Long term vision for LArSoft: Overview

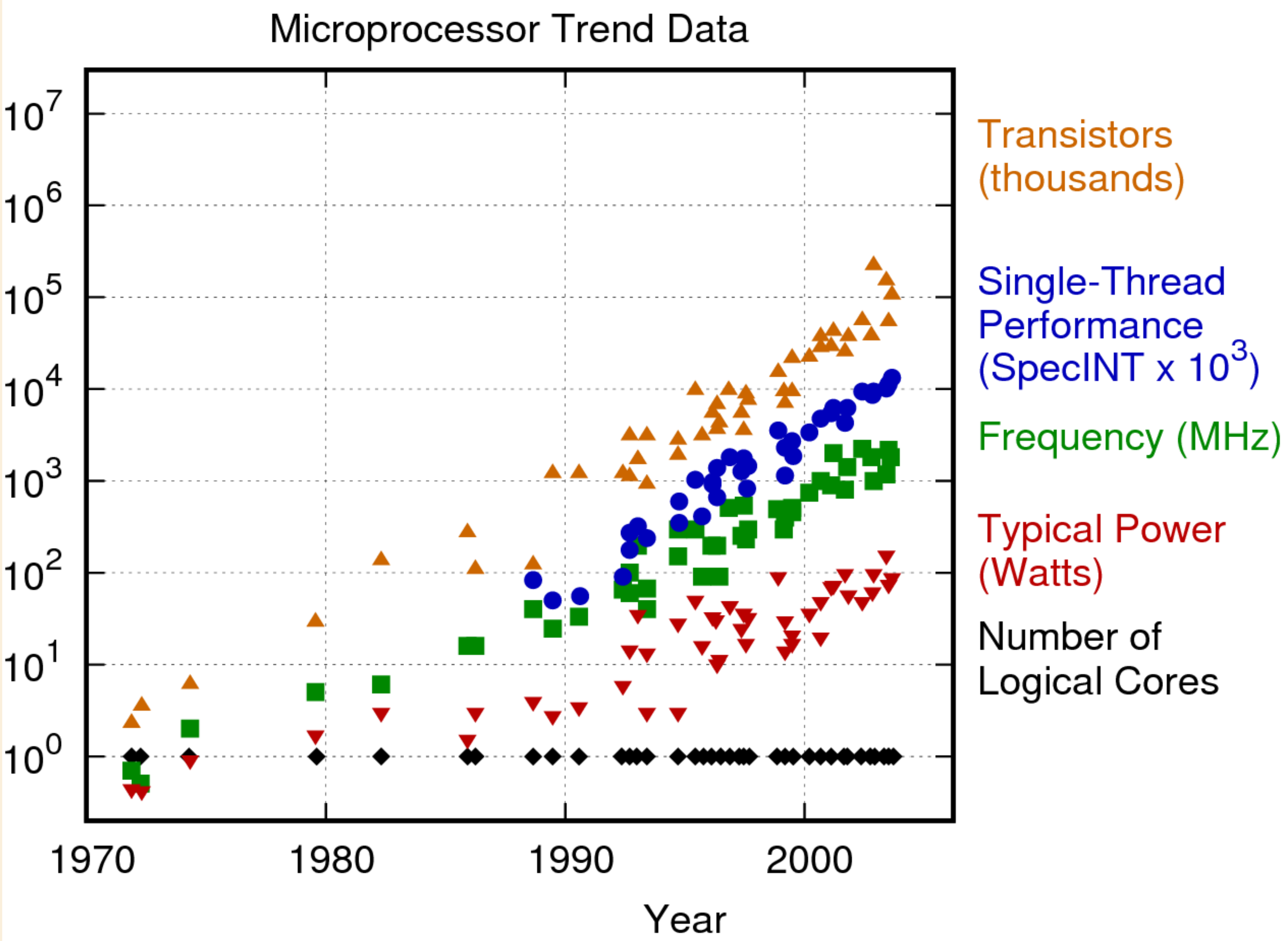
Adam Lyon

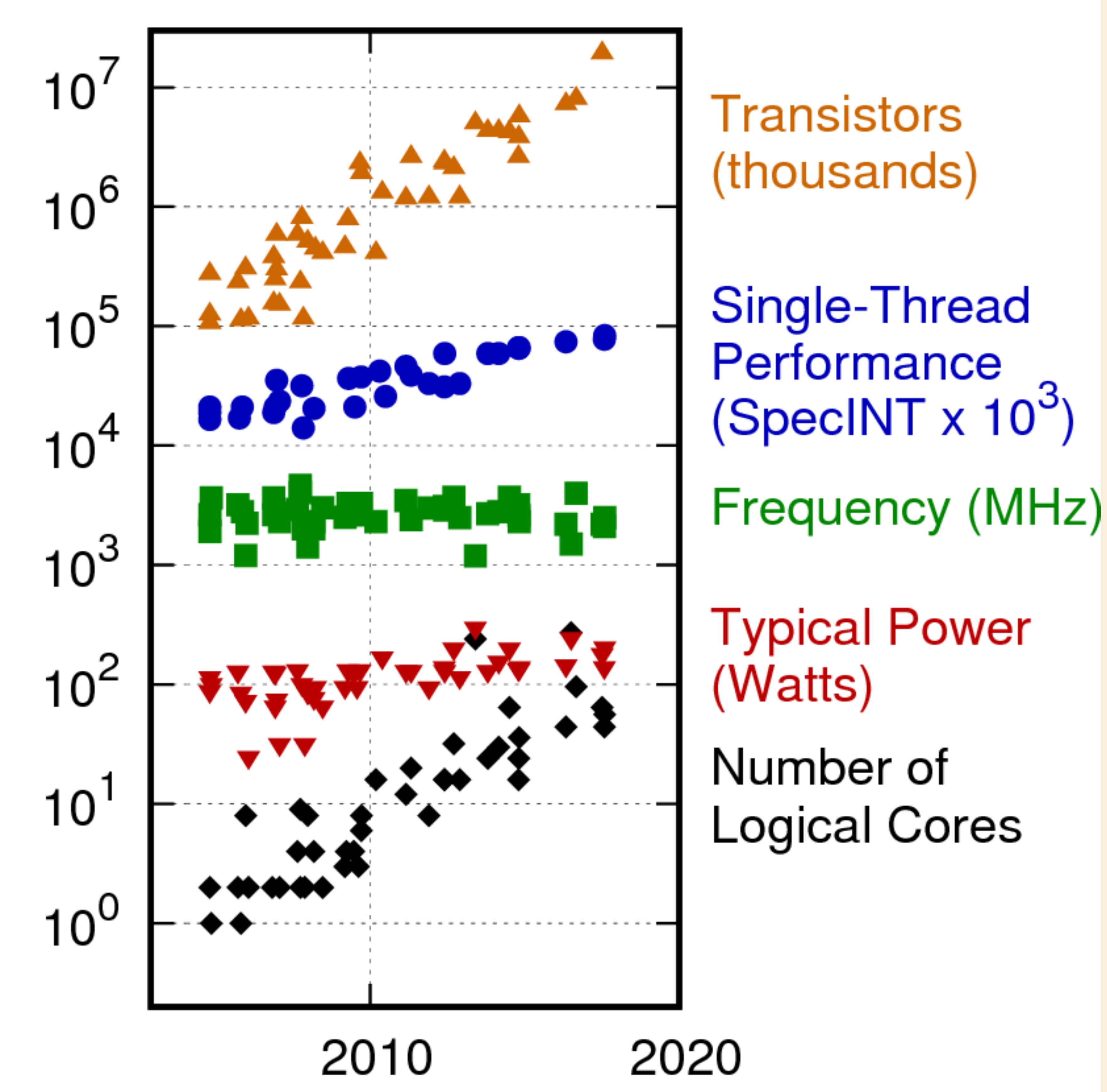
LArSoft Workshop 2019

25 June 2019

Long term computing vision

