# BLUE WATERS

SUSTAINED PETASCALE COMPUTING

## What's new in HPC?

Gregory Bauer


NCSA NSF GREAT LAKES CONSORTIUM FOR PETASCALE COMPUTATION CRAY

# To keep up-to-date on HPC

- HPC Guru - https://twitter.com/HPC_Guru
- Glenn Lockwood - http://www.glennklockwood.com/
- http://www.nextplatform.com

# What's old is new again?

All aspects of HPC are (again) rapidly changing.

- Return of Ethernet to HPC
- Revisiting (relaxed) POSIX I/O semantics

- New accelerators
- New CPUs

# HPC in the US

NSF and DOE

- NCSA Blue Waters (AMD CPU and NVIDIA GPU) 2013 **14 PF**
- ORNL Titan (AMD CPU and NVIDIA GPU) 2012 **27 PF**



- NERSC Cori (Intel Xeon Phi) 2016 **28 PF**
- ANL Theta (Intel Xeon Phi) 2017 **12 PF**
- TACC Stampede2 (Intel Xeon Phi and Intel CPU) 2017 **18 PF**




- ORNL Summit (IBM P9 + NVIDIA V100) 2018 **200 PF**
- LLNL Sierra (IBM P9 + NVIDIA V100) 2018 **125 PF**


- TACC Frontera (Intel CPU + GPU) 2019 **35-40 PF**

- NERSC Perlmutter (AMD EPYC + Nvidia GPU) 2020 **100 PF**
- ANL Aurora (Intel CPU and Xe GPU) 2021 **1 EF**
- ORNL Frontier (AMD EPYC Zen 4 and Radeon GPU) 2022 **1.5 EF**

Commercial HPC

- DUG McCloud (Xeon Phi) 2019 **125 PF** (DP)

# Changes to the landscape

- Mergers & Acquisitions
  - HPE
    - CRAY – accelerator **OpenMP support**
    - Long history: Convex, Compaq (DEC/Alpha), SGI, …
  - NVIDIA
    - Mellanox
    - PGI (2013) – **OpenACC support**
  - Intel
    - Altera FPGA (2015)
- “New” integrator
  - DownUnder Geosolutions

# Changes to the landscape

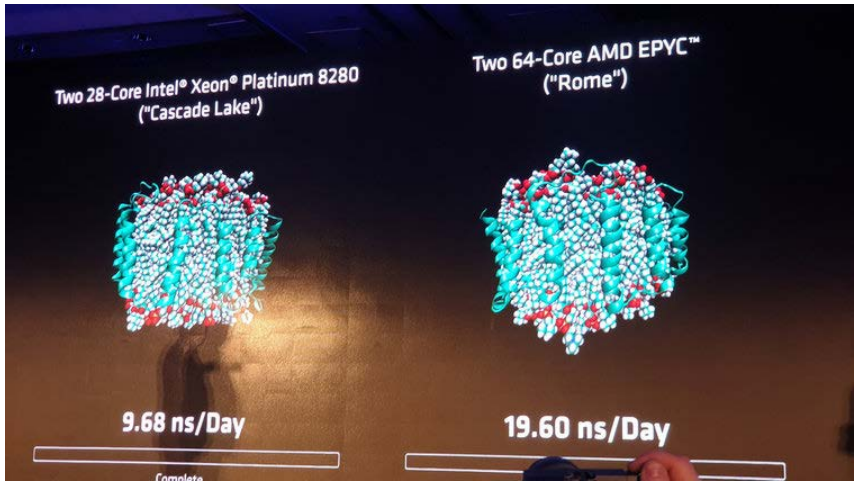
- ARM (Softbank)
  - Fujitsu A64FX
  - Marvell (Cavium) ThunderX2
- Intel
  - Xe GPU
- Google
  - TPU
- Tachyum
  - Prodigy CPU

