

Számítógép architektúrák

Korszerű architektúrák

Mai program

- Pentium P6 processzor (esettanulmány)
- Párhuzamosítások a CPU-n kívül

Az Intel P6 család

- IA instrukciókat feldolgozó (x86, MMX, Katmai Iset), háromutas szuperskalár, sok fokozatos (12-14 stage) futószalagelvű processzorok
- Spekulatív végrehajtás, elágazás előrejelzés
 - Pl.
 - Pentium Pro
 - Pentium II., Pentium II Xeon
 - Pentium Celeron
 - Pentium III., Pentium III Xeon
- Nézzük ezt ...



IA-64: Intel Itanium architecture

Jellemzők

- Szuperskalár CISC processzor, RISC maggal
- Ciklusonként három **RISC μ -operáció kibocsátására**, öt **RISC μ -operáció kiküldetésére** képes
- 20 bejegyzéses **várakoztató állomása** van
- A **szigorúan soros konzisztenciát** átrendező puffer (**ROB**) biztosítja
- A regiszterátnevezést is a ROB biztosítja

CISC processzor — RISC maggal

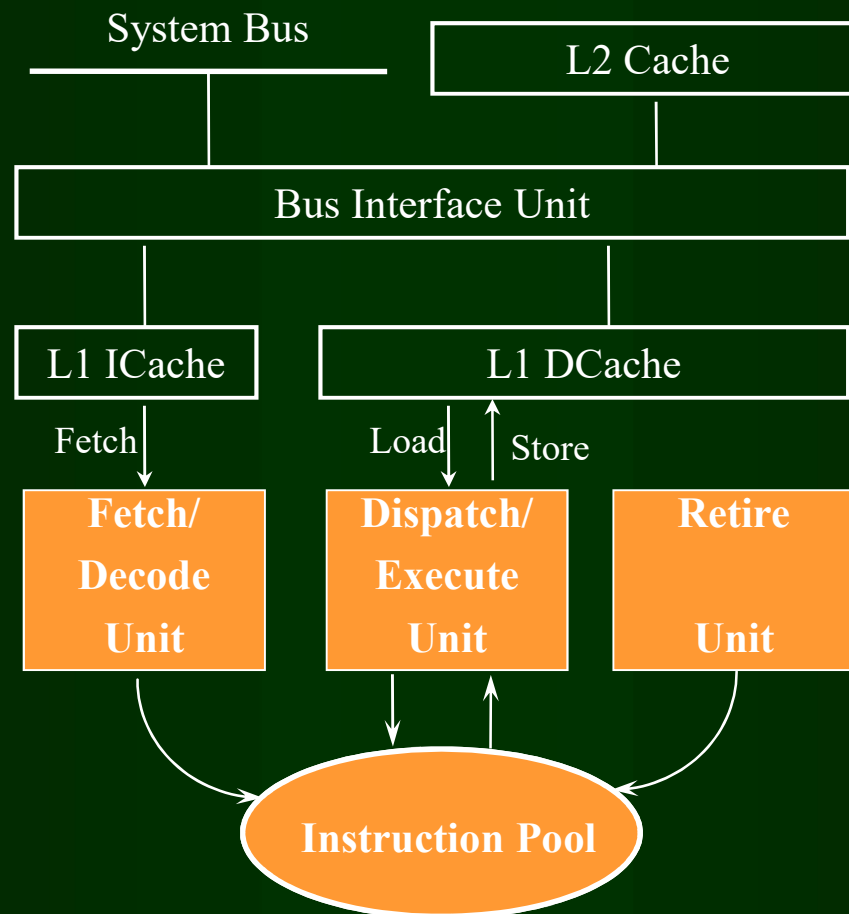
- Az ilyen processzorokban a **CISC instrukciókat RISC instrukciókká konvertálják**
 - A RISC mikro-instrukciókat (μ ops) egy RISC mag hajtja végre
 - A μ ops-ok három operandusúak (triadikus jelleg): egy „kimenő” és két „bemenő” operandus a szokásos
 - Egy CISC instrukció egy vagy több μ ops-sá is konvertálódhat

Pl. a köv. CISC instrukció kettővé:

```
SUB  EAX, [EDI]    // CISC
      MOV r1,[EDI]    // r1 ← [EDI]
      SUB  EAX,r1     // EAX ← EAX – r1
```

