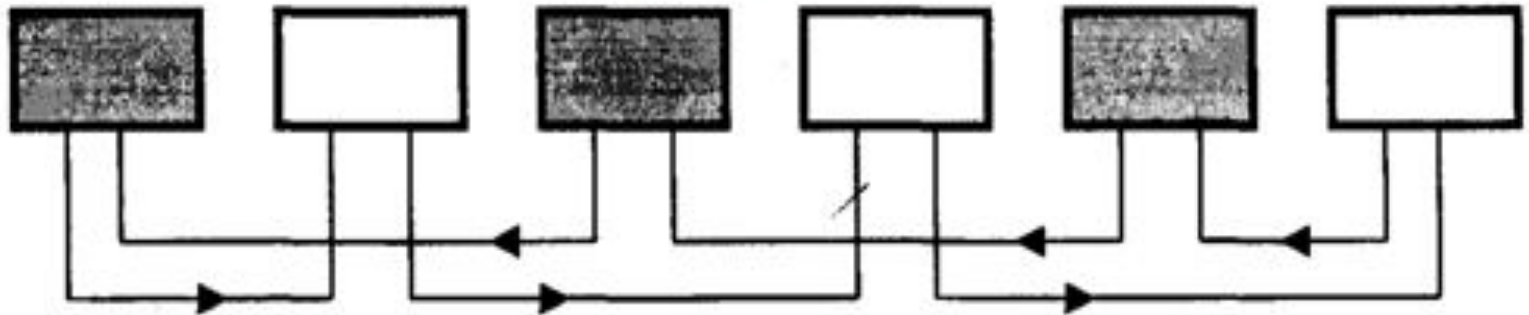
3. Комунакацiйне середовище на основi iнтерфейсу SCI.

