OAK RIDGE National Laboratory

Frontier: The world’s most powerful supercomputer

Bronson Messer
Director of Science
Leadership Computing Facility
Oak Ridge National Laboratory

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

U.S. DEPARTMENT OF ENERGY

Outline

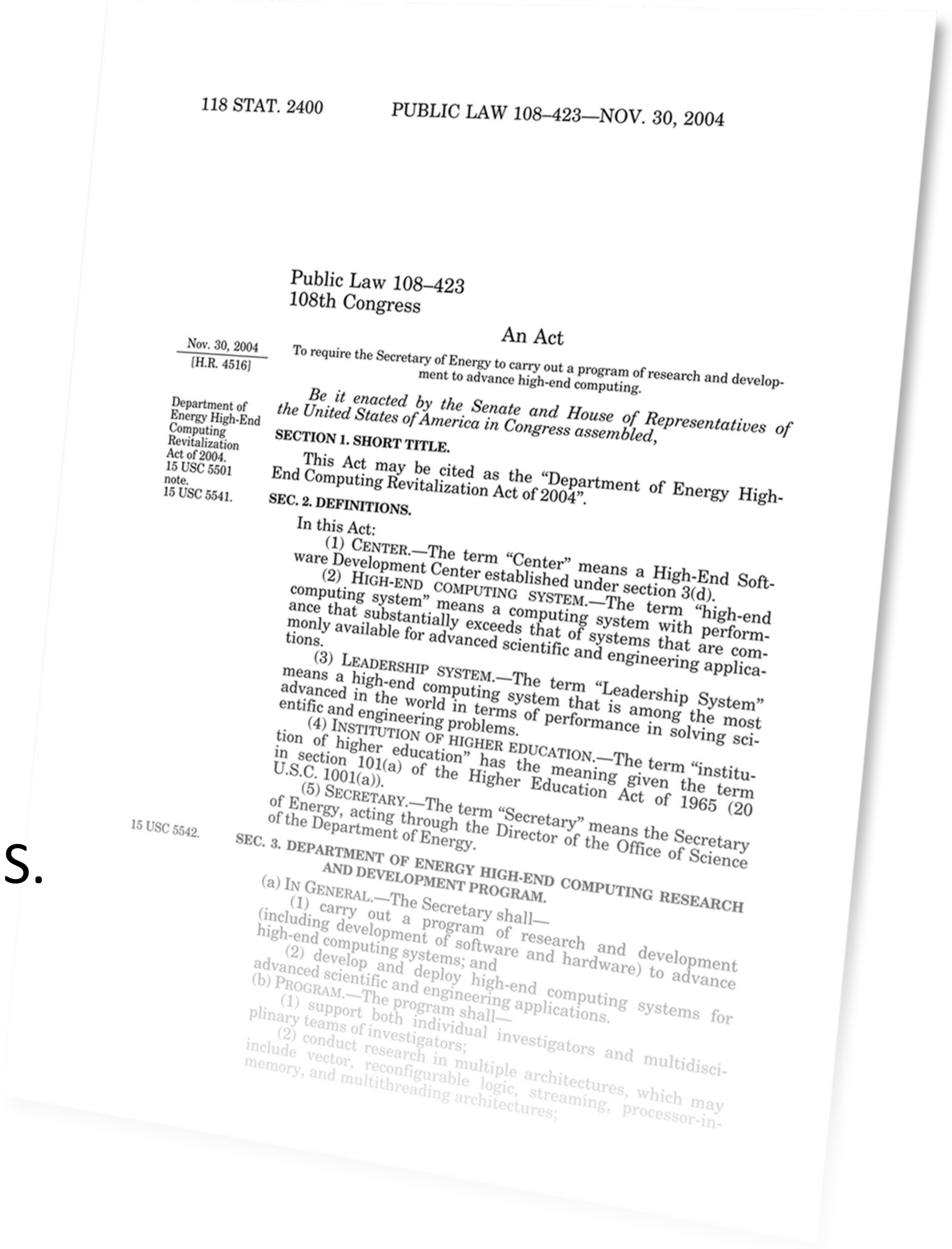
- What is the Leadership Computing Facility and why all the big computers at ORNL?
- An overview of Frontier
- Tales from building the world’s most powerful supercomputer
- What science will be done on the machine?
- An exascale “killer app” (that’s not machine learning).

Leadership Computing Facilities

Department of Energy High-End Computing Revitalization Act of 2004 (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.