- You already know this…





https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/

The response - Multicore processors

Examples…

Intel Xeon “Haswell”:
16 cores @ 2.3 GHz; 32 threads; Two 4-double vector units

Intel Xeon Phi “Knights Landing (KNL)”:
68 cores @ 1.4 GHz; 272 threads; Two 8-double vector units

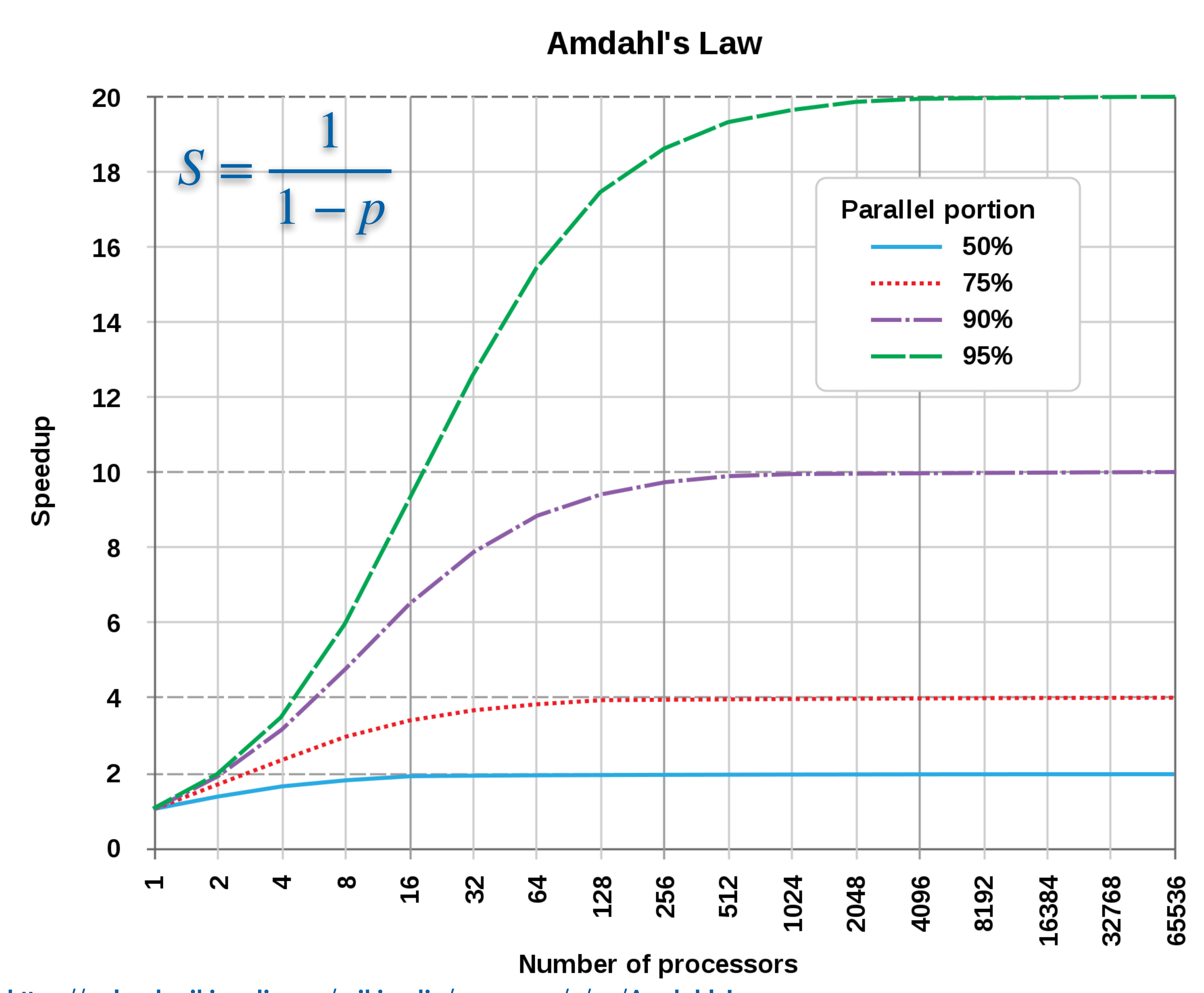
Nvidia Volta “Tesla V100” GPU:
5120 CUDA cores; 640 Tensor cores @ ~1.2 GHz

