Beyond the CPU: Is Accelerated Computing for Everyone?

Jack Wells

Director of Science
National Center for Computational Sciences
Oak Ridge National Laboratory

VP OpenACC Consortium

SC19, Denver

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725

The U.S. Department of Energy's Oak Ridge National Laboratory

OAK RIDGE National Laboratory | OAK RIDGE LEADERSHIP COMPUTING FACILITY

Oak Ridge National Laboratory is the largest US Department of Energy (DOE) open science laboratory

Oak Ridge Leadership Computing Facility (OLCF)

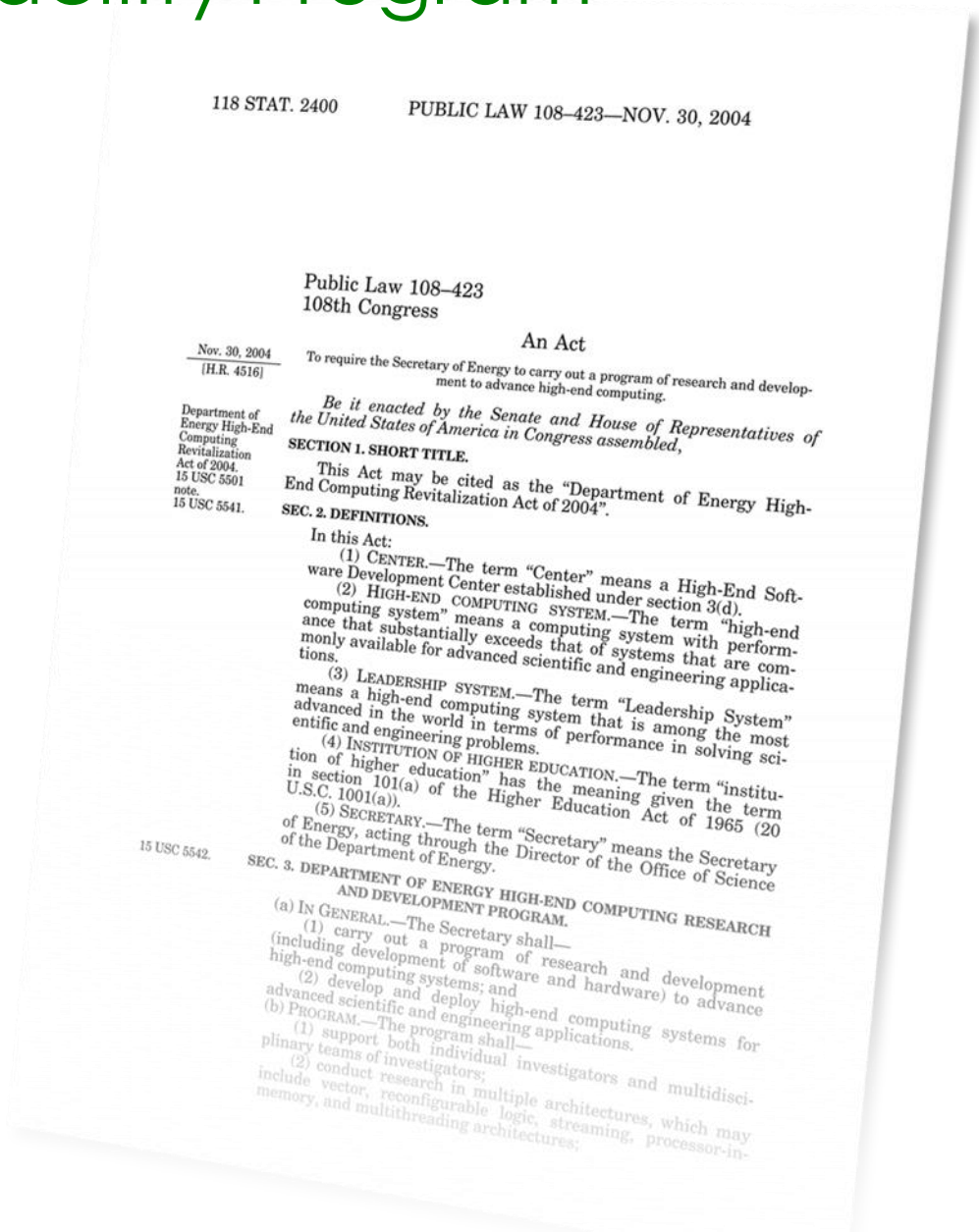
Mission: Deploy and operate the computational resources required to tackle global challenges

- Provide world-leading computational and data resources and specialized services for the most computationally intensive problems
- Provide stable hardware/software path of increasing scale to maximize productive applications development
- Operate world-class user programs accessible to industry, universities, national laboratories and federal agencies, with access is based on the merit of the proposed work.
- With our partners, deliver transforming discoveries and engineering breakthroughs across a breadth of applications.

Origin of Leadership Computing Facility Program

Department of Energy High-End Computing Revitalization Act of <u>2004</u>
(Public Law 108-423):
"The Secretary of Energy, acting through the Office of Science, shall

- Establish and operate Leadership Systems Facilities
- Provide access [to Leadership Systems Facilities] on a competitive, merit-reviewed basis to researchers in U.S. industry, institutions of higher education, national laboratories and other Federal agencies."



118 STAT. 2400 PUBLIC LAW 108–423—NOV. 30, 2004

Public Law 108–423
108th Congress

An Act

Nov. 30, 2004
[H.R. 4516]

To require the Secretary of Energy to carry out a program of research and development to advance high-end computing.

Department of Energy High-End Computing Revitalization Act of 2004.
15 USC 5501 note.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "Department of Energy High-End Computing Revitalization Act of 2004".

15 USC 5541.

SEC. 2. DEFINITIONS.

In this Act:

(1) CENTER.—The term "Center" means a High-End Software Development Center established under section 3(d).

(2) HIGH-END COMPUTING SYSTEM.—The term "high-end computing system" means a computing system with performance that substantially exceeds that of systems that are commonly available for advanced scientific and engineering applications.

(3) LEADERSHIP SYSTEM.—The term "Leadership System" means a high-end computing system that is among the most advanced in the world in terms of performance in solving scientific and engineering problems.

(4) INSTITUTION OF HIGHER EDUCATION.—The term "institution of higher education" has the meaning given the term in section 101(a) of the Higher Education Act of 1965 (20 U.S.C. 1001(a)).

(5) SECRETARY.—The term "Secretary" means the Secretary of Energy, acting through the Director of the Office of Science of the Department of Energy.

15 USC 5542.

SEC. 3. DEPARTMENT OF ENERGY HIGH-END COMPUTING RESEARCH AND DEVELOPMENT PROGRAM.

(a) IN GENERAL.—The Secretary shall—

(1) carry out a program of research and development (including development of software and hardware) to advance high-end computing systems; and

(2) develop and deploy high-end computing systems for advanced scientific and engineering applications.

(b) PROGRAM.—The program shall—

(1) support both individual investigators and multidisciplinary teams of investigators;

(2) conduct research in multiple architectures, which may include vector, reconfigurable logic, streaming, processor-in-memory, and multithreading architectures;

Who's Been Working With Us?

