

## Курс «Паралельні алгоритми: побудова та аналіз».

### Лекція 3. Особливості комунікаційного середовища паралельних обчислювальних систем.

#### План.

1. Структурна схема комунікаційного середовища.
2. Мережеві адаптери.
3. Комунікаційні мережі: адресація в мережах, мережеві протоколи, топологія комунікаційної мережі, мережеві концентратори.
4. Основні типи топологій комунікаційних мереж, діаметр комунікаційної мережі, мережі з фіксованою топологією та реконфігуровні, регулярні та нерегулярні мережі.
5. Мережеві комутатори: прості та складні комутатори.
6. Основні характеристики комунікаційних мереж.

**1. Структурно комунікаційне середовище** паралельної обчислювальної системи складається із **трьох** таких **компонент**:

- мережеві адаптери;
- комунікаційна мережа;
- мережеві комутатори.

Схематично це виглядає так, як подано на рис. 1. Зауважимо, що **комутатор** – це пристрій для комутації (встановлення зв'язку) двох і більше процесорів (або обчислювальних вузлів).

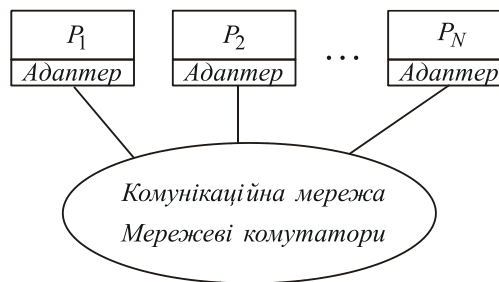


Рис. 1

Тут  $P_i$  –  $i$ -й вузол мережі (процесор або декілька процесорів, які утворюють обчислювальний вузол).

**2.** Під час обміну даними між вузлами мережі, що не є ближніми сусідами, відбувається передача (транспортування) даних через проміжні вузли. Очевидно, що у цих вузлах повинні знаходитися якісь **апаратні засоби**, що звільняють процесори вузлів від участі в передачі (транспортуванні) даних. Однією із основних функцій **мережевих адаптерів** є передавальна, тобто забезпечення обмінів даними в мережі.

Мережевий адаптер складається із:

- приймальнопередавальної частини, зв'язаної з вузловим процесором мережі;
- приймальнопередавальної частини, зв'язаної з мережею;
- узгоджувальних буферів (пристроїв, що використовуються для узгодження швидкостей обміну, розмірів блоків даних тощо).

Схематично структуру мережевого адаптера зображено на рис. 2.

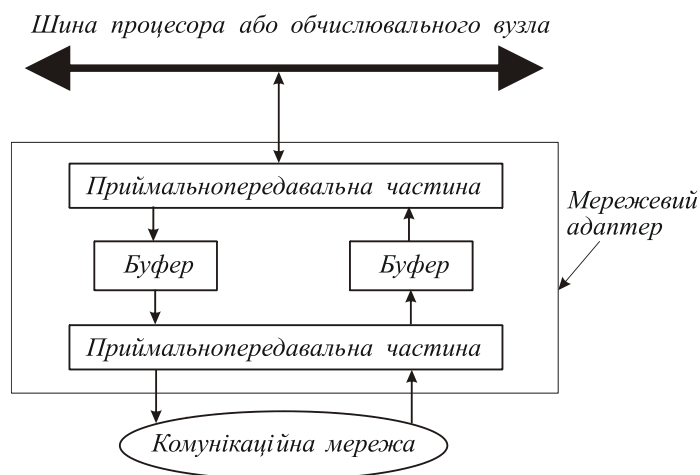


Рис. 2

Основною функцією **приймальнопередавальної** частини адаптера, зв'язаної з **комунікаційною мережею**, є реалізація протоколу мережі, а головною функцією такої ж частини адаптера, зв'язаної з **шиною вузла**, є реалізація протоколу цієї шини.

**Функцією** мережевого адаптера може бути також забезпечення **когерентності пам'яті** вузла мережі.

Якщо на вузловому процесорі одночасно функціонують декілька прикладних процесів, то мережевий адаптер може стати проблемою (**вузьким місцем**) для системи, оскільки за доступ до нього можуть змагатися прикладні процеси і процеси операційної системи. Тому інколи вузловий процесор забезпечують **двома мережевими адаптерами**: один – для процесів операційної системи; другий – для прикладних процесів.