Grid computing uses one or more “cores” (really threads) per job

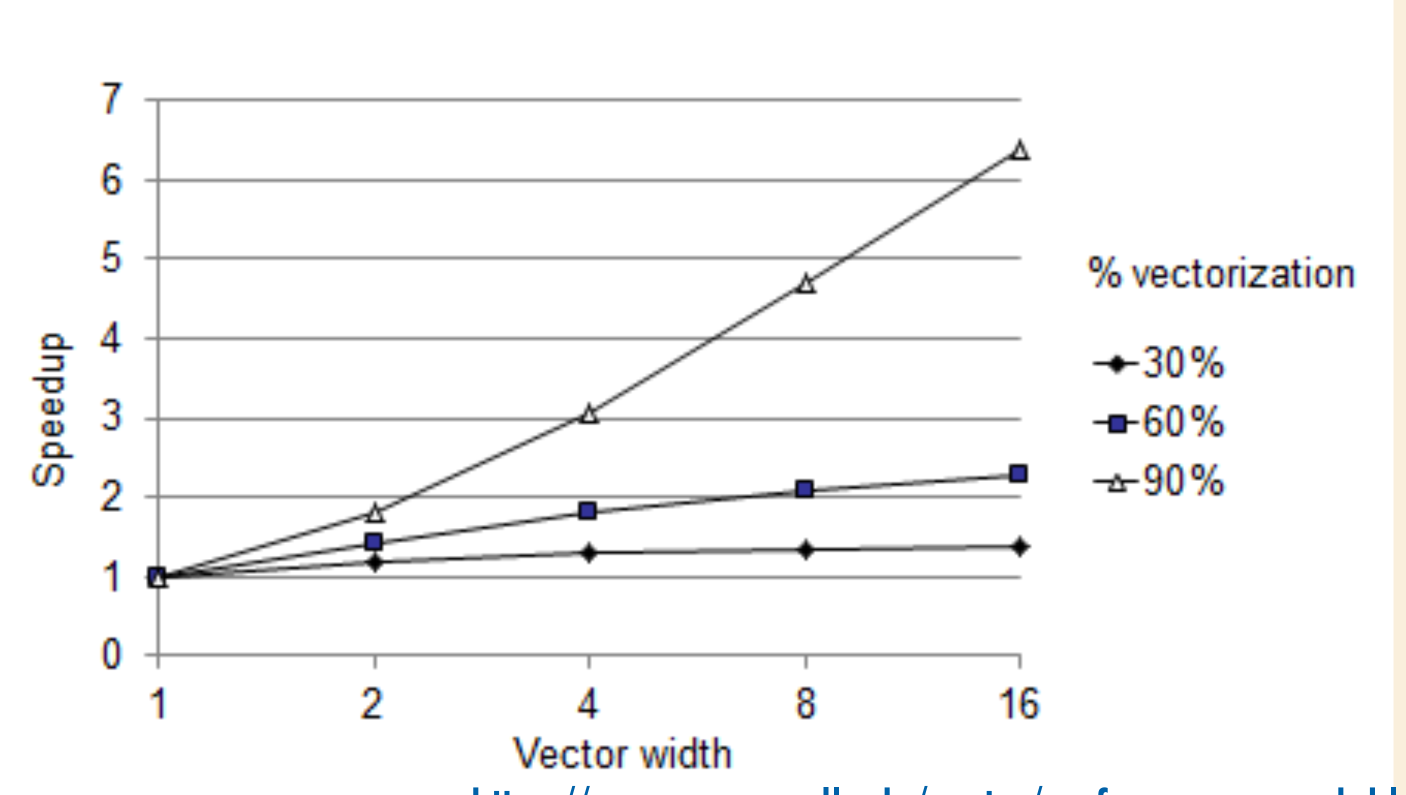
Advantages of multi-threading…

- Main advantage is memory sharing
- If you are looking for speedup, remember Amdahl’s law