Одновимірний тор (кільце) SCI

Послідовне кільце

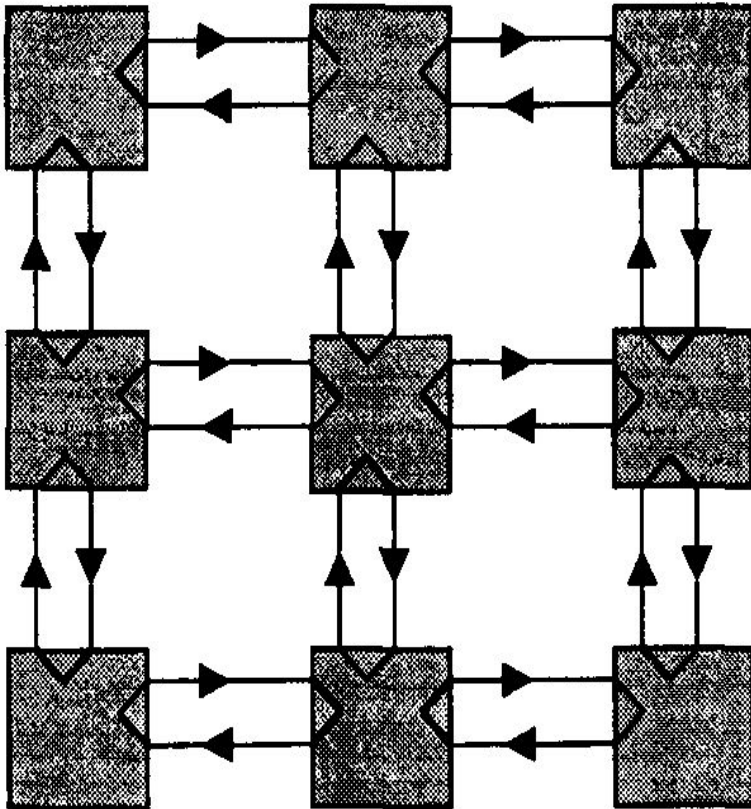


Кільце з чергуванням

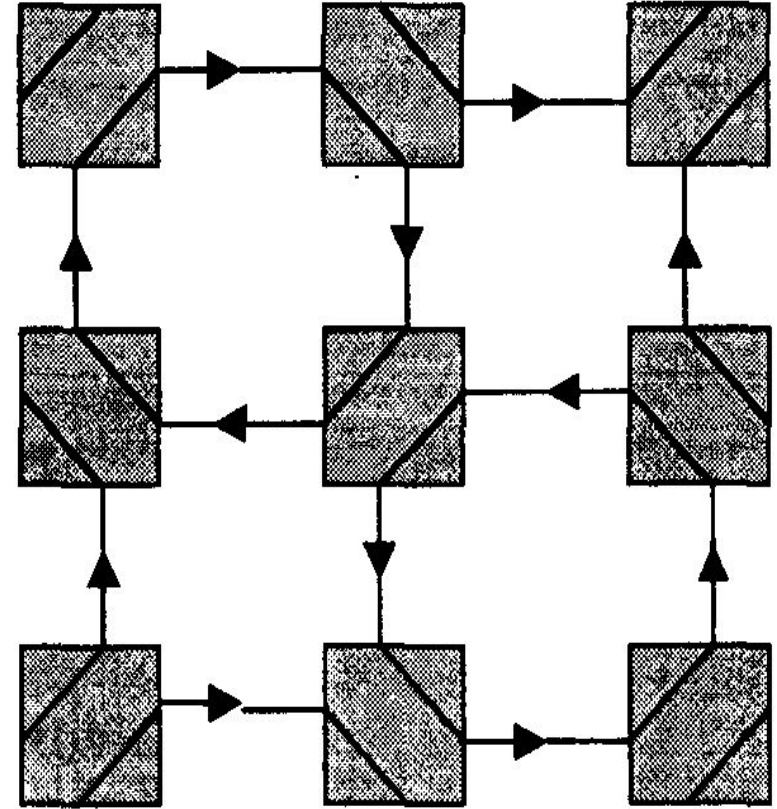


Двовимірна решітка SCI

Двунаправленна



Однонаправленна



4. КОМЕРЦІЙНІ МЕРЕЖІ МІЖЗ'ЄДНАНЬ.

Комерційні мережі міжз'єднань

Прикладом такої технології є система Ethernet. Існує три версії цієї системи: classic Ethernet, fastEthernet і gigabit Ethernet.

Вони працюють зі швидкістю 10, 100 і 1000, 10000 Мбіт/с (1,25, 12,5 і 125 Мбайт/с), відповідно. Всі вони сумісні щодо середовища, формату пакетів і протоколів.

ATM (Asynchronous Transfer Mode - асинхронний режим передачі). Технологія ATM має швидкості передачі даних 155 Мбіт/с (пізніше 622 Мбіт/с).

Багато компаній зараз випускають знімні плати для персональних комп'ютерів зі швидкістю передачі даних 155 Мбіт/с або 622 Мбіт/с. Друга швидкість, OC-12, добре підходить для мультикомп'ютерів.

