Hewlett Packard Enterprise

HPE in Weather & Climate

Ilene Carpenter

Ilene.carpenter@hpe.com

Earth Sciences Segment Manager

November 2020

ONE YEAR AGO...

Hewlett Packard
Enterprise

+

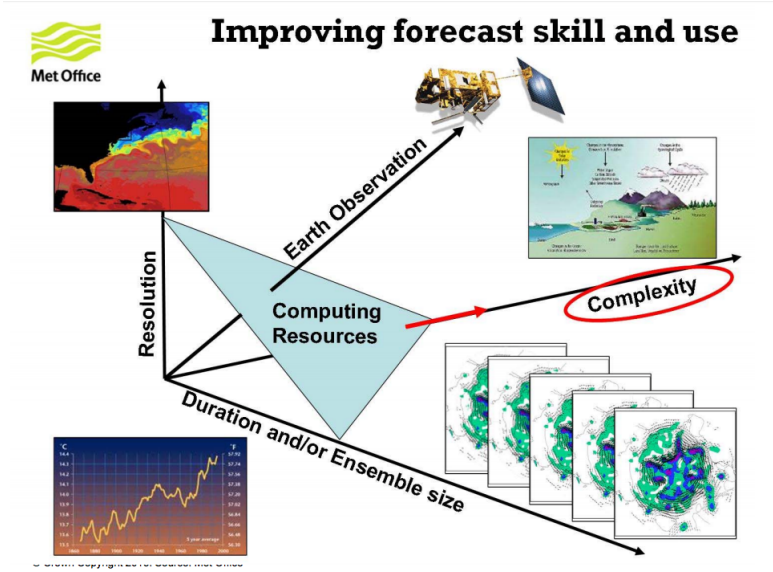
CRAY

EARTH SCIENCES: CORE DRIVERS

Performance: Drive continued improvements in fidelity of weather & climate simulations

Reliability: Maintain performance, reliability & serviceability as systems grow in size & complexity

Maximize value of environmental data



$2+ BILLION CONTRACTS WON[1]