Pratt & Whitney
A United Technologies Company

microsurgeonbot

FCA
FIAT CHRYSLER AUTOMOBILES

NUMECA INTERNATIONAL

GM

DRESSER-RAND
A Siemens Business

CAT

TOTAL
COMMITTED TO BETTER ENERGY

United Technologies Research Center

BOEING

P&G

Ford

GE

TENSCO

GENERAL ATOMICS

SMART TRUCK

Appentra

FM Global

CASCADE TECHNOLOGIES

SPACEX

NVIDIA

سابك
sabic

TVA

SIMULIA

Orbital ATK

3M Science. Applied to Life.™

KatRisk

GLOBALFOUNDRIES

ANSYS

ARKEMA
INNOVATIVE CHEMISTRY

Westinghouse

Rolls-Royce

TECH-X

TRI ALPHA ENERGY
THE POWER OF INGENUITY

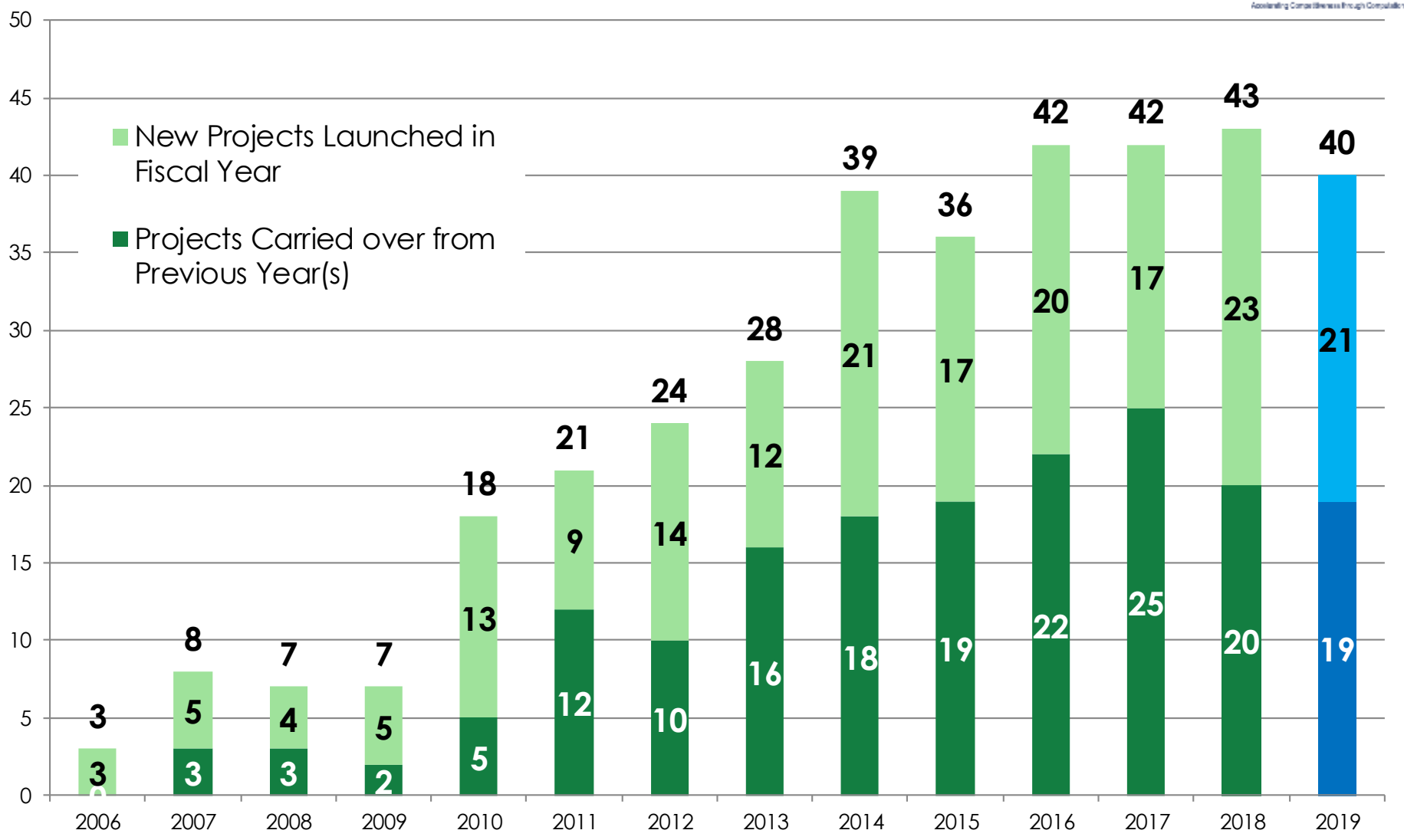
SILICON Therapeutics

BOSCH

Pinnacle Engines

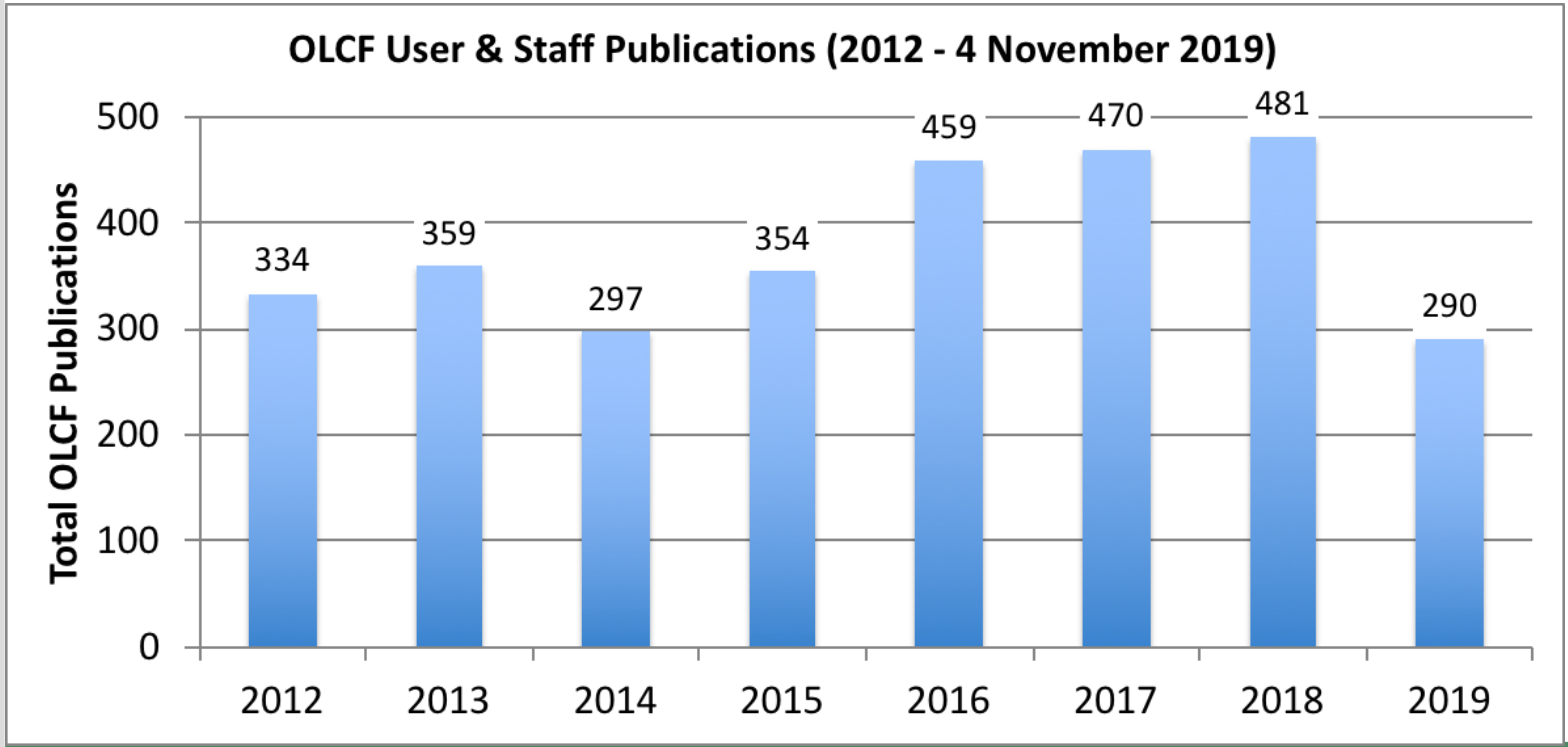
OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