Vectorization is another source of speedup … maybe

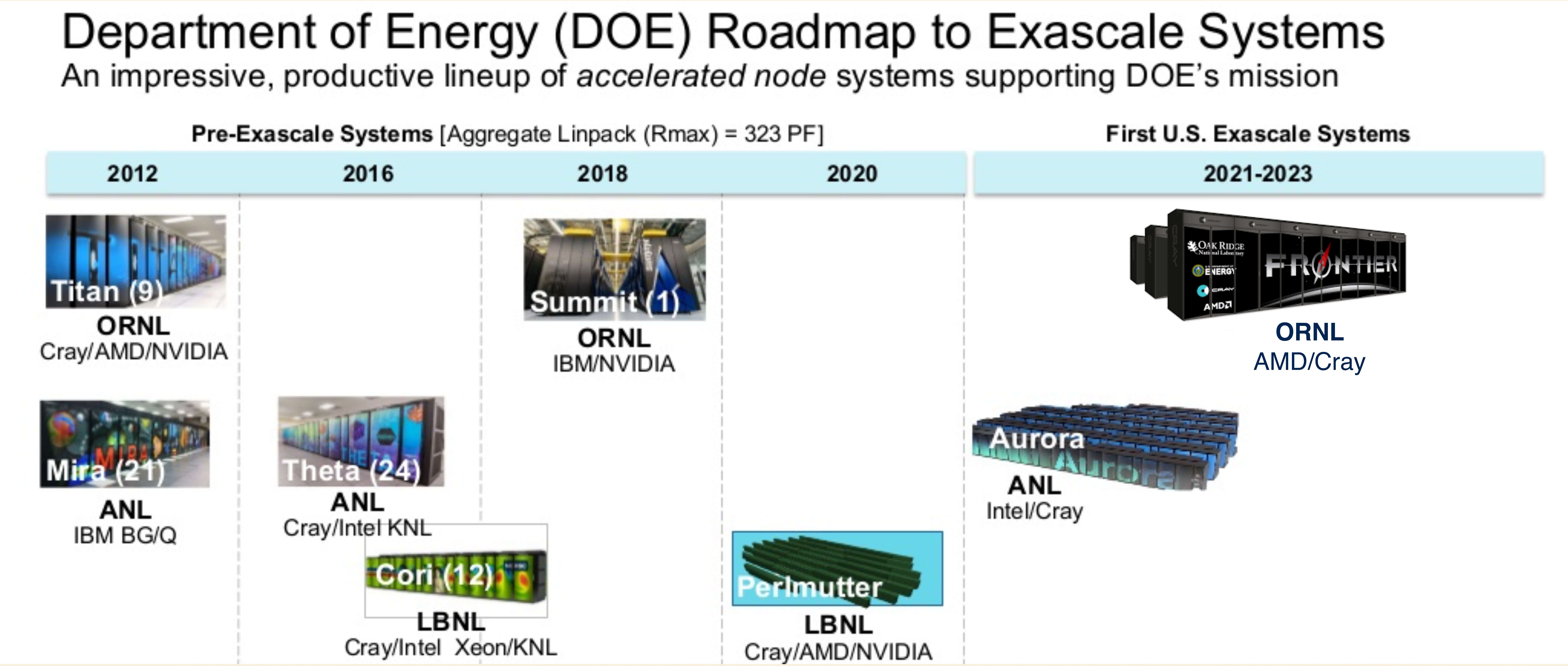


https://upload.wikimedia.org/wikipedia/commons/e/ea/AmdahlsLaw.svg



https://cvw.cac.cornell.edu/vector/performance_amdahl

High Performance Computing (next 5 years)