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

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

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;

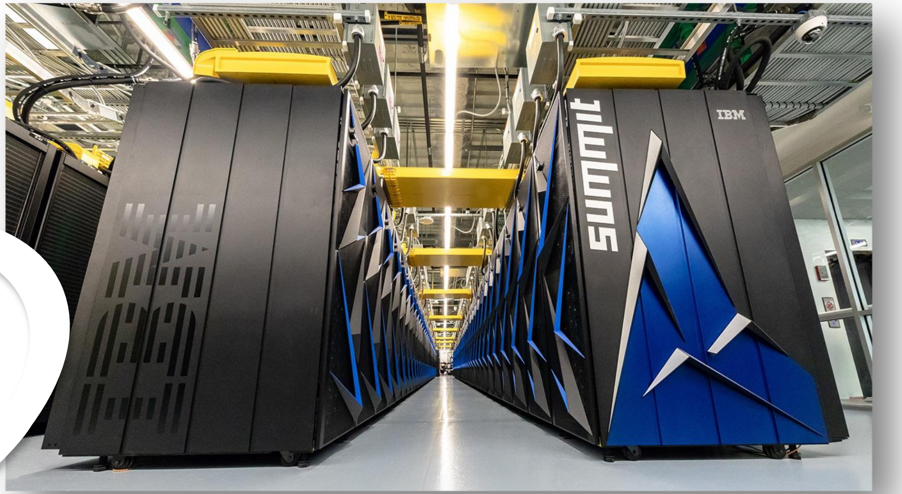
OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

What is the Leadership Computing Facility (LCF)?

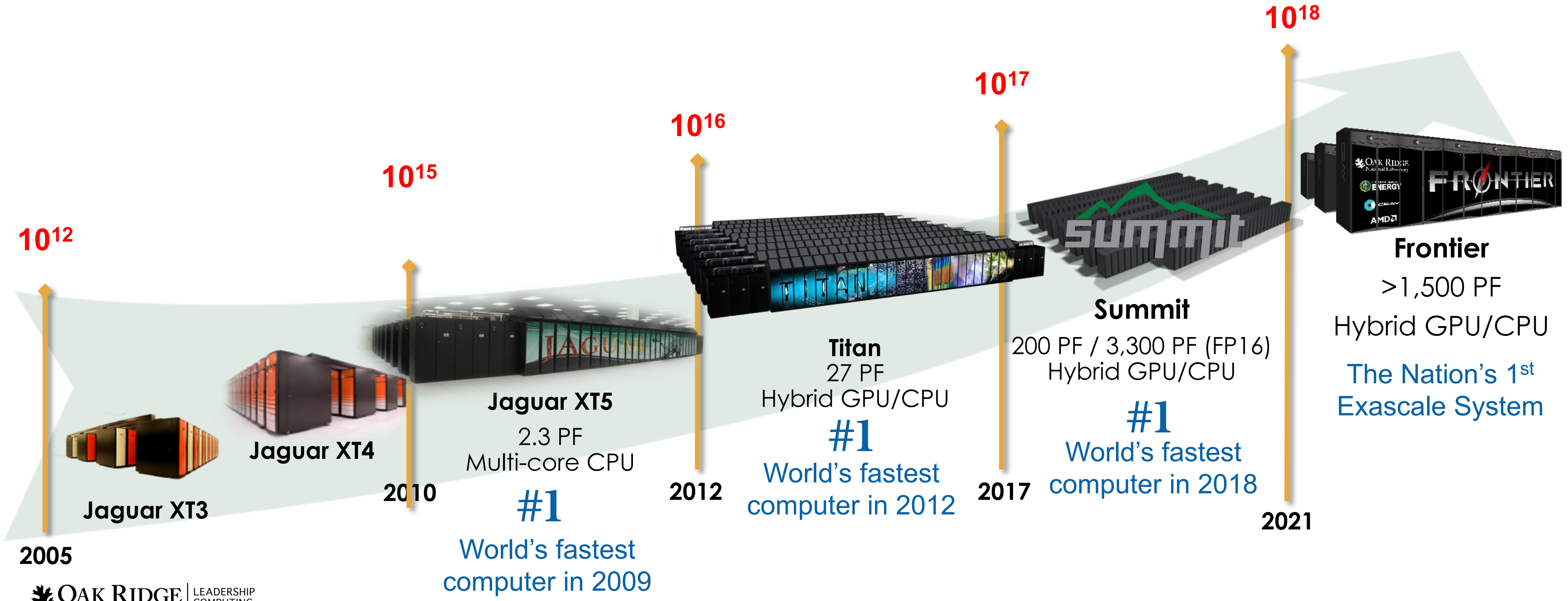
- Collaborative DOE Office of Science user-facility program at ORNL and ANL
- Mission: Provide the computational and data resources required to solve the most challenging problems.
- 2 centers/2 architectures to address diverse and growing computational needs of the scientific community
- Highly competitive user allocation programs (INCITE, ALCC).
- Projects receive 10x to 100x more resources than at other generally available centers.
- LCF centers partner with users to enable science and engineering breakthroughs (Liaisons, Catalysts).

OAK RIDGE National Laboratory

Argonne NATIONAL LABORATORY



ORNL has had a Top 10 supercomputer in every year since the Leadership Computing Facility was founded in 2005. Jaguar, Titan, and Summit are the only DOE/SC systems to be ranked #1 on the TOP500 list of fastest computers.