16-процесорна КС на основі АТМ комутатори

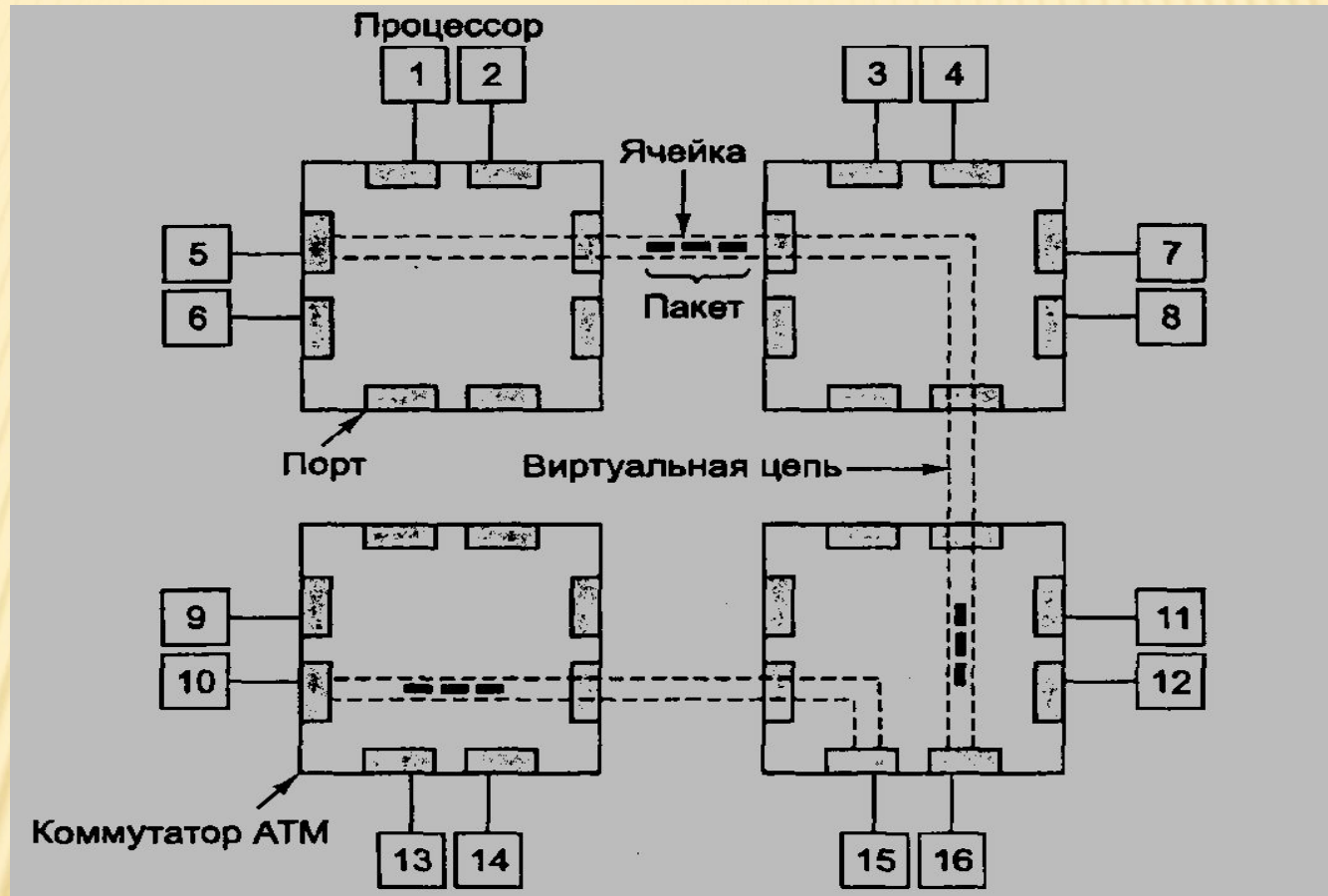


Рисунок – 16 процесорів, зв'язаних чотирма перемикачами АТМ. Пунктиром показано два віртуальні ланцюги (канали)

ВИСНОВОК

Конкретна реалізація з'єднання комп'ютерів між собою, з пам'яттю і з зовнішніми пристроями називається **комунікаційним середовищем мультикомп'ютера**. Одна з найпростіших реалізацій - це використання загальної шини, до якої підключаються як процесори, так і пам'ять.

Збільшення, навіть незначного числа пристроїв на шині викликає помітні затримки при обміні з пам'яттю і катастрофічне падіння продуктивності системи в цілому.

В даний час існує ряд технологій для побудови сполучних мереж мультипроцесорів і мультикомп'ютерів: 1) Масштабований когерентний інтерфейс SCI; 2) Мережева технологія Myrinet; 3) Комунікаційне середовище Racoba; 4) Комунікаційні середовища на базі трансп'ютероподібних процесорів.

ВИСНОВОК

SCI Призначений для досягнення високих швидкостей передачі з малим часом затримки, при цьому забезпечуючи масштабуючу архітектуру, що дозволяє будувати системи, що складаються з безлічі блоків

Технологія Myrinet дає великі можливості масштабування мережі і в даний час широко використовується при побудові високопродуктивних обчислювальних кластерів

Комунікаційне середовище Raceway забезпечує пропускну здатність на рівні 1 Гбайт / с; середовище передачі створюється за допомогою комутаторів фірми Supress і відповідних мережевих адаптерів.

ВИСНОВОК

Комунікаційні мережі на основі трансп'ютерів знаходять застосування при побудові обчислювальних кластерів, наприклад, МВС-100.

Важливою характеристикою комунікаційної мережі є час латентності - затримка початку передачі даних після її ініціалізації.

Для побудови мультикомп'ютерів і кластерів використовуються сучасні мережеві технології Fast Ethernet, Gigabit Ethernet, АТМ.