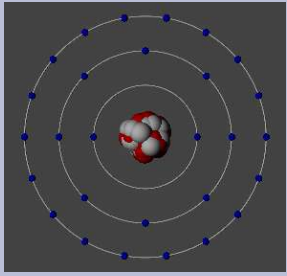


Паралелно програмиране с MPI



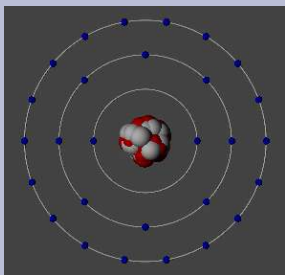
Предаване на съобщения



Съобщения



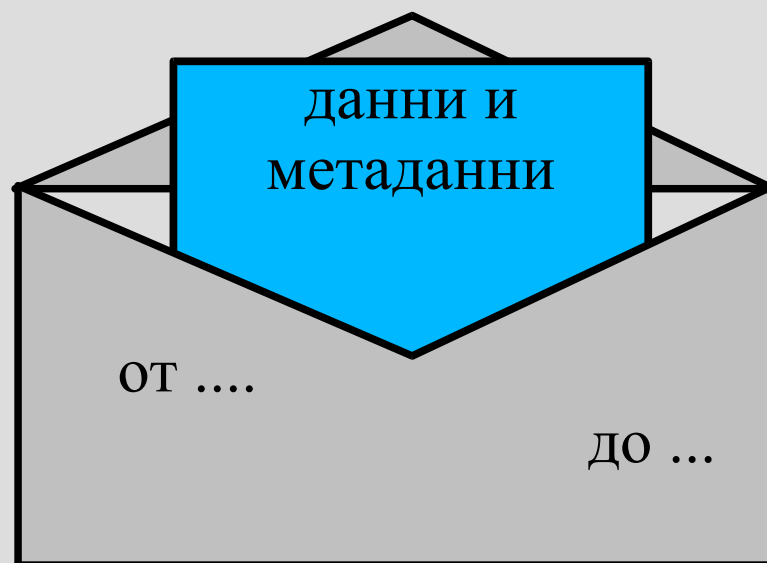
- Съобщенията са базовият метод за споделяне на данни между процесорите в MIMD архитектурите с разпределена памет.
- Явното управление на потока на информацията позволява контрол до краен предел на паралелното изпълнение на задачата.
- Предаването на съобщения е „асемблерът на паралелизма“.

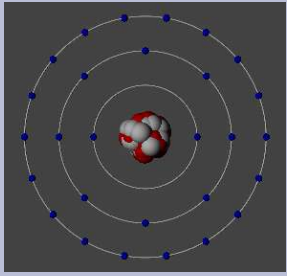


Анатомия на съобщението



- Подобно на пощенска пратка:
 - обвивка - плик с адреси на изпращача и на получателя
 - съдържание - данни и метаданни (данни за самите данни)

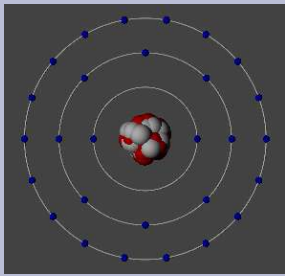




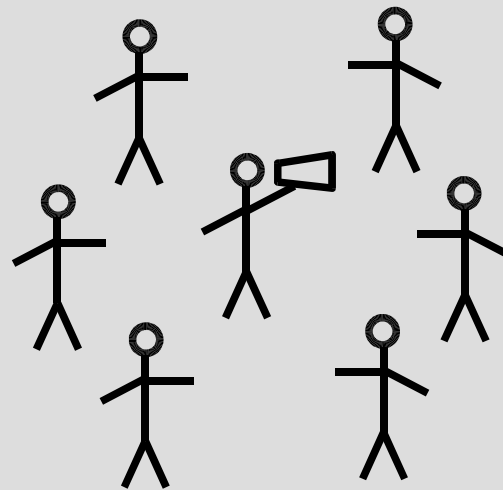
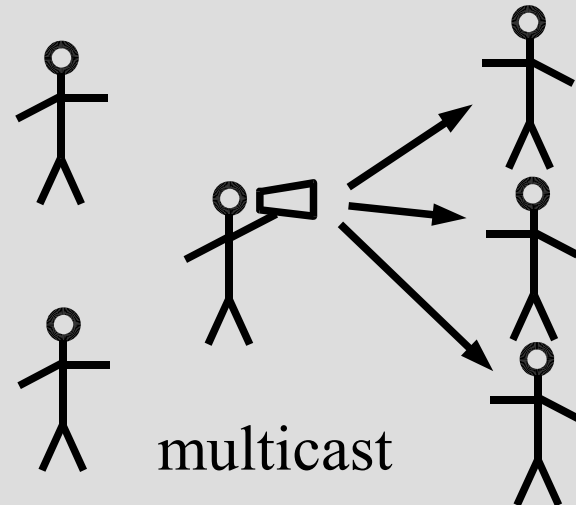
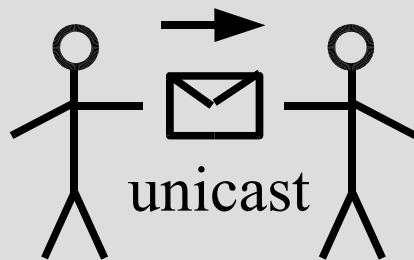
Адресация