# CPU peak feeds and speeds

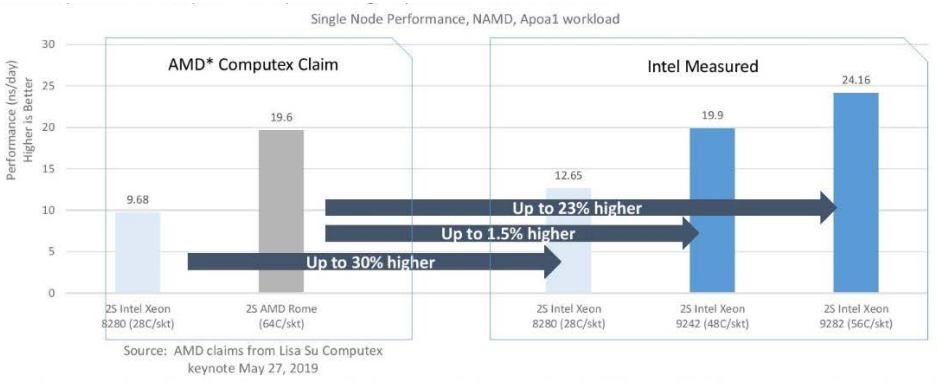
| Vendor/Processor | cores/node | clock rate (GHz) | FP64 rate (TFLOPS) | Memory Bandwidth (TB/s) | Bytes/flop ratio | Notes |
|---|---|---|---|---|---|---|
| AMD Interlagos | 2x8 | 2.3 | 0.313 | 0.102 | 0.33 | |
| Intel Sandybridge | 2x8 | 2.6 | 0.333 | 0.102 | 0.31 | |
| | | | | | | |
| Intel Skylake | 2x20 | 2.4 | 3.07 | 0.256 | 0.08 | |
| ARM ThunderX2 | 2x32 | 2.1 | 1.13 | 0.32 | 0.28 | NEON |
| | | | | | | |
| Intel Cascade Lake | 2x28 | 2.1 | 3.76 | 0.282 | 0.08 | AVX 512 |
| AMD Rome | 2x64 | 1.7 | 3.5 | 0.380 | **0.11** | AVX2 16 FP/clock |
| | | | | | | |
| Fujitsu ARM A64FX | 2x48 | ? | 2.7 | 2 | **0.74** | SVE 512 , HBM2 |
| Tachyum Prodigy | 2x64 | ? | 8 | 0.614 | 0.08 | DDR5 4800 512 bit vector 4 inst/clock |
| Intel Ice Lake | | | | | | |

# Benchmarketing

## AMD says



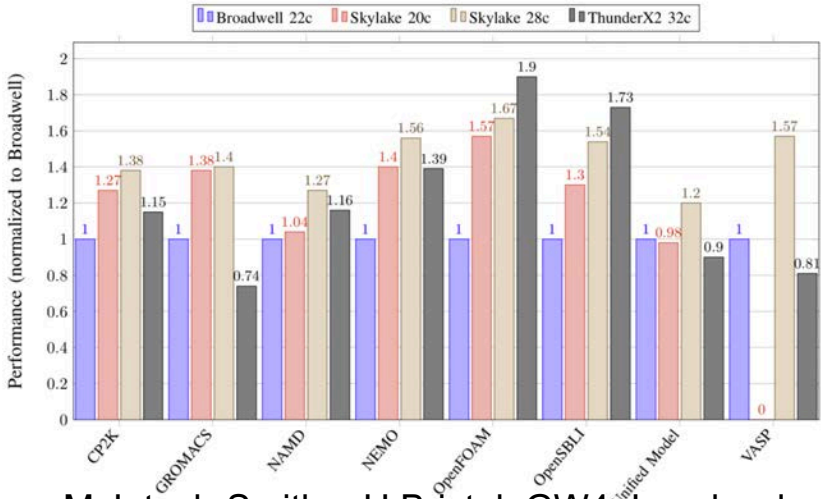

## Intel says



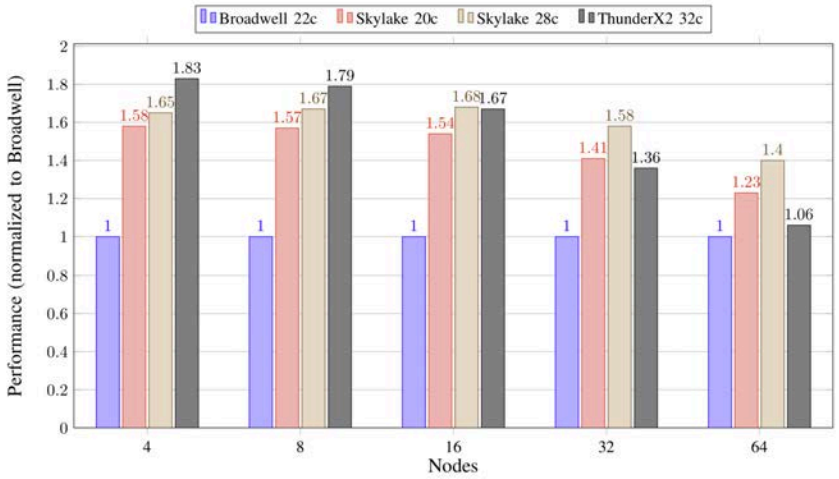

Source: AMD claims from Lisa Su Computex keynote May 27, 2019

# Thunderx2 on Cray XC50 Isambard

## Single node performance




## OpenMulti-node Scaling OpenFOAM




Simon McIntosh-Smith – U Bristol, GW4, Isambard
Comparative Benchmarking of the First Generation of HPC-Optimised Arm Processors on Isambard CUG 2018
Scaling Results From the First Generation of Arm-based Supercomputers CUG2019

