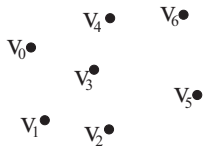
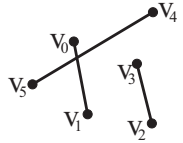
Grafikus csővezeték

Grafikus csővezeték

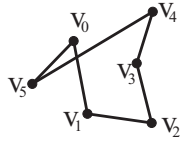
- Vertex feldolgozás
  - A vertexek egyenként a képernyő térbe vannak transzformálva
- Primitív feldolgozás
  - A vertexek primitívekbe vannak szervezve
- Raszterizálás
  - Primitívenként
    - Fragmensek
- Fragmens textúrázás és színezés
  - Fragmensenként



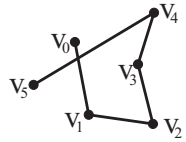
Pontok



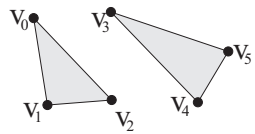
Vonalak



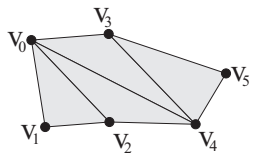
Vonal hurok



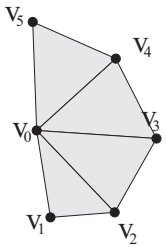
Töredezett vonal



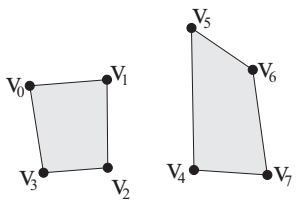
Háromszögek



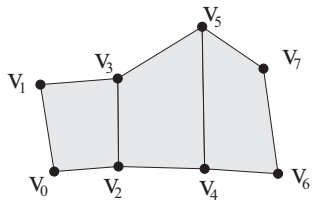
Háromszögsáv



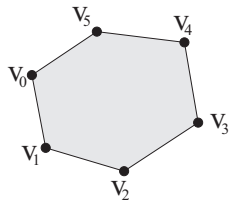
Háromszög-legyező



Négyszögek

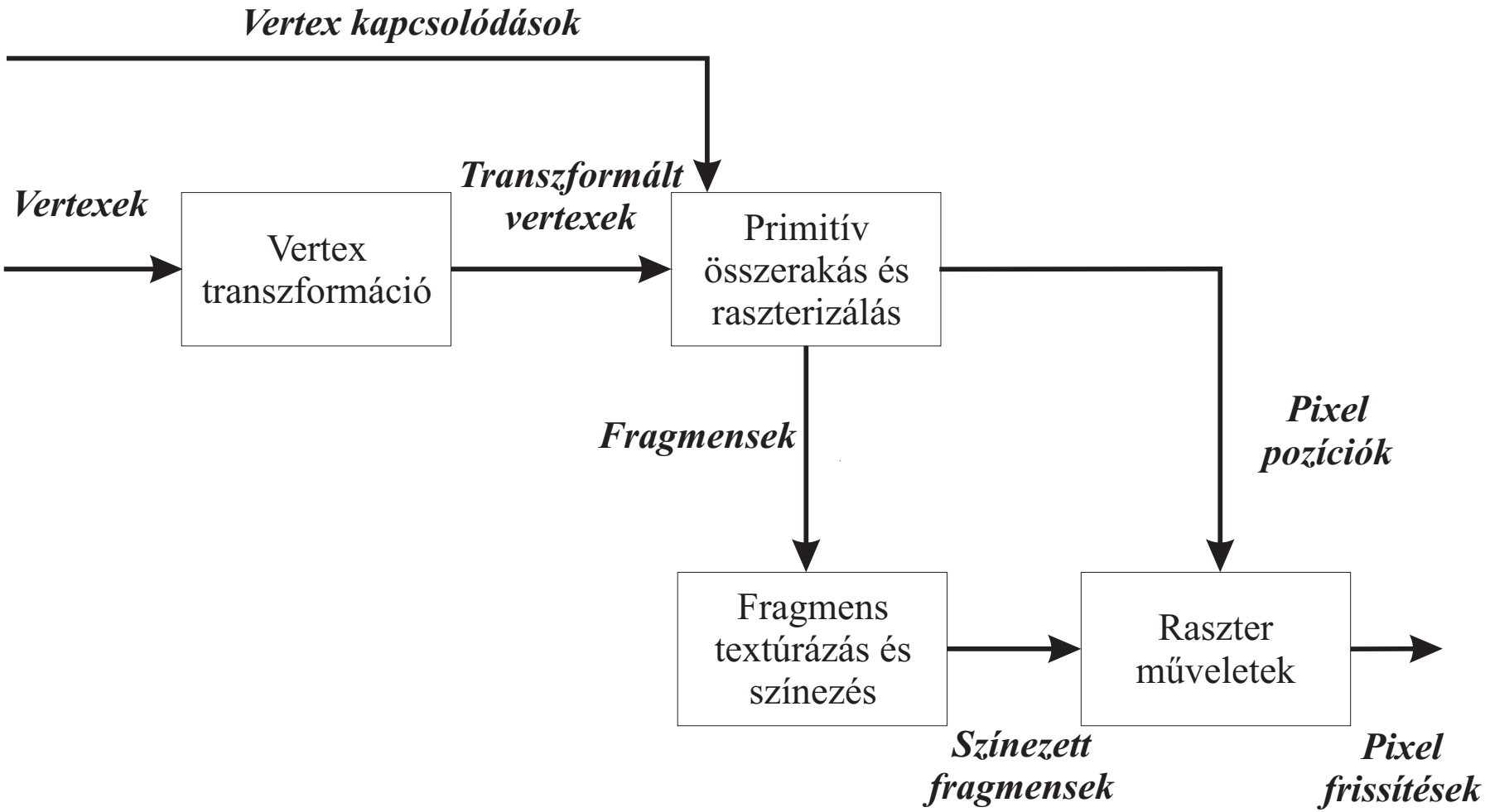


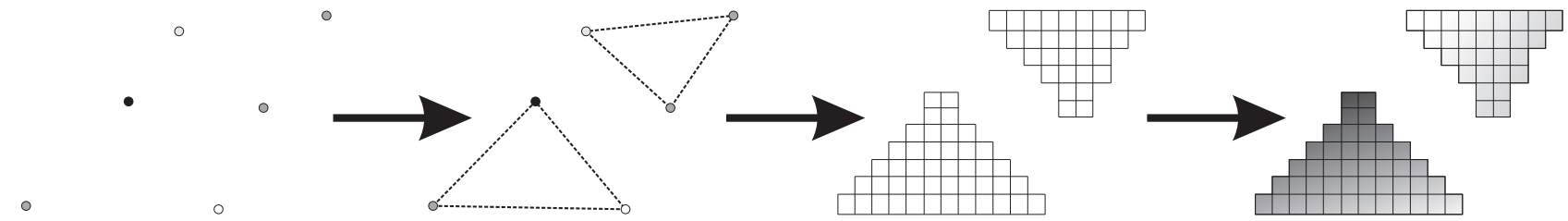
Négyszögsáv



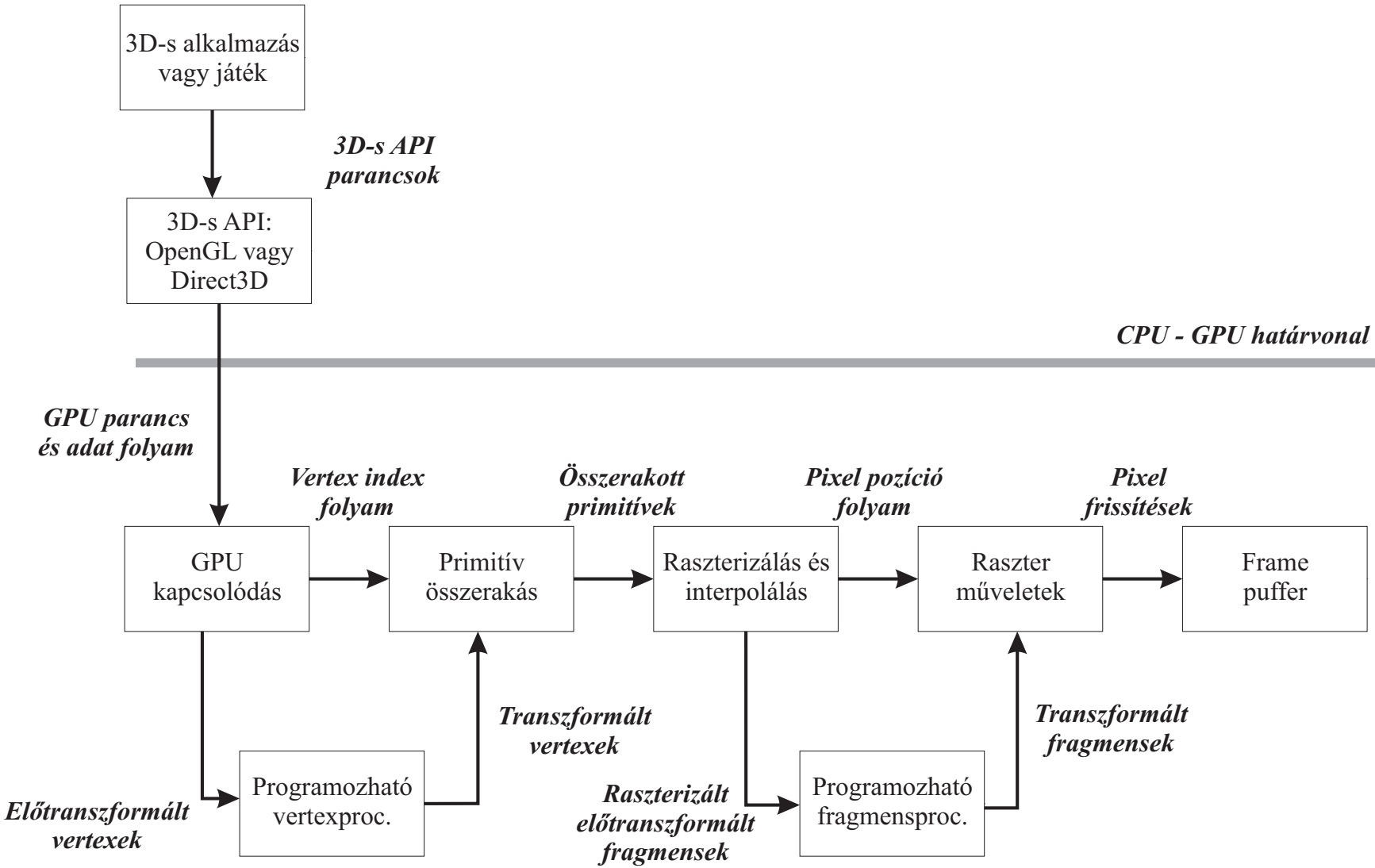
Poligon

Grafikus csővezeték

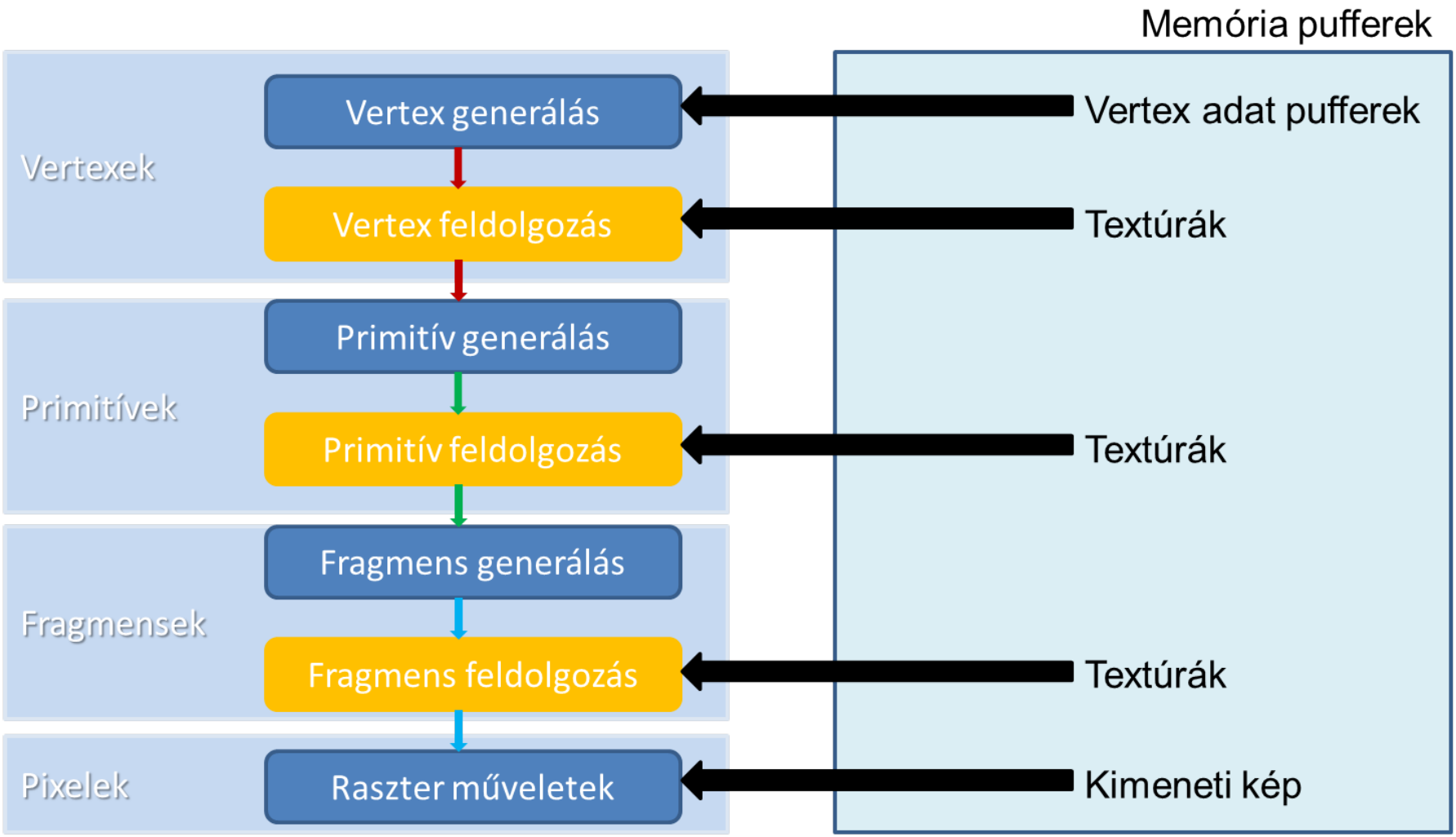



Színezett vertex
vertex transzformáció után
Primitív
összerakás
Raszterizálás
Interpoláció,
textúrázás és színezés