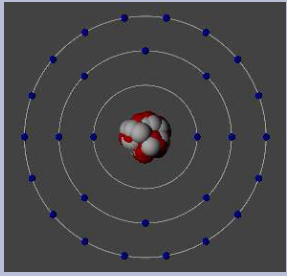
- Всеки процесор има свой уникален адрес (номер) в системата.
- Възможно е обединяването на процесори в групи, между които може да има припокриване.
- Източник на съобщението е винаги един процесор.
- Получател може да е един (unicast) или група от процесори (multicast, broadcast).



Адресация в картинки



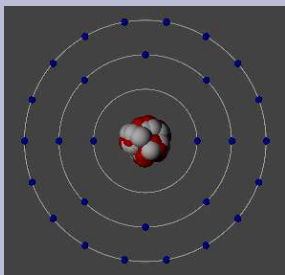
broadcast



Транспортни механизми



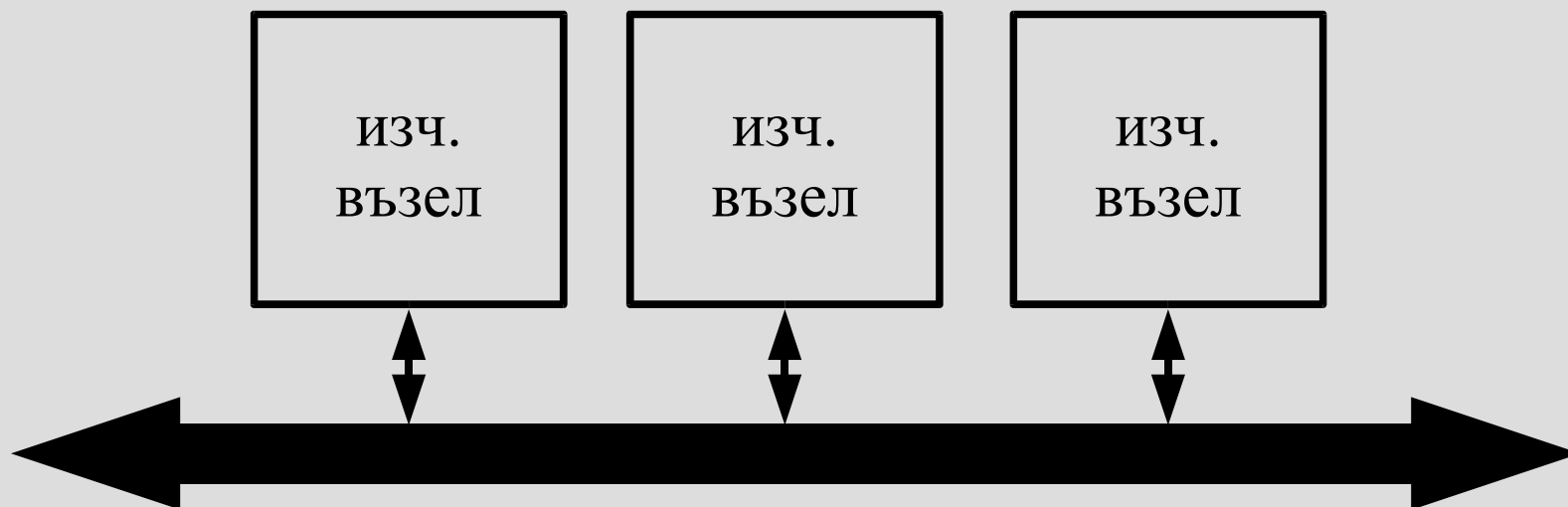
- Транспортът на съобщенията е пряко свързан с типа и топологията на свързващата подсистема.
- Различните топологии имат различна мащабируемост, латентност и сложност на реализиращото ги оборудване.
- Най-често свързващата подсистема определя до голяма степен цената на HPC системата, дори и при кластерите от работни станции (\$495 за Myrinet карта)

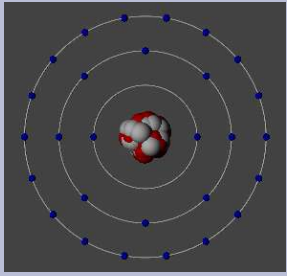


Топология - шина



- Най-елементарна и евтина за реализиране преносна среда със състезателен достъп:
 - само два възела комуникират в даден момент
 - или broadcast
- Латентността нараства с нарастване на броя на възлите - лоша мащабируемост.