Industrial Use of OLCF by Fiscal Year

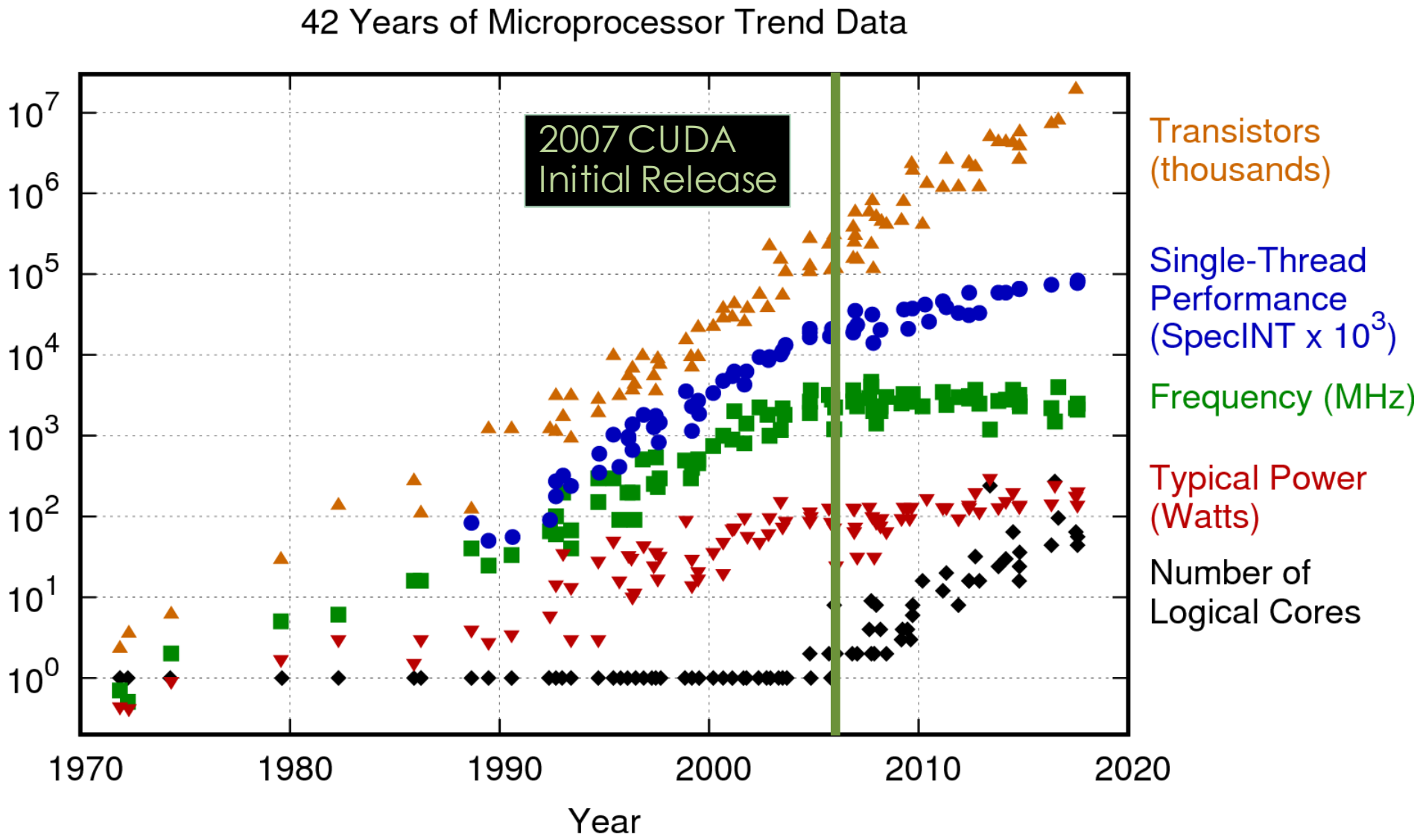


Science Impact – Publications Growth

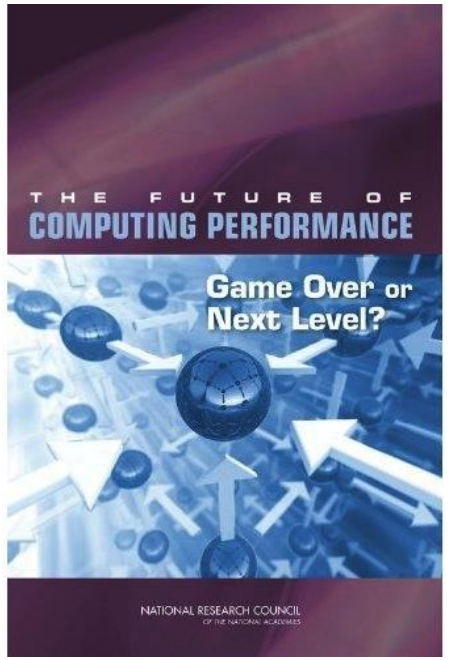
- In 2018, OLCF staff and users have published 472 peer-reviewed articles. This is an OLCF maximum yield.
- Annual publications (User + Staff) at all-time high for the past three successive years (CY16, CY17, CY18.



No more free lunch: *Moore's Law continues, Denard Scaling is over*



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp

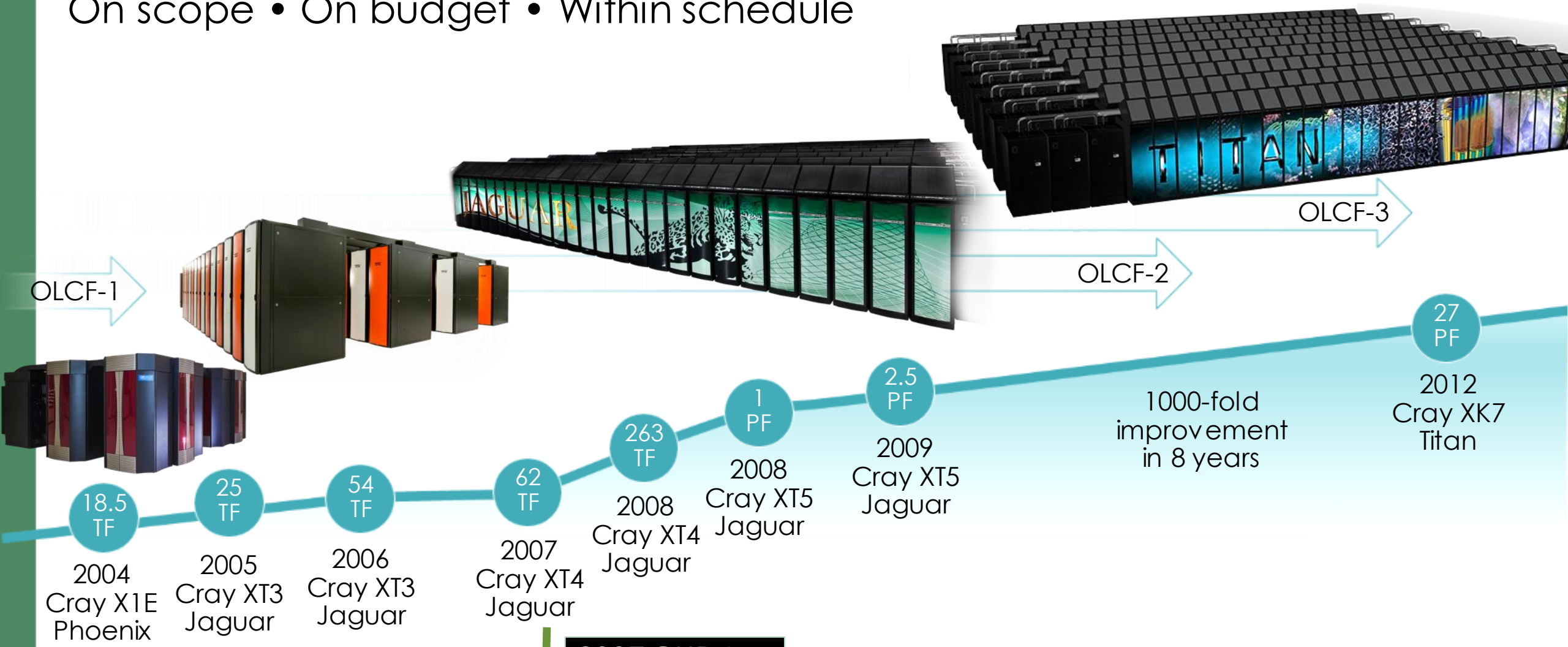


National Research Council (NRC) – Computer Science and Telecommunications Board (2012)

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

ORNL has systematically delivered a series of leadership-class systems

On scope • On budget • Within schedule

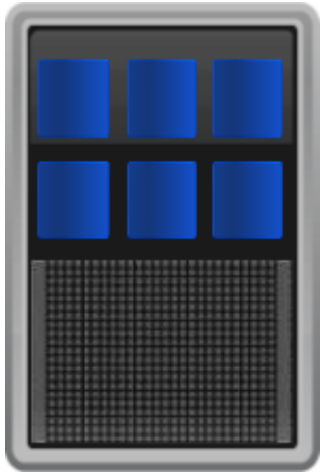


2007 CUDA Initial Release

GPUs provided a path forward using Hierarchical Parallelism

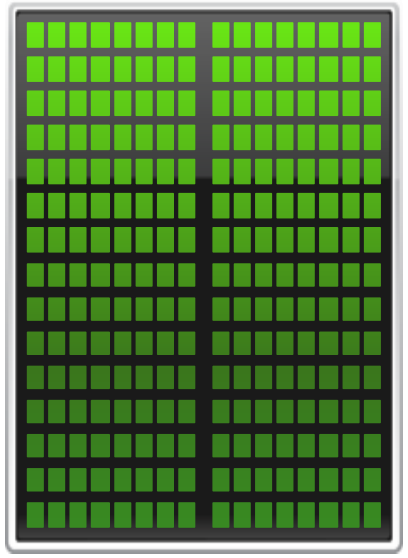
- **Expose more parallelism through code refactoring and source code directives**
  - Doubles CPU performance of many codes
- **Use right type of processor for each task**
- **Data locality: Keep data near processing**
  - GPU has high bandwidth to local memory for rapid access
  - GPU has large internal cache
- **Manage data movement between CPU and GPU memories**