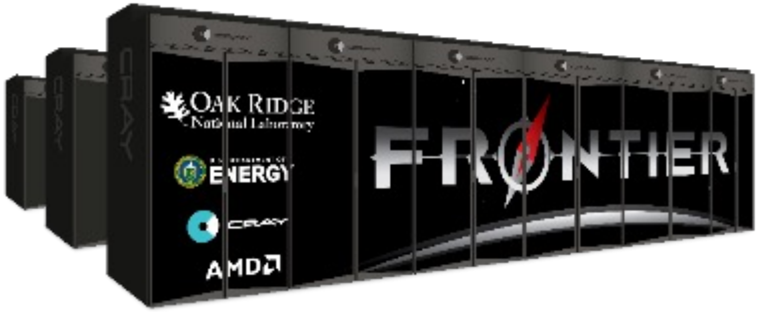
Frontier


FRONTIER
OAK RIDGE National Laboratory
U.S. DEPARTMENT OF ENERGY
Hewlett Packard Enterprise
AMD


FRONTIER
CRAY


CRAY

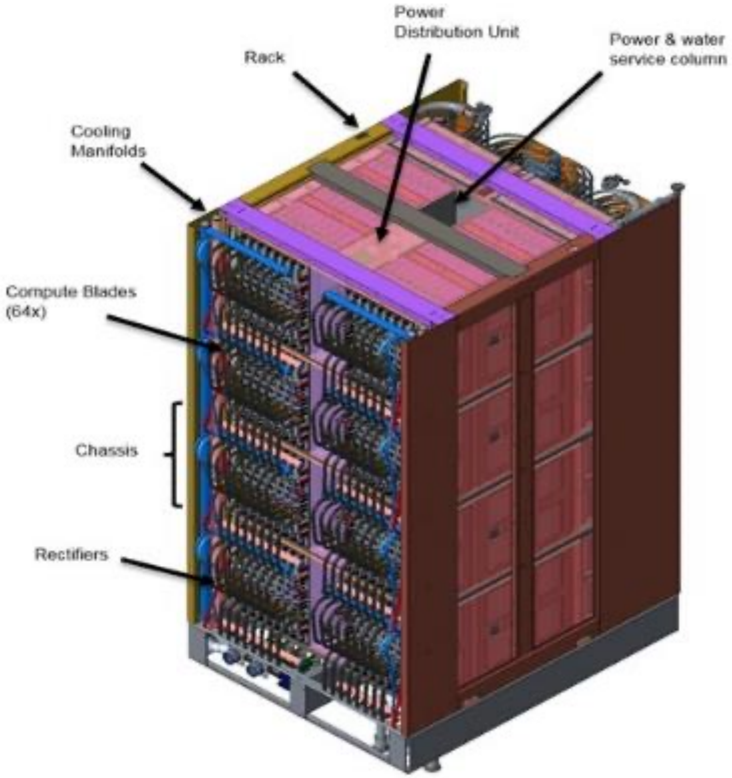
Frontier Overview



System
- 2 EF Peak DP FLOPS
- 74 compute racks
- 29 MW Power Consumption
- 9,408 nodes
- 9.2 PB memory
  (4.6 PB HBM, 4.6 PB DDR4)
- Cray Slingshot network with dragonfly topology
- 37 PB Node Local Storage
- 716 PB Center-wide storage
- 4000 ft^2 foot print

Built by HPE

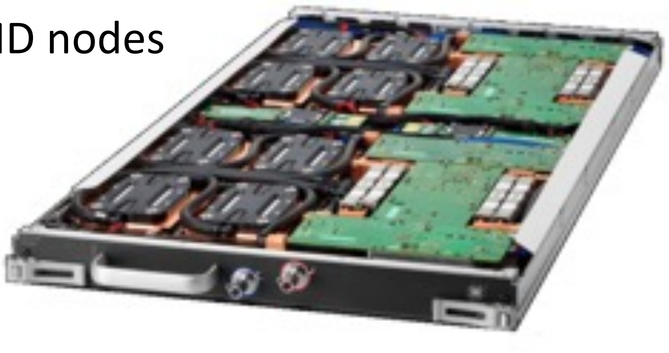
Olympus rack
- 128 AMD nodes
- 8,000 lbs
- Supports 400 KW



Powered by AMD

AMD node
- 1 AMD "Trento" CPU
- 4 AMD MI250X GPUs
- 512 GiB DDR4 memory on CPU
- 512 GiB HBM2e total per node
  (128 GiB HBM per GPU)
- Coherent memory across the node
- 4 TB NVM
- GPUs & CPU fully connected with AMD Infinity Fabric
- 4 Cassini NICs, 100 GB/s network BW