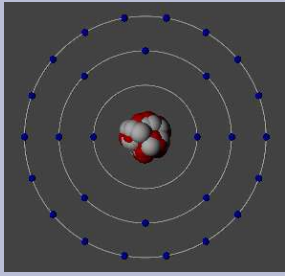




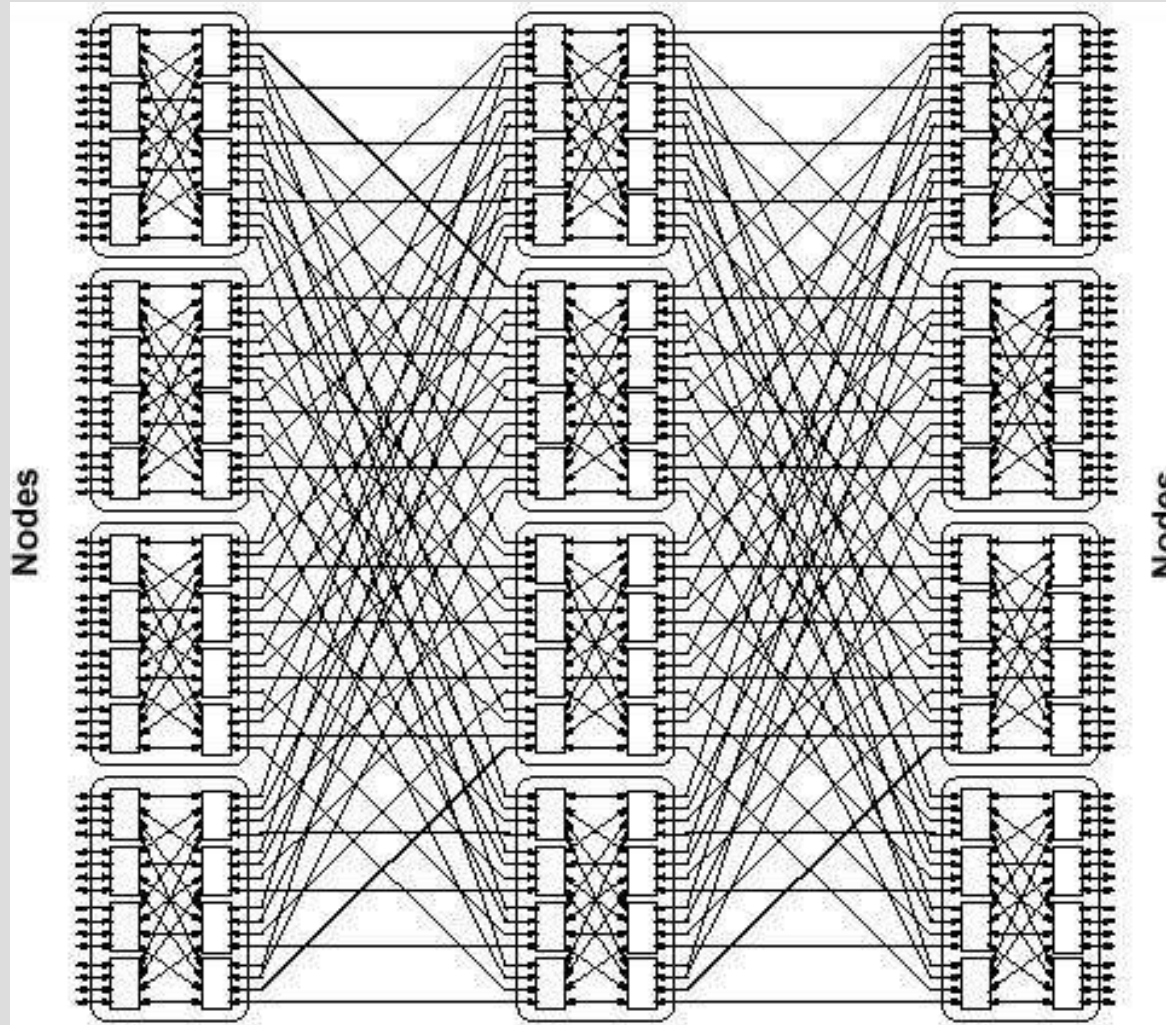
Топология - кросбар



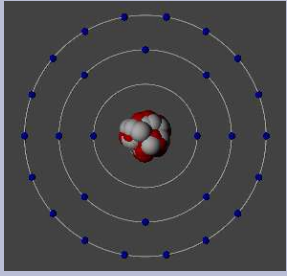
- Като телефонна централа:
 - комутация на канали
 - множество независими конкурентни връзки
 - голяма пропускателна способност
 - ниска латентност
 - могат да се свързват каскадно в многостепенни мрежи (MIN)
- По-скъпи от шината, но с по-добра мащабируемост откъм пропускателна способност
- Трудно се мащабират към по-голям брой процесори