- **Bonyolult CISC instrukciók sok μ ops-sá**

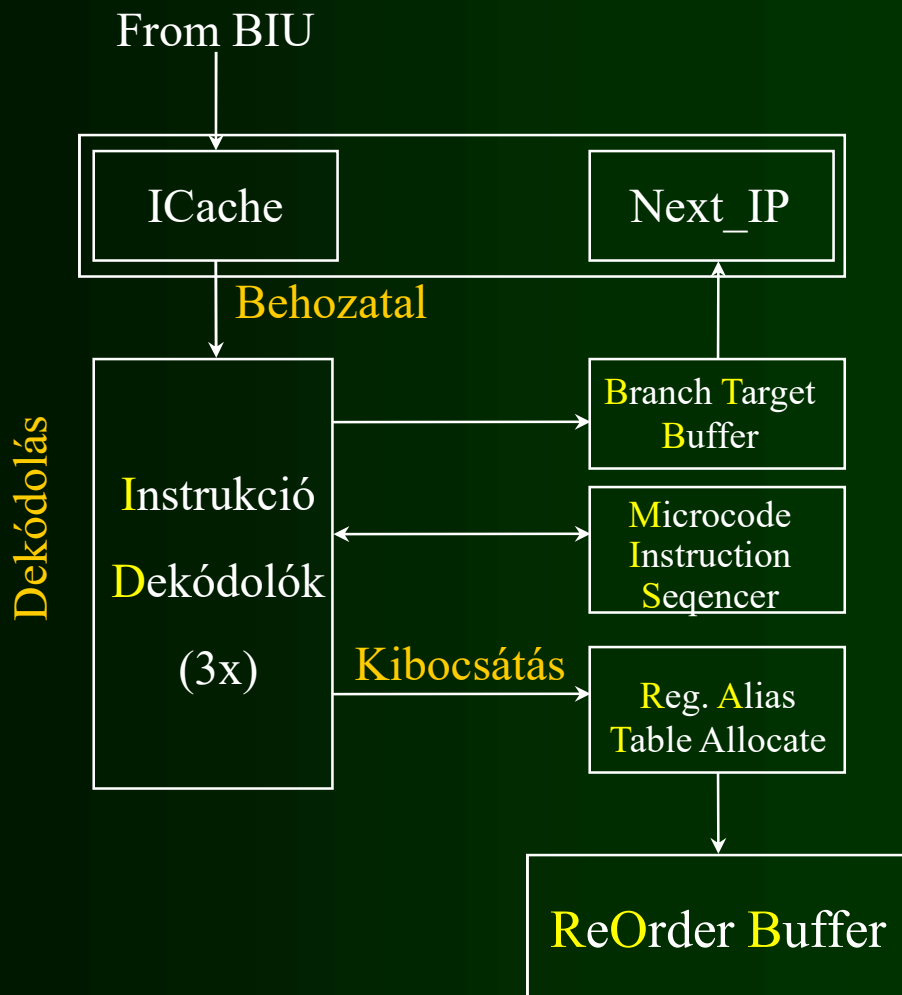
3 egység kommunikál közös instrukció mezőn keresztül



3 egység kommunikál közös instrukció mezőn keresztül



A behozó és dekódoló egység



Next_IP: L1 ICache index.

Az ID-be 16 byte-os egységekben IA instrukciókat hoznak be (**behozatal**, **fetch**)
BTB – **elágazásjelző**: merre olvasson előre (IP előtt 20-30 IA és 5 ugrás).

ID elődekódolás választ: 3 dekóder (2 egyszerű, 1 komplex) az IA instrukciókat μ ops-okká alakítja (1 v. 4 μ ops-sá). Amit nem tudnak: **a mikroutasítás sorvezérlő (MIS)** alakít át (sok μ ops-sá).

A μ ops-ok sorban átkerülnek a **regiszterátnevezés leképző-táblába (RAT)**: a hamis függőségek feloldására (az IA regiszter hivatkozások fizikai regiszter hivatkozásokra képződnek). Ez az ún. **instrukció kibocsátás (dispatch)**.