NERSC

SIZE

$100M +

ANNOUNCED

30-OCT-2018

Argonne NATIONAL LABORATORY

SIZE

$100M +

ANNOUNCED

18-MAR-2019

OAK RIDGE National Laboratory

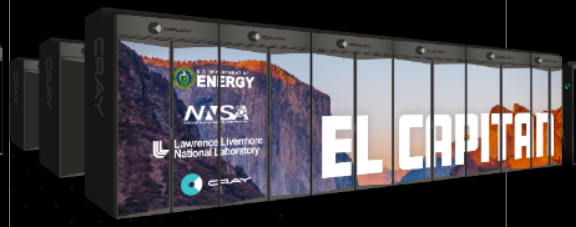
SIZE

$600M +

ANNOUNCED

7-MAY-2019

Lawrence Livermore National Laboratory



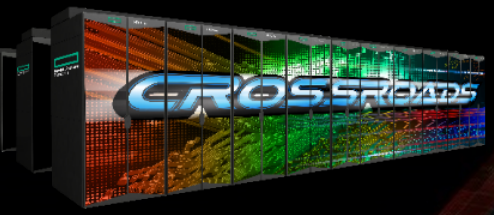
SIZE

$600M +

ANNOUNCED

5-MAR-2020

Los Alamos NATIONAL LABORATORY EST. 1943



SIZE

$100M +

ANNOUNCED

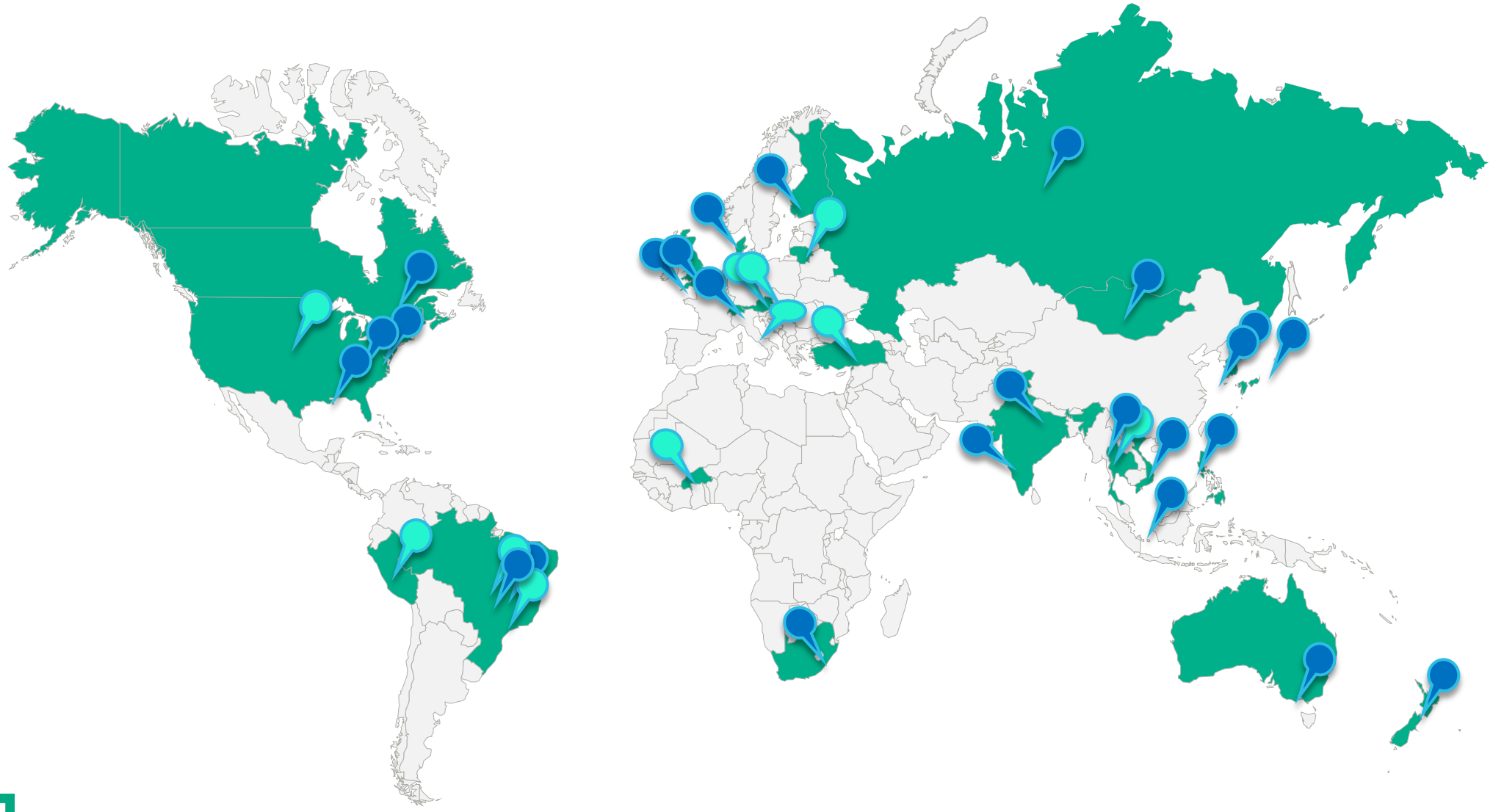
1-OCT-2020

1. Listed are the publicly disclosed won-to-date contracts

ANNOUNCED SLINGSHOT SYSTEMS

- Indiana University's BigRed 200
- Air Force Weather's HCP11
- UK's Archer2
- Lawrence Berkeley National Lab's Perlmutter
- Oakridge National Lab's Frontier
- Argonne National Lab's Argonne
- Lawrence Livermore National Lab's El Capitan
- Los Alamos National Lab's Crossroads
- AWE
- Pawsey Supercomputing Center
- EuroHPC JU's LUMI
- And others…

HPE IN WEATHER FORECASTING – GLOBAL PRESENCE, ALL GEOGRAPHIES



NEW CRAY SYSTEMS IN WEATHER, CLIMATE & OCEANOGRAPHY

IBS Center for Climate Physics

CSCS

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Ministry of Earth Sciences
Government of India

Australian Government
Bureau of Meteorology

NIWA
Taihoro Nukurangi

Meteorological Service Singapore
Centre for Climate Research Singapore

South African Weather Service
ISO 9001 Certified Organisation

Met Office

Japan Meteorological Agency

FINNISH METEOROLOGICAL INSTITUTE

Korea Meteorological Administration

AWI

ECMWF

THE NEED FOR SCALABILITY

Dennard scaling ended 10 years ago

- Clock speeds stopped increasing
- # cores/processor increases each generation

⇒ Need **application scalability** so that using more cores delivers more performance

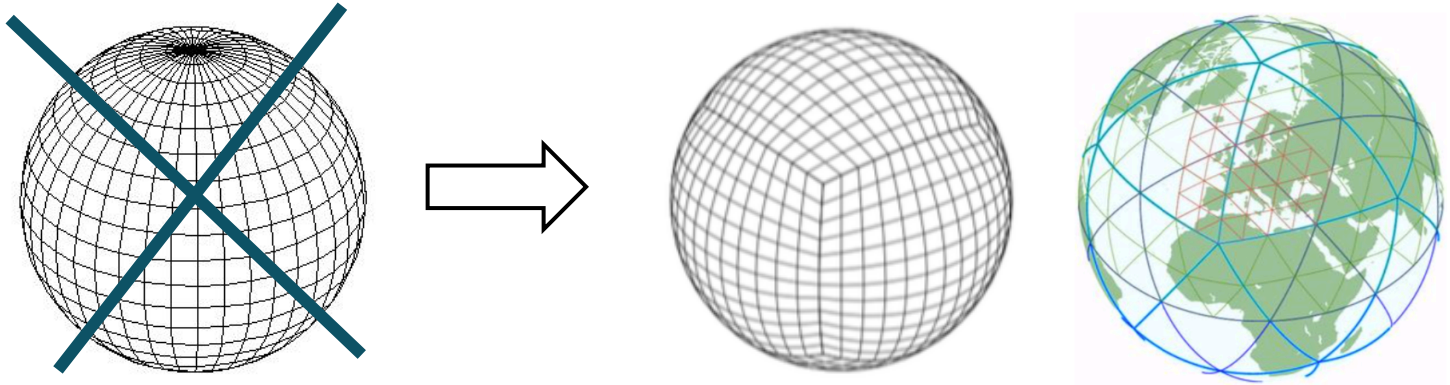
Impact on model algorithms and structure:

- Algorithms must
  - Use lots of cores
  - Minimize global communication

⇒ Finite volume, finite element, spectral element

Quasi-uniform grids

- Avoid the “pole problem” of lat-lon grids
- Cubed sphere, icosahedral



EMERGING TECHNOLOGIES
INCREASING DIVERSITY IN PROCESSORS AND MEMORY

End of Dennard scaling led to the use of **architectural specialization** to get performance - different processors optimized for different workloads

- CPUs from AMD, Intel, ARM
- Vector processors - Fujitsu A64FX, NEC Aurora SX
- GPUs from NVIDIA, AMD, Intel
- FPGAs
- AI chips
- On-chip HBM

WHAT ABOUT HBM?
HIGH MEMORY BANDWIDTH AND ENERGY EFFICIENCY

Accelerators:

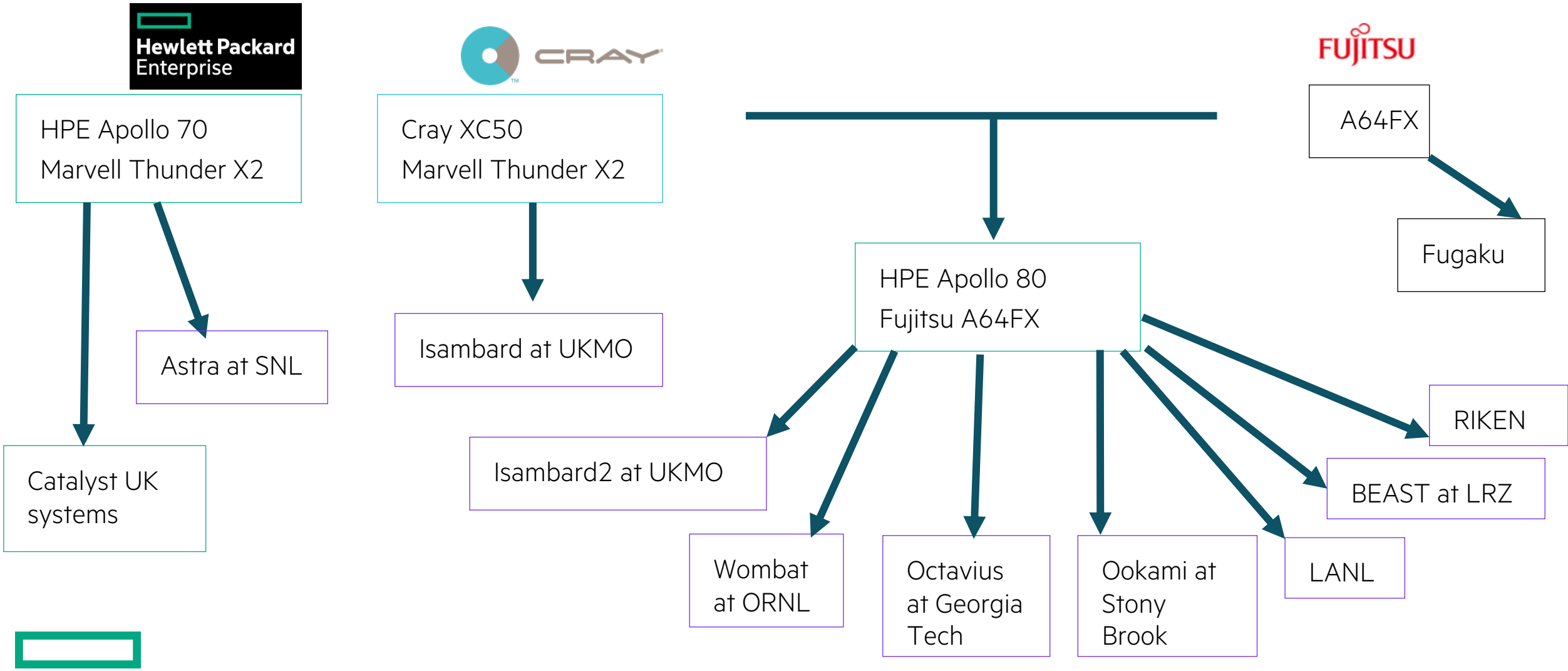
- NVIDIA GPUs
- NEC SX Aurora

Processors:

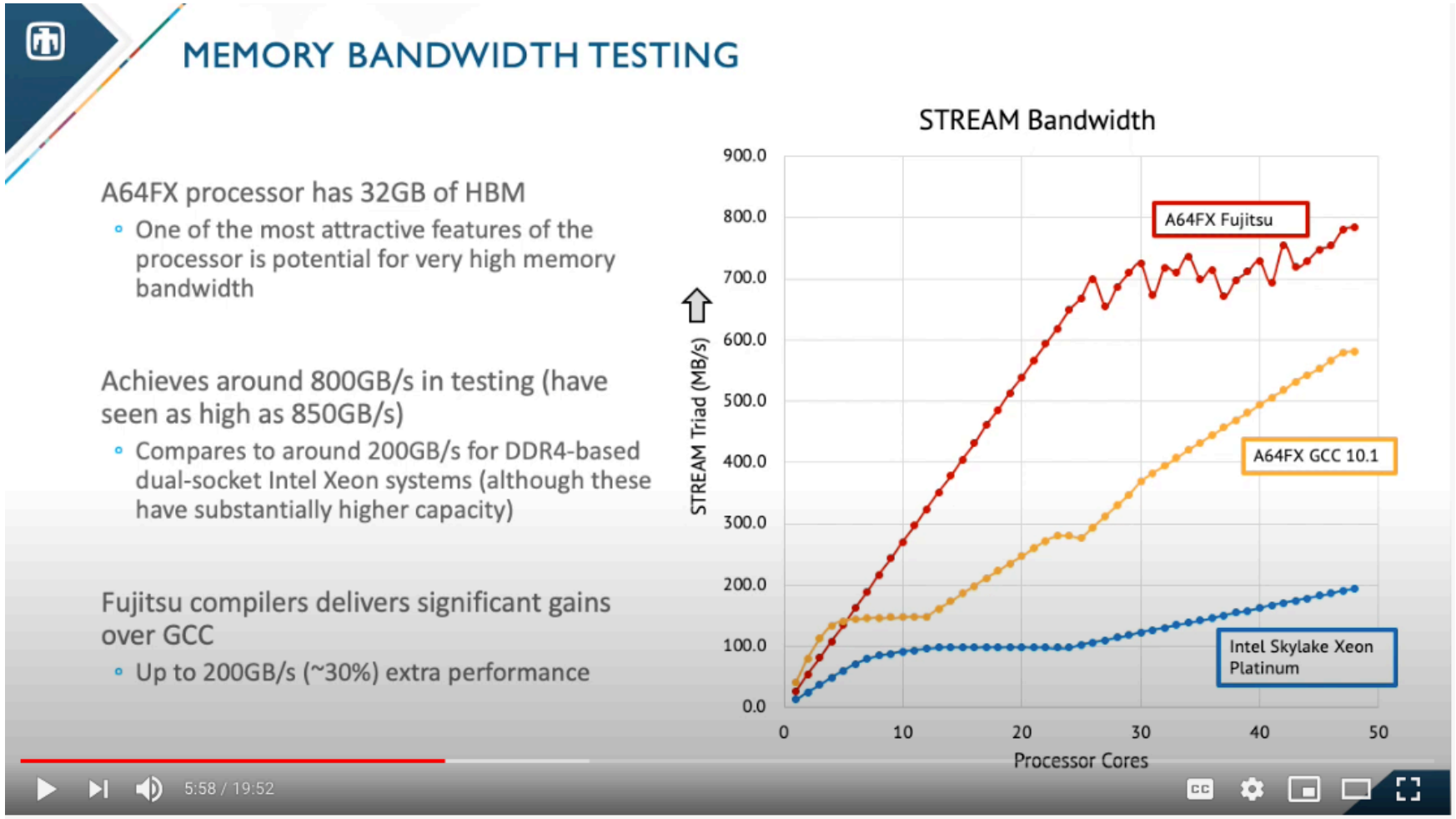
- Intel Xeon Phi – HBM + DIMMs
- Fujitsu A64FX
  - Armv8-A, designed for HPC workloads
  - 48 cores
  - 32 GB HBM2
  - 2 512-bit wide SIMD SVE units per core, designed for HPC
  - But … no DRAM DIMMs

