

Výpočetní prostředí s distribuovanou pamětí.

Thursday, May 30, 2013 8:33 AM

MPMD

Systém pro paralelní výpočet s distribuovanou pamětí se skládá z výpočetních uzlů a komunikačních kanálů.

- Univerzální počítačová síť (softwarový multipočítač, SETI)
- Univerzální paralelní počítač (stavěný přímo pro paralelní počítač, Cluster)
- Jednoúčelový paralelní počítač (jedna konkrétní aplikace s maximální optimalizací)

protože neexistuje sdílená paměť, používá se pro komunikaci mezi procesy především zasílání zpráv

protože systémy s distribuovanou pamětí nemají žádný úzký profil ve formě sběrnice, přes kterou by procesory přistupovaly ke sdílené paměti, hodí se pro úlohy vyžadující tzv. masivní paralelismus (stovky až tisíce procesorů)

Obecně systém s distribuovanou pamětí umožňuje větší urychlení než systém se sdílenou pamětí díky paralelizaci komunikace

- Zatímco se data přenášejí kanálem, uzel může počítat
- Urychlení ovšem závisí na dalších parametrech
 - Objemu interakce
 - Celkovém objemu zpracovávaných dat
 - Konkrétní hw architektura
 - Jak dalece je použitý programový kód optimální pro danou architekturu

HW pohled

- Topologie obecně
 - **Pravidelná:** krychle, mřížka, hvězda
 - **Nepravidelná:** Internet

Směrování

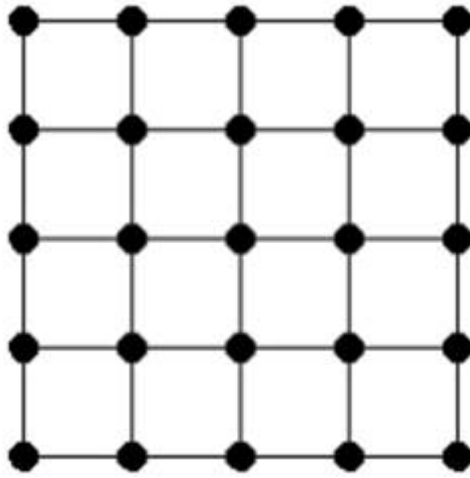
Pevná topologie – můžete poslat vzkaz pouze sousedovi

Podle příjemce – směrování jak ho známe např. z IP

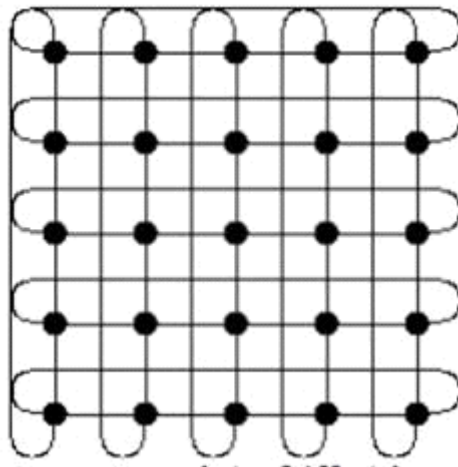
Podle odesílajícího – odesílající si určí cestu

- IP „strict source and record route“ (SSRR)
 - IP „loose source and record route“ (LSRR)
 - Často blokováno z bezpečnostních důvodů
- address spoofing (podvrhnutí adresy)

- Fyzická topologie
 - Pevná: procesory jsou propojeny komunikačním kanálem
 - **Adresa může reflektovat polohu v síti**
 - Každý s každým
 - 2d mřížka ($d_{max} = 2(n - 1)$)



- Toroid ($d_{max} = n-1$)



- 3d mřížka = krychle
- n-rozměrné krychle...
- Flexibilní: přepínání okruhů, přepínání paketů
 - Circuit switching
 - Packet switching

Parametry

- N: Celkový počet uzlů v síti
- d_{ij} : vzdálenost mezi dvěma uzly (sousedé mají 1)
- d_{max} : nejhorší varianta, kolika uzly musí projít zpráva, než je doručena (nejdelší cesta zprávy v celém systému)
- počet sousedů: s kolika dalšími uzly je daný uzel spojen přímo
- přenosová kapacita:
 - agregovaně – kolik uzlů může najednou posílat zprávu
 - odolnost proti chybám – kolik komunikačních kanálů musí selhat, než se z jedné sítě stanou dvě

Snahou je dosáhnout

- Co největšího počtu uzlů v síti
 - Škálovatelnost
- Co nejmenší komunikační vzdálenosti – d_{max}
 - Tj. omezit komunikační zpoždění
- Co nejmenší počet sousedů
 - Aneb, i komunikační kanál něco stojí
- Dosáhnout co největší přenosové rychlosti

- Cílem je namapovat virtuální topologii na síťovou tak, aby docházelo k co nejmenšímu zpoždění
 - Ideálně se fyzická, síťová i virtuální topologie shodují 1:1
 - „Úplně ideálně“ jde zároveň i o optimální metodu výpočtu (který používá danou topologii)
 - Např. mřížka na mřížku o stejných, nebo větších, rozměrech
 - Úlohy tzv. geometrické dekompozice v rovině
 - Model Farmer-worker/Master-slave
 - Farmer úkoluje workery
 - Virtuální topologie je typu hvězda

SW pohled

- **Alokování uzlů**
 - 1 proces na 1 uzel
 - Např. pevně daná u paralelního počítače
 - 1 proces dokáže plně využít celý uzel, takže nemá smysl jich na jednom uzlu spouštět několik
 - OS uzlu neumí spustit více jak jeden proces najednou
 - Potenciálně nula až několik procesů na jeden uzel
 - Přidělení celé sítě pro jeden výpočet
 - Celkový čas výpočtu je pak dán
 - Dobou k zavedení programů, spuštění procesů a distribuce dat do uzlů
 - Vlastním výpočtem
 - Získáním výsledků z uzlů
 - Přidělení části sítě jednomu výpočtu
 - Několik paralelně běžících výpočtů
 - Na jednom uzlu může běžet několik procesů
 - Nelze se spoléhat na odvozená urychlení, protože ta nepočítala se zátěží, kterou vygeneruje neznámý kód
 - Nehodí se pro synchronní/lockCstepped algoritmy – na společném uzlu by dva spolupracující procesy na sebe musely čekat dobu výpočtu jednoho kroku
- **Identifikace procesů**
 - Jedinečná ID procesů
 - Interakce send/receive (vše ostatní je na nich postaveno)
 - Podle přidělení na uzly:
 - 1 uzel – 1 proces
 - Více procesů na uzlu
 - Více procesů na uzlu a procesy mohou migrovat (tabulka umístění procesů)
- **Komunikační schéma**
 - **Fyzická topologie** = *jak je to sdrátováno*
 - **Síťová topologie** = *jak vidí fyzickou topologii software*
 - **Virtuální topologie** = *komunikační vazby procesů*
 - Ideálně 1:1 (aby docházelo k nejmenším zpožděním)
- Virtuální topologie může mít:
 - Pravidelnou strukturu - mřížka, hvězda...
 - Nepravidelnou strukturu
- Může být
 - Statická - např. mřížka
 - Dynamická - procesy mohou procházet různými fázemi, mohou vznikat i zanikat

From <<https://d.docs.live.net/e3534876709763a3/Dokumenty/ZCU/Statnice/Statnice.docx>>