CPU



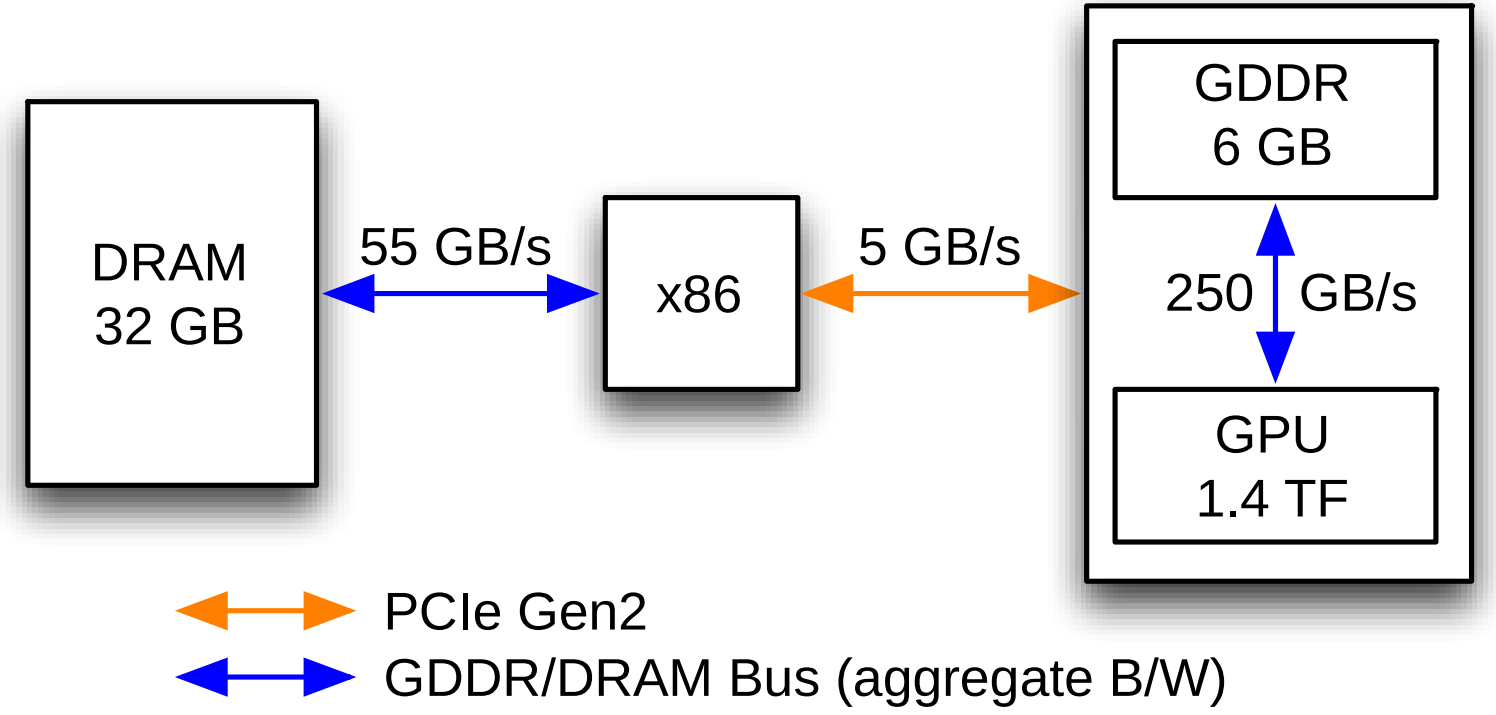
- Optimized for latency and sequential multitasking

GPU Accelerator



- Optimized for throughput and many simultaneous tasks
- 10× performance per socket
- 5× more energy-efficient systems

Titan's Node Architecture



We are building on this record of success to enable exascale in 2021

Titan was decommissioned August 2019.

- From Titan to Summit, 8-fold improvement
- From Titan to Frontier, 50-fold improvement
- From Jaguar to Frontier, 500-fold improvement

ORNL Debuted Summit Supercomputer: June 8, 2018

- ORNL launched the Summit IBM AC922, the world's most powerful and smartest scientific supercomputer, on June 8. Energy Secretary Rick Perry, Tennessee Governor Bill Haslam, and the CEOs of IBM and NVIDIA were present to celebrate the announcement.
- Attending the June 8 launch were (left to right) Tennessee Governor Bill Haslam, US Secretary of Energy Rick Perry, IBM CEO Ginni Rometty, ORNL lab director Thomas Zacharia, and NVIDIA CEO Jensen Huang.

ORNL has launched its newest supercomputer: the IBM AC922 Summit, located at the OLCF.

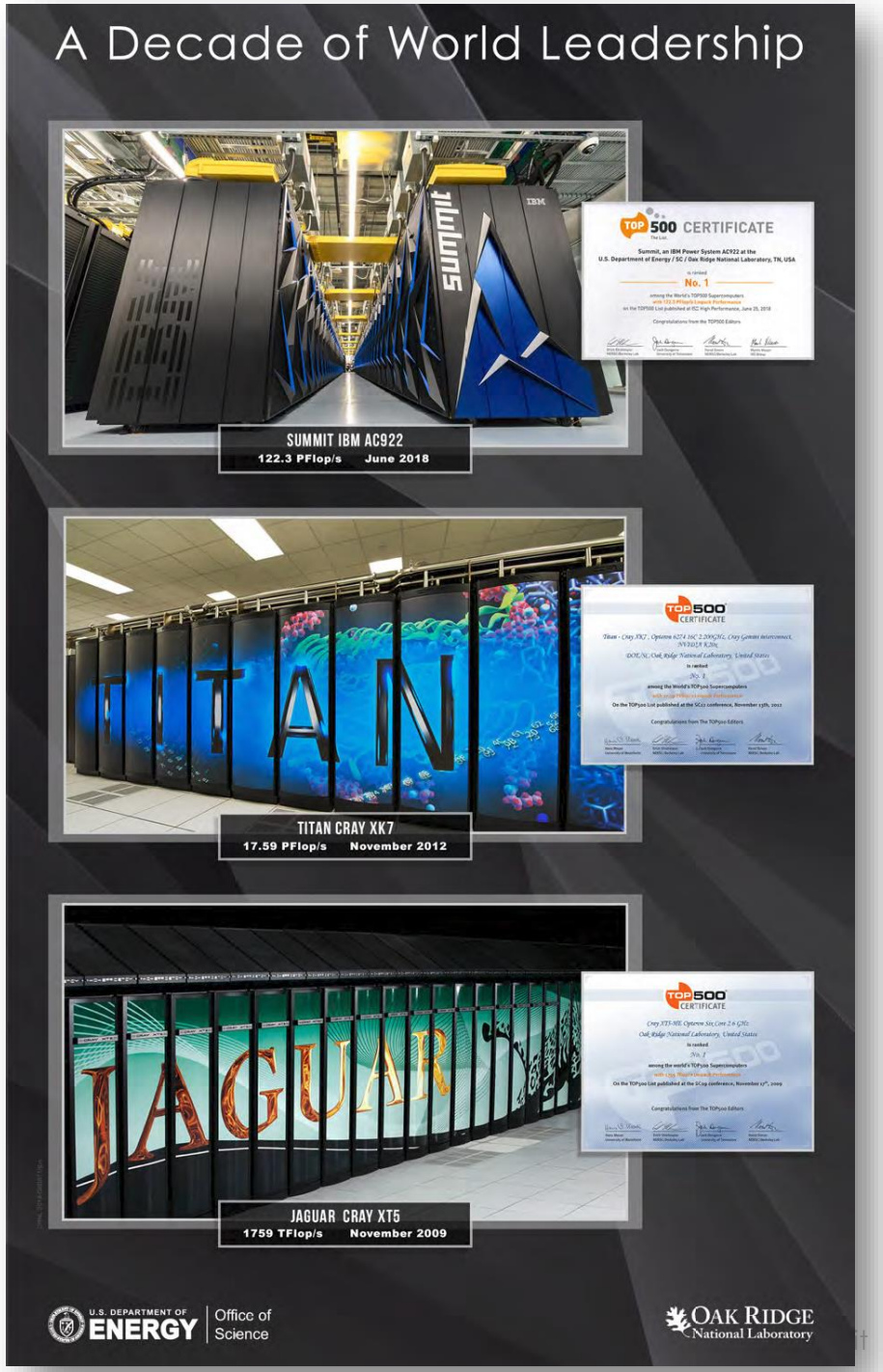
Summit is latest DOE #1 system on Top500