"GPUs have two key technologies, wide vectors like SVE and a lot of memory bandwidth. Fujitsu stuck both onto A64FX", Simon McIntosh-Smith

ARM BASED SUPERCOMPUTERS



HIGH MEMORY BANDWIDTH – EARLY A64FX RESULTS



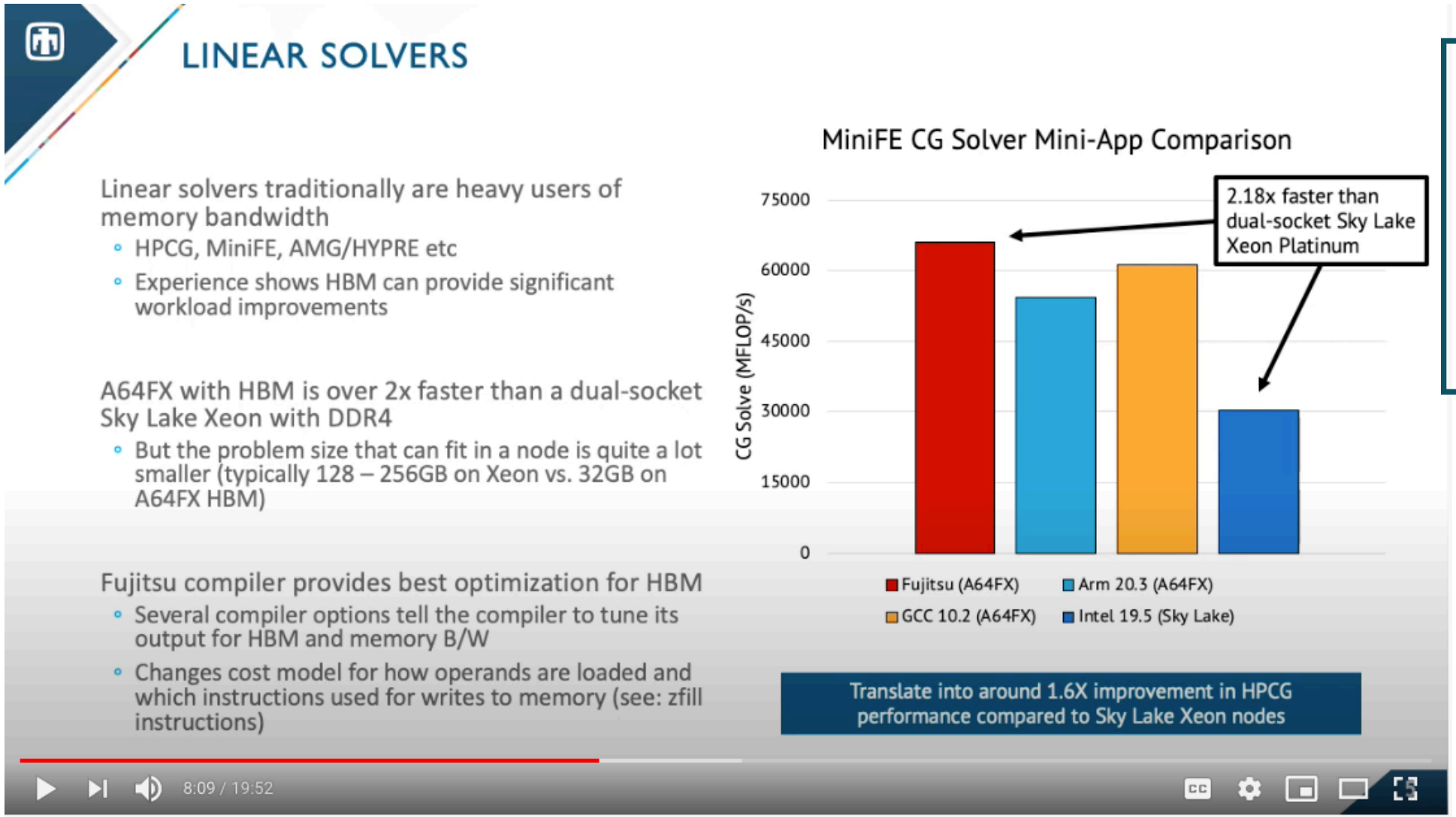
Early Experiences with A64FX, AHUG SC'20

System delivered by Penguin

Si Hammond

- Very efficient memory subsystem but small capacity
- Lots of flags on Fujitsu compiler to tweak