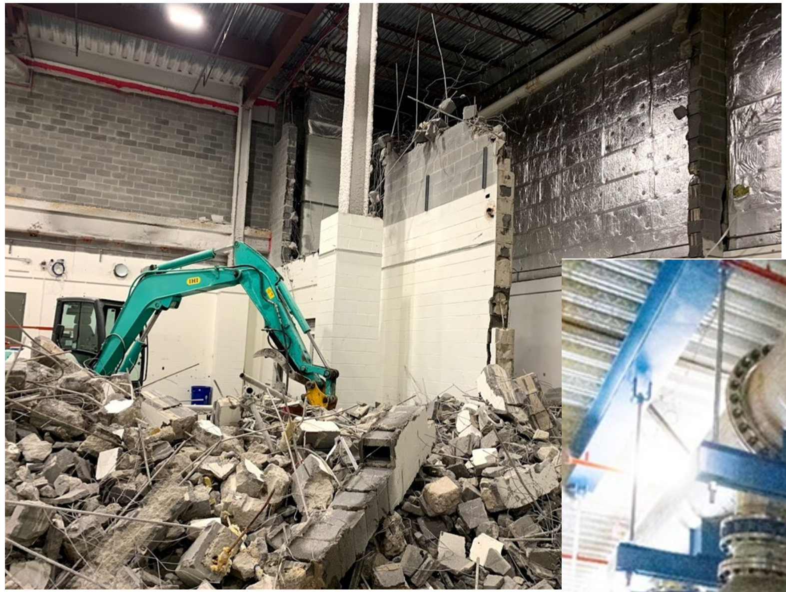
Compute blade
- 2 AMD nodes



All water cooled, even DIMMS and NICs

Power, space, and cooling – (one of) the hard part(s)

- 30 offices, 8 laboratories, and a 20,000 s.f. data center were repurposed



40 MW of power

A new data center (recall the 8,000 pound cabinets…)

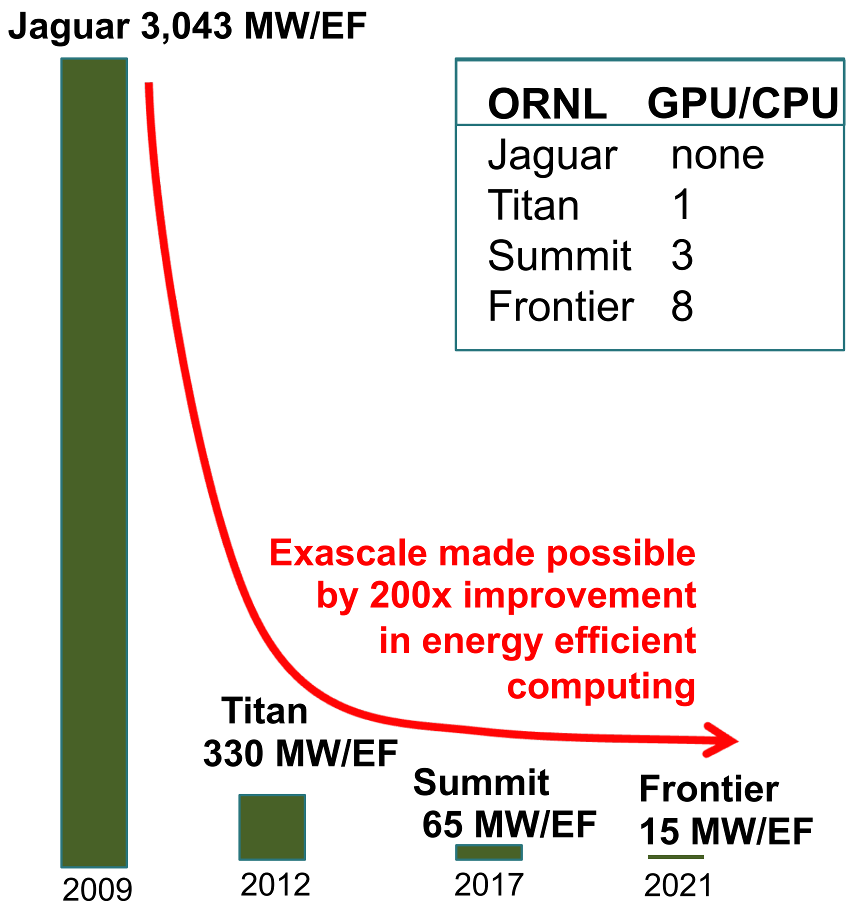
Energy-efficient computing – Frontier achieves 14.5 MW per EF

Since 2009 the biggest concern with reaching Exascale has been energy consumption

- **ORNL pioneered GPU use in supercomputing** beginning in 2012 with Titan thru today with Frontier. Significant part of energy efficiency improvements.
- **ASCR [Fast, Design, Path] Forward vendor investments** in energy efficiency (2012-2020) further reduced the power consumption of computing chips (CPUs and GPUs)..
- **200x reduction in energy per FLOPS** from Jaguar to Frontier at ORNL
- ORNL achieves additional energy savings from using warm water cooling in Frontier (32 C).
  ORNL Data Center PUE= 1.03

Frontier first US Exascale computer

Multiple GPU per CPU drove energy efficiency



ORNL	GPU/CPU
Jaguar	none
Titan	1
Summit	3
Frontier	8

During Frontier build -- the chip shortage hit in earnest!