Programozható grafikus csővezeték



Grafikus csővezeték



GPU-k fejlődése

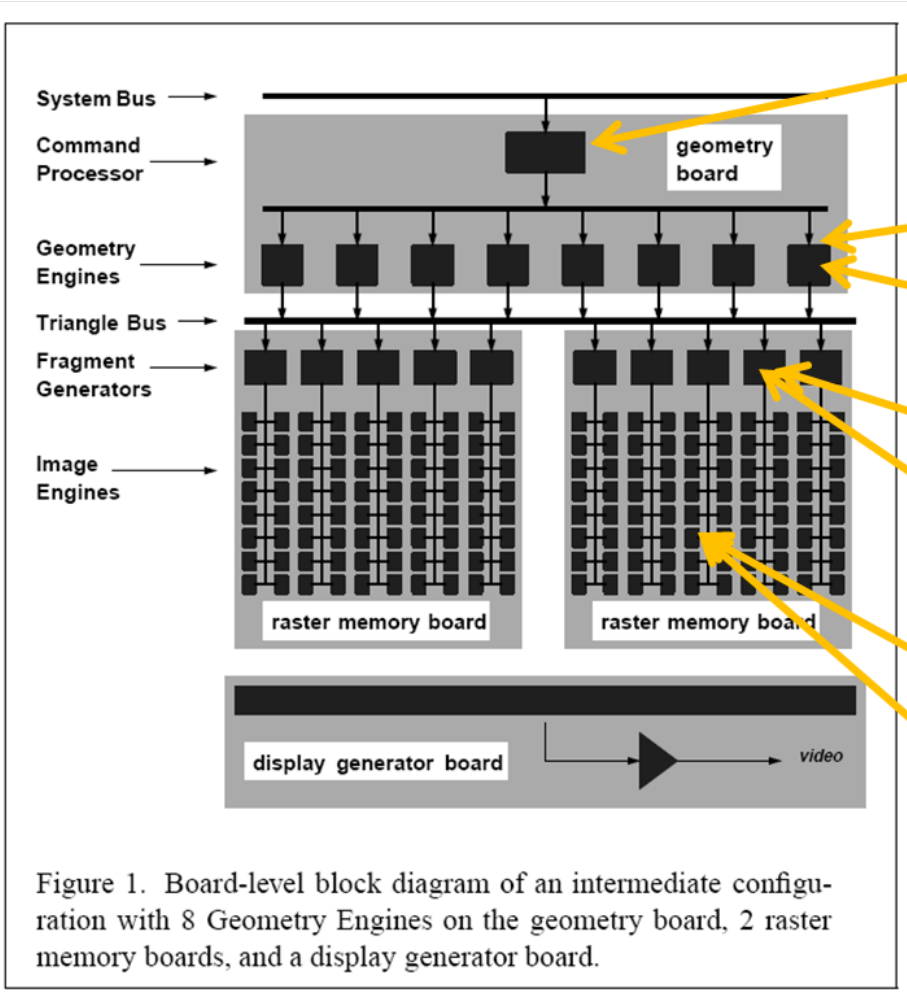
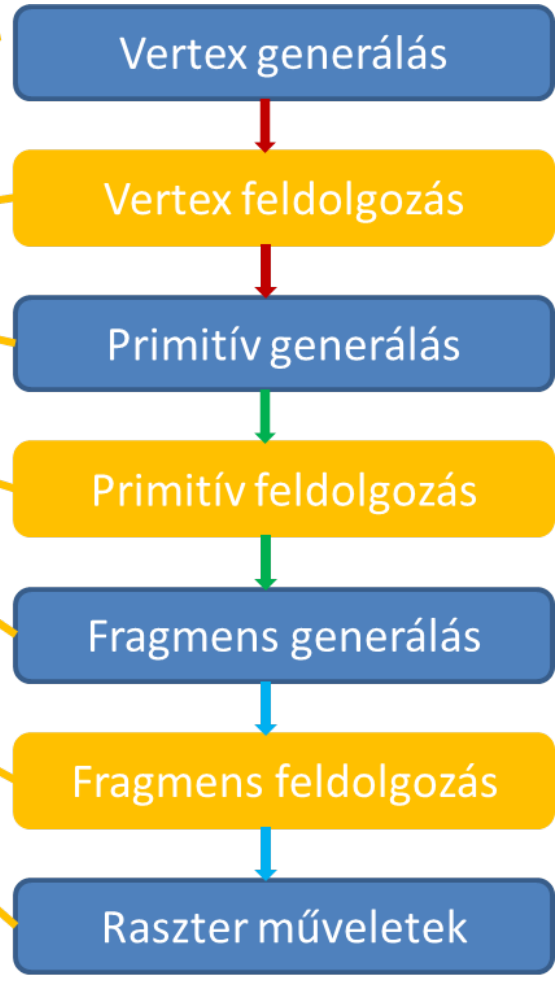


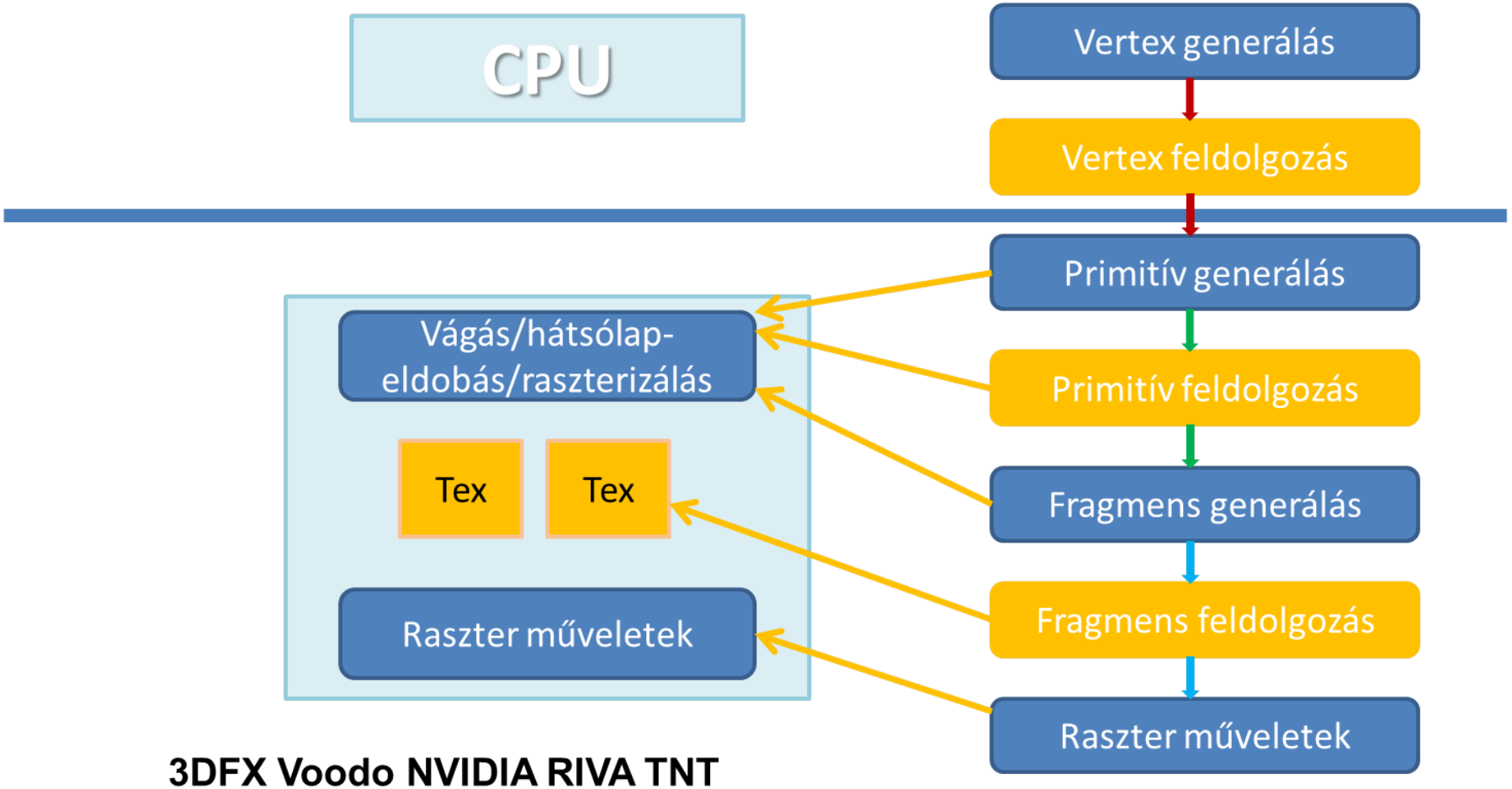
Figure 1. Board-level block diagram of an intermediate configuration with 8 Geometry Engines on the geometry board, 2 raster memory boards, and a display generator board.



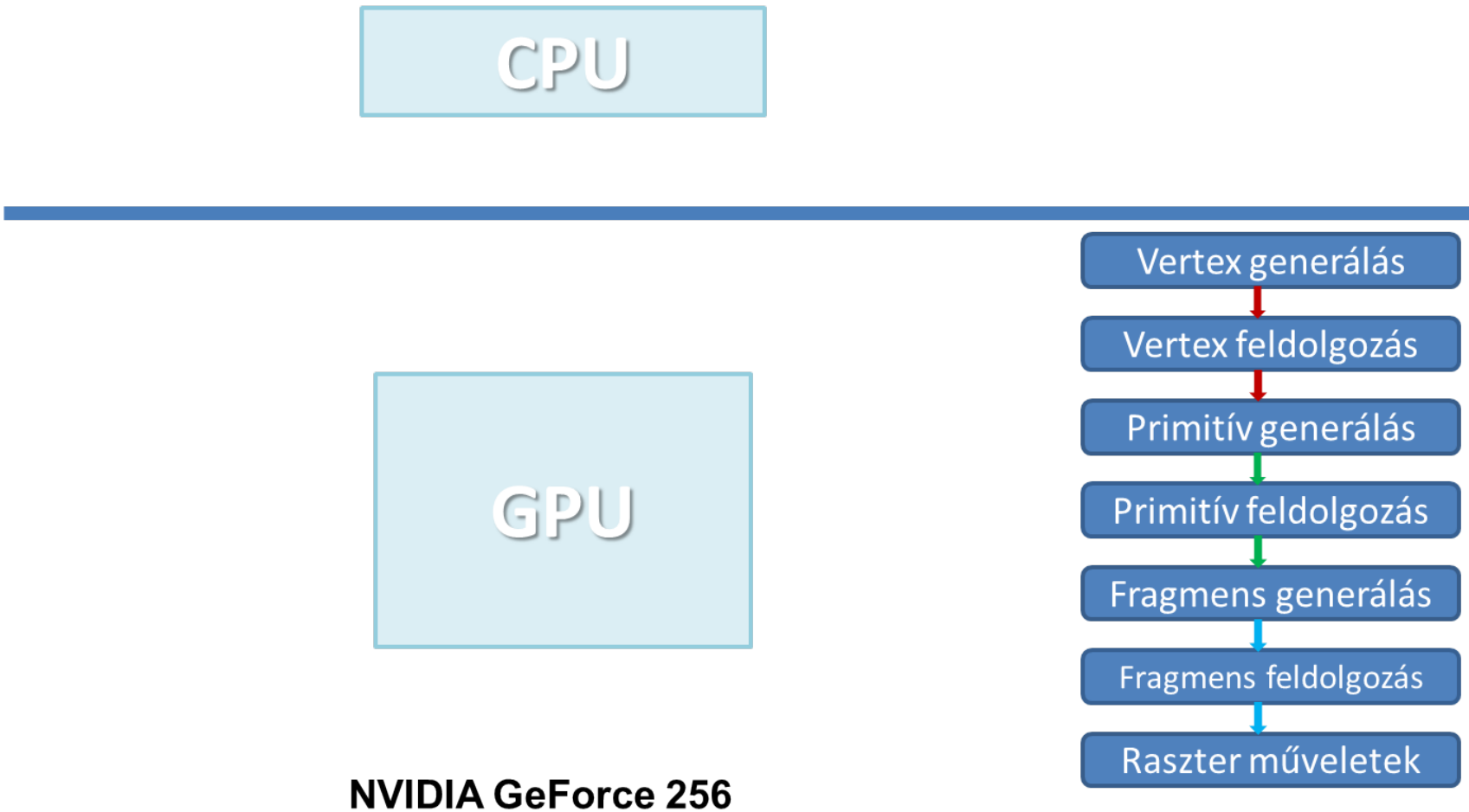
Akeley, Kurt. "RealityEngine Graphics". Proceedings of SIGGRAPH '93, pp. 109-116.