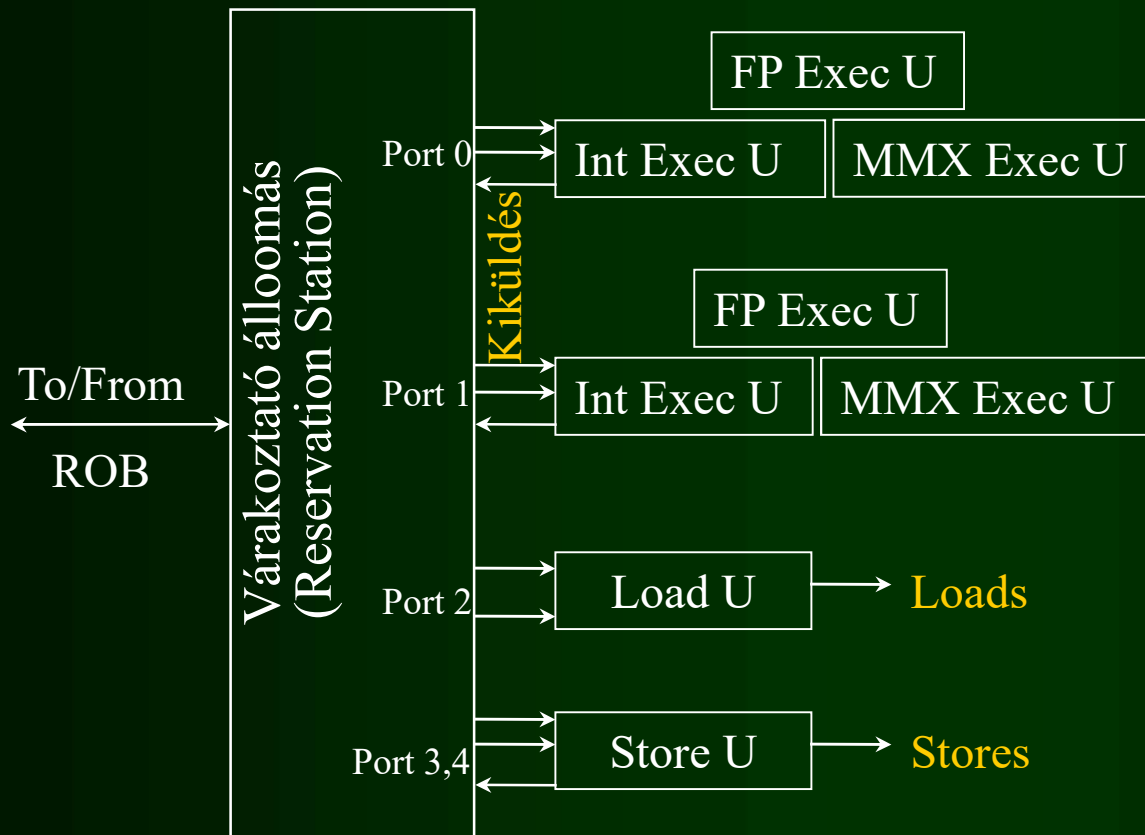
ReOrder Buffer: várakoztató puffer

Kibocsátáshoz (dispatch) kötött operandus-behívás

- A kibocsátás és regiszterátnevezés során megtörténik a (bemenő) operandusok „behívása”
(*kibocsátáshoz kötött operandus-behívás*)
 - Az átnevezett (bemenő operandust tartalmazó) regiszterek értéket kapnak (a megfelelő regiszterektől),
 - A ROB-beli μ ops-okba beolvasódnak a regiszterértékek (regiszterfájl olvasás futószalag fázis)
- Létezik kiküldéshez kötött operandus-behívás politika is
- A kibocsátott μ ops-ok a **várakozó állomásban** (Reservation Station) várakoznak



A kiküldés (issue)



RAT (Register Alias Table)
ROB (ReOrder Buffer)
Reservation Station

} Várakoztatás

A várakoztató állomás "kiegyenlítő tartály".

Minden ALU számára közös (központi várakoztató állomás)
Bekerülnek a μ ops-ok, és ha minden operandusuk előállt (és megfelelő ALU rendelkezésre áll), a megfelelő porton a megfelelő ALU-ba kerülnek végrehajtásra. Onnan az eredményükkel együtt visszajutnak az I Poolba.

Ciklusonként max 5 μ ops kiküldése lehetséges, átlag 3 a tipikus.

A várakozató állomás elv

- A *várakoztató állomás elv* azt jelenti, hogy a VÁ várakoztatásra, utasításátrendezésre és regiszterátnevezésre szolgál egyidejűleg.
- Minden ALU számára közös a VÁ, ezért a Pentiumok ún. *központi várakoztató állomás* típusúak.
- *Elosztott várakoztató állomás* (Distributed reservation stations), ha az ALU-knak (ALU csoportoknak) külön várakoztató állomása van.
- **Dispatching** – associating instructions with FU types
- **Issuing** – initiation of execution in FUs (ALUs)

A várakozató állomás elv

- Centralized issue + dispatch
 - Good utilization of entries
 - Large associative search
 - Multi-ported buffer
- Distributed issue/reservation stations
 - Quicker access time of reservation stations
 - Pipelining into separate dispatch and issue
 - Less balanced utilization of RS entries
 - Potentially less balanced utilization of FUs

A kiküldési politika

- A „**kiküldés**” („**issue**”) feladatai:
 - A VÁ-ban pufferelt μ ops-ok közül **a végrehajthatók kiválasztása** (**kiválasztási szabály**: aminek **minden forrás-operandusa már rendelkezésre áll**: spekulatív politika, adatfolyam elvű a kiválasztás),
 - A megfelelő **ALU kiválasztás** (melyik port)
 - **Döntés a kiküldési sorrendről** (ha a küldhetőnél több μ ops létezik). (A Pentiumnál az Intel nem specifikálta pontosan, FIFO.)
- Létezik sorrendben kiküldés, vagy részben sorrendben kiküldési politika is, de nem a Pentiumoknál.

A végrehajtó egységek (ALU-k)



- A kiküldött μ ops-ok valamely kapun (port) jutnak a VE-ek valamelyikébe
- Szokásos ALU-k:
 - **Fixpontos ALU** (IU, legalább 2, egy-egy porton)
 - **Fixpontos osztó** (IDIV), **fixpontos eltoló** (ISHF)
 - **Lebegőpontos összeadó** (FADD), **osztó** (FDIV), **szorzó** (FMUL)
 - **MMX ALU** (SIMD)
 - **Elágazás kezelő** (JEU, elágazási cím számító)
 - **Címkszámító adatbetöltéshez** (load AU)
 - **Címkszámító adat tároláshoz** (store AU)
 - Utóbbiakból a kiküldött μ ops-ok a **Bus Interface Unit**-ba kerülnek

A Bus Interface Unit (BIU)

- **Az adat betöltő (load) és kiíró (store) μ ops-ok ide kerülnek**
 - **A betöltéshez kell: memória cím, a szélesség és a céloperandus.** Egyetlen μ ops-sá dekódolható. *Spekulatíven végrehajtható.*
 - **Kiíráshoz kell cím, szélesség és az adat.**
Két μ ops kell, egyik generálja az adatot, másik a címet és szélességet. *Nem hajthatók végre spekulatíven.*
Ezeket nem „zavarhatják” load-ok, régi store-ok.