When HPE began ordering parts, suppliers said the lead time on orders was increasing an additional 6-12 months.

60 Million parts needed for Frontier

685 Different part numbers used in Frontier

167 Frontier part numbers affected by the chip shortage
(more than 2 million parts from dozens of suppliers worldwide)

12 Part numbers blocked building the first compute cabinet

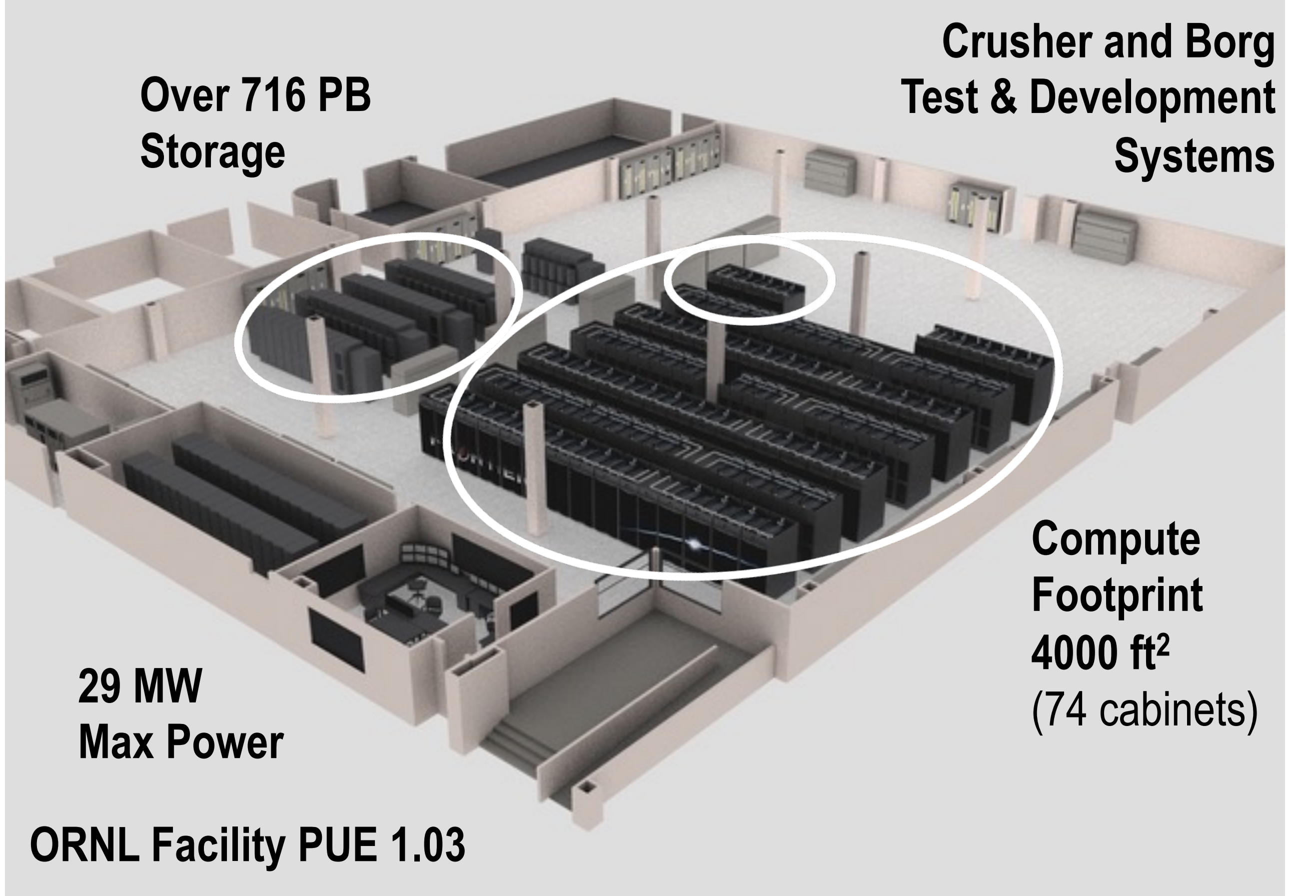
15 Part numbers shortage for AMD building all the MI200 cards for Frontier

It wasn't exotic parts like CPUs or GPUs, rather parts needed by everyone – in cars, TVs, electronics, such as voltage regulators, oscillators, power modules, etc.