HIGH MEMORY BANDWIDTH – EARLY A64FX RESULTS

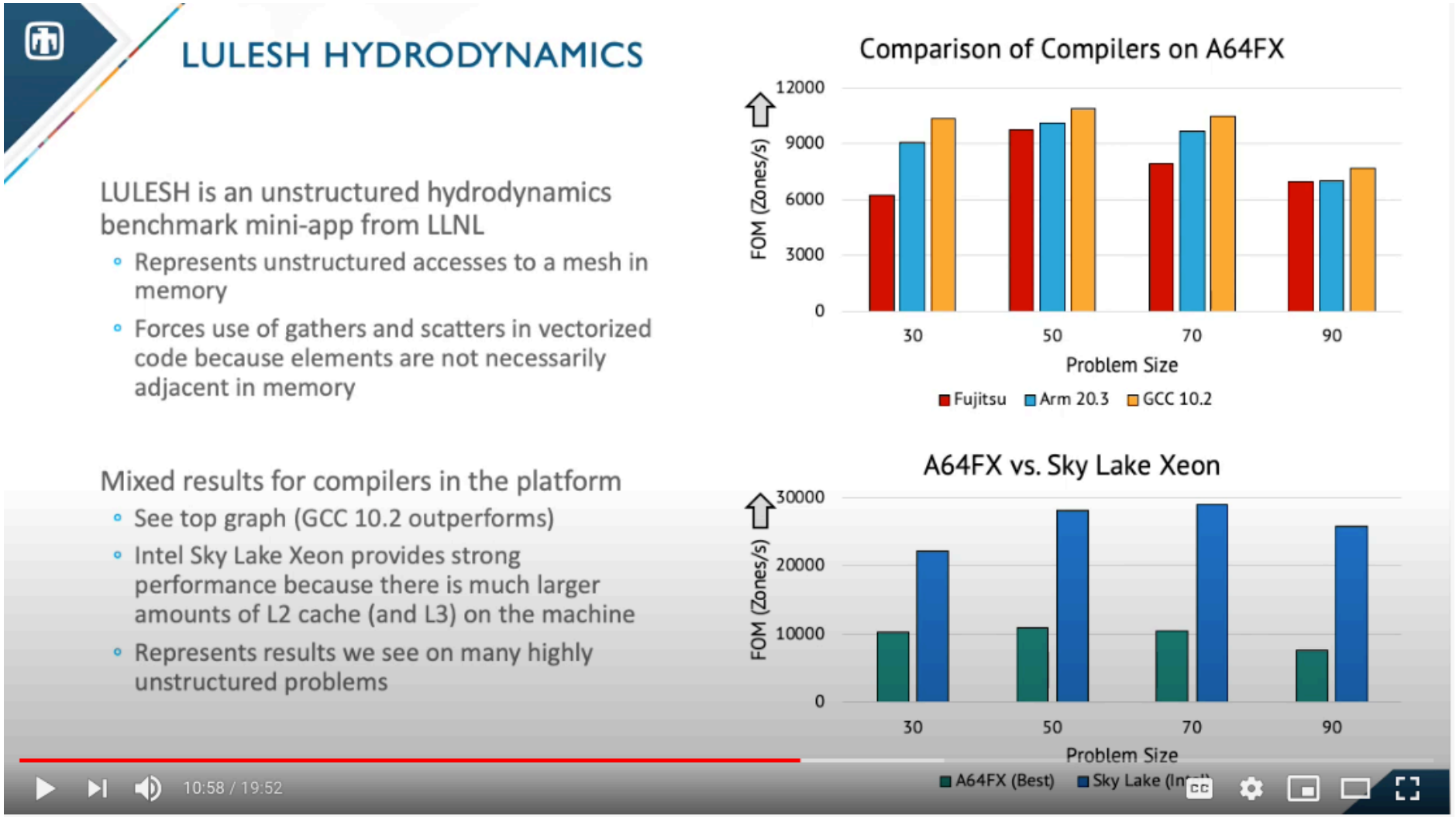


Early Experiences with A64FX, AHUG SC'20

Si Hammond

- Performance translates well to linear solvers.
- The compiler and compiler flags used matter.