143.5 PF HPL
Shows math performance

2.9 PF HPCG
Shows fast data movement

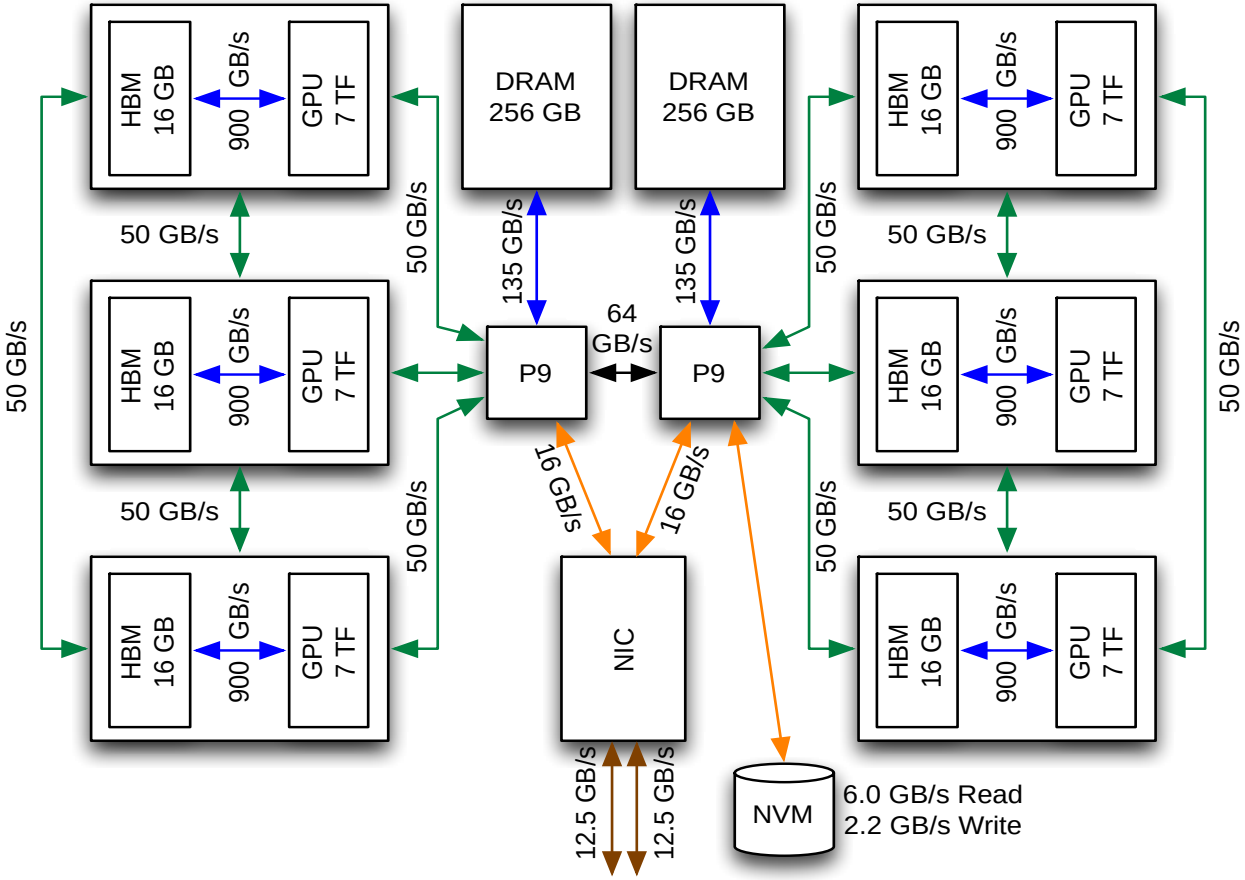
14.7 GF/W
Shows energy efficiency

#1 on the IO-500
Shows file system performance



Summit Node Schematic

- Coherent memory across entire node
- NVLink v2 fully interconnects three GPUs and one CPU on each side of node
- PCIe Gen 4 connects NVM and NIC
- Single shared NIC with dual EDR ports

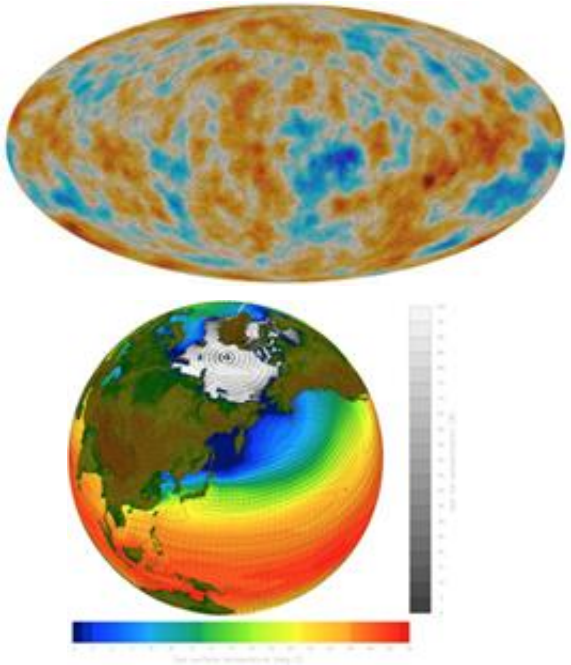


TF	42 TF (6x7 TF)
HBM	96 GB (6x16 GB)
DRAM	512 GB (2x16x16 GB)
NET	25 GB/s (2x12.5 GB/s)
MMsg/s	83

- HBM/DRAM Bus (aggregate B/W)
- NVLINK
- X-Bus (SMP)
- PCIe Gen4
- EDR IB

HBM & DRAM speeds are aggregate (Read+Write).
All other speeds (X-Bus, NVLink, PCIe, IB) are bi-directional.

Common Themes Across DOE Science Offices
Data: Large-scale data storage and analysis



Experimental and simulated data set volumes are growing exponentially. Examples: High luminosity LHC, light sources, climate, cosmology data sets ~ 100s of PBs.

Methods and workflows of data analytics are different than those in traditional HPC. Machine learning is revolutionizing field. Established analysis programs must be accommodated.

DOE exascale ecosystem requirements reviews: http://exascaleage.org/

OLCF Training Program: Tom Papatheodore, leader

Introduction to HPC

intro topics + traditional classroom-style

https://www.olcf.ornl.gov/calendar/introduction-to-hpc/

Advanced Topics

Summit Programming Environment

https://www.olcf.ornl.gov/calendar/summit-training-workshop-february-2019/

App Teams

hands-on + application readiness

https://www.olcf.ornl.gov/for-users/training/gpu-hackathons/

Special Topics & Tutorials

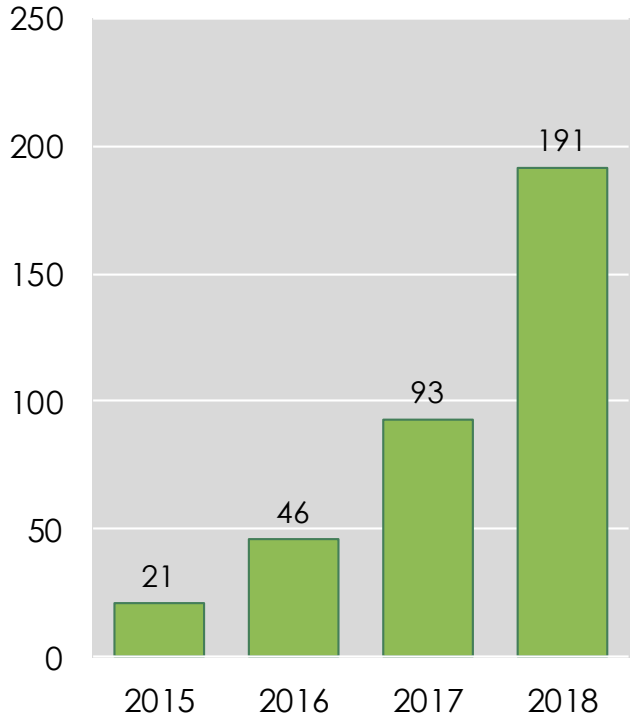
current topics + scientific domains

https://github.com/olcf-tutorials

https://www.olcf.ornl.gov/for-users/training/training-archive/

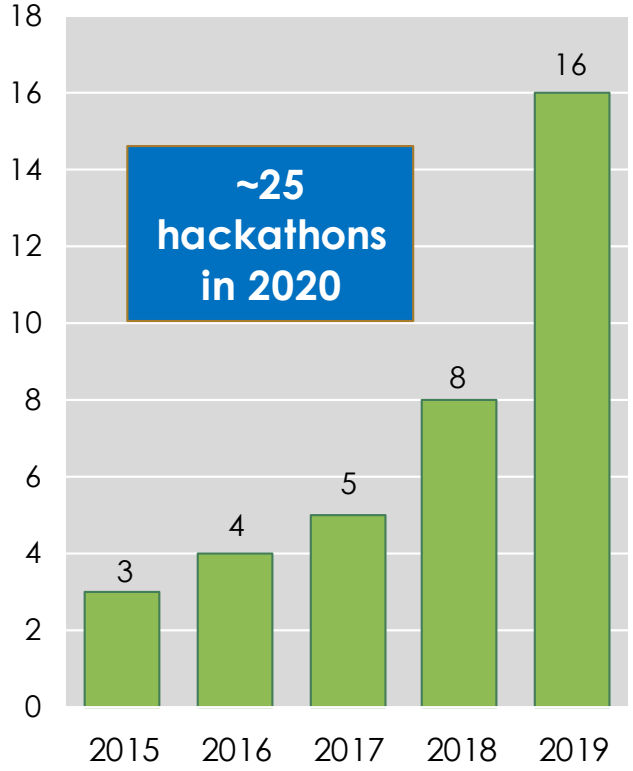
GPU Hackathons Impact

Number of Codes Accelerated at Hackathons (cumulative)



