

Lecture 25: Parallel archs

Friday, March 8, 2019 10:49 AM

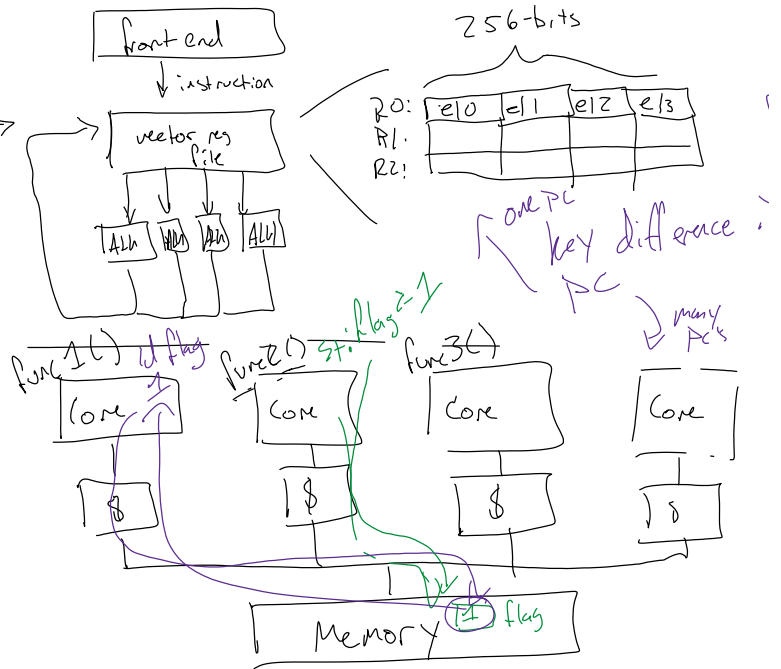
Outline

- Example parallel systems
- Types of parallel systems
- Flynn's Taxonomy

Do course evals!
eval.ucdavis.edu

Exmpk parallel system

SIMD → Single instruction multiple data
↳ ex: AVX/SSE extensions to x86
NEON (ARM)
3DNOW (AMD)



Multi-core processor

How to program?

- ↳ threads execute concurrently
- ↳ independent execution streams
- ↳ "Normal" ISA / no need to add instructions
- ↳ locks / synchronization
- ↳ communicating through shared memory
- ↳ diff processes
- ↳ share virtual memory

Proc 1
ld 0x10000

Proc 2
str 0x0000

MIMD → multiple inst. multiple data

Graphics processing units

- programming?

CUDA API

- ↳ run a diff program on GPU controlled by CPU
- ↳ explicitly copy data

"Accelerator interface"