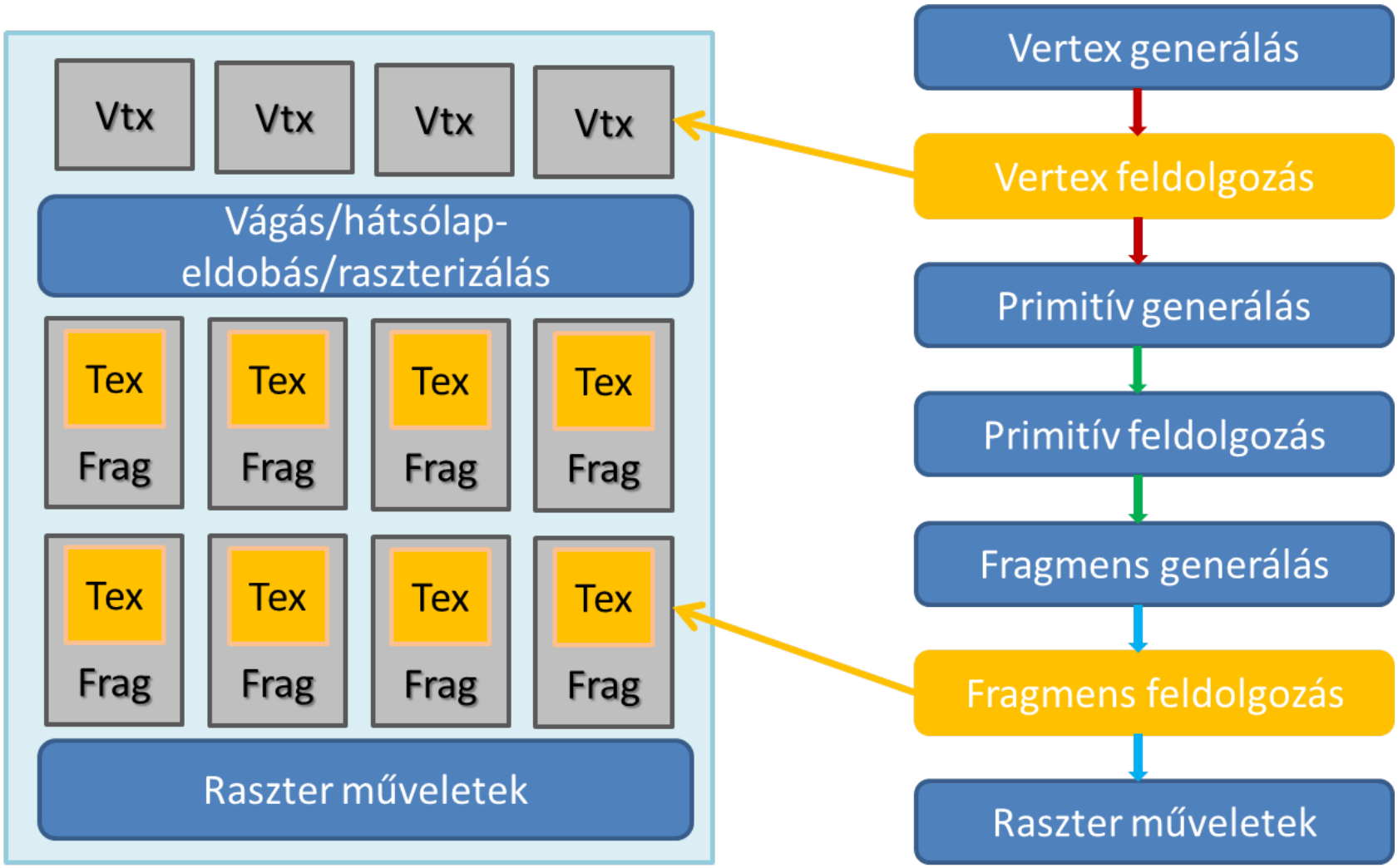
Grafikus csővezeték

3D-s grafikus gyorsító 1999 előtt

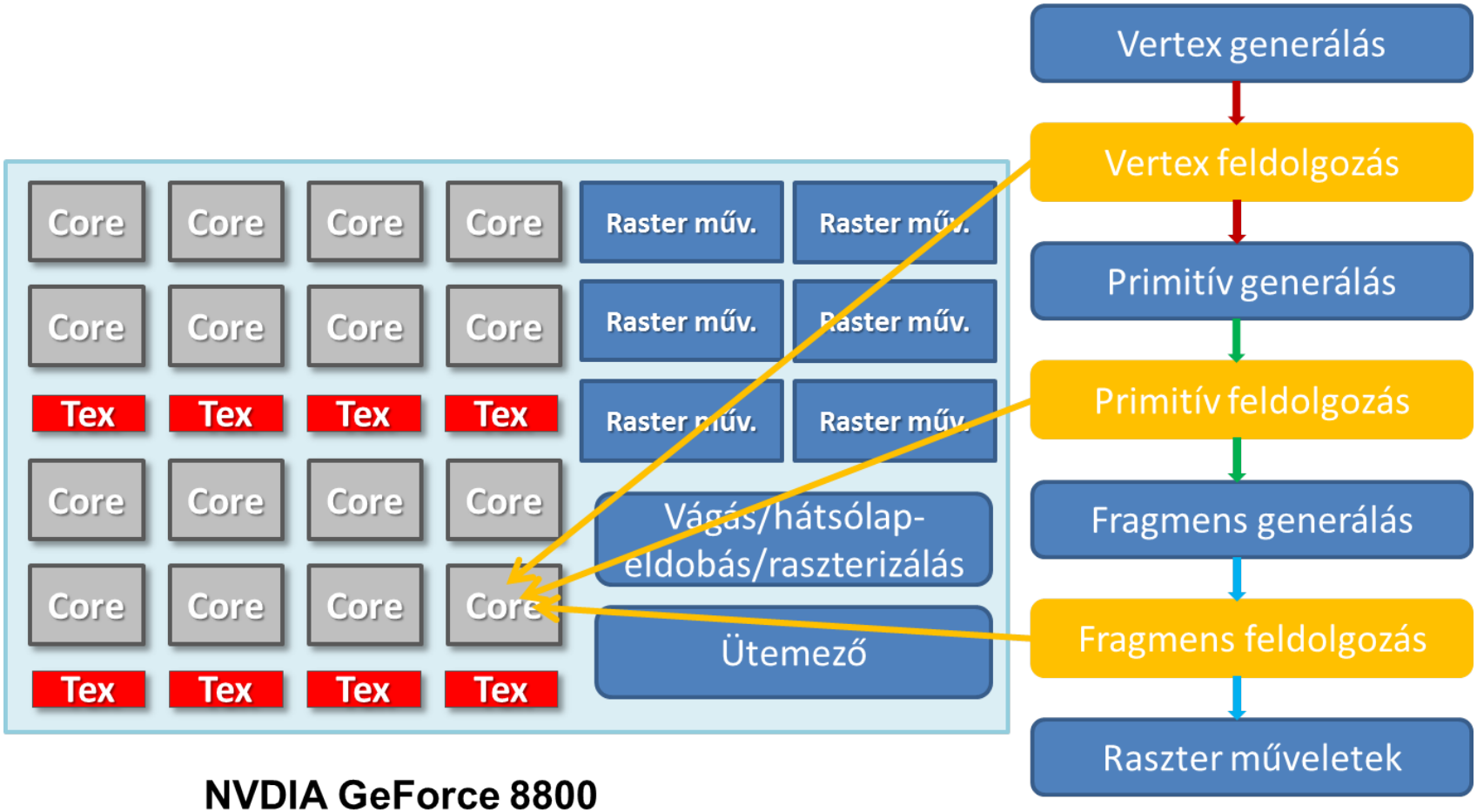


3DFX Voodo NVIDIA RIVA TNT


CPU
GPU
Vertex generálás
Vertex feldolgozás
Primitív generálás
Primitív feldolgozás
Fragmens generálás
Fragmens feldolgozás
Raszter műveletek
NVIDIA GeForce 256



ATI Radeon 9700

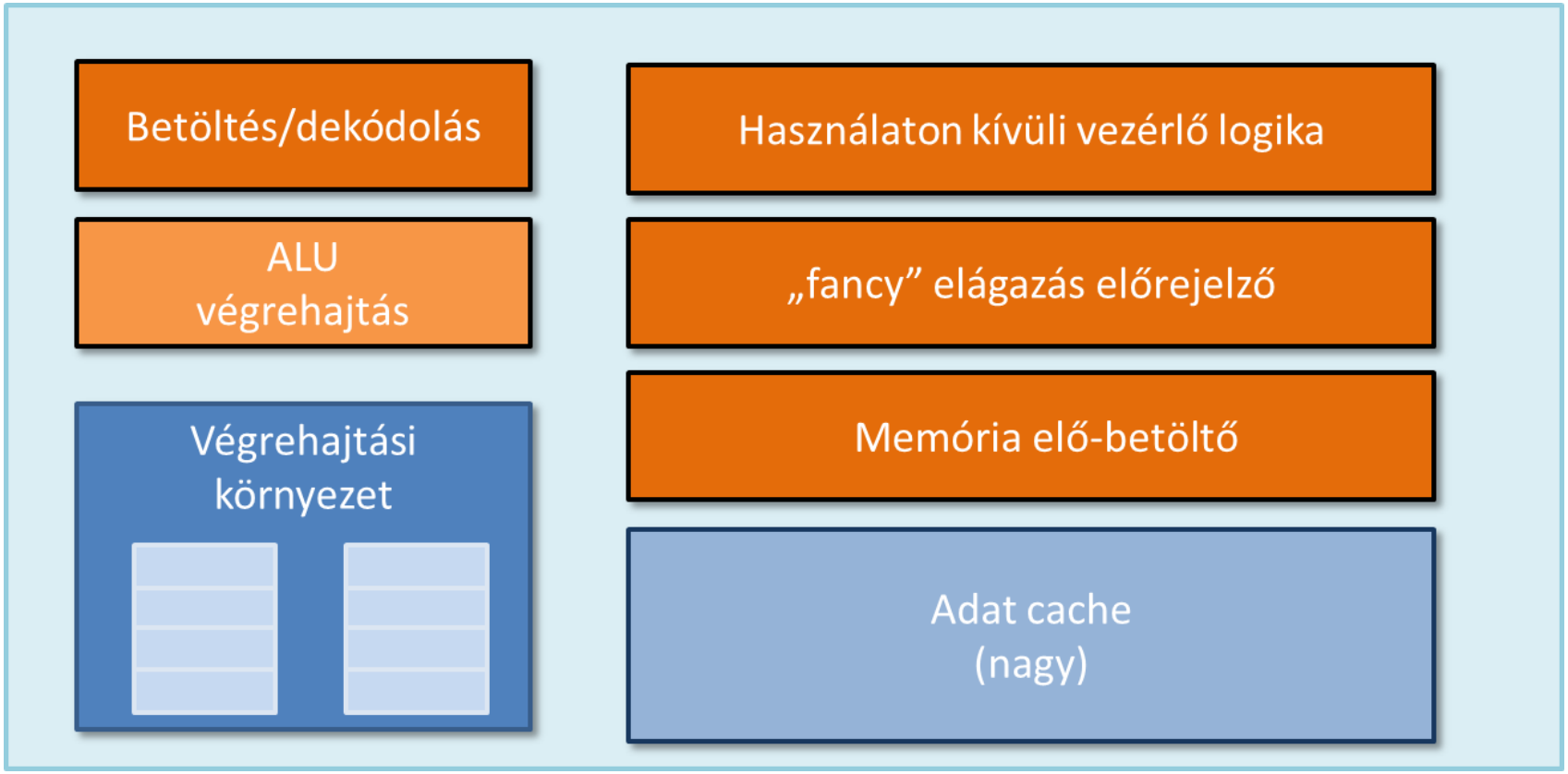


NVDIA GeForce 8800
egységes shader GPU

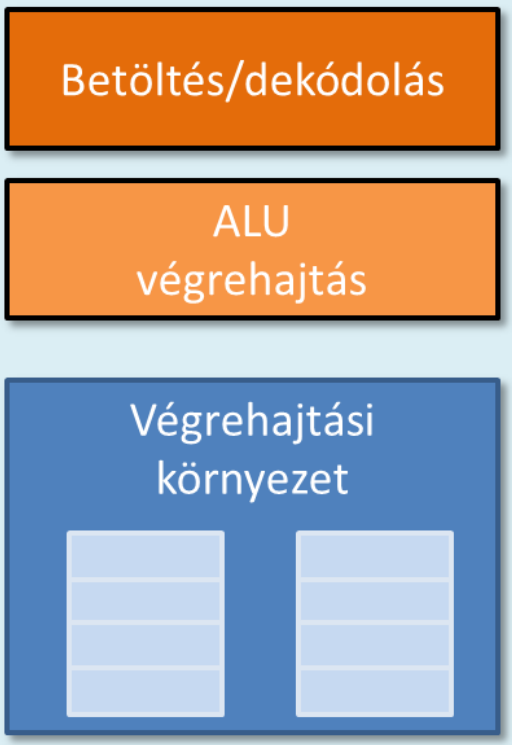
- Egységes shader modell
  - Shaderek megvalósítása közelebb került egymáshoz
    - Egyszerű
    - Kevés utasítás
    - Több száz általános célú végrehajtóegység
    - Hatalmas számítási kapacitás