Last Cabinet of Frontier Delivered to ORNL October 18, 2021

Thanks to Heroic Efforts of the HPE and AMD teams

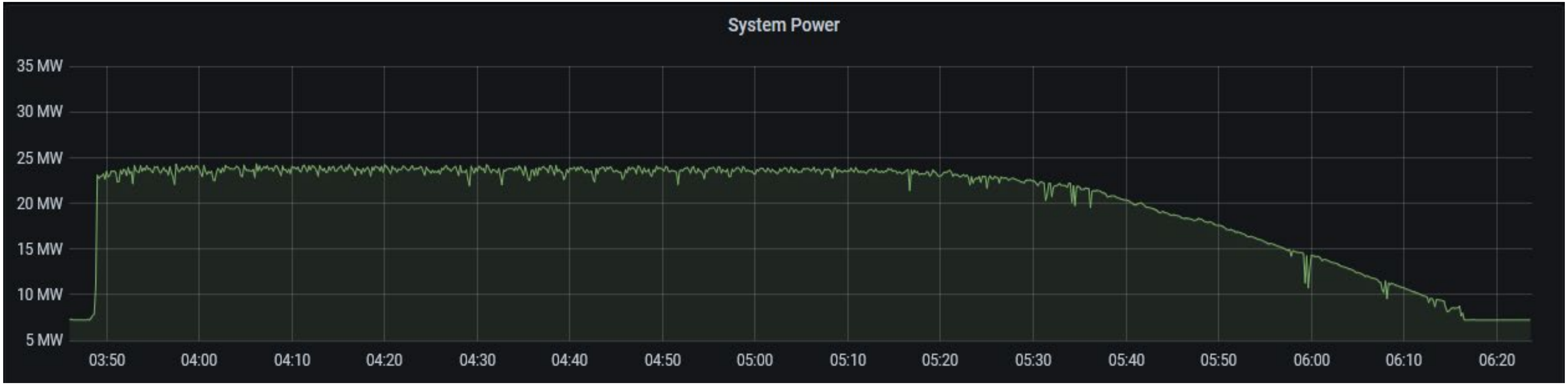
Last cabinet being rolled into place.
(Each cabinet weighs 8,000 lbs.)



After the cabinets arrived they had to be connected. There are 81,000 cables between all the Frontier nodes

Then system debug and tuning began

- We fell into a pattern of repairing hardware, updating software, and tuning the system by day
- And running benchmarks like HPL at night



- In May, as time was running out for the June Top500, we had a successful exascale HPL run:

9,248 nodes of Frontier achieved 1.1 EF
#1 TOP500 list
#2 Green500 achieving over 52 Gflop/W

OAK RIDGE NATIONAL LABORATORY'S FRONTIER SUPERCOMPUTER

- 74 HPE Cray EX cabinets
- 9,408 AMD EPYC CPUs, 37,632 AMD GPUs
- 700 petabytes of storage capacity, peak write speeds of 5 terabytes per second using Cray Clusterstor Storage System
- 90 miles of HPE Slingshot networking cables

TOP500

#1

1.1 exaflops of performance on the May 2022 Top500.

GREEN500

#1, #2

62.04 gigaflops/watt power efficiency on a single cabinet.

52.23 gigaflops/watt power efficiency on the full system.

