Storage and Computing in CryoEM - a Roundtable

Steve Ludtke
Michael Cianfrocco
Pawel Penczek
David Mastronarde

Relative Scaling

	1996	2006	2016
CPU (workstation)	0.1 GFLOPS/s	20 GFLOPS/s	1000 GFLOPS/s
GPU		50 GFLOPS/s	10,000 GFLOPS/s
Disk Capacity	1 GB	100 GB	10,000 GB
Disk I/O	8 MB/s	80 MB/s	600 - 1500 MB/s
Network	10 Mb/s	100 Mb/s	1000 Mb/s

100x in 10 years

10x in 10 years

Storage 2018

- 1 Drive = 12 TB
- 8 Drive RAID5 part of workstation = 84 TB $3.5k + CPU
- 12 Drive RAID6 under Desk = 120 TB for ~$7k
- 4U Rackmount, 36 drive + CPU = 360 TB RAID6 ~$20k
- 1 PB (usable) ~$50k

Compute Choices

- CPU
  - Xeon - 28 cores/chip @ 2.5 Ghz (max), 20 @ 2.4 Ghz (typ wkst) - ~1.5 TFLOPS
  - Phi - 72 cores/chip @ 1.5 Ghz (max) - 3.5 TFLOPS
  - AMD - EPYC 32 @ 2.2 Ghz (max) - 1.3 TFLOPS? (but cheaper)
  - ARMv8 - ThunderX2 54 @ 3 Ghz - competitive with Xeon

Compute Choices

- CPU
  - Xeon - 28 cores/chip @ 2.5 Ghz (max), 20 @ 2.4 Ghz (typ wkst) - ~1.5 TFLOPS
  - Phi - 72 cores/chip @ 1.5 Ghz (max) - 3.5 TFLOPS
  - AMD - EPYC 32 @ 2.2 Ghz (max) - 1.3 TFLOPS? (but cheaper)
  - ARMv8 - ThunderX2 54 @ 3 Ghz - competitive with Xeon
- GPU
  - NVIDIA - Titan V - 5120 cores, 640 tensor @ 1.2 Ghz - 12 TFLOPS SP
  - AMD - Vega - ? @ 1.6 Ghz - 11 - 13 TFLOPS SP (OpenCL)

Compute Choices

GPU

- Pros:
  - Fast
  - Cost effective (at least short-term)
- Cons:
  - Tricky to program well
  - Performance varies widely among algorithms
  - Rapid changes require frequent reprogramming
  - Short useful lifetime

Compute Choices

- CPU
  - Xeon - 28 cores/chip @ 2.5 Ghz (max), 20 @ 2.4 Ghz (typ wkst) - ~1.5 TFLOPS
  - Phi - 72 cores/chip @ 1.5 Ghz (max) - 3.5 TFLOPS
  - AMD - EPYC 32 @ 2.2 Ghz (max) - 1.3 TFLOPS? (but cheaper)
  - ARMv8 - ThunderX2 54 @ 3 Ghz - competitive with Xeon
- GPU
  - NVIDIA - Titan V - 5120 cores, 640 tensor @ 1.2 Ghz - 12 TFLOPS SP
  - AMD - Vega - ? @ 1.6 Ghz - 11 - 13 TFLOPS SP (OpenCL)
- FPGA - Field Programmable Gate Array (Altera -> Intel), 10 TFLOPS + GA

Compute Choices

- CPU
  - Xeon - 28 cores/chip @ 2.5 Ghz (max), 20 @ 2.4 Ghz (typ wkst) - ~1.5 TFLOPS
  - Phi - 72 cores/chip @ 1.5 Ghz (max) - 3.5 TFLOPS
  - AMD - EPYC 32 @ 2.2 Ghz (max) - 1.3 TFLOPS? (but cheaper)
  - ARMv8 - ThunderX2 54 @ 3 Ghz - competitive with Xeon
- GPU
  - NVIDIA - Titan V - 5120 cores, 640 tensor @ 1.2 Ghz - 12 TFLOPS SP
  - AMD - Vega - ? @ 1.6 Ghz - 11 - 13 TFLOPS SP (OpenCL)
- FPGA - Field Programmable Gate Array (Altera -> Intel), 10 TFLOPS + GA
- Neural Net - Intel Nervana, Apple A11, Google (180 TFLOPS), Titan V 120 TFLOPS