Ötletek egy GPU felépítéséhez

CPU-stílusú core-ok



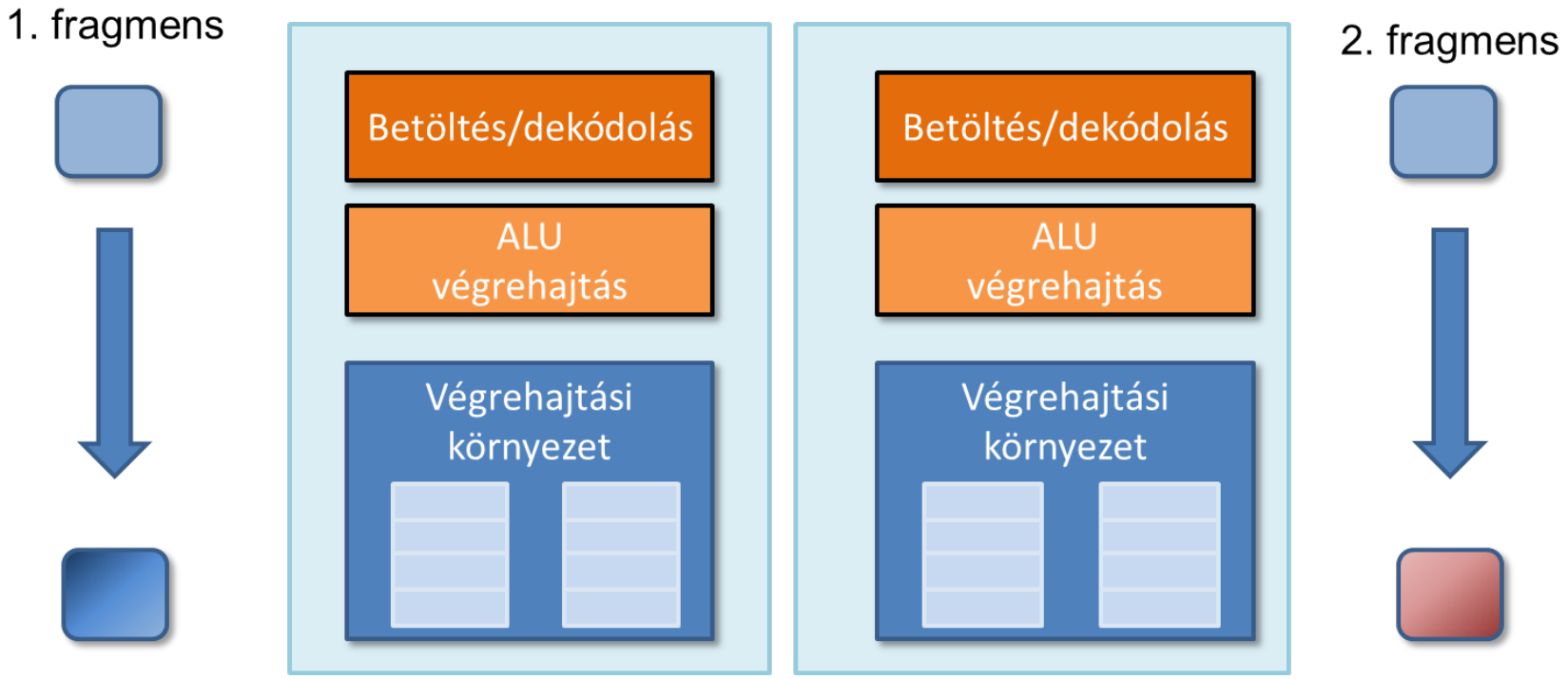
Első ötlet



Tüntessük el azokat a komponenseket, amelyek az egyetlen utasítás folyam gyors futását segítik!

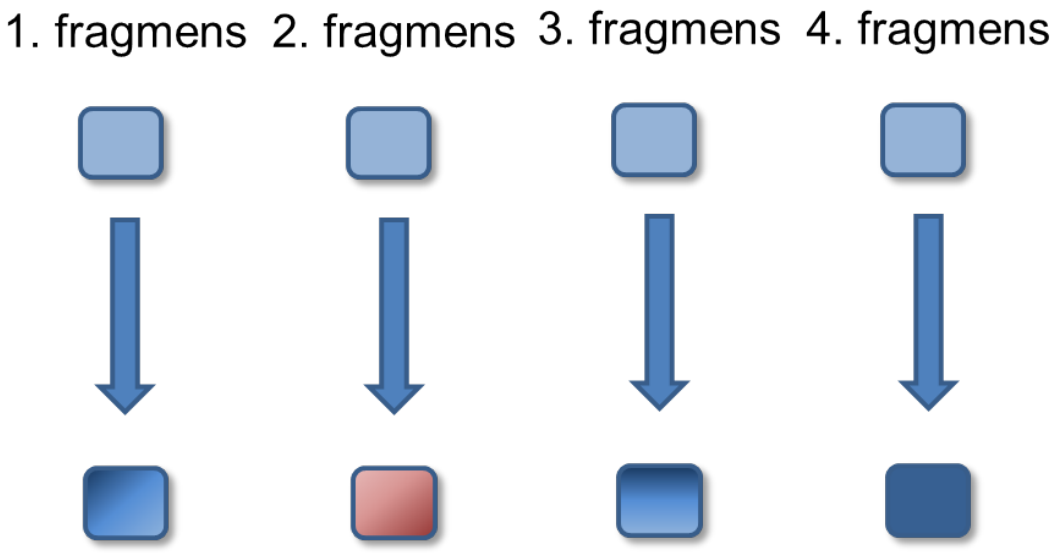
Két mag (core)

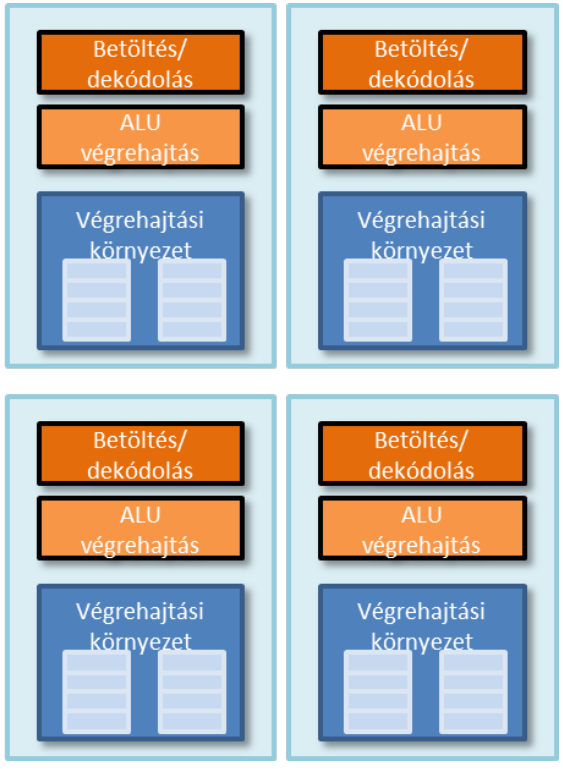
Két fragmens párhuzamos feldolgozása



Négy mag (core)

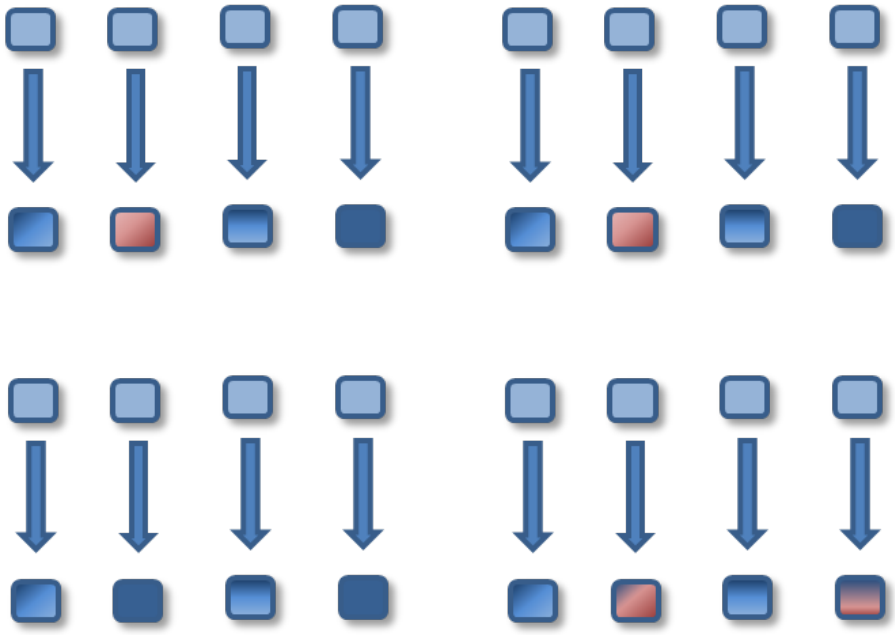
Négy fragmens párhuzamos feldolgozása





Tizenhat mag (core)

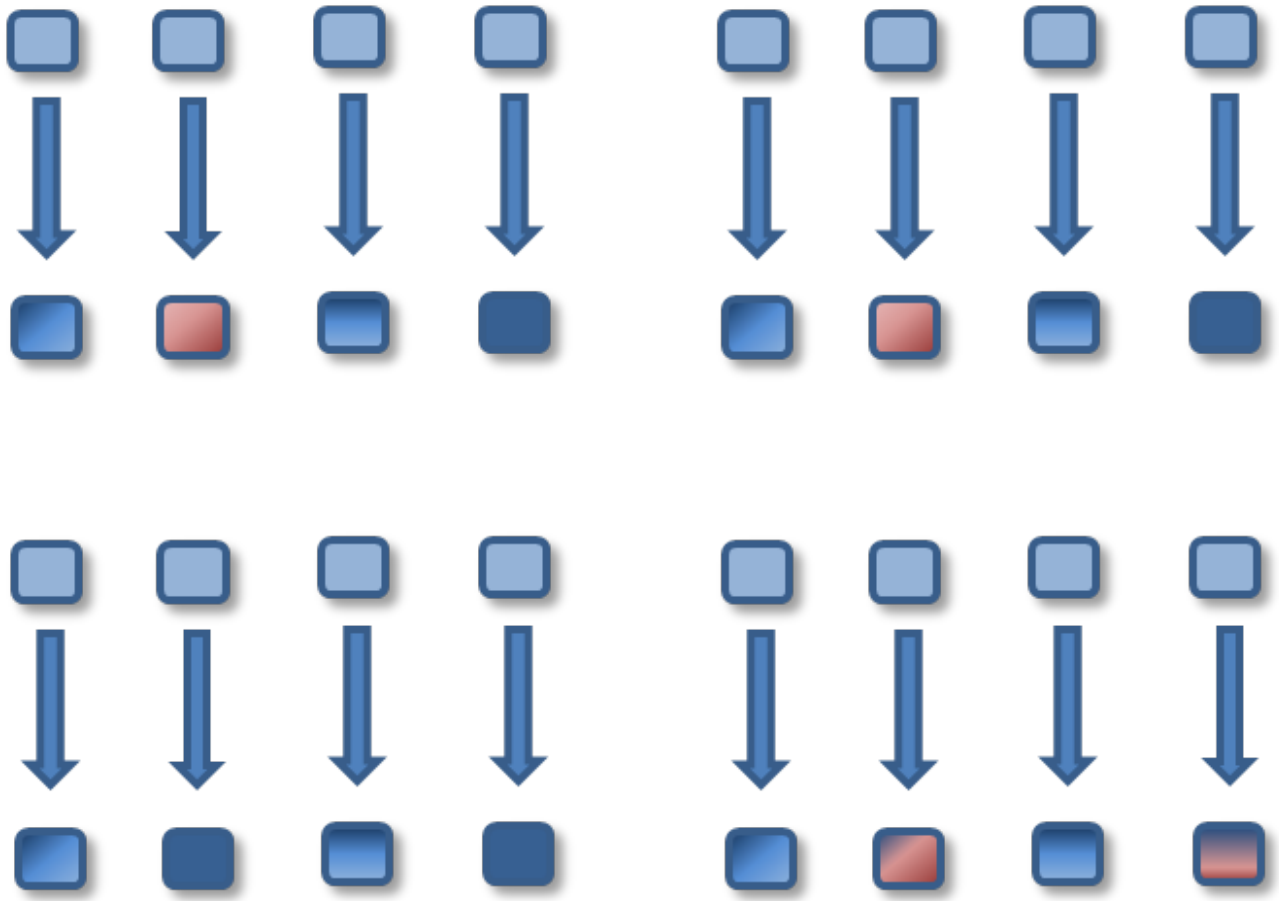
Tizenhat fragmens párhuzamos feldolgozása



16 mag = 16 egyidejű utasítás folyam

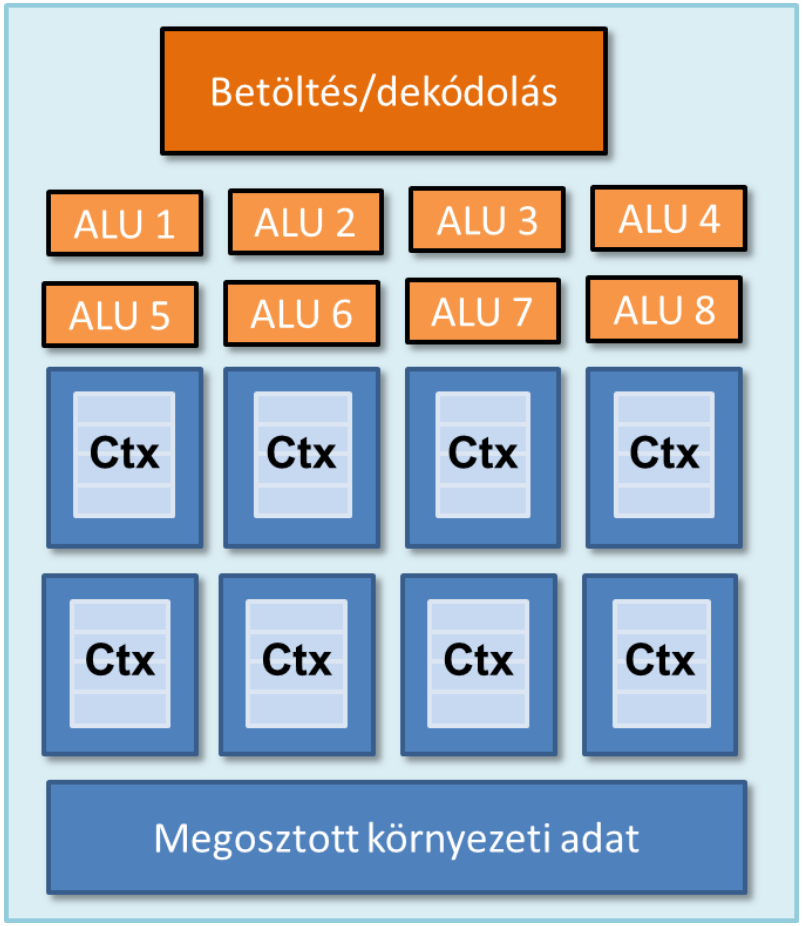
Második ötlet

Fragmensek közötti utasítás folyam megosztása



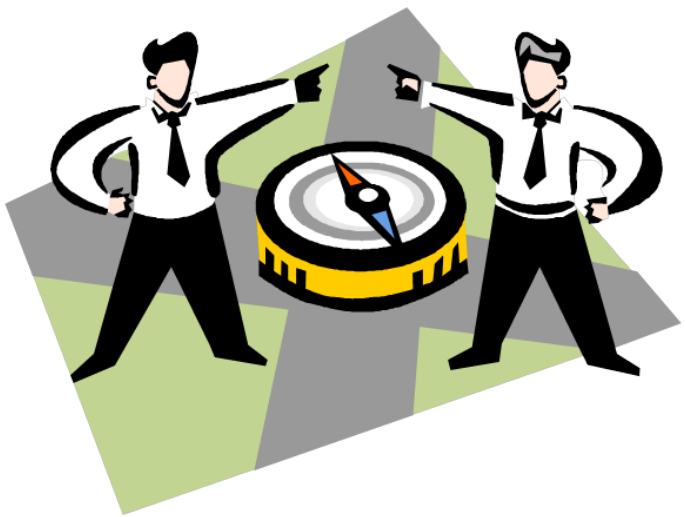
Második ötlet
Single Instruction Multiple Data (SIMD)