HPL-AI

#1

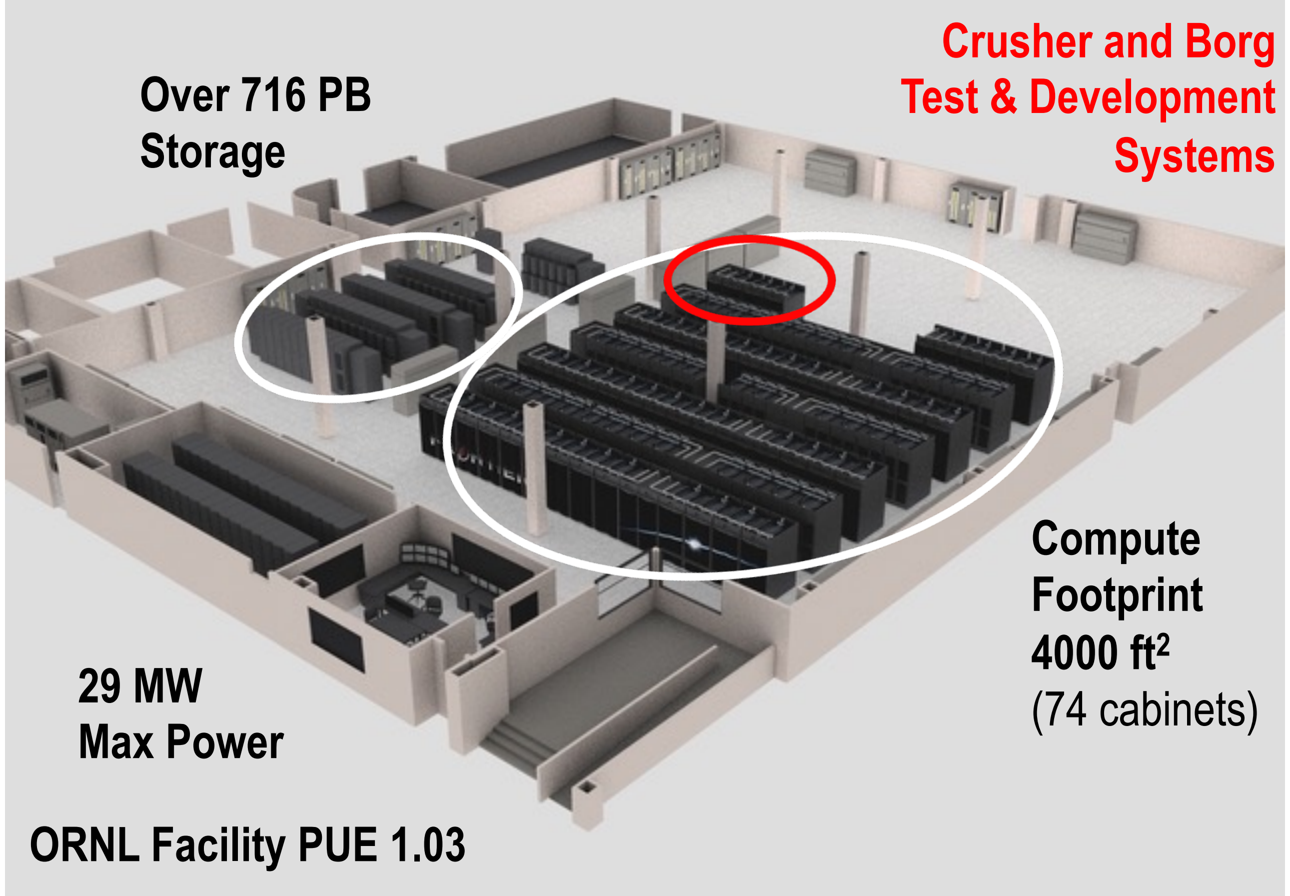
6.88 exaflops on the HPL-AI benchmark.

Sources: May 30, 2022 Top500 release

Last Cabinet of Frontier Delivered to ORNL October 18, 2021

Thanks to Heroic Efforts of the HPE and AMD teams

Last cabinet being rolled into place.
(Each cabinet weighs 8,000 lbs.)



After the cabinets arrived they had to be connected. There are 81,000 cables between all the Frontier nodes

Crusher (Frontier Test and Development System)

- 2 cabinets, the first with 128 compute nodes and the second with 64 compute nodes, for a total of 192 compute nodes. ~40PF (!!)
  - *Crusher is about as powerful as 1.5 Titans!*
- Each node
  - One 64-core AMD EPYC 7A53 CPU
  - 512 GB of DDR4 memory.
  - Four AMD MI250X, each with 2 Graphics Compute Dies (GCDs) for a total of 8 GCDs per node
  - Connected with 4 HPE Slingshot 200 Gbps NICs
- Kept in rough sync with Frontier SW stack

CAAR

The **Center for Accelerated Application Readiness (CAAR)** is the primary OLCF program to achieve and demonstrate application readiness

- ***Build on the experience from the successful CAAR programs for OLCF-3 (Titan) and OLCF-4 (Summit)***

- ***CAAR project resources***
  - Dedicated collaboration with OLCF staff
  - Support and consultation from other project personnel, particularly from the Programming Environment and Tools area, and the vendor Center of Excellence
  - OLCF Postdoctoral fellows (both during application readiness and early science)
  - Allocations to available compute resources (Summit, early access systems)

Characteristics of CAAR Projects

Application	Programming languages	Scientific libraries used	I/O	Algorithms	Initial parallelization
Cholla	C++	None.	HDF5	Finite volume hydrodynamics	MPI, CUDA
NAMD	C++	FFTW (node-level)	VMD (custom)	MD, PME	CHARM++, CUDA
LSMS	F90/C++	BLAS, LAPACK, FFTW	HDF5	Dense Linear Solvers, Coupled ODE, Poisson Eq., Monte Carlo	MPI+CUDA
CoMet	C++	cuBLAS, MAGMA	None	2-way and 3-way Proportional Similarity Method and Custom Correlation Coefficient	MPI+OpenMP, CUDA
GESTS	F90	FFTW	HDF5	Fourier pseudo-spectral methods	MPI+OpenMP 4.5
NUCCOR	F90 + F2008; C	BLAS, LAPACK	HDF5	CCSD + CCSDT, Hartree-Fock, Sparse and dense linear algebra (eigensolvers)	MPI+OpenMP, CUDA
PIConGPU	C++	Alpaka, SOLLVE	ADIOS	PIC	MPI+OpenMP, CUDA/HIP/TBB thru Alpaka
LBPM	C++	Zlib	SILO, HDF5	Lattice Boltzmann methods	MPI, CUDA