HIGH MEMORY BANDWIDTH – EARLY A64FX RESULTS

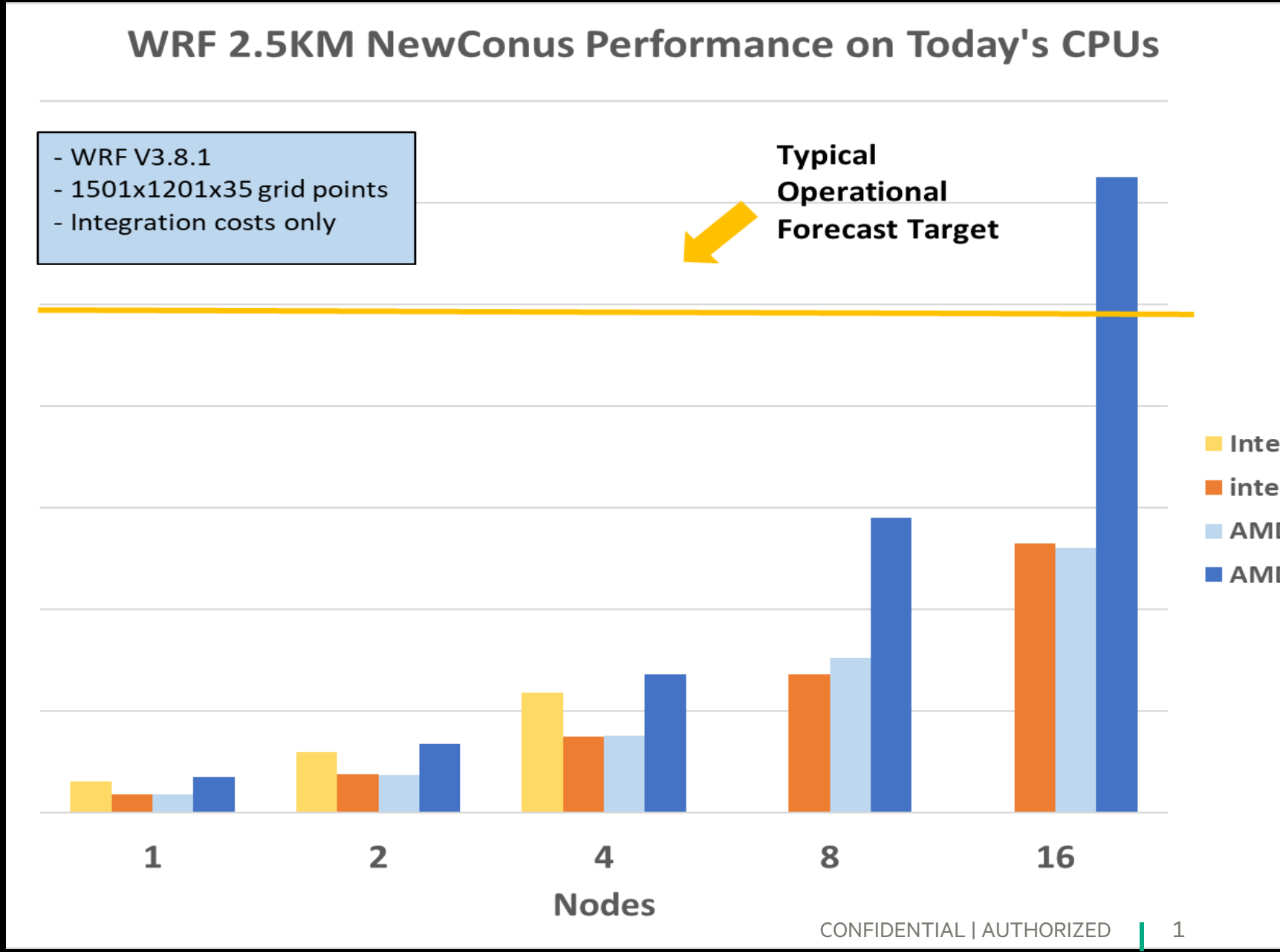


Early Experiences with A64FX, AHUG SC'20

- Gather/scatter heavy codes are sensitive to memory latency and caches provide good performance.
- Intel Sky Lake significantly outperforms A64FX because of large L2 (+L3) caches.

WHICH PROCESSOR?

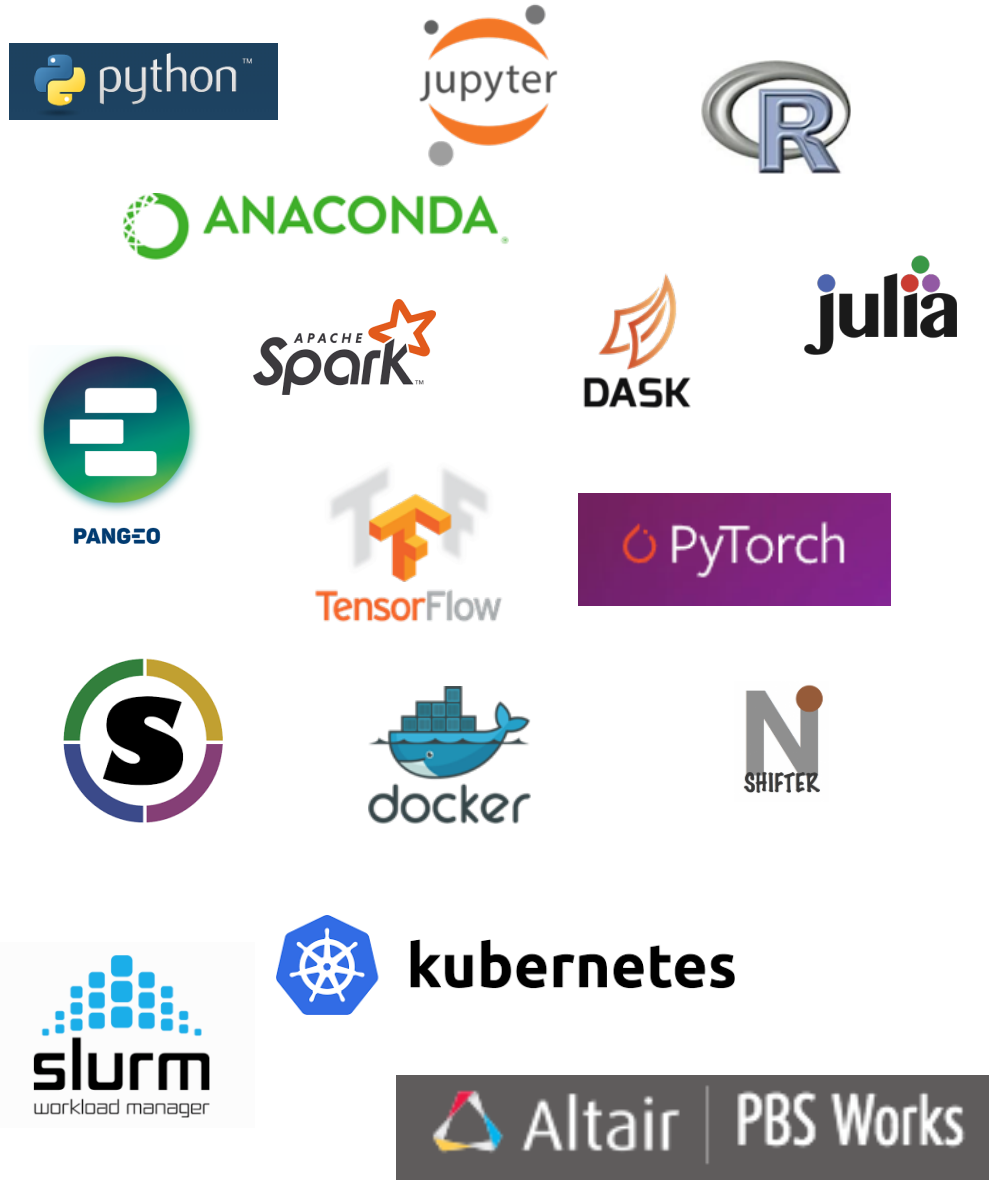
- Intel Platinum 9242 2.3GHz Cascade Lake AP 48 core 350W
- Intel Platinum 8260 2.4GHz Cascade Lake 24 core 165W
- AMD EPYC 7502 2.5GHz Rome 32 core 180W
- AMD EPYC 7742 2.25GHz Rome 64 core 225W