- Csökkentsük az ALU-k közötti utasítás folyam
  - Kezelésének költségét
  - Összetettségét



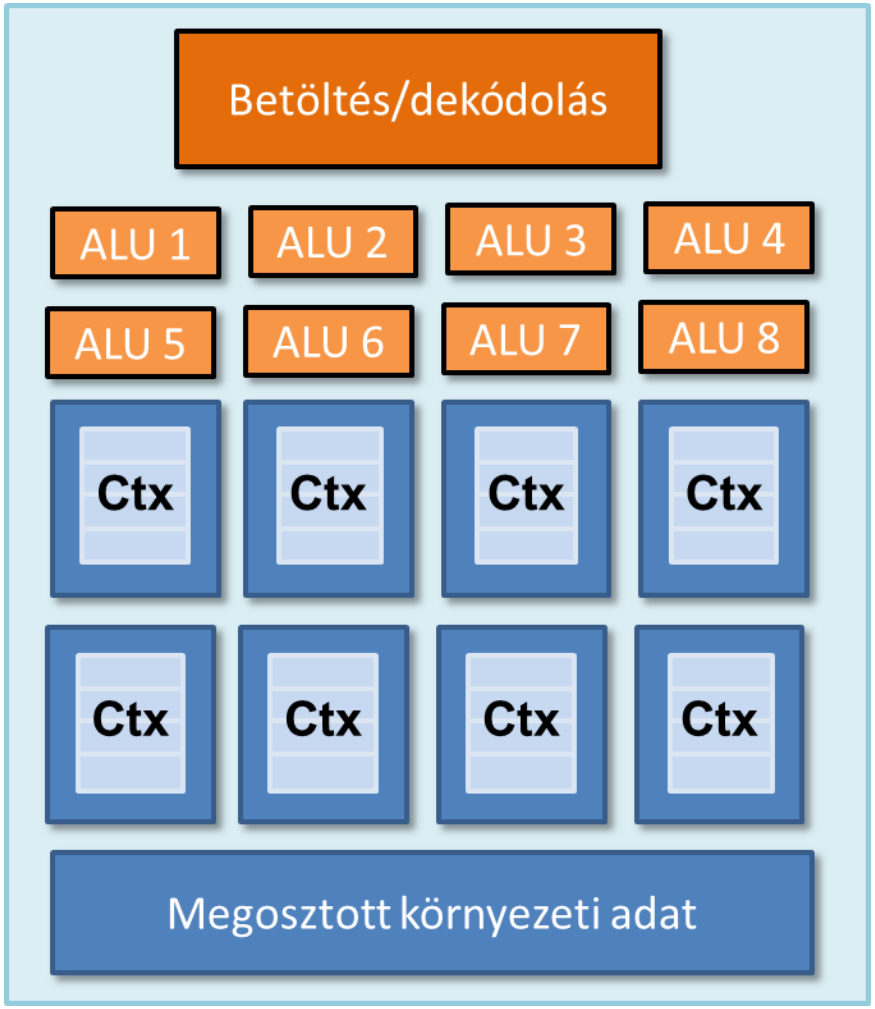
Második ötlet
Shader módosítása

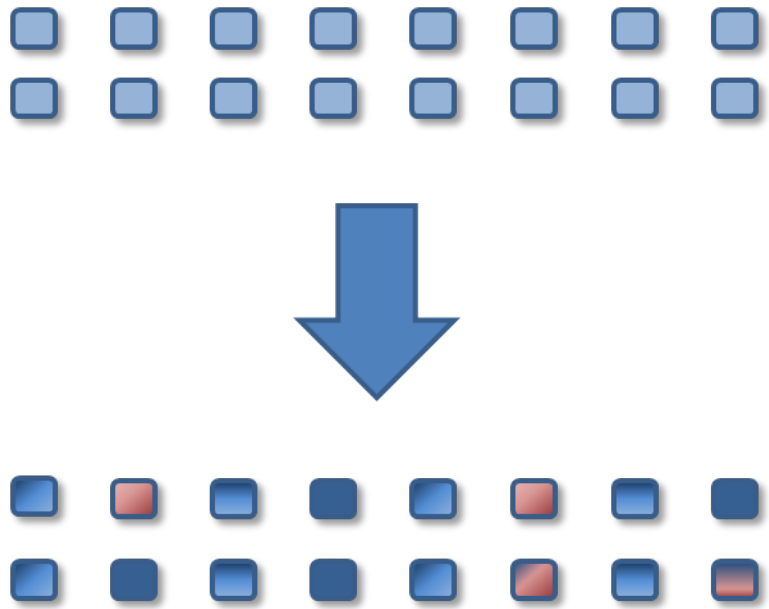
- Előző shader egy fragmenst dolgozott fel
  - Skalár műveletek
  - Skalár operandusok
- Új shader nyolc fragmenst dolgoz fel
  - Vektor műveletek
  - Vektor operandusok



Második ötlet

Fragmensek közötti utasítás folyam megosztása

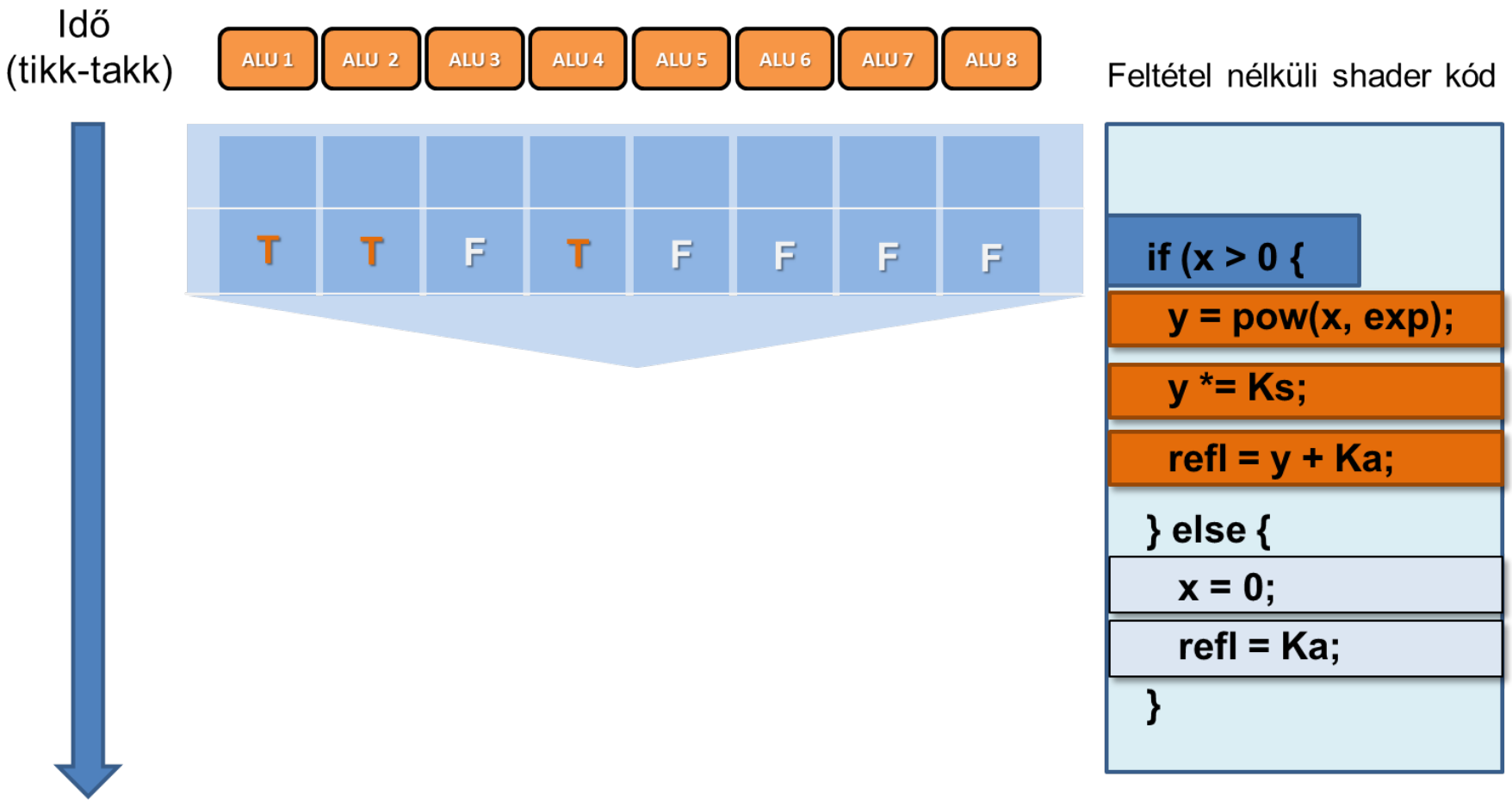




128 fragmens feldolgozása párhuzamosan

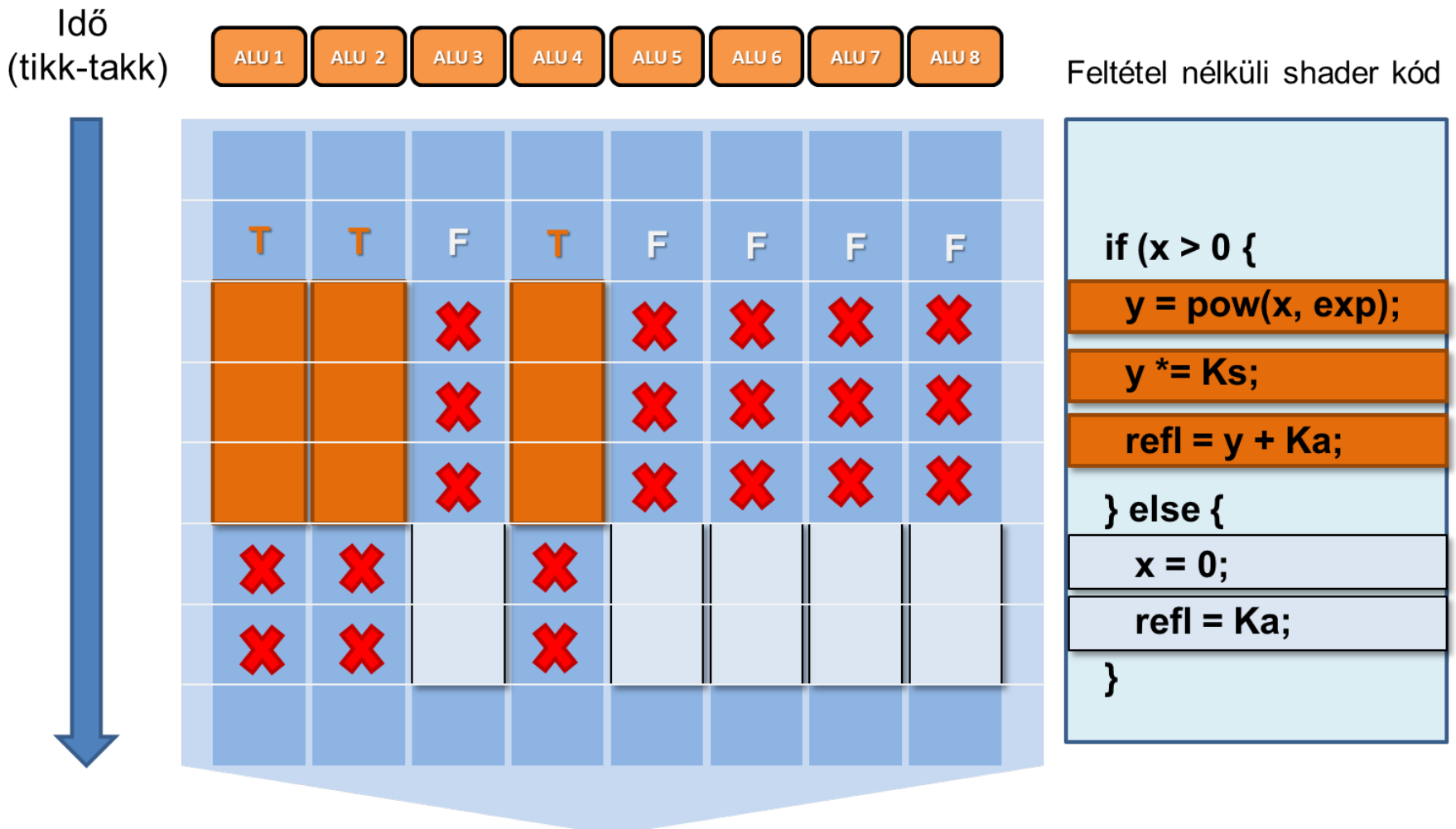
- 16 mag = 128 ALU
- 16 egyidejű utasításfolyam
- 128
  - Vertex
  - Primitív
  - Fragmens

Mi van az elágazásokkal?



Mi van az elágazásokkal?