# Hardware factors

- Cache speed
  - AMD and ARM are typically slower than Intel; impacting strong scaling.
- Memory bandwidth
  - 8 channels (ARM) better than 6.
- Vector widths
  - Intel vector wider but at a clock speed cost
  - ARM SVE catching up

# GPUs

- NVIDIA Ampere
  - Better than V100
  - V100 performance
    - 7.5/15/120 TF (DP/SP/HP) 900 GB/s 16 GB HBM2
- AMD Radeon Instinct
  - 6.7/13.4/26.8 TF (DP/SP/HP) 1 TB/s 16 GB HBM2
- Intel GPU (Xe)
  - not much generally available

# Software

- Now need to support 3 GPUs (NVIDIA, AMD, Intel)
- Possibly 3 different vector engines

- “frameworks” like Kokkos, Raja, etc. can provide portability and performance for CPU, GPU targets.
- Intel “OneAPI”
- AMD ROCm, HIP

# Software

- Compiler performance with TSVC loop suite
  - 151 loops
- Blue Waters
- Intel Skylake

**Table 6 - Vectorization effectiveness on Cray-XC**

| Compiler | Mean Speedup | Number of Optimized Loops |
|---|---|---|
| Cray | 4.32 | 108 |
| Intel | 3.28 | 102 |
| PGI | 3.82 | 84 |
| GNU | 1.59 | 47 |

**Table 7 - Mean Speedup across optimized loops**

| Compiler | Blue Waters | Cray-XC |
|---|---|---|
| Cray | 2.91 | 5.63 |
| Intel | 2.32 | 4.38 |
| PGI | 4.77 | 6.08 |
| GNU | 2.58 | 2.92 |

Evaluating Compiler Vectorization Capabilities on Blue Waters, CUG2019

# Quantum Computing

November 12, 2007 11:33 AM Eastern Standard Time

RENO, Nevada--(BUSINESS WIRE)--D-Wave Systems puts the world's first commercial quantum computer on display in an online demonstration here this week at the prestigious SCO7 Conference -- an international gathering of technologists and computer scientists focused on high performance computing, networking, storage and analysis.

"Our product roadmap takes us to 512 qubits in the second quarter of 2008 and 1024 qubits by the end of that year. At

The company introduced its revolutionary 16 qubit machine last February in Silicon Valley, California. "Advancing the machine to 28 qubits in such a short space of time lends credibility to our claim of having a scaleable architecture," stated Herb Martin, D-Wave's CEO. "Our product roadmap

- Disruptive technology at SC'07
- D-Wave, Fujitsu, Google, Honeywell, Lockheed-Martin, Microsoft, NEC, Toshiba, …
- Various ways to provide qubits: trapped ions, quantum dots, superconductors, ..
- "Proven" for certain types of problems: encryption, discrete event modeling, …
- Accessible via cloud computing with various SDKs etc.

# Things to play with

- Google Edge TPU – only runs TensorFlow lite for inference currently but …
  - https://www.sparkfun.com/products/15318 $156.95

# Current trend

- Additional tiers
  - NVMe > SSD > Spinning disk > ???
- I/O Accelerators
  - Burst buffers

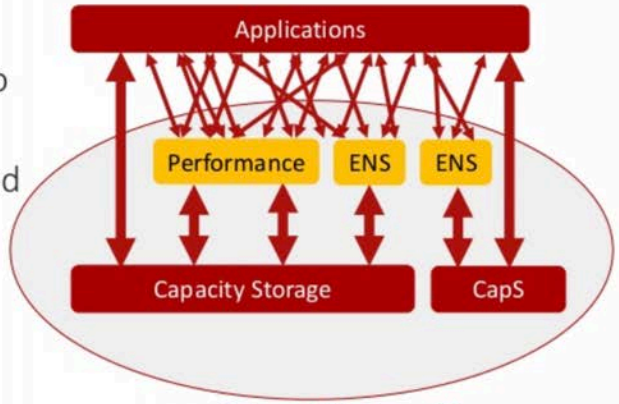
# One view about changes to storage

Long Live Posix - HPC Storage and the HPC Datacenter

Clip slide

## Next Decade of HPC Storage: Back to the Future?

- Still PFS...
  - ... but with increasing portions that are dynamically allocated and integrated into the compute platform
  - e.g. a file server turns into a containerized process run anywhere
- Still POSIX...
  - ... but relaxed where needed
- Still Tiers...
  - ... but relaxed to reflect actual application workflows




**Relaxed Tiering with Dynamically-Managed, System-integrated Storage**

https://insidehpc.com/2019/04/long-live-posix-hpc-storage-and-the-hpc-datacenter/