Department of Energy (DOE) Roadmap to Exascale Systems

An impressive, productive lineup of *accelerated node* systems supporting DOE's mission

Pre-Exascale Systems [Aggregate Linpack (Rmax) = 323 PF]

First U.S. Exascale Systems

2012

Titan (9)
ORNL
Cray/AMD/NVIDIA

Mira (21)
ANL
IBM BG/Q

2016

Theta (24)
ANL
Cray/Intel KNL

Cori (12)
LBNL
Cray/Intel Xeon/KNL

2018

Summit (1)
ORNL
IBM/NVIDIA

2020

Perlmutter
LBNL
Cray/AMD/NVIDIA

2021-2023

ORNL
AMD/Cray

Aurora
ANL
Intel/Cray

https://www.slideshare.net/insideHPC/exascale-computing-project-software-activities

Heterogenous Computing

- Future: multi-core, limited power/core, limited memory/core, memory bandwidth increasingly limiting
- The old days are not coming back

- The DOE is spending $2B on new “Exascale” machines (10^{18} floating point operations/sec) …

- OLCF: Summit	IBM CPUs & 27K NVIDIA Volta GPUs (#1 supercomputer in the world)
- NERSC: Perlmutter	AMD CPUs & NVIDIA Tensor GPUs (2020)
- ALCF: Aurora	Intel CPUs & Intel Xe GPUs (early 2021) — first US Exascale machine
- OLCF: Frontier	AMD CPUs & AMD GPUs (later 2021) - Exascale

- Notice a pattern above? GPUs are winners. Intel has discontinued Phi processors

- These machines offer massive computing capacity … much much more than what we’re used to
- How do we use these machines efficiently?
- GPUs will be everywhere … can we use them?
- Machine Intelligence (MI) will be the “killer app” … Do we need to make everything we do look like MI?
- What’ll be hot… GPU enabled code; What’ll be not… perhaps vectorization (would not have guessed this)

- GPU multithreading has different issues than CPU multithreading
- Starting to explore parallel execution abstraction libraries, like OpenMP, Kokkos (Sandia) and Raja (LLNL)

What of LArSoft's future?

- The Fermilab Scientific Computing Division is committed to LArSoft for current and future LAr experiments
  - Fermilab SCD developers will continue to focus on infrastructure and software engineering
  - Continue to rely on developers from experiments
  - Continue to interface to neutrino toolkits like Pandora
  - Need to confront the HPC evolution
  - Reduce dependency on the framework

- What about the framework?
  - Evolving two major frameworks (*CMSSW* and *art*) into the Dune/HL-LHC era is difficult to defend
  - *art* is feature frozen so developers can focus on LArSoft and multi-threading
  - SCD is exploring options to move ahead with one framework
  - Things to keep in mind
    - We recognize that framework features used by LArSoft need to continue
    - The voice of neutrino experiments in guiding the framework, like you do now with art, will not diminish
    - Stay tuned!

- Making development and builds easier
  - Integrated GitHub, CI, Spack, SpackDev

Summary

Computing is changing (and the change has changed - GPUs over KNLs)

Keep adapting. Parallelization abstractions may make things easier

Don’t let Amdahl’s law discourage you … speedup is just one reason to go parallel (other reasons: better memory use; efficient use of HPC)

LArSoft is here to stay. Thanks to your help in making it a success