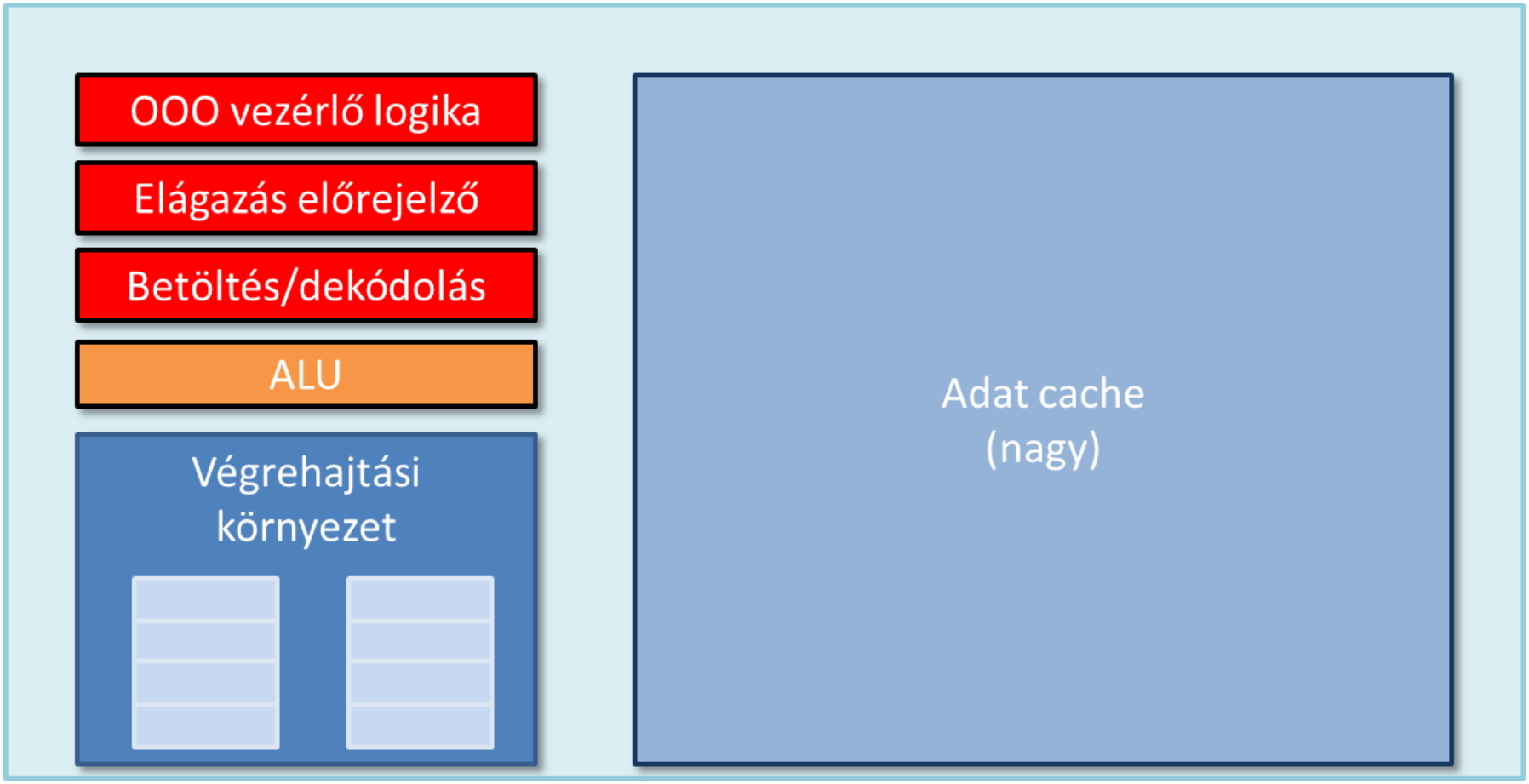
Nem mindegyik ALU végez hasznos munkát



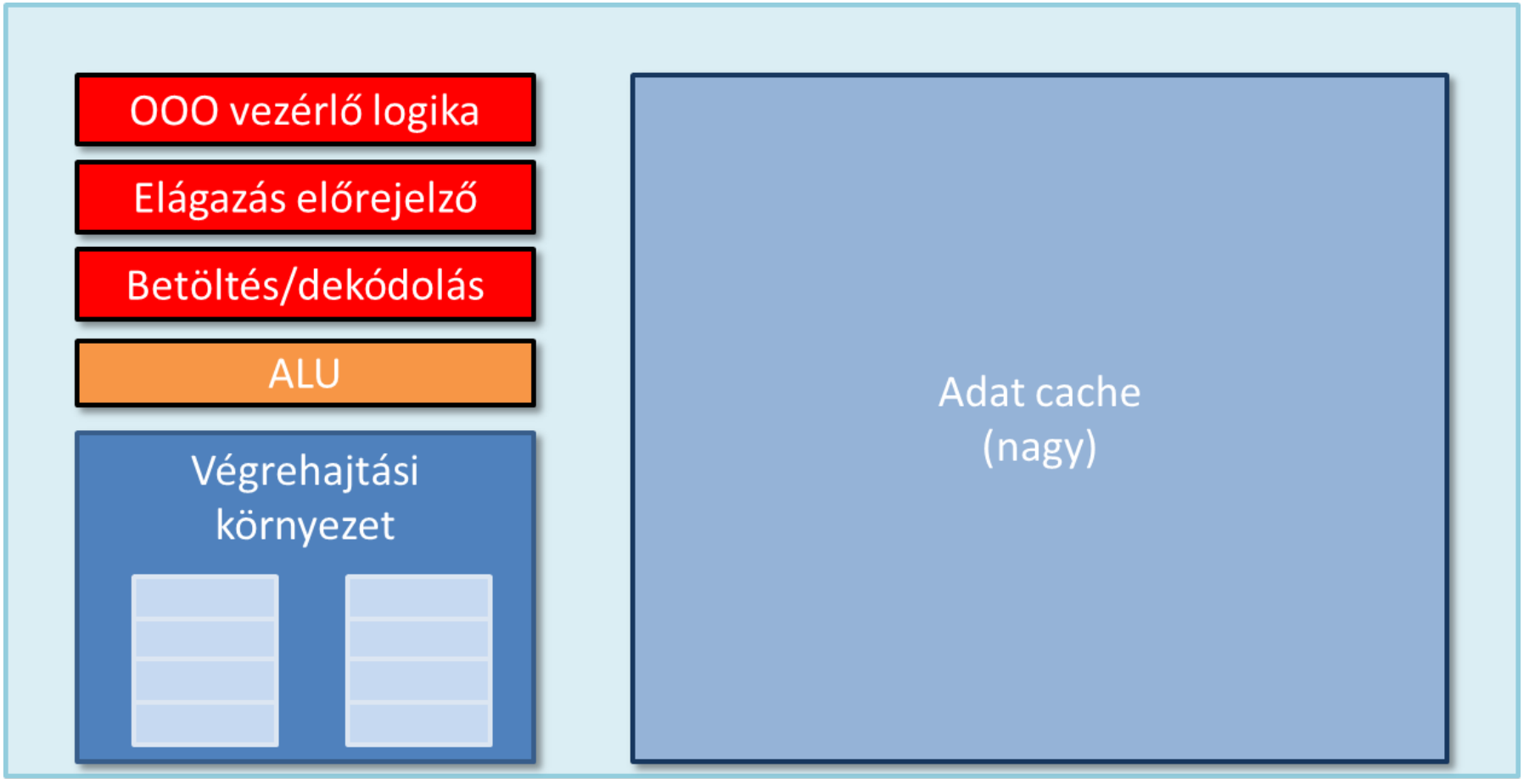
Állások (stalls)

- Állás akkor következik be, amikor egy mag (core) nem tudja futtatni a következő shader utasítást, mivel egy előző utasításra várakozik
  - Függőségek vannak az utasítás folyamban
    - Pl. ADD függ a LOAD befejezésétől
- Késleltetés
  - Adat elérése a memóriából sokszor 1000-nél több ciklust igényel
  - Rossz ötlet volt az első egyszerűsítés?
    - Az eltávolított részek segítenének az állások megoldásában
  - A GPU-k sok független feladatot tételeznek fel
    - Független SIMD csoportok

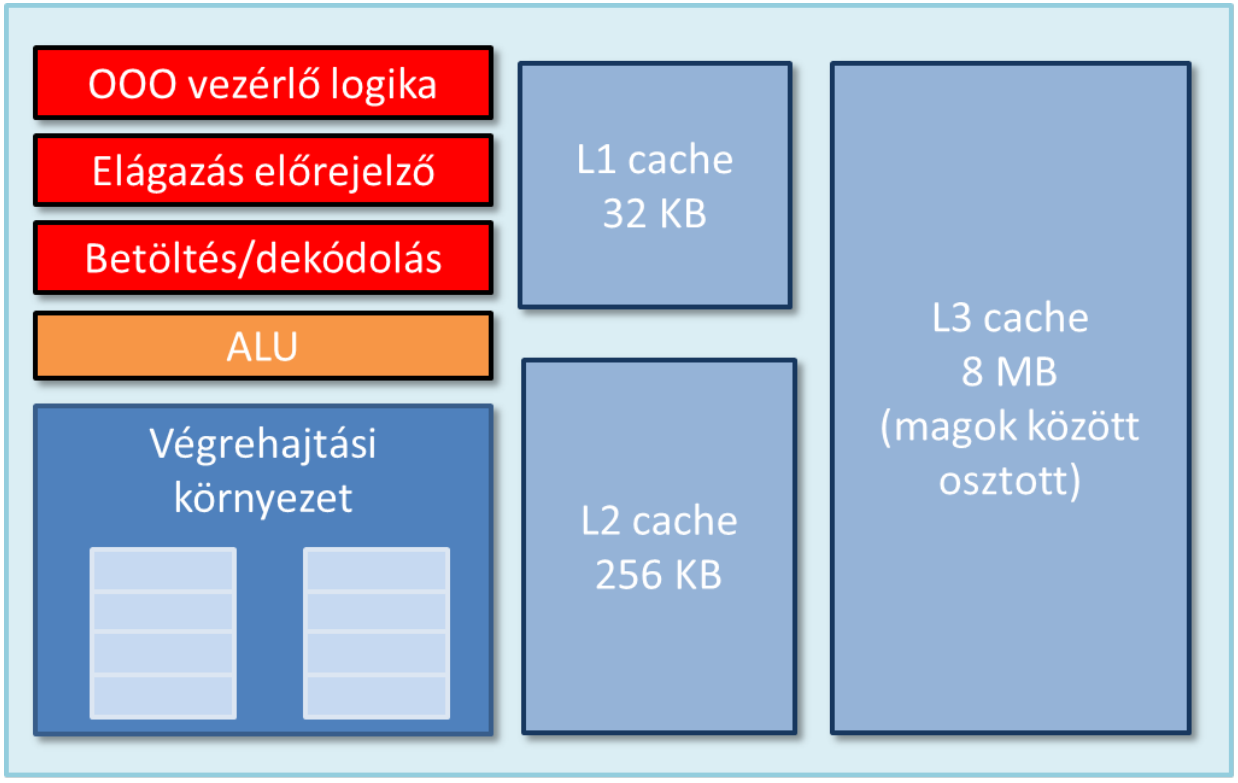
CPU-stílusú magok (core)



CPU-stílusú magok (core)



CPU-stílusú memória felépítés



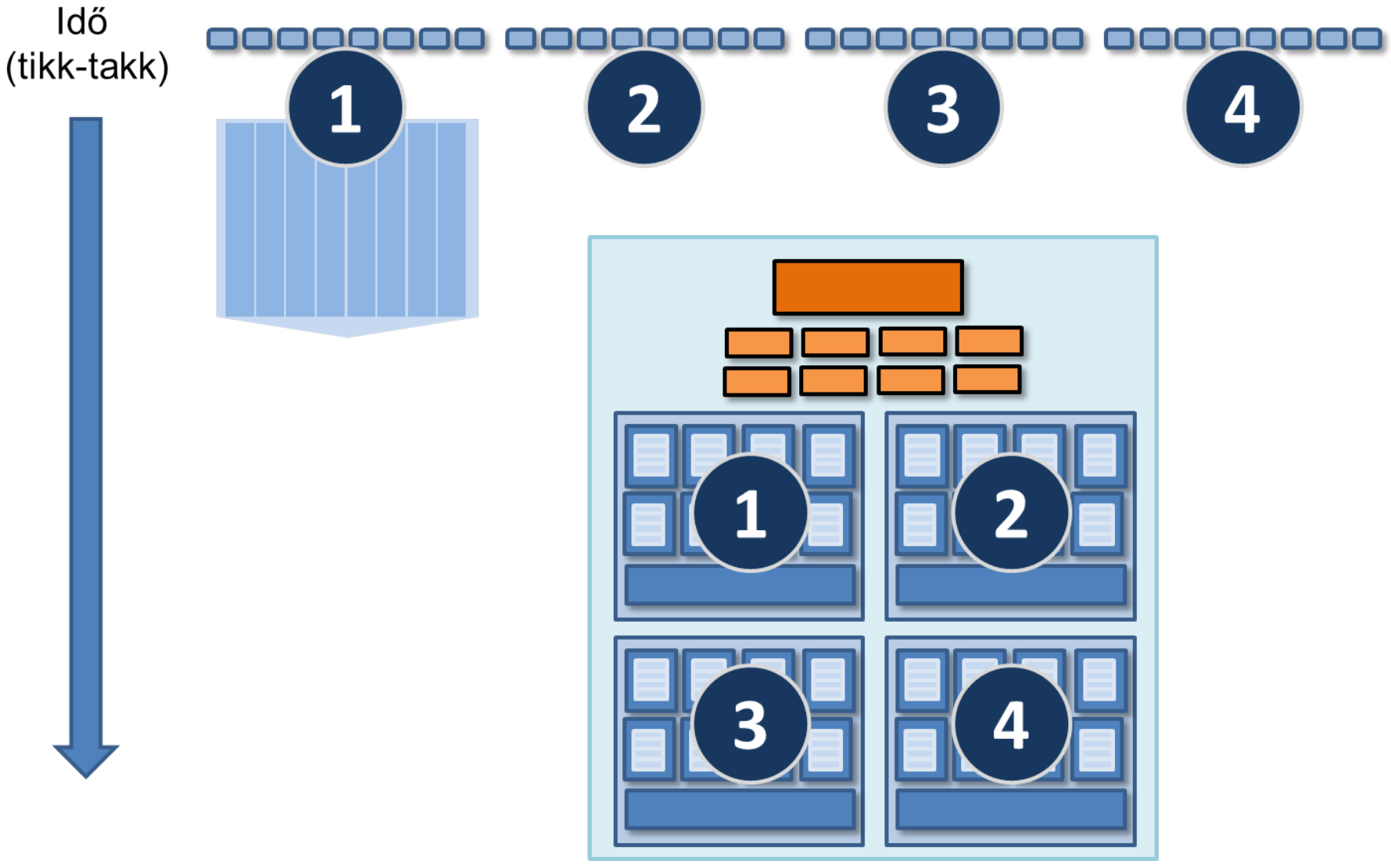
25 GB/sec elérés a memóriához

Magok hatékony kihasználása cache-ben lévő adatok esetén (késleltetés csökkentése, nagy sávszélesség)