**3. Комунікаційні мережі** діляться на **широкомасштабні** комунікаційні мережі (Wide Area Networks) – WAN-мережі та **локальні** комунікаційні мережі (Local Area Networks) – LAN-мережі.

**Адресація в мережах.** Множина адрес, допустимих у межах деякої схеми адресації, називається **адресним простором** комунікаційної мережі. Адресний простір може мати **лінійну або ієрархічну** організацію.

У **лінійному** адресному просторі комунікаційної мережі множина адрес не є структурованою, всі адреси є рівноправними. Прикладом лінійного адресного простору є простір MAC (Media Access Control)-адрес, які призначені для однозначної ідентифікації мережеских інтерфейсів у локальних мережах. Ці адреси розподіляються між виробниками спеціальним комітетом IEEE (Інститут інженерів з

електротехніки і електроніки) і вбудовуються в апаратуру (адаптери та інше мережеве обладнання) під час її виготовлення. Такі адреси ще називають апаратними, формат їх – 6 байт і позначаються вони двійковим або шістнадцятковим кодом (наприклад, 11 A0 17 3B FD 01).

Типовим представником **ієрархічного** адресного простору комунікаційної мережі є простір мережевих 4-байтових IP (Internet Protocol)-адрес, що використовуються для адресації вузлів у WAN-мережах. IP-адреси є дворівневими: адреса ділиться на старшу частину (номер мережі) і молодшу частину (номер вузла в мережі). Така організація IP-адрес дозволяє передавати повідомлення між мережами лише на підставі номера мережі. Номер вузла використовується після доставки повідомлення в середині мережі. IP-адреса однозначно ідентифікує окремий вузол, групу вузлів або цілу мережу та подається у десятковій (ОС), двійковій (адміністрування) або шістнадцятковій (дуже рідко) формі. Десяткова IP-адреса 102.56.187.5 означає вузол з номером 56.187.5, який знаходиться в мережі з номером 102.

**Мережеві протоколи.** Правила, що визначають послідовність і формат повідомлень, якими обмінюються вузли мережі, називаються **мережевим комунікаційним протоколом**.

У **WAN-мережах** зазвичай використовуються протоколи **комутації пакетів**. Комутація пакетів: у пункті відправлення повідомлення розбивається на порції (пакети, декілька кілобайт); пакети висилаються за адресою призначення незалежно один від одного; у пункті призначення здійснюється збір повідомлення із пакетів. Кожен пакет має свій заголовок, в якому вказується, як мінімум, адреса вузла-отримувача і номер пакета.

У паралельних обчислювальних системах використовуються LAN-мережі і як **протоколи комутації пакетів**, так і протоколи **комутації каналів**. Зв'язок з комутацією каналів – це зв'язок між двома вузлами мережі, між якими спочатку встановлюється з'єднання, перш ніж вони почнуть обмін інформацією (на протязі всього сеансу це з'єднання використовується лише тими вузлами).

Зазвичай локальні мережі будують на базі найбільш поширеної технології Ethernet. Технології Ethernet, Fast Ethernet, Gigabit Ethernet, 10GE забезпечують швидкість передачі даних відповідно 10, 100, 1000 і 10000 Mbit/s. Оскільки комунікаційне середовище є розділюваним ресурсом, то можливі ситуації, коли два вузли намагаються одночасно передати кадр даних. Така ситуація призводить до «викривлення» інформації і називається **комунікаційною колізією**. У технології Ethernet передбачено механізм для вирішення таких ситуацій.

Головний **недолік мереж**, побудованих за технологією **Ethernet**, полягає в тому, що в них за одиницю часу може передаватися лише **один пакет** даних. Для забезпечення високої продуктивності таких мереж їх не слід завантажувати більш, ніж на 30–40 % від їх максимальної пропускної здатності.