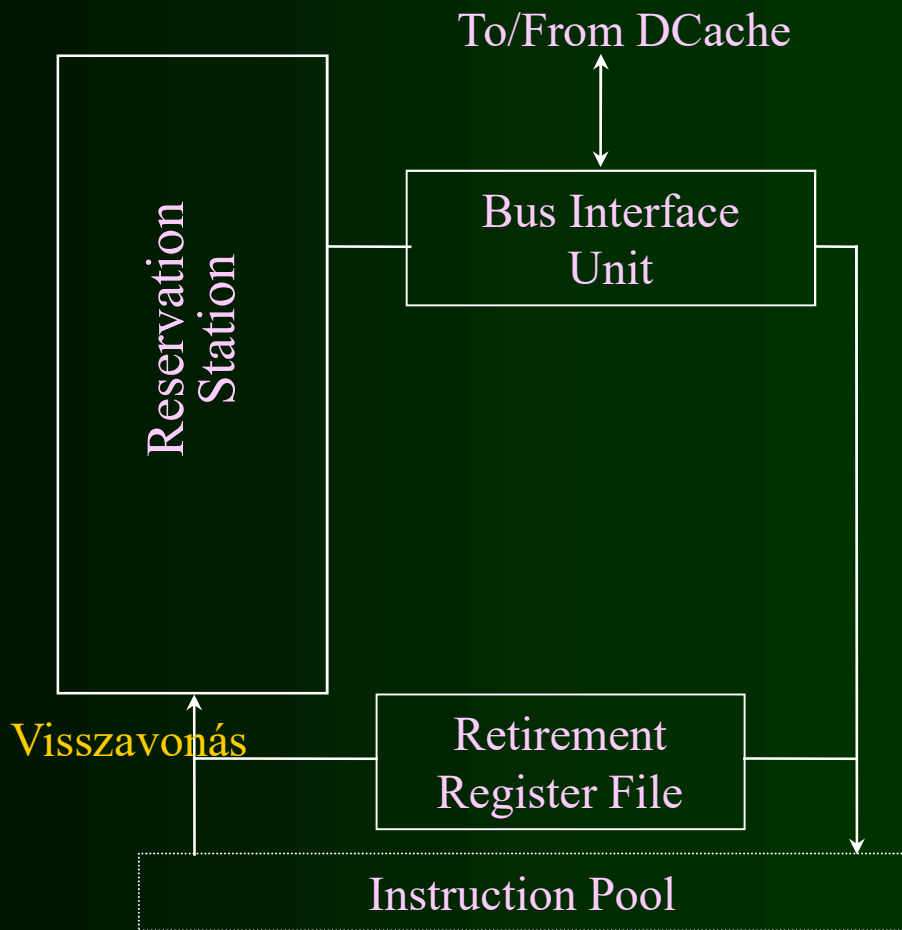
Memory Reorder Buffer (MRB)

- A BIU-ban van Memory Reorder Buffer:
„kiegyenlítő tartály” (ROB-hoz hasonlít):
 - Engedi, hogy load-ok előzzenek load-okat;
 - Blokkol függő load-okat és store-okat;
 - Újraütemezi a blokkoltakat, mikor a blokkoló feltételek (függőség, erőforrás hiány) megszűnnek.

A VÁ és a ROB „frissítése”

- A végrehajtó egységekből (ill. a BIU MRB-ből) a végrehajtott μ ops-ok visszakerülnek a várakozó állomásba (Reservation Station)
 - Asszociatív kereséssel keres olyan μ ops-okat, melyek valamely forrás-operandusa az éppen végrehajtott cél-operandussal egyezik. Ezeket frissíti: az eredményt beírja a ROB-ba, érvényességi biteket beállítja.
 - A frissítés kell a függőségek megszüntetésére, illetve a visszaírások előkészítésére.

A Retire Unit (Visszaíró)



Ez is nézi a μ ops-okat az VÁ-ban: olyanokat keres, amik **már végrehajtottak és kivehetők** (can be removed). Nézi **az eredeti sorrendet** (ROB), és a RAT (Reg. Alias Table) **leképezéseket**, hogy helyes eredményt kapjunk, **figyelemmel az interruptokra, trap-ekre, töréspontokra és a helytelen előrejelzésekre is**. A visszavont μ ops-ok eredményeit beíratja az IA regiszterekbe, vagy az L1 Dcache-be (M/BIU segítségével).

Ciklusonként 3 μ ops-ot tud visszavonni.

A load/store-ok a BIU-n át jutnak vissza az VÁ.ba.