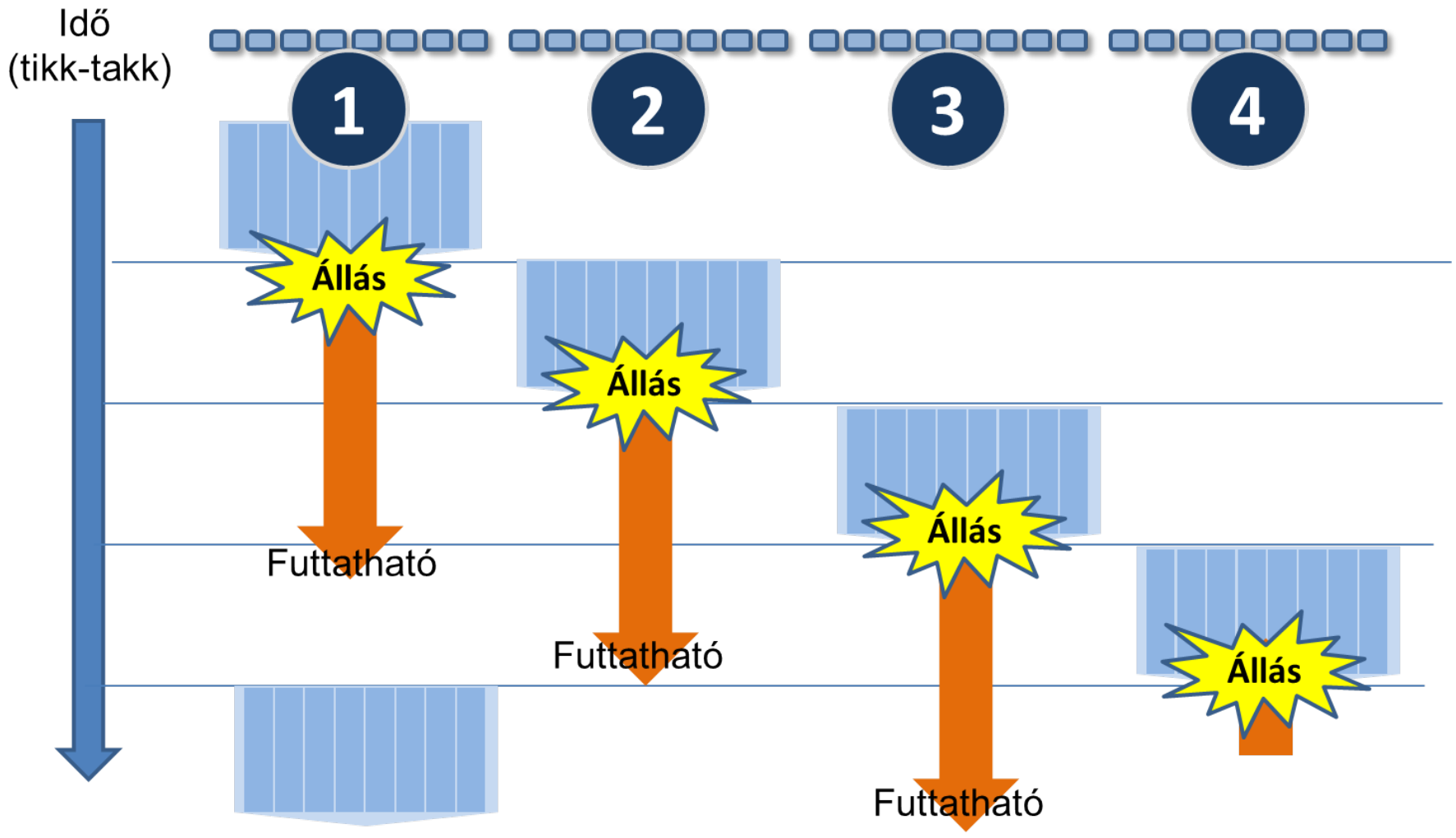
Harmadik ötlet

- Sok független fragmensünk van
- Sok fragmens összefésült feldolgozása egy magon
  - Utasítás folyam váltás egy másik (nem álló) SIMD csoportra, ha az aktív csoport áll
  - GPU hardveresen kezeli
    - Overhead mentesen
    - Ideális esetben teljesen láthatatlan
    - Maximális áteresztőképesség