Для побудови високошвидкісних комунікаційних мереж використовували, наприклад, технології

100 Gigabit Ethernet,  
InfiniBand (100 Gbit/s, після 2023 р. – 250 Gbit/s),

Intel Omni-Path (100 Gbit/s, у 2019 р. відміна створення технології з продуктивністю 200 Gbit/s).

Важливою характеристикою комунікаційної мережі є її **топологія**, під якою розуміють спосіб з'єднання процесорних вузлів між собою каналами передачі даних. Топологію комунікаційної мережі прийнято зображати у вигляді графу, вершини якого відповідають процесорним вузлам, а ребра – каналам зв'язку.

Практично кожна паралельна обчислювальна система має свою **оригінальну топологію** комунікаційної мережі. Це пов'язано з тим, що неможливо з'єднати багато вузлів швидкими каналами зв'язку за типом «кожен з кожним». Тому в паралельних системах швидкий обмін інформацією здійснюється не між усіма вузлами, а лише між деякими. Всі інші обміни здійснюються відносно повільно.

Вказаний розподіл зв'язків на швидкі та повільні, по-перше, породжує **різноманіття топологій** комунікаційних мереж, а, по-друге, є однією із головних причин обмеження класу алгоритмів, які можуть бути ефективно реалізовані на заданій паралельній обчислювальній системі.

**Швидкі обміни** здійснюються по некомутованих лініях зв'язку, а **повільні** – або через мережеві комутатори, або з допомогою послідовних швидких обмінів.

Під час побудови LAN-мереж широко використовують також мережеві **концентратори** (concentrator, hub). Концентратор на один із своїх портів приймає сигнали від одного із вузлів мережі і синхронно передає їх на всі свої решту портів, з'єднаних з іншими вузлами мережі (див. рис. 3).

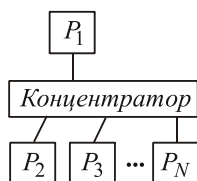


Рис. 3

**4. Діаметром  $d$**  комунікаційної мережі називається максимальна відстань між двома її довільними вузлами. **Відстань** між вузлами комунікаційної мережі – це кількість ребер, які утворюють найкоротший шлях між двома вузлами в графі мережі.

За топологією комунікаційні мережі **поділяються** на мережі з **фіксованою** топологією і на **реконфігуровні** мережі. З іншого боку ці мережі діляться на **регулярні і нерегулярні**. Регулярні мережі – це мережі, в яких кожен вузол (вершина) має однакову кількість зв'язків (ребер).

Найчастіше в паралельних системах використовують регулярні мережі з фіксованою топологією. Розглянемо найбільш розповсюджені топології мереж.

**Комунікаційна мережа з топологією типу «шина»** (див. рис. 4).

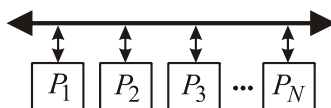


Рис. 4

Тут всі процесори з'єднані за допомогою (високошвидкісної) шини. Перевагою такої мережі є дуже мала кількість ліній зв'язку, але при цьому можливі затримки (конфлікти на шині) під час використання шини декількома процесорами. Це може стати серйозною проблемою під час збільшення кількості процесорів. **Діаметр** такої мережі  $d = 1$ . Тут і надалі  $N$  – це кількість процесорів у системі. Зазначимо, що **шину** можна інтерпретувати як **мережевий комутатор** з часовим розділенням.

**Комунікаційна мережа з топологією типу «лінійка»** (див. рис. 5).

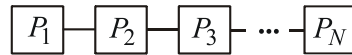


Рис. 5

У мережі з цією топологією кожен процесор зв'язаний з двома сусідами (за винятком кінцевих процесорів, які мають лише по одному з'єднанню). **Перевага** цієї топології – в її **простоті** (із кожного процесора виходить не більш як дві лінії зв'язку). З іншого боку, згадана топологія відповідає структурі передачі даних під час розв'язання багатьох обчислювальних задач. Серйозним **недоліком** є те, що дані, перш ніж досягти свого кінцевого призначення, можливо, повинні будуть пройти через декілька проміжних процесорів. **Діаметр** такої мережі  $d = N - 1$ .