Hackathon Hosting Institutions

Number of Hackathons



https://www.gpuhackathons.org/

So far in 2019…

60+ application teams
400+ attendees
80+ institutions

papatheodore@ornl.gov

TRAINING AND EDUCATION

ORNL Leader: papatheodore@ornl.gov
OpenACC Leader: jlevites@nvidia.com

GPU Hackathons and Bootcamps

Learn

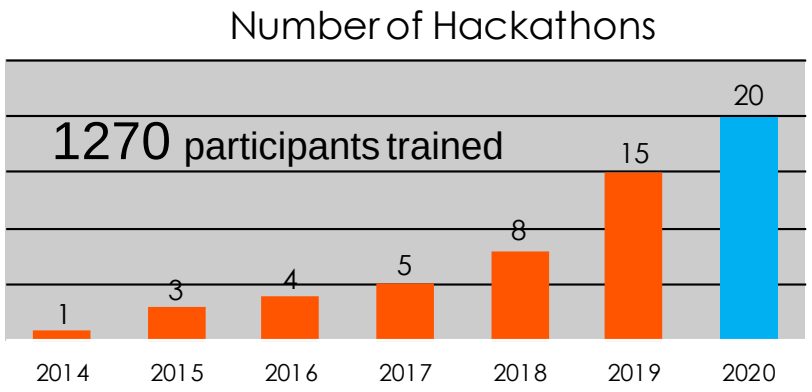
GPU Bootcamps

- 1 or 2-day event
- Labs and a mini-app challenge
- 50 people with 2-3 mentors
- Prepare for future hackathons and start collaborations

Hackathons

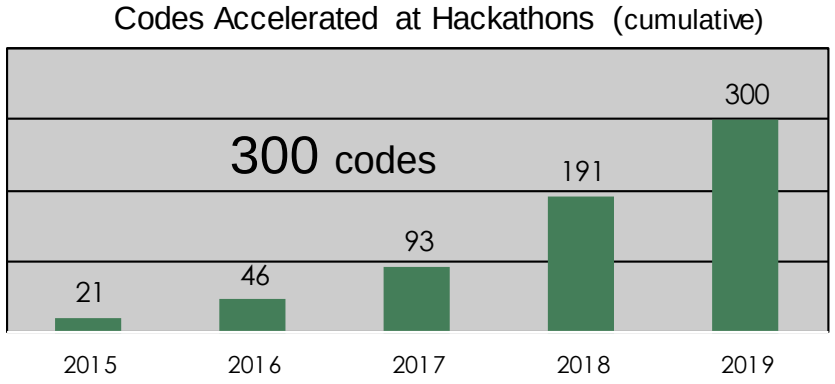
- 5-day event
- Users bring their own codes or data
- Up to 10 teams with 3+ people per team
- 2 mentors per team

Optimize Accelerate



23 Bootcamps in 2019

40 Bootcamps in 2020



www.gpuhackathons.org

SELECTED HACKATHONS HIGHLIGHTS

Code	Domain	Result in 5 days
CHARMM	MD	7.5x on a kernel
LAVA(NASA)	CFD	10x on a mini-app
HiFUN	CFD	3.2x on a full app, 100x on a kernel
Yambo	QC	6x on a full app
CASTRO	Astro	14x on mini-app
Garnet	Earth Sciences	40x on mini-app
CGYRO	Physics	10x on a mini-app
CASTEP	QC	2.7x and 11x on 2 parts
SPECFEM-X	Geophysics	205x vs single CPU core
GTS	Fusion	25x for the kernel and 3x for the overall app
Quantum Espresso	QC	6x on a part of the code
DFT-FE	Quantum Mechanics	40x on kernels

I'm a firm believer in **collaborative science**, and it was wonderful to see several of my graduate students and postdocs deeply engaged. The team made tremendous progress during the week, and that momentum has carried the project forward ever since. **In four days they managed to obtain a two-orders-of-magnitude increase in performance, and since then they've made further optimizations** that gained them another order of magnitude. This speedup has opened up an entirely new class of problems in quasistatic global geophysics. From my perspective, this was a very successful event!

***Jeroen Tromp**, Blair Professor of Geology and Professor of Applied & Computational Mathematics at Princeton University*

OPENACC ORGANIZATION MISSION

The OpenACC Organization is dedicated to helping the research and developer community advance science by expanding their accelerated and parallel computing skills. We have 3 areas of focus: participating in computing ecosystem development, providing training and education on programming models, resources and tools, and developing the OpenACC specification.

Ecosystem Development	Training/Education	OpenACC Specification

PILLARS OF OPENACC ORGANIZATION

ECOSYSTEM DEVELOPMENT

Work with Language Standards committees

- Apply collective lessons learned
- Bridge gaps in base languages
- Develop roadmap for interoperability

Strive for performant interoperability

TRAINING & EDUCATION

Hackathons/Bootcamps

- Focus on accelerated computing

Training Materials:

- Courses
- Containers
- Workshops

Programs for educators, students, mentors

OPENACC SPECIFICATION

Develop OpenACC Specification by introducing new features and functionality

ECOSYSTEM DEVELOPMENT

APPLICATIONS

BASE AND DOMAIN SPECIFIC LANGUAGES	HIGH LEVEL MODELS & LIBRARIES	DIRECTIVES	PLATFORM-SPECIFIC ENVIRONMENTS

Enabling Performant Interoperability

OPENACC DIRECTIVES

a directive-based parallel programming model designed for usability, performance and portability