INCREASINGLY HETEROGENEOUS, DATA-CENTRIC WORKLOADS

Diverse software stacks and usage styles

- Interactive computing
- High productivity interpreted languages
- New analytics and AI frameworks
- Containerized software



AI FOR WEATHER FORECASTING

There is a rapid increase in the use of Machine Learning in the weather enterprise all through the forecast chain.

AI has different needs –storage, software stack, sometimes processors – where should you run it?

Where is the training data?

- If from NWP models, run on (potentially heterogeneous) HPC system where the data is generated
- If training data is in public cloud, move computation to where the data is?
- If training data is in multiple locations, run locally and pull in data as needed?

MODELING & SIMULATION + ARTIFICIAL INTELLIGENCE + BIG DATA ANALYTICS

RUNNING ON ONE MACHINE IN MISSION-CRITICAL WORKFLOWS

EXASCALE ERA

"The **confluence of AI with traditional simulations** is going to transform the very nature of high performance computing. That's the thing that I think is going be yet another sea change in how we do science."

"We're not doing our grandfather's HPC here."

Rick Stevens, ANL Associate Lab Director

HPE CRAY EX (SHASTA) SYSTEMS SELECTED FOR WEATHER CENTERS –

- All will use AMD EPYC processors

- NOAA NCEP
  - 2 systems in different physical locations (geo-redundant)
  - GDIT is the new managed service provider for NOAA

- US Air Force 557th Weather Wing
  - 2 systems in a hall A/hall B configuration
  - Acquired and run by Oak Ridge National Laboratory, to provide HPC-as-a-service to the Air Force

- US Naval Meteorology and Oceanography
  - To be housed at the Navy DSRC at Stennis Space Center

- A NMHS outside the US
- An Air Force outside US

First Cray EX systems sold for weather forecasting outside of the US! Not announced yet.

CRAY PRODUCTS SOLUTIONS RESOURCES COMPANY SUPPORT CONTACT

Home > Company > Investors > Press Release

CRAY SHASTA SUPERCOMPUTER TO POWER WEATHER FORECASTING FOR THE U.S. AIR FORCE

Strategic Partnership with Oak Ridge National Lab Highlights First Cray Shasta Supercomputer for Operational Weather Forecasting

SEATTLE, Aug. 07, 2019 (GLOBE NEWSWIRE) -- Global supercomputer leader Cray Inc. (Nasdaq:CRAY) today announced that the first Cray Shasta™ supercomputing system for operational weather forecasting and meteorology will be acquired by the Air Force Life Cycle Management Center in partnership with Oak Ridge National Laboratory. The powerful high-performance computing capabilities of the new system, named HPC11, will enable higher fidelity weather forecasts for U.S. Air Force and Army operations worldwide. The contract is valued at $25 million.

"We're excited with our Oak Ridge National Laboratory strategic partner's selection of Cray to provide Air Force Weather's next high performance computing system," said Steven Wert, Program Executive Officer Digital, Air Force Life Cycle Management Center at Hanscom Air Force Base in Massachusetts, and a member of the Senior Executive Service. "The system's performance will be a significant increase over the existing HPC capability and will provide Air Force Weather operators with the ability to run the next generation of high-resolution, global and regional models, and satisfy existing and emerging warfighter needs for environmental impacts to operations planning."

Oak Ridge National Laboratory (ORNL) has a history of deploying the world's most powerful supercomputers and through this partnership, will provide supercomputing-as-a-service on the HPC11 Shasta system to the Air Force 557th Weather Wing. The 557th Weather Wing develops and provides comprehensive terrestrial and space weather information to the U.S. Air Force and Army. The new system will feature the revolutionary Cray Slingshot™ interconnect, with features to better support time-critical numerical weather prediction workloads, and will enhance the Air Force's capabilities to create improved weather forecasts and weather threat assessments so that Air Force missions can be carried out more effectively.

"The HPC11 system will be the first Shasta delivery to the production weather segment, and we're proud to share this milestone with ORNL and the Air Force," said Peter Ungaro, president and CEO at Cray. "The years of innovation behind Shasta and Slingshot and the success of prior generations of Cray systems continue to demonstrate Cray's ability to support demanding 24/7 operations like weather forecasting. This is a great example of the upcoming Exascale Era bringing a new set of technologies to bear on challenging problems and empowering the Air Force to more effectively execute on its important mission."

HPC11 will be ORNL's first Cray Shasta system, as well as the first supercomputing system with 2nd Gen AMD EPYC™ processors for use in operational weather forecasting. HPC11 will join the 85% bastion of weather centers that rely on Cray, and will feature eight Shasta cabinets in a dual-hall configuration.

"We are incredibly excited to continue our strategic collaboration with Cray to deliver the first Shasta supercomputer to the U.S. Air Force, helping to improve the fidelity of weather forecasts for U.S. military operations around the globe," said Forrest Norrod, senior vice president and general manager, Datacenter and Embedded Systems Group, AMD. "The 2nd Gen AMD EPYC processors provide exceptional performance in highly complex workloads, a necessary component to power critical weather prediction workloads and deliver more accurate forecasts."

The system is expected to be delivered in Q4 2019 and accepted in early 2020.

ORNL will provide supercomputing-as-a-service on the HPC11 system to the Air Force 557th Weather Wing.

OAK RIDGE National Laboratory

First Shasta system for production weather forecasting.