**Комунікаційна мережа з топологією типу «кільце» (коло)** (див. рис. 6).

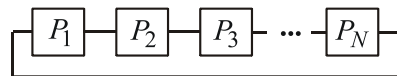


Рис. 6

Ця топологія одержується з топології типу «лінійка» унаслідок з'єднання між собою першого і останнього процесора, при цьому **діаметр**  $d \approx N/2$ .

**Комунікаційна мережа з топологією типу «двовимірна решітка» (гратка)** (див. рис. 7).

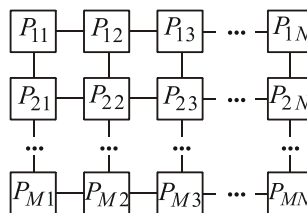


Рис. 7

**Перевагою** такої мережі є її **простота** реалізації, а **недолік** полягає в тому, що під час обміну між віддаленими процесорами дані повинні пройти через низку проміжних процесорів. **Діаметр** такої мережі  $d = N + M - 2$ . Із цієї топології легко одержати **топологію типу «тор»**. Для цього досить з'єднати між собою граничні процесори так: по «горизонталі» –  $P_{11}, P_{1N}; P_{21}, P_{2N}; \dots; P_{M1}, P_{MN}$  і по «вертикалі» –  $P_{11}, P_{M1}; P_{12}, P_{M2}; \dots; P_{1N}, P_{MN}$ .

**Комунікаційна мережа «повністю зв'язана»** (див. рис. 8).

Розглянемо приклад такої мережі для  $N = 6$ .

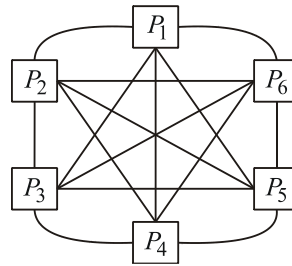


Рис. 8

У цій мережі між будь-якою парою процесорів існує безпосередній зв'язок. Повна зв'язаність мережі для  $N$  процесорів вимагає, щоб із кожного процесора виходили  $N - 1$  ліній зв'язку, що є **непрактичним** при великому  $N$ . Для цієї мережі  $d = 1$ .

**Мережа з топологією типу «бінарне дерево»** (див. рис. 9).

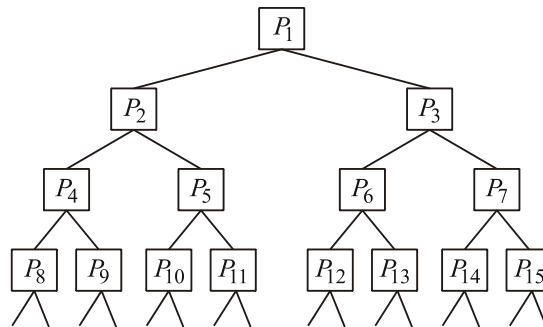


Рис. 9

Діаметр цієї мережі  $d \approx 2(\log_2 N - 1)$ .

**Комунікаційна мережа з топологією типу «зірка»** (див. рис. 10).

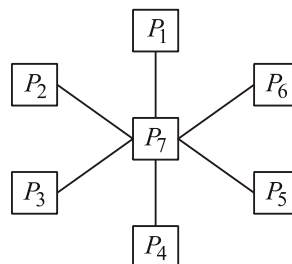


Рис. 10

Мережа з такою топологією є **ефективна**, наприклад, під час організації централізованих схем паралельних обчислень, при цьому  $d = 2$ .

**Комунікаційна мережа з топологією типу «гіперкуб» (трьохвимірний гіперкуб)** (див. рис. 11).

**Діаметр** такої мережі  $d = \log_2 N$ ,  $N = 2^p$ , де  $p = d$  – кількість вимірів гіперкуба. У цій мережі її діаметр **повільно** збільшується із збільшенням кількості процесорів.

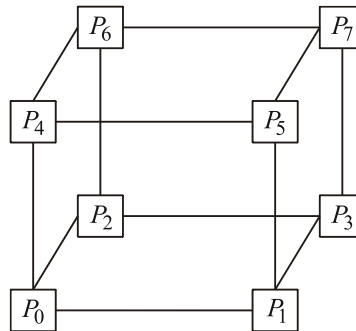


Рис. 11

Наприклад, для  $p = 6$ ;  $N = 2^6 = 64$  діаметр  $d = 6$ , а для  $p = 10$ ;  $N = 2^{10} = 1024$  –  $d = 10$ .