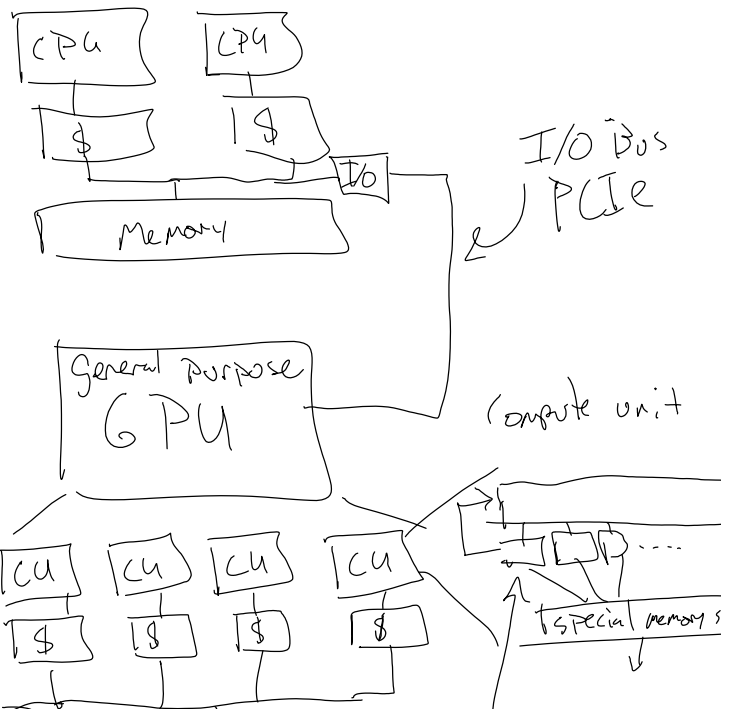
key diff. from SIMD

- ↳ instead of one PC each processing element has its own PC

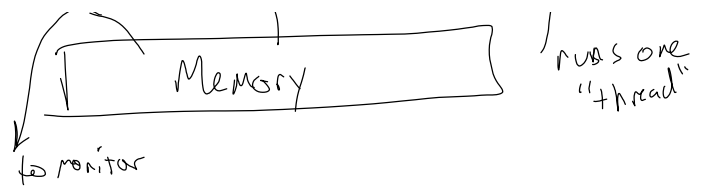
↳ programmed like MIMD

2000 + concurrent threads

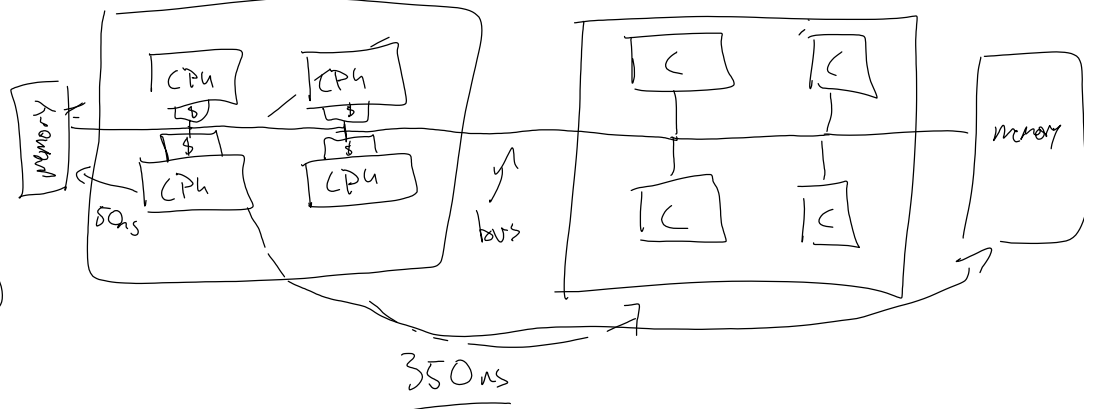
compute units
8-64



↳ 10,000's of waiting threads
no shared memory between CPU-GPU

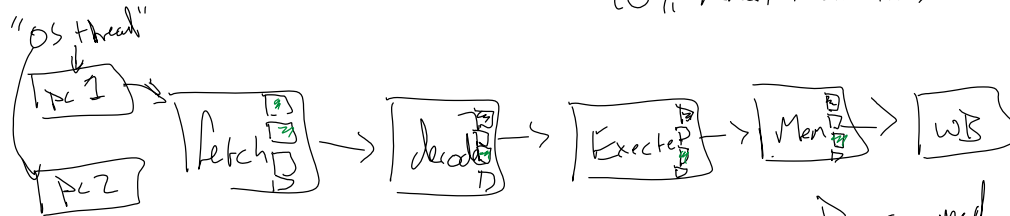


Multichip systems
↳ MIMD
high sync. overhead
Think about where data is
Non-uniform memory architecture (NUMA)



Many modern workloads spend
lots of time waiting for memory
executing no instructions

AMAT → 3 cycles
6 inst per cycle
40% memory instructions
} low utilization of pipeline



when thread 1 hit cache miss start fetching thread 2

↳ Coarse-grained multithreading

Fine-grained multithreading → each cycle pick from diff thread

Simultaneous multithreading → in one cycle choose from many threads

↳ Intel hyperthreading

Programmed like MIMD
↳ to use loads like
2 cores