Топология кросбар IBM SP switch



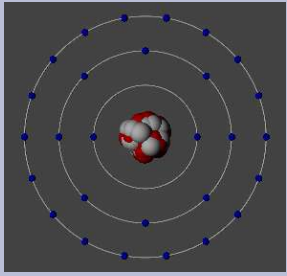
128-ВЪЗЛОВ IBM SP switch



Топологии мрежа, тор, хиперкуб



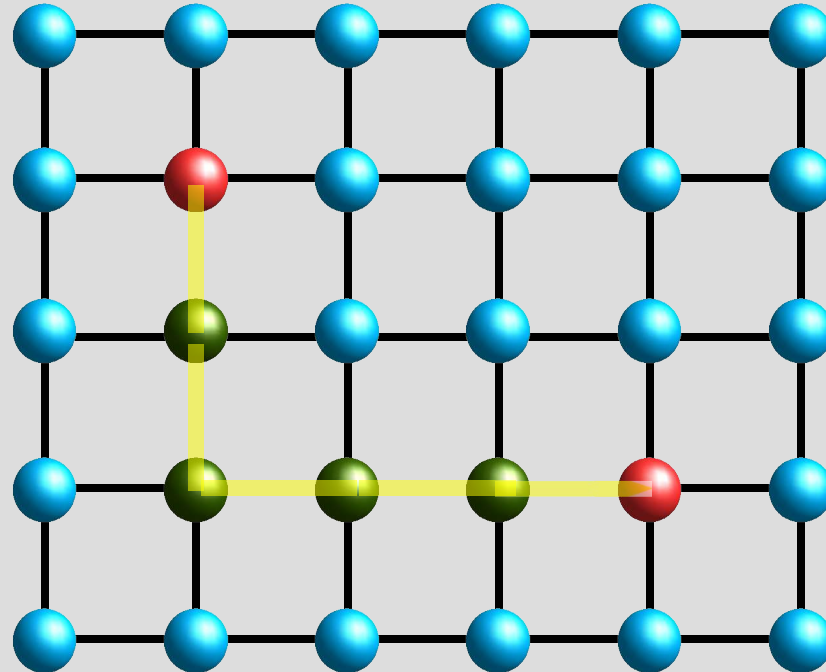
- Разклонени топологии, с множество връзки между изчислителните елементи.
- Използва се механизъм на маршрутизация на съобщенията между отдалечените възли.
- Латентността зависи от разстоянието и натоварването на мрежата.
- Изключително мащабируеми, поради което намират приложение в масивно паралелните HPC системи.

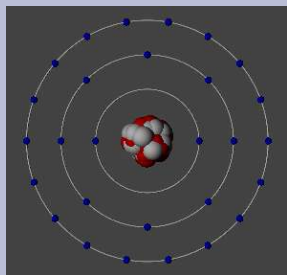


Топологии – мрежа, тор



- 2D свързаност от тип „най-близък съсед“
- Още се нарича Декартова (Cartesian)
- Периодични гранични свързаности \Rightarrow тор

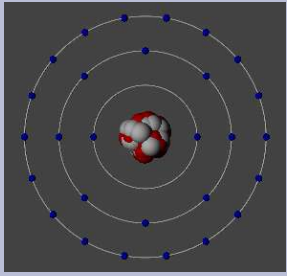




Топология хиперкуб



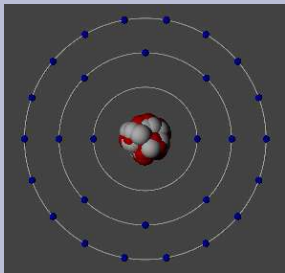
- Компромис между сложност на свързване и дължина на пътя на съобщението
- n -мерен куб, където $n > 3$
- Максимална дължина на пътя между два възела в система с 1024 процесора:
 - топология мрежа - 64 възела (hops)
 - топология 10-мерен хиперкуб - 10 възела
- Трудно мащабируеми
- Изместени от другите топологии след изобретяването на т.нар. маршрутизация с преходи на дървояди (wormhole)



Латентност и производителност



- Времето, необходимо на съобщението да достигне до местоназначението си + евентуално потвърждение.
- Латентността дава принос към серийната част на програмата и намалява паралелната ефективност.
- Различната латентност при различните архитектури води до нееднаква производителност на кода.



Латентност (2)