3 OF TOP 5 HPC

Intersect360 RESEARCH

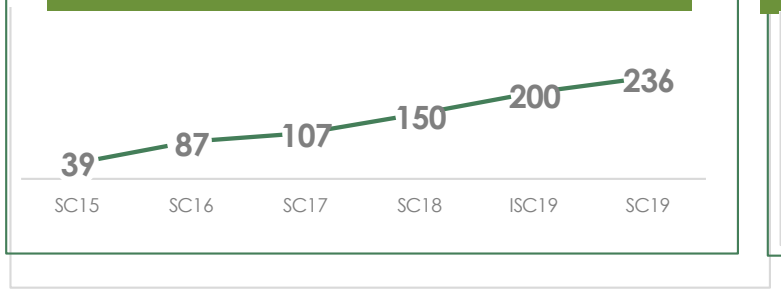
18% OF INCITE AT SUMMIT

summit

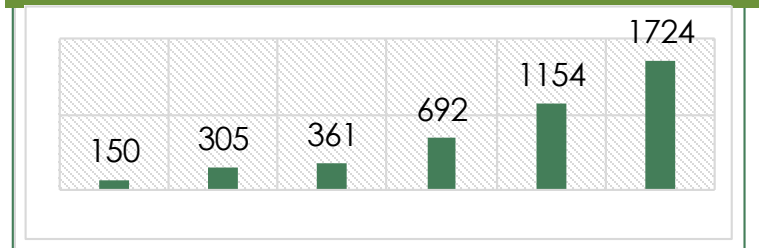
PLATFORMS SUPPORTED

NVIDIA GPU
X86 CPU
POWER CPU
Sunway
ARM CPU
AMD GPU

OPENACC APPS



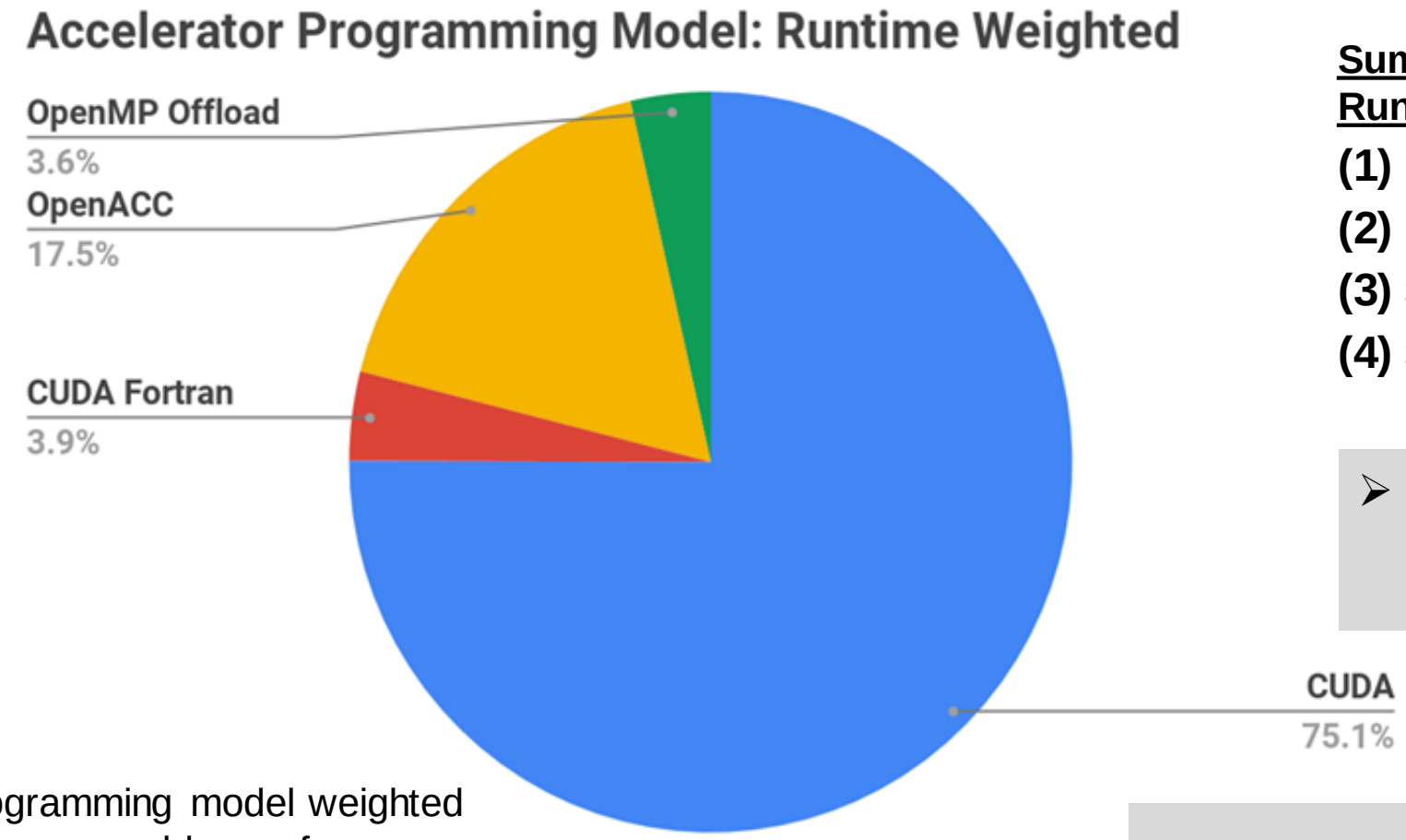
OPENACC SLACK MEMBERS



>200K DOWNLOADS

PGI Community EDITION

Accelerator Programming Models on Summit- 2019 INCITE



Programming model weighted by consumed-hours for INCITE up to September 2019

Source: XALT /Reuben Budiardja, NCCS

Summit Accelerator Programming Models, Runtime Weighted

(1) 75.1% CUDA (includes RAJA, KOKKOS)

(2) 17.5% OPENACC

(3) 3.9% CUDA Fortran

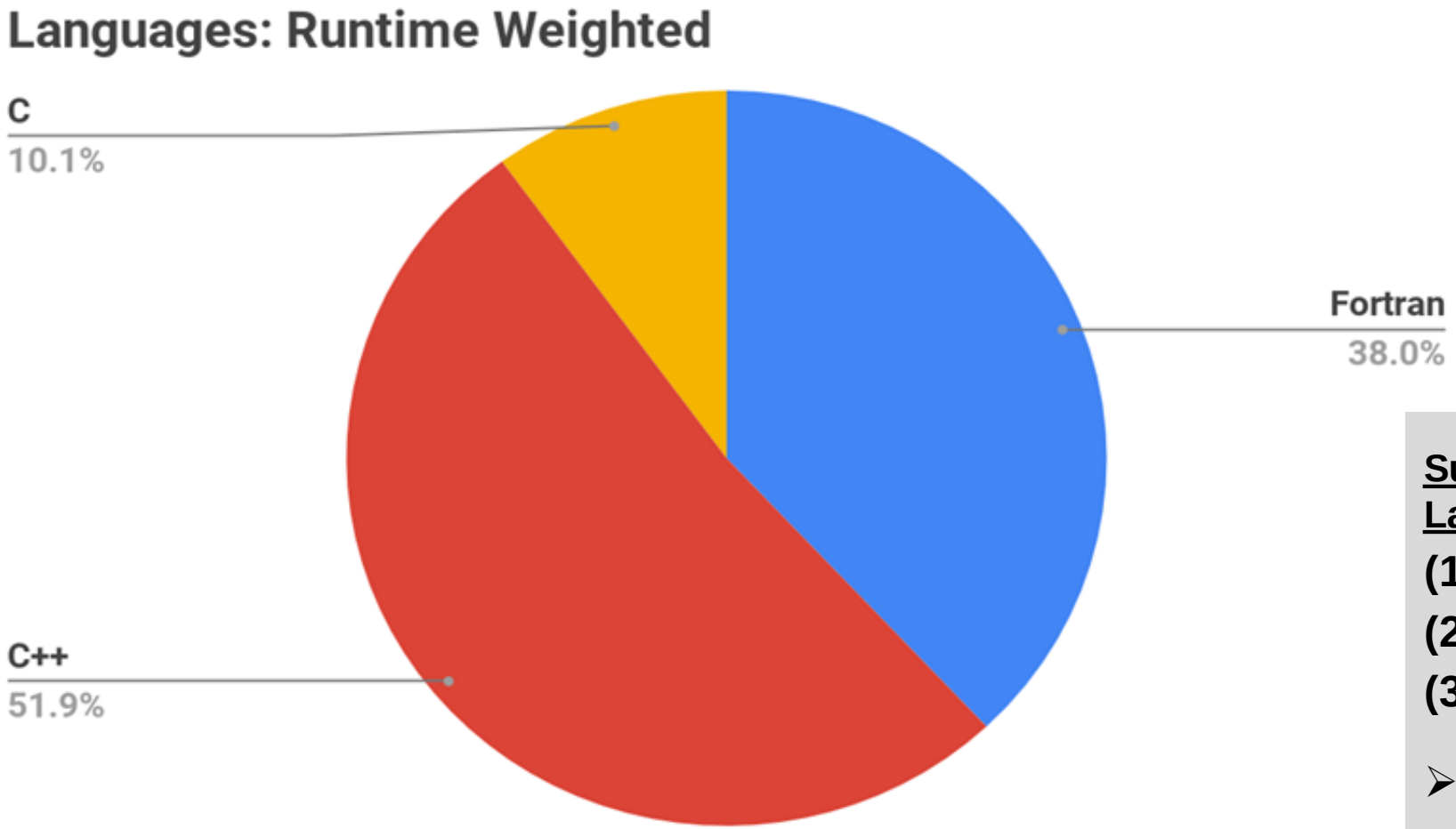
(4) 3.6% OPENMP

- **OpenACC programming model is critically important for Summit's user programs**

NOTES:

- Programming model weighted by consumed-hours.
- Using XALT, RAJA & KOKKOS use show up as CUDA

Base languages in 2019 INCITE on Summit



Summit: Base Programming Languagues, Runtime Weighted

(1) 52% C++

(2) 38% Fortran

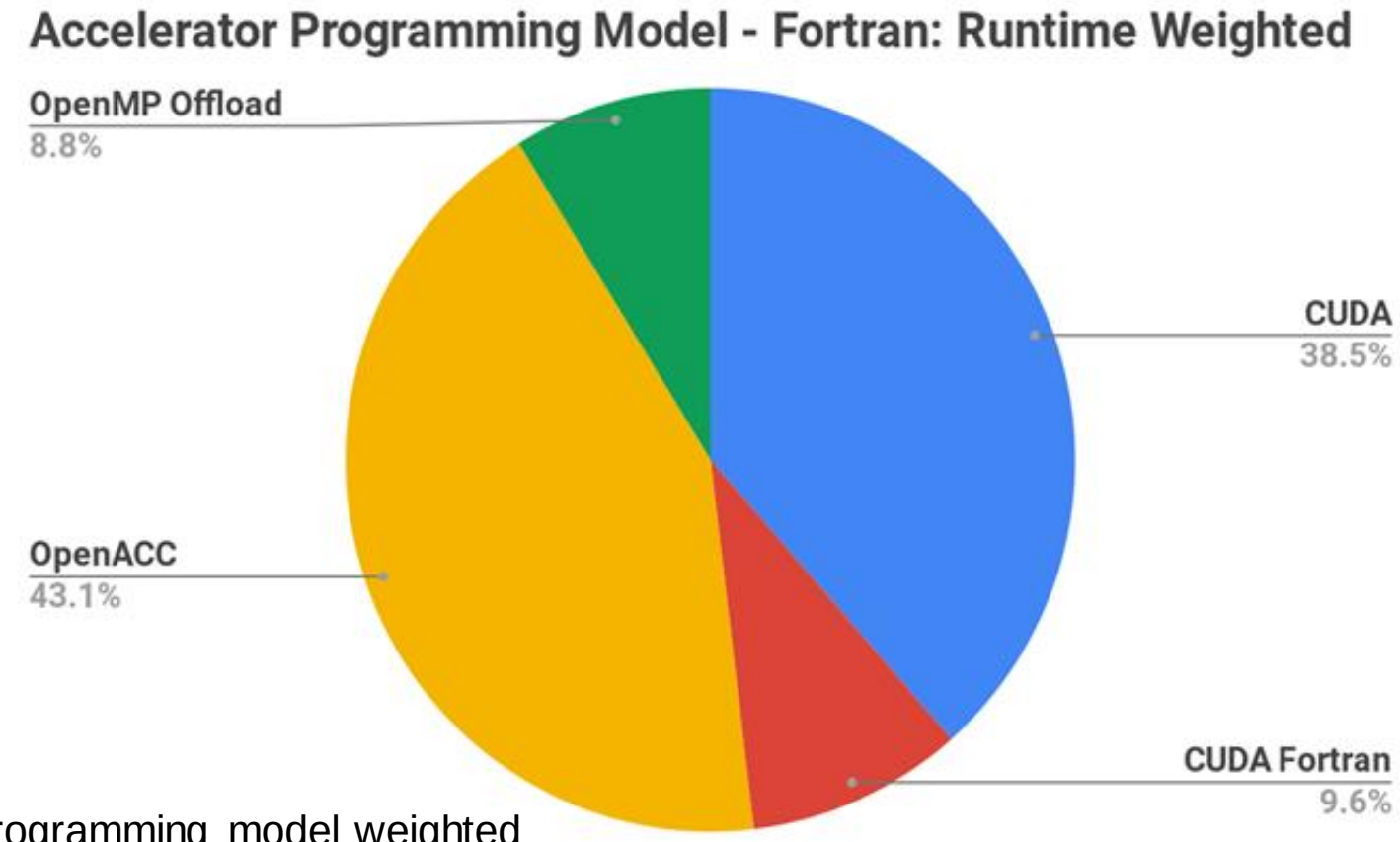
(3) 10% C

- **Modern FORTRAN required as a "first-class" citizen in our user programs.**

Source: XALT /Reuben Budiardja, NCCS

Based on usage up to September 2019, results will continue to change

Accelerator Programming Models for **Fortran Codes** on Summit - 2019 INCITE



Programming model weighted by consumed-hours for INCITE up to September 2019

Source: XALT /Reuben Budiardja, NCCS

Summit Accelerator Programming Models, Fortran Code - Runtime Weighted

(1) 43.1% OPENACC

(2) 38.5% CUDA (includes RAJA, KOKKOS)