A konzisztencia

- **A Pentiumoknál erős processzor konzisztenciát biztosítanak**
 - **A ROB-ba beírt sorrend szerint, azaz az IA instrukció sorrendje szerint történik a visszaírás**
- **A memória konzisztencia nem erős.**
- **A kivételkezelés konzisztenciája erős.**



Intel Pentium 4

- **NetBurst Micro-Architecture**
 - **Hyper-Pipelined Technology (20-stage)**
 - **400 MHz System Bus (3,2 Gbytes/sec)**
 - **Advanced Dynamic Execution**
 - 126 instrukció a spekulatív végrehajtás időablakában (P6-nál csak 42)
 - 4K a Branch Target Buffer: részletesebb história az elágazásokról, ezekből jobb előrejelzési algoritmus, kb. 33%-kal kevesebb a téves előrejelzés a P6-hoz képest
 - **Rapid Execution Engine**
 - Az ALU-k a processzoron belül kétszeres frekvenciával dolgoznak

<ftp://download.intel.com/design/pentium4/papers/24943801.pdf>

Intel Pentium 4 (folyt)

- **NetBurst Micro-Architecture (folyt)**
 - **Advanced Transfer Cache**
 - 256 KB L2 Cache, 256 bites interfésszel
 - 48 GB/s érhető el (P6-nál ez csak 16GB/s volt)
 - **Execution Trace Cache**
 - Új L1 **Icache: képes dekódolt μ ops-okat pufferelni** (nem kell újra dekódolni)
 - Az elágazások instrukciószekvenciái ugyanazon a cache vonalon: gyorsabb elérés.
 - **Streaming SIMD Extension 2 (SSE2)**
 - SIMD párhuzamosságok mind a fix, a lebegőpontos és MMX aritmetikában
 - 144 új instrukció
 - **Hardware Prefetcher**
 - Konkurencia a memória elérés és a számítások között

Intel® Desktop Processor Roadmap

2005. 2. félév	2006. 1. félév
Intel® Pentium® Processor Extreme Edition 840 • 2x1MB L2 cache, 3.20 GHz, 800 MHz FSB • Intel® 955X Express Chipset	Intel® Pentium® Processor Extreme Edition 840 or greater • 2x1MB L2 cache, 3.20 GHz, 800 MHz FSB • Intel® 955X Express Chipset
Intel® Pentium® D Processor 840 • 2x1MB L2 cache, 3.20 GHz, 800 MHz FSB • Intel® 955X Express Chipset • Intel® 945G/P Express Chipsets	Intel® Pentium® D Processor 840 or greater • 2x1MB L2 cache, 3.20 GHz, 800 MHz FSB • Intel® 955X Express Chipset • Intel® 945G/P Express Chipsets

<http://www.intel.com/products/roadmap/>

Intel® Server Processor Roadmap

2005. 2. félév	2006. 1. félév
Intel® Xeon® Processor 7000^{Δ††} sequence • Dual-core processing, 64-bit • 2x2M L2 cache, up to 3 GHz, 800 MHz FSB • Intel® E8501 Chipset^{††} (expected to be available in early 2006) / Enabled Chipsets	Intel® Xeon® Processor 7000^{Δ††} sequence • Dual-core processing, 64-bit • 2x2M L2 cache, up to 3 GHz, 800 MHz FSB • Intel® E8501 Chipset^{††} (expected to be available in early 2006) / Enabled Chipsets
Intel® Xeon® Processor MP • 64-bit • 8 MB iL3 cache, 3.33 GHz, 667 MHz FSB • Intel® E8500 Chipset / Enabled Chipsets	Intel® Xeon® Processor MP • Intel® Xeon® Processor MP Family Next Generation • 64-bit • Intel® E8500 Chipset / Enabled Chipsets

<http://www.intel.com/products/roadmap/>

Korszerű architektúrák © Vadász,
2007.

23

Intel Core i3/i5/i7

- **Ivy Bridge**

32KB + 32KB L1 Cache / mag

256 KB L2 Cache / mag

8MB L3 Cache (össz.)

3.5 GHz max órajel frek. (3.9 GHz Turbo Boost)

14 / 19 stage pipeline

GPU integrálva!

Hardware véletlenszám generátor

VT-x – Intel virtualizációs HW támogatás (később MSc...)

AVX utasításkészlet (Advanced Vector eXtensions)

Titkosítást elősegítő algoritmusokhoz instrukciók (AES)

pl. Xeon E7 8890v2

15 mag / 30 szál, 37.5 MB L3 Cache

kezelhető memória: 1.5 TB

IBM POWER

- **POWER 8**

64KB + 32KB L1 Cache / mag

512 KB L2 Cache / mag

96 MB L3 Cache (össz.)

L4 lehetőség (Centaur chipset)

5 GHz max órajel frek.

12 mag, magonként 8 szál → 96 szál

1 TB memória CPU socketenként

Külön felügyeleti microcontroller, amely dinamikus állít a CPU-ban 1764 fesz. szabályozót az energiahatékonyság érdekében

16 végrehajtó egység (execution unit)

2 fixed-point unit, 4 floating point unit

2 instruction fetch unit, 2 load-store unit

2 vmx unit, 1 cryptográfiai egység, 1 decimális lebegőpontos egység

1 condition register unit, 1 branch register unit

Párhuzamosítás a CPU-n kívül

- **Fix feladat szétosztás**
 - Aritmetikai társprocesszorok,
 - Grafikus, képfeldolgozó társprocesszorok,
 - Hangfeldolgozó ... stb.
- **Változó feladatszétosztás:** több processzor, mind általános célú ...