Зауважимо, що в  $p$ -вимірному гіперкубі кожен процесор безпосередньо зв'язаний рівно з  $p$  сусідами.

Відзначимо, що обчислювальні системи з **великою кількістю** процесорів зазвичай мають **ієрархічну** комунікаційну мережу. У такій мережі всі процесори розбиваються на групи з однаковою і відносно невеликою кількістю процесорів ( $\approx 32 - 64$ ) в кожній. Фізично група процесорів монтується в **один блок**. Кожен з таких блоків має комунікаційну мережу однакової топології. Групи процесорів (блоки) об'єднуються між собою комунікаційною мережею такої ж або іншої топології.

Важливою **властивістю топології** комунікаційної мережі є можливість забезпечення **масштабованості** обчислювальної системи. З цього погляду **добре масштабуються** системи, що мають топологію мережі типу «лінійка», «кільце», «шина». Тут нарощування системи можливе по одному процесору. У мережі з топологією «повністю зв'язана» долучення одного процесора вимагає додавання  $N$  каналів зв'язку ( $N$  – кількість процесорів у вихідній мережі), що утруднює масштабування системи.

Нарощування кількості процесорів у обчислювальних системах з мережею топології «плоска  $M \times N$  двовимірна решітка» або «тор» можливе лише квантами по  $M$  або  $N$  процесорів.

**5. Мережеві комутатори** поділяють на **прості і складні**. Перші потребують менше часу на комутацію, але можуть бути використані лише для побудови систем із малою кількістю вузлів. **Складні** комутатори будують на підставі **об'єднання простих** комутаторів.

**Прості комутатори** поділяють на комутатори з **часовим** розділенням і комутатори з **просторовим** розділенням.

Комутаторами з **часовим розділенням** є шини. Оскільки шина є розділюваним ресурсом, то процесорні вузли вимушені конкурувати за доступ до цього ресурсу. Для розв'язання **конфліктів** між вузлами використовують різні алгоритми



**арбітражу.** Шина є доволі дешевим і надійним комутатором. Однак, під час комутації великої кількості коротких повідомлень накладні витрати на арбітраж стають значними, що суттєво обмежує пропускну її здатність.

Комутатор з **просторовим розділенням** ще називають координатним комутатором. Свою назву «**координатні комутатори**» прості комутатори одержали тому, що схему комутації з допомогою цього  $3 \times 3$  комутатора можна подати так, як на рис. 12.

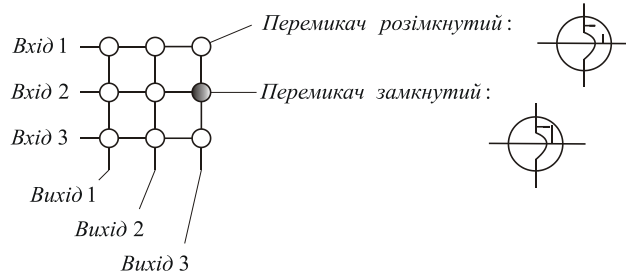


Рис. 12

Такі комутатори потребують мінімального часу на комутацію, але мають **високу складність**, а тому і **не досить високу надійність**. Кількість перемикачів, що дозволяють з'єднувати  $p$  процесорів з  $p$  пристроями пам'яті, в такому комутаторі дорівнює  $p^2$ , тому при великих  $p$  така схема з'єднань стає непрактичною.

**Складні комутатори.** Головна **ідея** створення таких комутаторів полягає в **об'єднанні простих** комутаторів каналами типу «пункт-пункт» («інстанція-інстанція»). Складний комутатор потребує менше обладнання, ніж простий з такою ж кількістю входів-виходів. Однак складному комутатору при цьому потрібно більше часу на комутацію, який росте пропорційно до кількості рівнів комутації.

Найчастіше складні комутатори будуються на основі  $2 \times 2$  простих комутаторів з просторовим розділенням.