(3) 9.6% CUDA Fortran

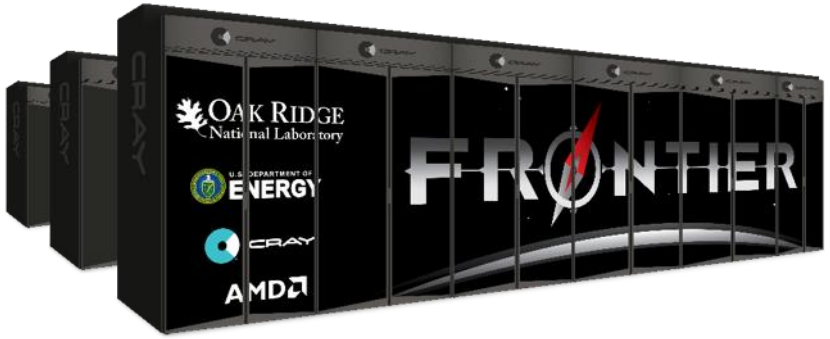
(4) 8.8% OPENMP

- **OpenACC programming model is critically important for Summit's user programs**

NOTES:

- Programming model weighted by consumed-hours.
- Using XALT, RAJA & KOKKOS use show up as CUDA

Frontier Overview



Partnership between ORNL, Cray, and AMD

The Frontier system will be delivered in 2021

Peak Performance greater than 1.5 EF

Composed of more than 100 Cray Shasta cabinets

- Connected by Slingshot™ interconnect with adaptive routing, congestion control, and quality of service

Node Architecture:

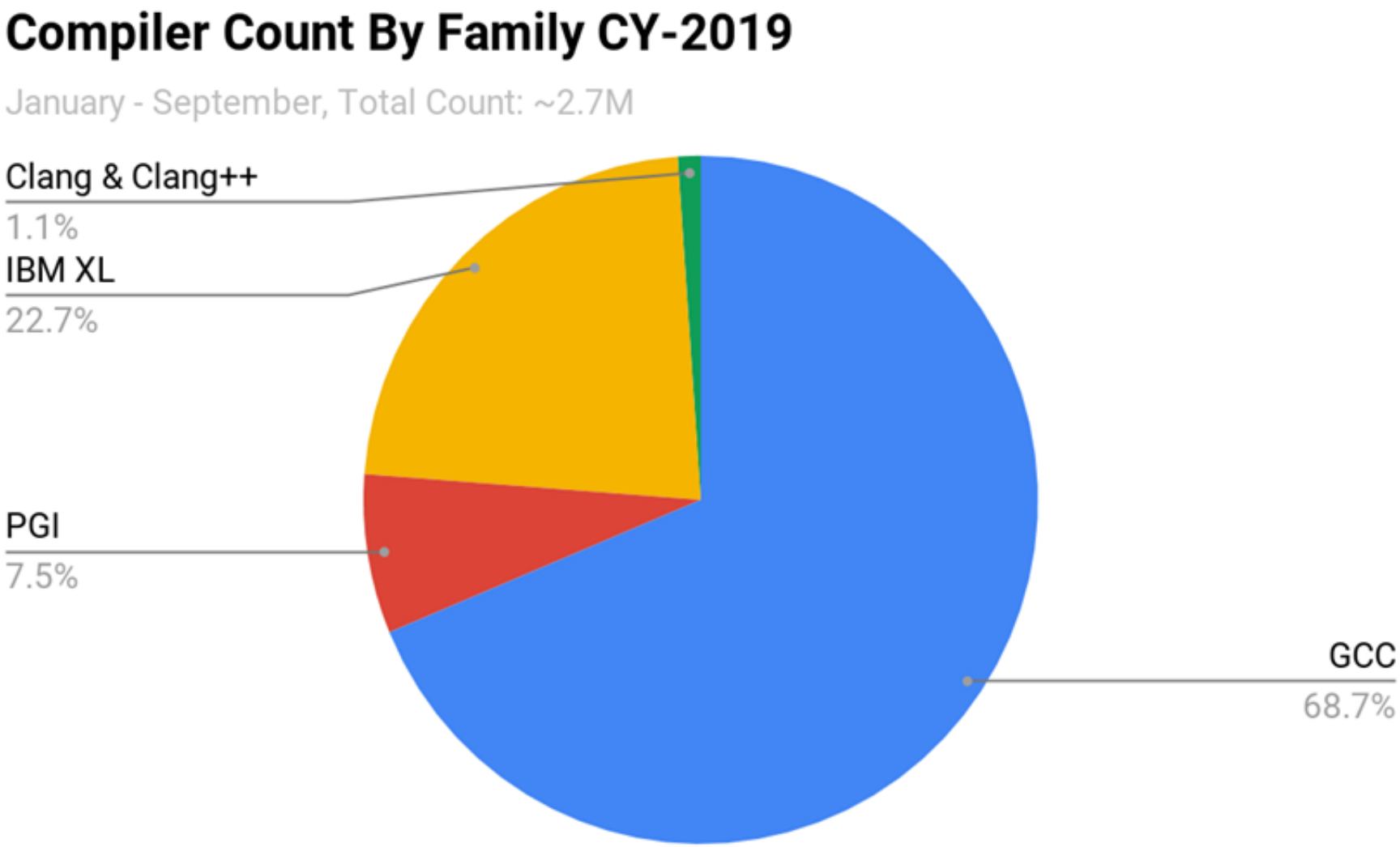
- An AMD EPYC™ processor and four Radeon Instinct™ GPU accelerators purpose-built for exascale computing
- Fully connected with high speed AMD Infinity Fabric links
- Coherent memory across the node
- 100 GB/s injection bandwidth
- Near-node NVM storage

Migration Path from Summit to Frontier

Summit	Frontier	Comments
CUDA C/C++	HIP C/C++	HIP provides tools to help port existing CUDA codes to the HIP layer.
OpenACC	OpenMP (offload)	OpenACC codes can be migrated to OpenMP (offload) for Frontier. ***Direct support for OpenACC on Frontier is still under discussion.***
OpenMP (offload)	OpenMP (offload)	Virtually the same on Summit and Frontier
Fortran w/CUDA C/C++	Fortran w/HIP C/C++	As with CUDA, this will require interfaces to the C/C++ API calls
CUDA FORTRAN	FORTRAN w/HIP C/C++	As with CUDA, this will require interfaces to the C/C++ API calls

https://www.olcf.ornl.gov/frontier/

GCC: Most Frequently Used Compiler Family



RFP for Work on GCC Compiler Suite

Ensure that OpenACC is performant on Frontier through GCC

- OLCF will enable up-to-date implementations of OpenACC and OpenMP standards within the GCC compiler suite up to the latest versions of the standards.
- A "Request for Proposals" was opened Aug. 29, 2019.
  - Support the GPUs of immediate interest to OLCF (NVIDIA & AMD) so that OpenACC becomes fully capable for OLCF users on Summit, Frontier, and other platforms.
- The expectation is that this work will be completed as quickly as reasonably possible, but prior to April 2022, in anticipation of the upcoming delivery of the new Frontier system.

Summary & Conclusions

- Titan is gone; Summit is here; Frontier is coming for users in 2022.
- Training and education is a critical component of enabling accelerated computing deliver on its opportunities.
- OpenACC directives approach is very important for OLCF user programs.
- Community-focused GPU Hackathons are having a big impact.
- OLCF is increasing investments in GCC to make OpenACC and OpenMP performant on Summit and Frontier.

Acknowledgment

XALT data is supported by a team effort behind the scene:

- Bill Renaud diligently keeps XALT operational data current
- Jason Kincl set up services on Kubernetes cluster / Openshift for analysis
- Ryan Prout helped set up XALT data flows through Apache Kafka

Questions & Discussion

Jack Wells, wellsjc@ornl.gov