Flynn osztályozásához alapfogalmak

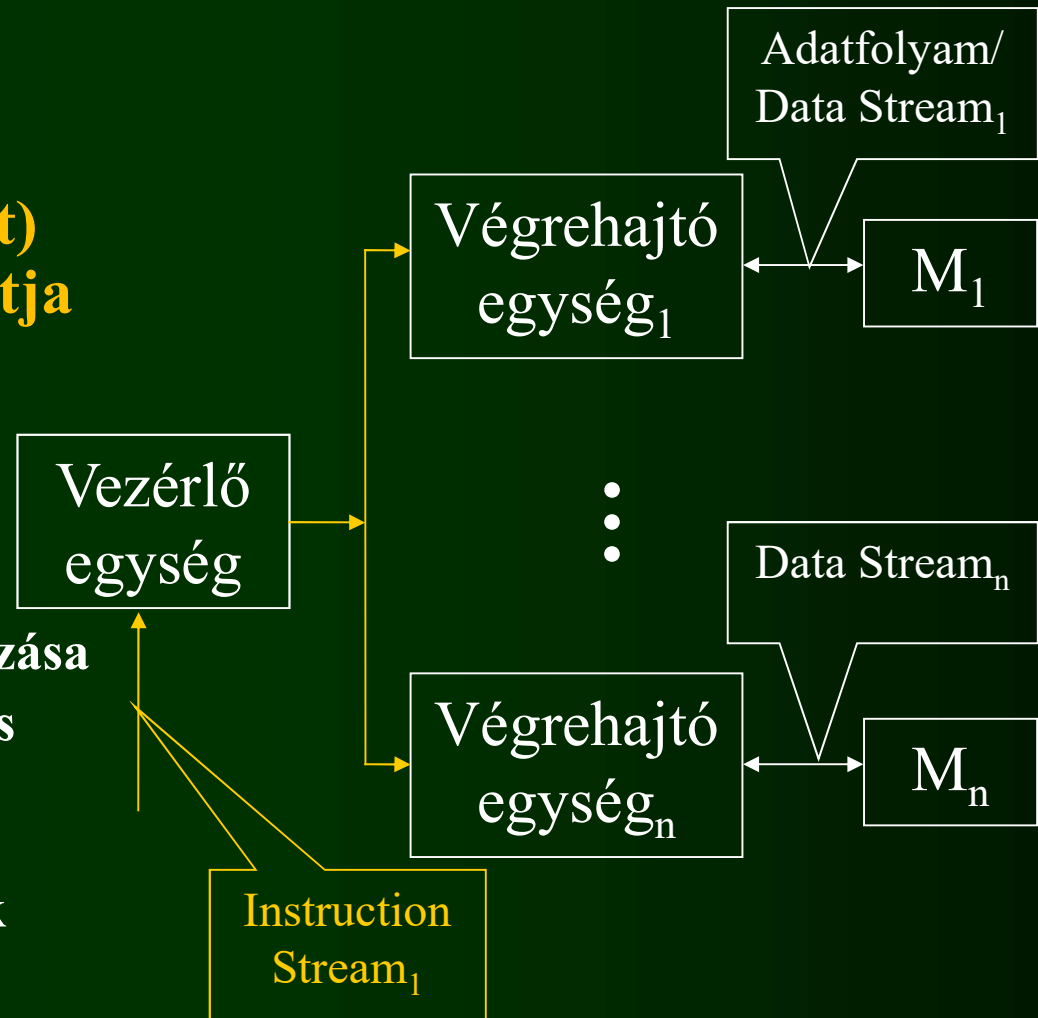
- Egy processzor által feldolgozott *instrukció folyam* (IS, Instruction Stream) fogalma
- A memória és a processzor közötti *adat folyam* (DS, Data Stream) fogalma
- Mindkettőben lehet párhuzamosság

Flynn osztályozása

- **Lehetséges osztályok:**
 - **Single Instruction Single Data stream (SISD)**
 - A Neuman gép ilyen
 - **Single Instruction Multiple Data stream (SIMD)**
 - Adatpárhuzamos processzorok
 - **Multiple Instruction Single Data stream (MISD)**
 - Nem valósult meg, bár egyesek a futószalag elvű gépeket és a szisztolikus tömböket ebbe az osztályba sorolják
 - **Multiple Instruction Multiple Data stream (MIMD)**
 - Oszott memóriás rendszerek
 - Saját erőforrásos, üzenetváltásos rendszerek

Az adatpárhuzamosság

- **Az összes processzor (végrehajtó egység) ugyanazt az algoritmust különböző adatokon hajtja végre**
- **Gyakori feladatok**
 - Mátrixok és vektorok elemeinek feldolgozása
 - Kép pixelpontok feldolgozása
 - AB rekordok párhuzamos feldolgozása
 - Hierarchikus fa- és piramisszerkezetű adatok feldolgozása