Як складний комутатор розглянемо **комутатор Клоза**. Цей комутатор складається із трьох рівнів комутації: вхідного, проміжного та вихідного. Комутатор Клоза має однакову кількість входів та виходів  $n = m \times l$  і складається із:

- $m$  вхідних  $l \times l$  простих комутаторів;
- $m$  таких же вихідних комутаторів;
- $l$  проміжних  $m \times m$  комутаторів.

Схема такого комутатора для  $m = 3$ ;  $l = 4$  зображена на рис. 13.

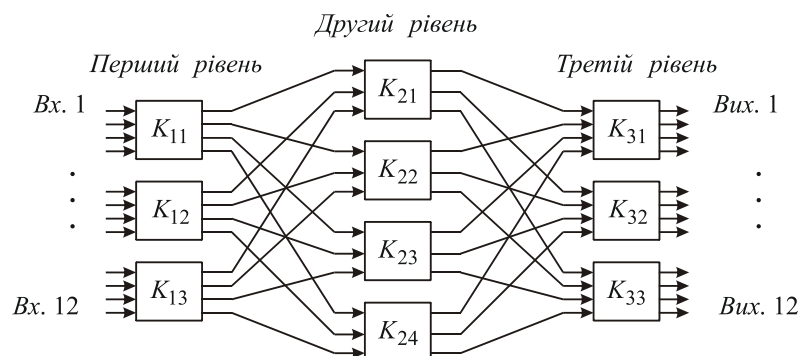


Рис. 13



Тут  $K_{ij}$  ( $i = \overline{1,3}; j = \overline{1,3}$ );  $K_{24}$  – прості комутатори. З'єднання простих комутаторів в комутаторі Клоза виконується за таким правилом (див. рис. 13):

- $k$ -й вихід комутатора  $K_{1i}$  з'єднується з  $i$ -м входом комутатора  $K_{2k}$ ;
- $k$ -й вхід комутатора  $K_{3j}$  з'єднується з  $j$ -м виходом комутатора  $K_{2k}$ .

Загальна кількість простих комутаторів у комутаторі Клоза дорівнює  $2m + l$ . Цей комутатор є **блокуючою** комунікаційною **мережею**. Вхід тут може одержати відмову від з'єднання із-за того, що який-небудь простий комутатор є зайнятий.

**6. Продуктивність** комунікаційної мережі визначається двома такими основними характеристиками: **латентністю** комунікаційної мережі та її **пропускною здатністю**. Латентність – це час підготовки до передавання інформації по каналу мережі. Пропускна здатність – це швидкість передачі інформації по каналу мережі або, більш строго, кількість інформації, що передається між вузлами мережі за одиницю часу.

Зазначимо одразу, що **наявність латентності** означає наступне: максимальна швидкість передачі по мережі не може бути досягнута на повідомленнях із невеликою довжиною. При цьому швидкості передачі даних на повідомленнях з великою і малою довжиною можуть відрізнятися на порядок. Звідси випливає важливий **практичний висновок**: під час розробки паралельних алгоритмів та їх програмної реалізації слід намагатися скоротити не лише загальний обсяг передаваних даних, але й кількість пересилань «коротких» даних.

Загалом для будь-якої комунікаційної технології справджується така **оцінка**: на реальних задачах швидкості передачі даних є на 20–40%, а інколи і на всі 100% гіршими, ніж максимально можливі.

Повний час, необхідний для передачі повідомлення довжиною  $L$ , визначається так:

$$T = S + L/R,$$

де  $S$  – латентність,  $L$  – довжина повідомлення (від декількох байт до мегабайт), а  $R$  – пропускна здатність каналу зв'язку.

Якщо паралельна обчислювальна система призначена для розв'язання задач, в яких є **великою частка комунікаційних операцій**, то «грубо» необхідну пропускну здатність каналів комунікаційної мережі можна визначити за таким **правилом**: «швидкість міжпроцесорних обмінів між двома вузлами, що вимірюється в Мбайт/с, повинна бути не меншою за 1/10 від пікової (теоретично максимально можливої) продуктивності обчислювального вузла, яка вимірюється в Mflops».