- Slingshot interconnect
- 2nd gen AMD EPYC processors
- 8 cabinets in a hall A/hall B configuration

2 HPE Cray EX systems
Each with 800 AMD Rome (64c) nodes and 256 GB/node

US Navy to deploy Cray Shasta supercomputer with AMD and Nvidia chips

The system will aide in hurricane and engineering research

February 18, 2020 By: Alex Alley

The US Navy will get a supercomputing system installed as part of the Department of Defense (DoD) High Performance Computing Modernization Program (HPCMP).

The new 12.8 petaflops Cray Shasta supercomputer will be the first machine in the program to provide over ten petaflops of computing power to DoD scientists, researchers, and engineers.

The system will be housed and operated at the Navy DSRC at Stennis Space Center, Mississippi. Were it to launch today, it would be among the top 25 supercomputers in the world, according to the ranking site, Top500. The system is expected to be online by early 2021.

Rear Admiral John Okon, the head of Navy Meteorology and Oceanography Command, said: "The investment and increase in supercomputing power at the Navy DSRC at Stennis Space Center is absolutely critical to Naval Oceanography.

"Delivering future capability upgrades to global and regional ocean and atmospheric prediction systems, to include later this year the Navy's first Earth Systems Prediction Capability.

"Naval Oceanography's ability to be the Department of Defense's authoritative source for characterizing and applying data of the physical battlespace into a decisive advantage for naval, joint and allied forces hinges on the continual upgrade and advancements in high-performance computing from the HPCMP."

Battle/Hurricane computer

The Cray Shasta supercomputer will feature 290,304 AMD Epyc 7002-series processor cores and 112 Nvidia Volta V100 GPUs, interconnected by a 200 gigabit per second Cray Slingshot network.

Screenshot

The Cray Shasta supercomputer will feature 290,304 AMD EPYC 7002-series processor cores and 112 NVIDIA Volta V100 General-Purpose Graphics Processing Units (GPGPUs), interconnected by a 200 gigabit per second Cray Slingshot network. The system will also feature 590 total terabytes (TB) of memory and 14 petabytes (PB) of usable storage, including 1 PB of NVMe-based solid state storage.

In particular, Navy DSRC supercomputers support climate, weather, and ocean modeling by NMOC, which assists U.S. Navy meteorologists and oceanographers in predicting environmental conditions that may affect the Navy fleet. Among other scientific endeavors, the new supercomputer will be used to enhance weather forecasting models; ultimately, this improves the accuracy of hurricane intensity and track forecasts.

HPE Cray EX system with

- 2176 AMD Rome (64 core) nodes, 256 GB/node
- 112 NVIDIA V100 GPUs

Cray to Provide NOAA with Two AMD-Powered Supercomputers

By Tiffany Trader

Each identical Cray Shasta system spans 2,560 dual-socket nodes — housed in 10 cabinets — powered by second-gen AMD Epyc 'Rome' 64-core 7742 processors, connected by Cray's Slingshot network. The total system memory per machine is 1.3 petabytes. Cray's ClusterStor systems provide 26 petabytes of storage per site (a flash storage system with 614 terabytes of usable space and two HDD file systems with 12.5 petabytes of usable storage).

The peak theoretical performance of each Cray system is 12 petaflops, which combined with NOAA's research and development machines brings the agency's aggregate operational and research capacity to 40 peak petaflops.

- Two HPE Cray EX systems, each with
  - 2560 dual-socket AMD (64c) Rome nodes, 512 GB/node
  - HPE Cray ClusterStor Lustre – flash and HDD
- NOAA R&D system – Cray CS500 (2019)
  - 1200 Intel SkyLake nodes
  - 100 nodes with NVIDIA P100
- NOAA Climate research system – Cray XC hosted by ORNL

The twin Cray systems are perfectly symmetrical between geographically-segregated sites (Manassas, Virginia, and Phoenix, Arizona), and take turns acting as the primary or backup system. Michaud explained that on any given day, NOAA can run at production, its full operational 24×7 modeling suite on one of the systems. The backup system is used for transition to operations and other development work while it's not in use as the primary, and NOAA can switch the orientation of the primary and the backup site in operations within a 15 minute period, and does so regularly, at least on a monthly basis.

The arrangement assures redundancy, as data is always mirrored to the backup system, offering advantages from a troubleshooting and maintenance perspective and providing an added layer of protection for the mission- and safety-critical work of weather prediction. "If we make a change to one, we know we can test it, and then we can apply the change to the back-up system as well," said Michaud. "We know if one's not behaving similar to the other, we can identify the differences and troubleshoot them. And then, the other thing that's really important is for the type of work that we do, given storm systems and other weather systems can be massive in scale and encompass hundreds of miles, it's really beneficial for us to have the separation of the sites, so that if we have any issues on one site, we can switch to the other site."

The transition to a new managed service provider coincides with a change in filesystem technology. After about 20 years of being on GPFS, NOAA is switching its systems over to Lustre. The move should not be seen as reflecting NOAA's preference for a given filesystem, rather the agency provided the specification for performance-based requirements for the contract and what it required in terms of availability (99 percent system availability) as part of the open bid process and let industry decide what the best fit was in terms of the total proposed solution. "We were essentially looking for what the best fit was for what the integrator could provide…[and] the best performance-per-dollar with the availability requirements that we require for operational use of the system," David Michaud, director, Office of Central Processing for NOAA's National Weather Service, told *HPCwire*.

EARTH SCIENCES: WHY HPE?

- Performance
  - Balance performance & throughput across workflow
  - Focus on TCO
- Reliability
  - Operationally proven, unrivalled experience
- Vision
  - World-leading software development environment, performance tools & application support experts
  - Long-term customer partnerships, workshops
- Experience
  - Through its acquisitions of SGI and Cray, HPE has systems at a large fraction of the worlds weather forecasting centers.

Hewlett Packard
Enterprise

THANK YOU