Adatpárhuzamos architektúrák

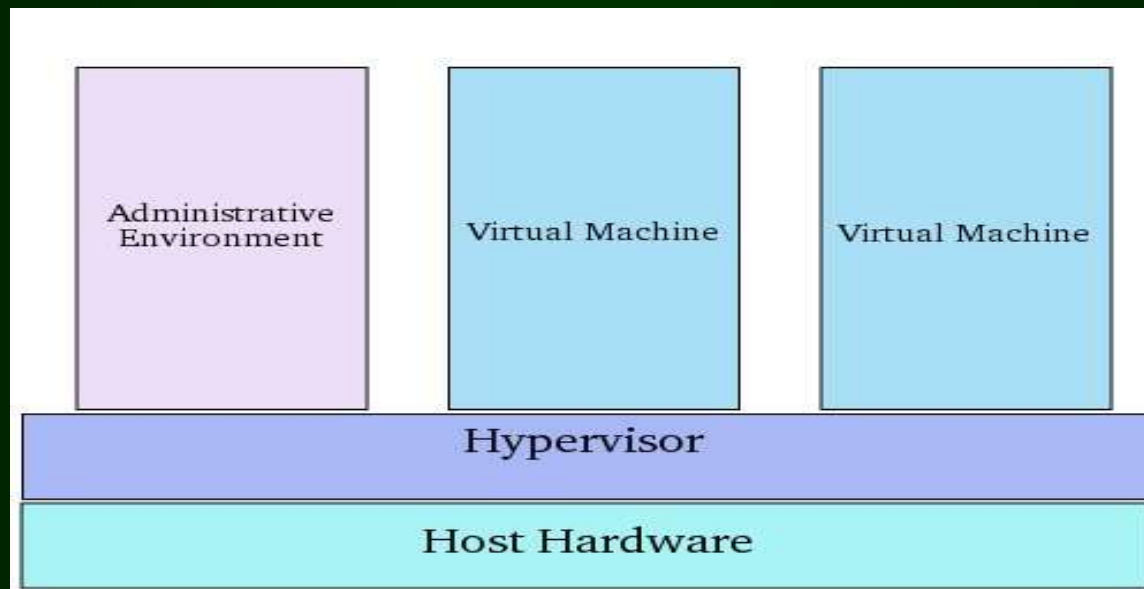
- **Asszociatív és neurális számítógépek**
 - Processzoron belül említettük az asszociatív memória kezelését. Az elven külső párhuzamosság is elképzelhető
 - Az adatok minden elemét párhuzamosan összevetjük egy AB megfelelő elemeivel, és az egyezőség mértékét (egyezés jóságát), vagy az ettől függő eredményt kapjuk meg
 - A neurális hálókról később tanulnak.
- **Szisztolikus és irányított tömbök**
 - A bevitt adatokat „átpumpálják” egy műveletvégző rendszeren (a szív szisztolikus pumpálása analógiájára)
- **Vektorelrendezésű architektúrák**
- **A klasszikus SIMD**
 - Finom, ill durva szemcsézettség

A MIMD lehet

- **Közös erőforrású struktúrák** (pl. többmagos proc.)
 - osztoznak a memórián, eszközökön,
- részben v. teljesen **saját erőforrású struktúrák** (pl. transzputerek)
 - processzoronkénti memória,
 - esetleg eszközök,
 - intenzív kommunikáció.
- A **párhuzamosság granulátuma lehet**
 - **durva szemcsés** (pl. processzek egy-egy CPU-hoz kötődnek, ilyen volt a zeus),
 - **finom szemcsés** (egy processz fonalai más CPU-kon).

Virtualizáció

- **OS Virtualizáció (OS tárgyakban majd...)**
 - Látszólag egymástól független és izolált operációs rendszer fut egy fizikai hardware-n
 - Hypervisor / VMM (Virtual Machine Monitor)
 - Sok előny még akkor is ha csak 1 virtuális gép van



Cloud Computing

- **Felhő (később részleteiben)**
- **Alapja a virtualizáció: Mindegy, hogy hol fut fizikailag**
- **Valójában csak egy szemlélet (üzleti modell...)**
- **SaaS, PaaS, IaaS**
 - **Software as a Service**
(alkalmazásokhoz és adatbázisokhoz hozzáférés "on-demand software")
 - **Platform as a Service**
(számítási platform: operating system, programming-language execution environment, database, and web server)
 - **Infrastructure as a Service**
(a (virtuális) gép és más erőforrások (hely, adat, backup) a szolgáltatás)

Két fogalom

- **Grid computing** (Computational grid ; peer-to-peer-computing) (grid = villamos elosztó hálózat)
 - Számítógépek összekapcsolása hálózaton keresztül, mely gépek ugyanazon probléma ugyanabban az időben történő megoldásán egységként dolgoznak
- **Cluster computing** (gépek klasztere)
 - Személyi gépek, munkaállomások együttes használata úgy, hogy egy felhasználó számára egy egységként tűnjenek
Single System Image.
- **High Throughput Computing** (minél többet)
- **High Performance Computing** (minél párhuzamosabban)