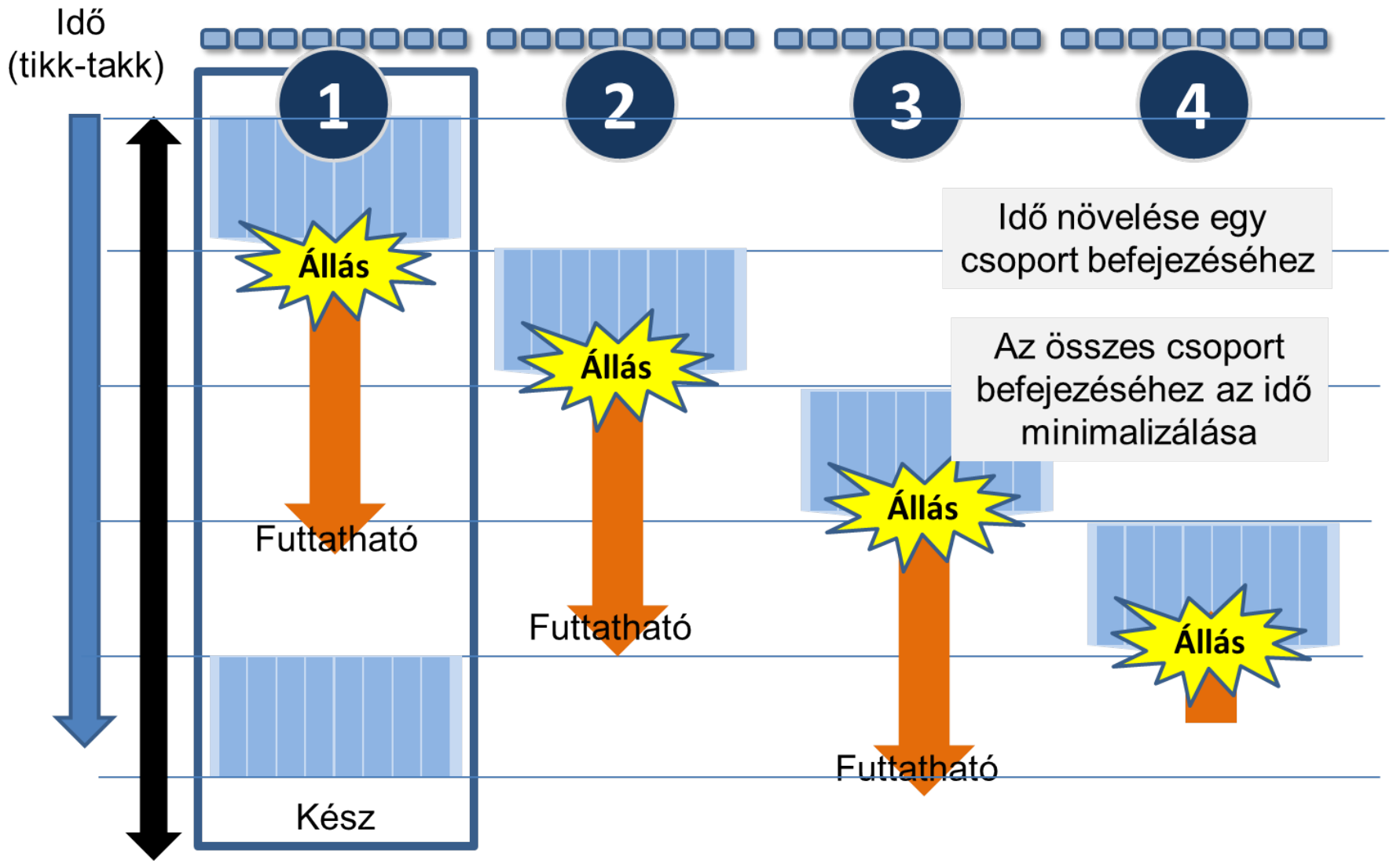
Shader állások elrejtése



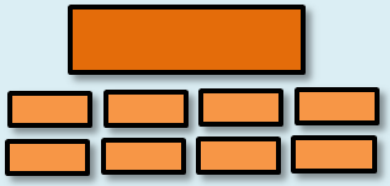
Shader állások elrejtése



Shader állások elrejtése



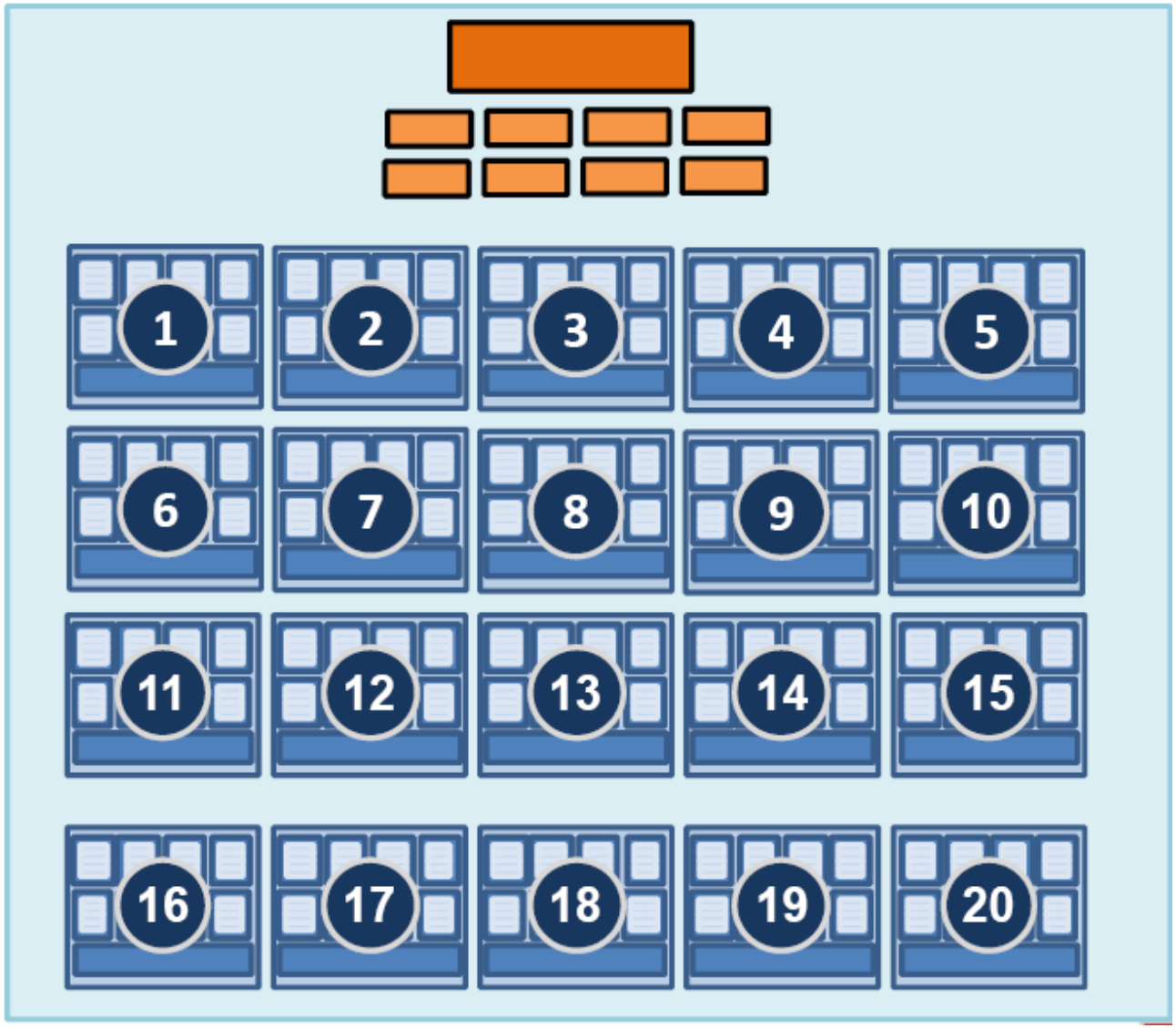
Közös környezet tárolása



Közös környezet készlet tárolás
64 KB

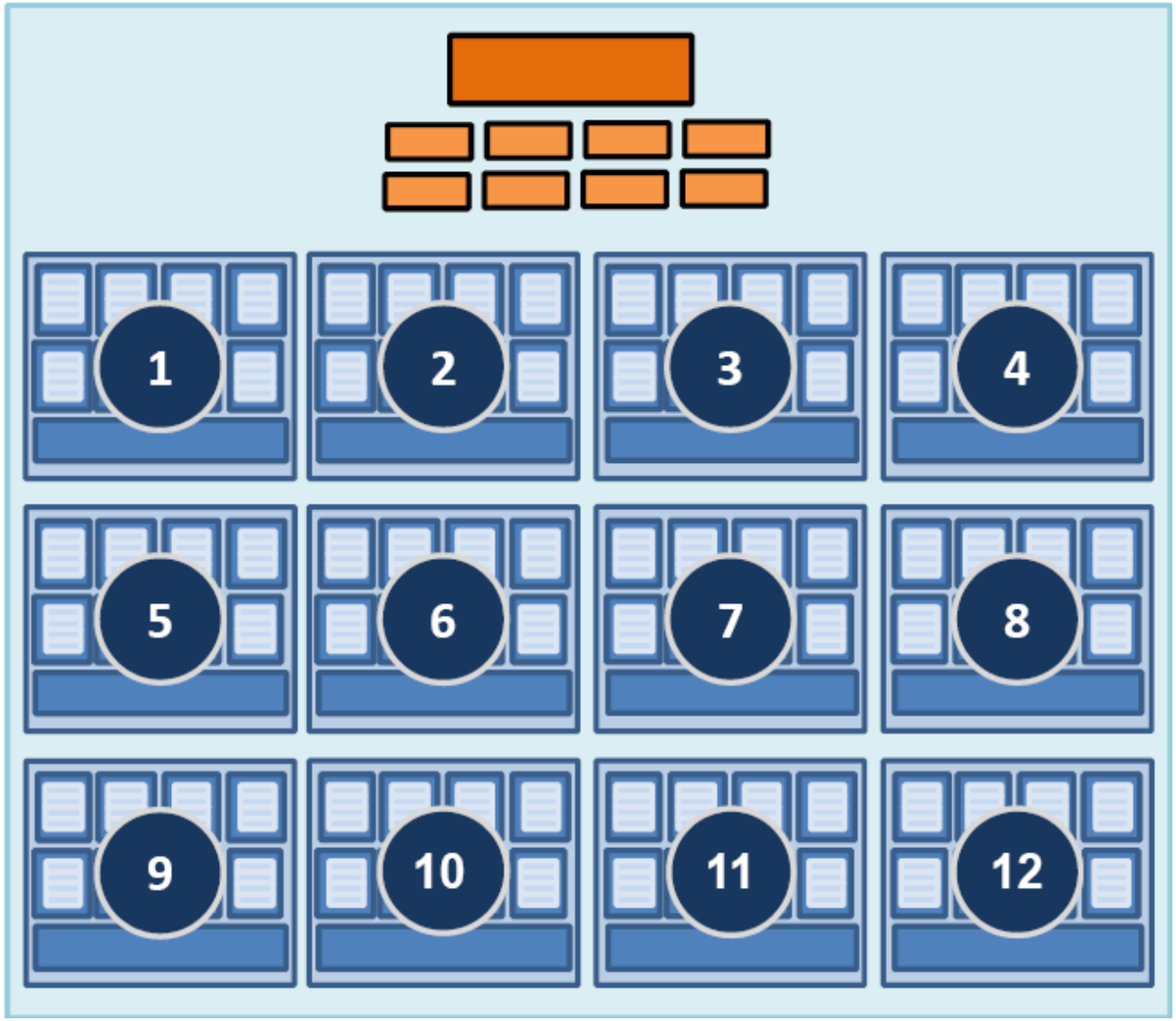
Maximális késleltetés elrejtési képesség

Húsz kicsi környezet



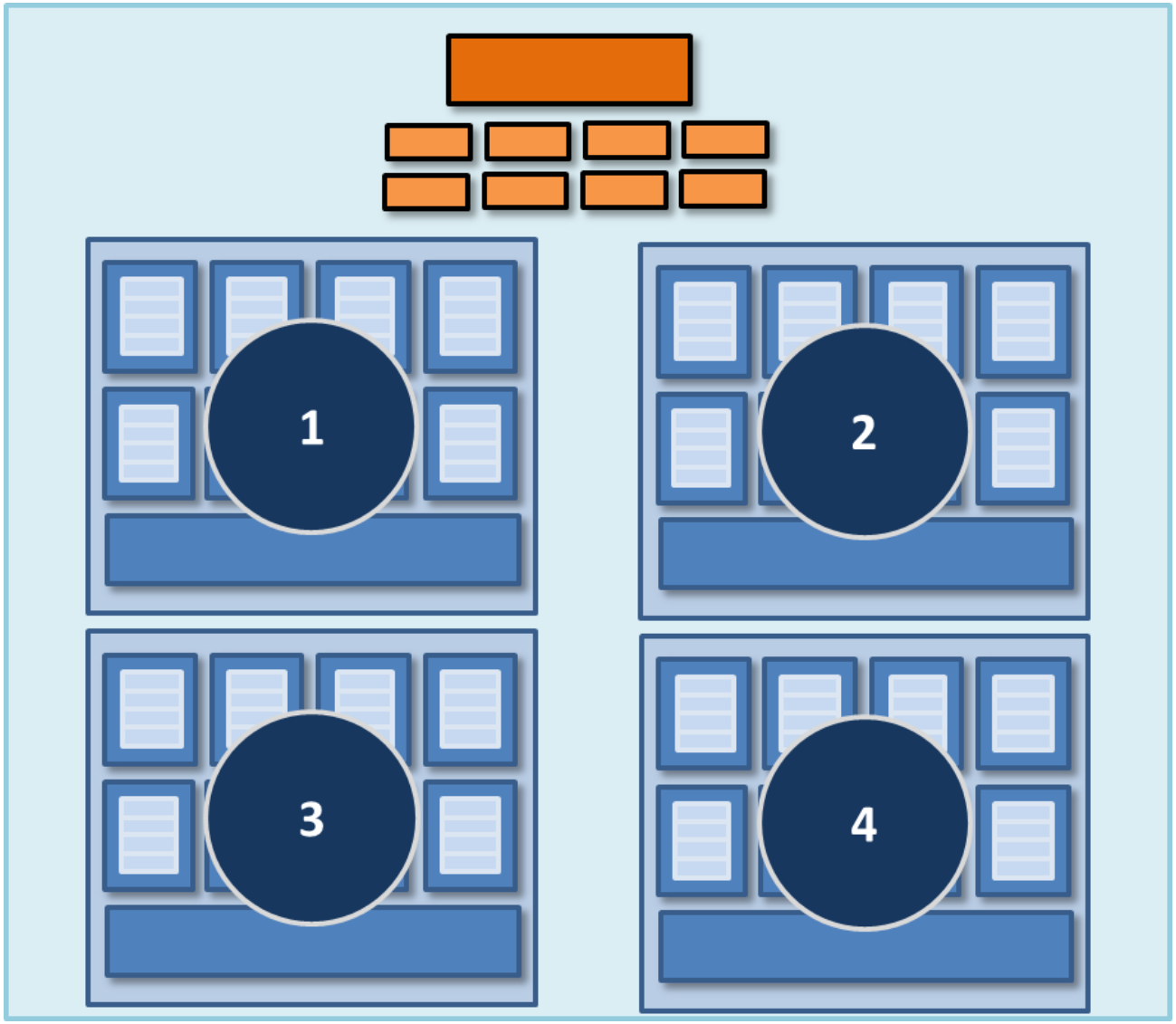
Közepes késleltetés elrejtési képesség

Tizenkét kicsi környezet



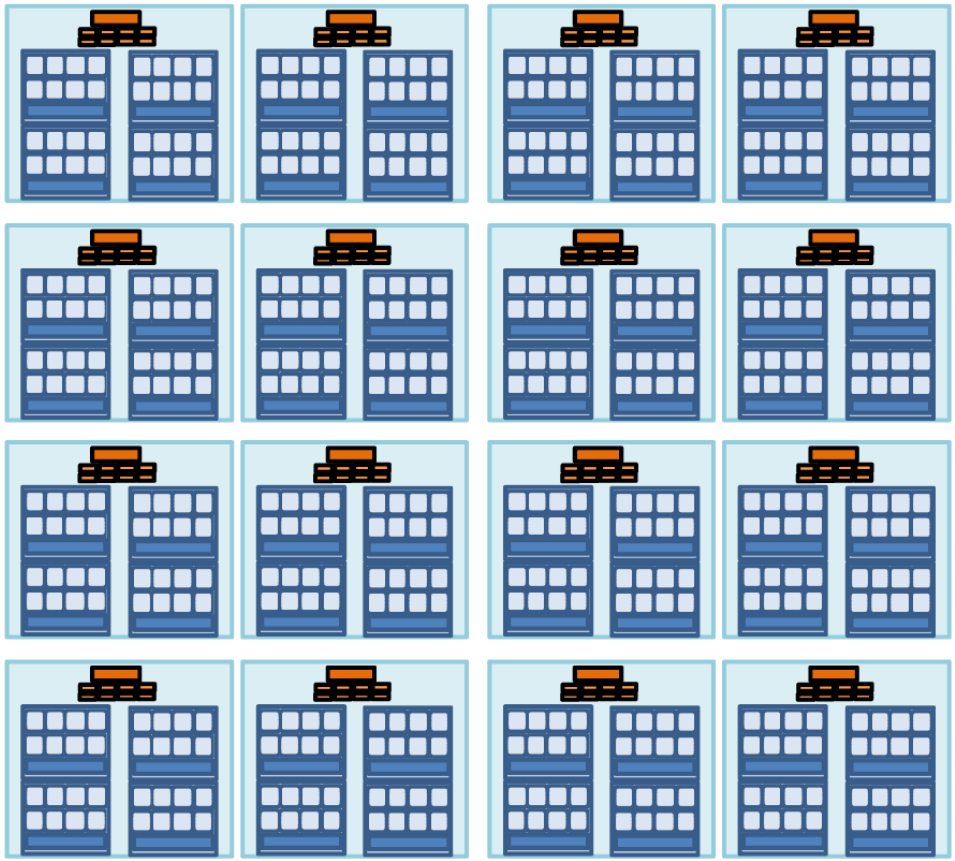
Kicsi késleltetés elrejtési képesség

Négy nagy környezet



GPU shading rendszer

- 16 mag
- 8 mul-add ALU magonként (128 összesen)
- 16 egyidejű utasítás folyam
- 64 konkurens (összefésült) utasítás folyam
- 512 konkurens fragmens



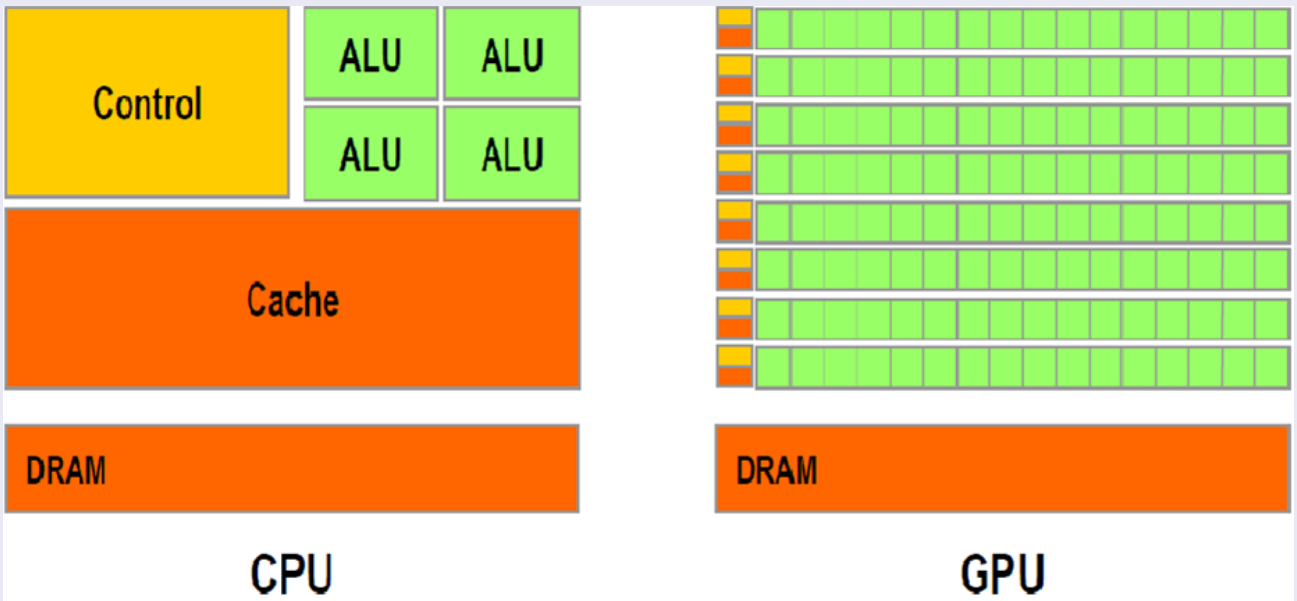
=256 GFLOPS (@1 GHz)

Összefoglalás

Három kulcs ötlet

- Használjunk sok, karcsúsított magot a párhuzamos futtatáshoz
- A magokat rakjuk tele ALU-kkal
  - Megosztott utasítás folyamok fragmensek csoportjainál
- Kerüljük el a késleltetett állásokat fragmens csoportok összefésült végrehajtásával
  - Amikor egy csoport áll, akkor dolgozzunk egy másik csoporton

CPU-GPU összehasonlítás



Összefoglalás

Emlékezzünk a következőkre!

- A GPU-ra egy több magos processzorként gondoljunk, amelyet arra optimalizáltak, hogy
  - A vertex és fragmens adatok maximális „áteresztéssel folynak át” a grafikus csővezetéken
  - Speciálisan támogatja
    - A grafikus csővezeték leképezését ezekre az erőforrásokra
    - A raszterizálást
    - Vágást
    - Hátsó oldal eltávolítást
    - Textúrázást
    - Stb.

- Nagyobb és gyorsabb
  - Több mag
  - Nagyobb FLOPS (manapság 2 TFLOP)
- Milyen fix-funkcióknak kell megmaradnia?
- Néhány CPU-hoz hasonló tulajdonság hozzáadása
  - Általános R/W cache (Fermi)
  - Szinkronizálás

- Alternatív programozási felületek támogatása
  - Általános célú programozás
    - CUDA
    - OpenCL
    - DirectCompute
  - Alkalmazások, amelyek a GPU-t egy több processzoros rendszernek tekintik
- Hogyan változik a grafikus csővezeték absztrakció?
  - Direct3D 11
    - 3 új csővezeték szakasz
- Sugárkövetés