Nvidia Tesla

- **Nvidia Tesla**
 - **Társprocesszor**
 - PCI-Express
 - **Sok-sok mag, kisebb teljesítmény, és nem általános, mint a CPU**
 - **Tesla K40**
 - GPGPU
 - 745 MHz
 - 2880 mag (CUDA)
 - ugyanaz a művelet, sok adaton
 - 12 GB RAM
 - 288GB/sec mem elérés
 - 1.43 TFLOPS



Intel Xeon Phi

- **Xeon Phi**

- **Társprocesszor**

- a. PCI-Express
 - b. CPU foglalatba helyezhető

- **„buta” magok, de sok**

- **Pentium 1 CPU (P5) + 512-bit AVX**

- 61 db mag, 1.238 GHz
 - 244 szál
 - 16 GB RAM
 - 1.2 TFLOPS

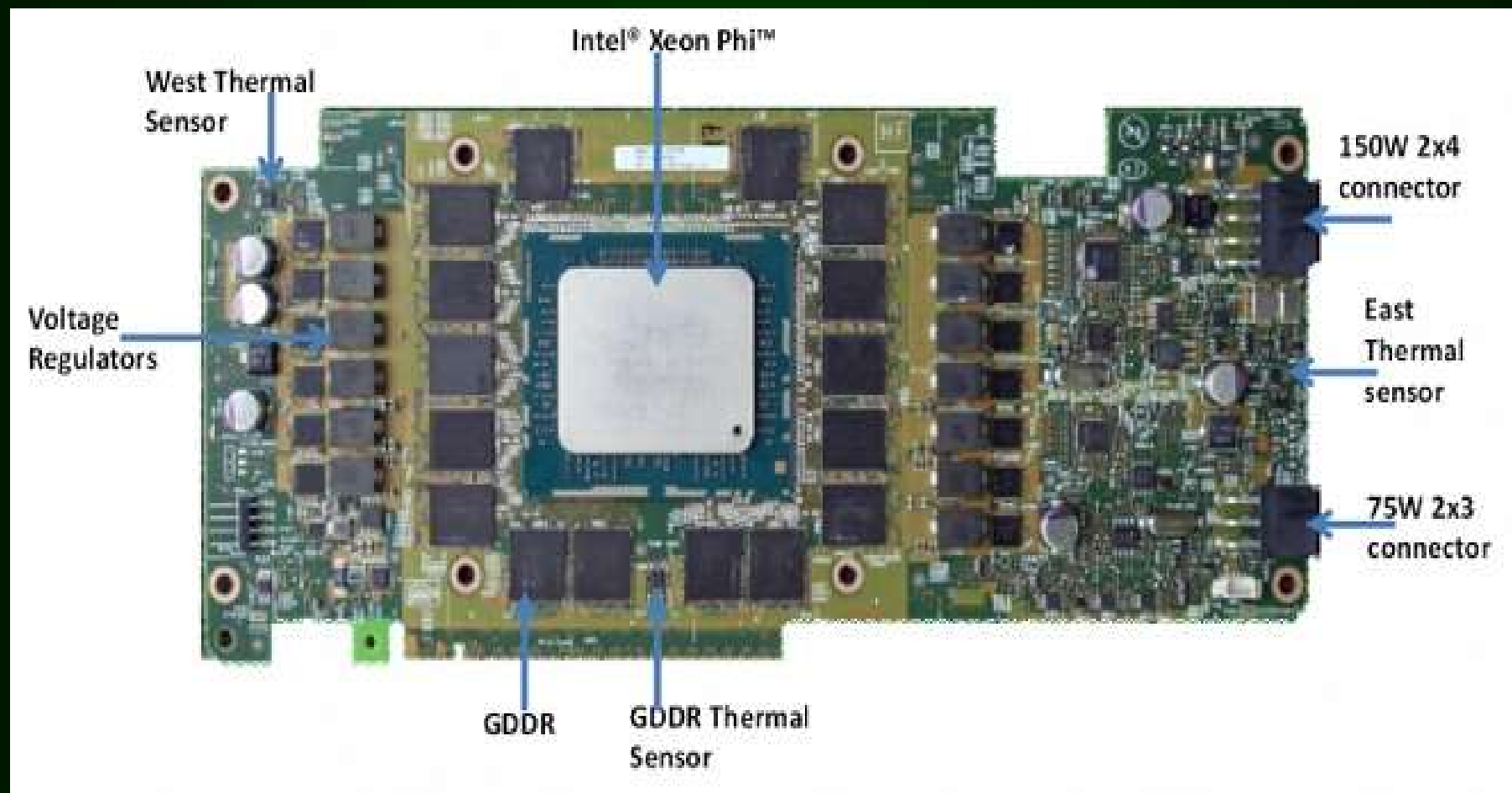
- **HPC**

- Tianhe-2 (www.top500.org)

- **33+ PFLOPS**



Xeon Phi





2014/04/22

Számítógép architektúrák

Korszerű architektúrák
VÉGE