Large Scale Density Functional Theory at the Exascale with LSMS

Workflows and high performance computations to predict materials properties

Research Topics

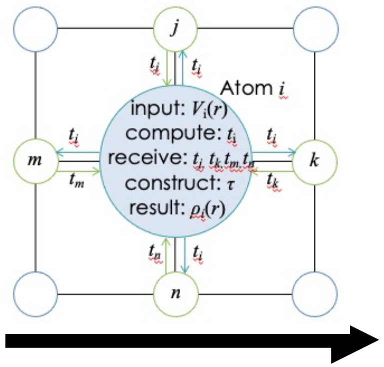
- Understanding the role of disorder and defects in materials for electronic and mechanical properties
- Complex magnetic order – topological magnetic structures (e. g. Skyrmions) and magnetism beyond ideal crystal

Recent Highlights

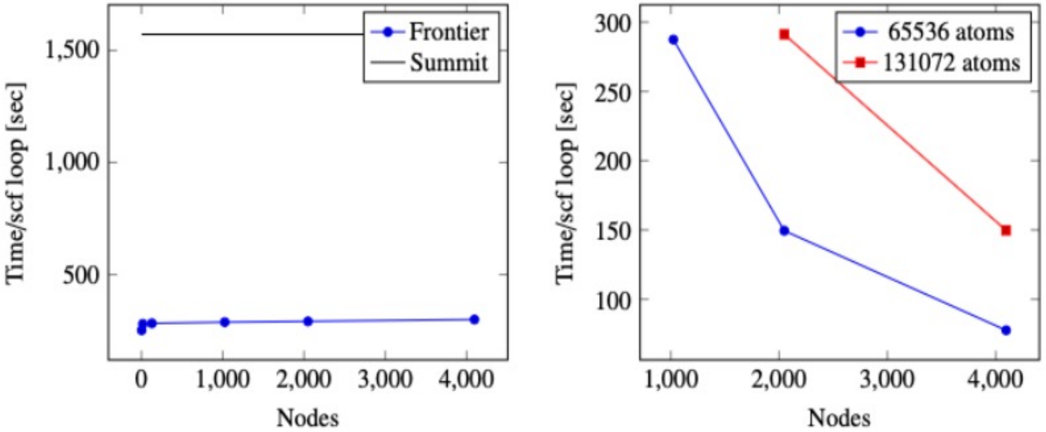
- Successful porting of the LSMS code (github.com/mstsuite/lsms) to Frontier for exascale materials simulations.
- <u>**Scaling of first principles calculations to O(100,000) up to O(1,000,000) atoms for the first time.**</u>
- Demonstrated scaling of LSMS on Frontier up to 1,048,576 atom FePt system on 8192 Frontier nodes.
- Speedup of LSMS from Summit to Frontier from combined hardware and software improvements is ~8x

Future work

- Capabilities for non-metallic quantum materials
- Calculation of forces for ab-initio relaxation and first-principles molecular dynamics.



Moving from CUDA to HIP



Weak (left) and strong (right) scaling results of LSMS for FePt calculations on Frontier

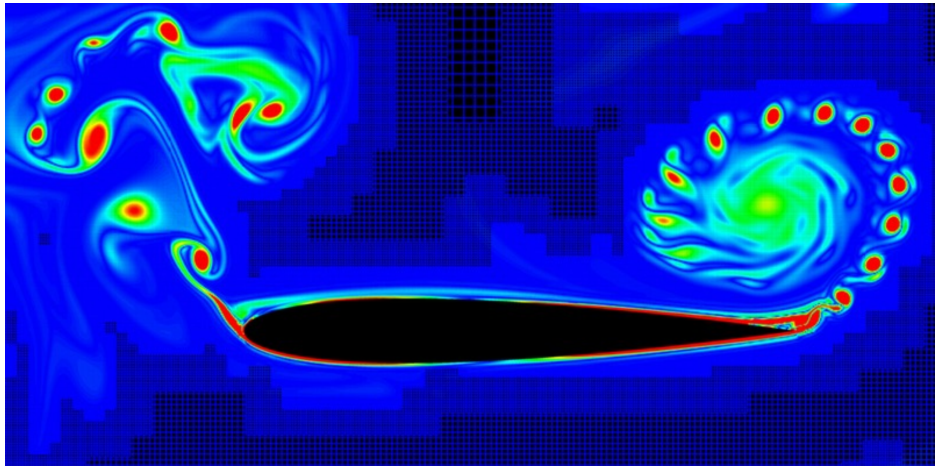
OAK RIDGE National Laboratory

An ill-posed question, but…

- **What is the "killer app" for exascale computing?**
- This question is kinda dumb. A lot of the value of unique supercomputing facilities is the ability to impact a huge variety of scientific problems.
- But, people ask it…

- Maybe there's not a killer app, but there is a ubiquitous physical problem that requires:
  - More memory (i.e. resolution)
  - Faster compute speed
  - Inclusion in multiphysics simulations…

Understanding turbulence



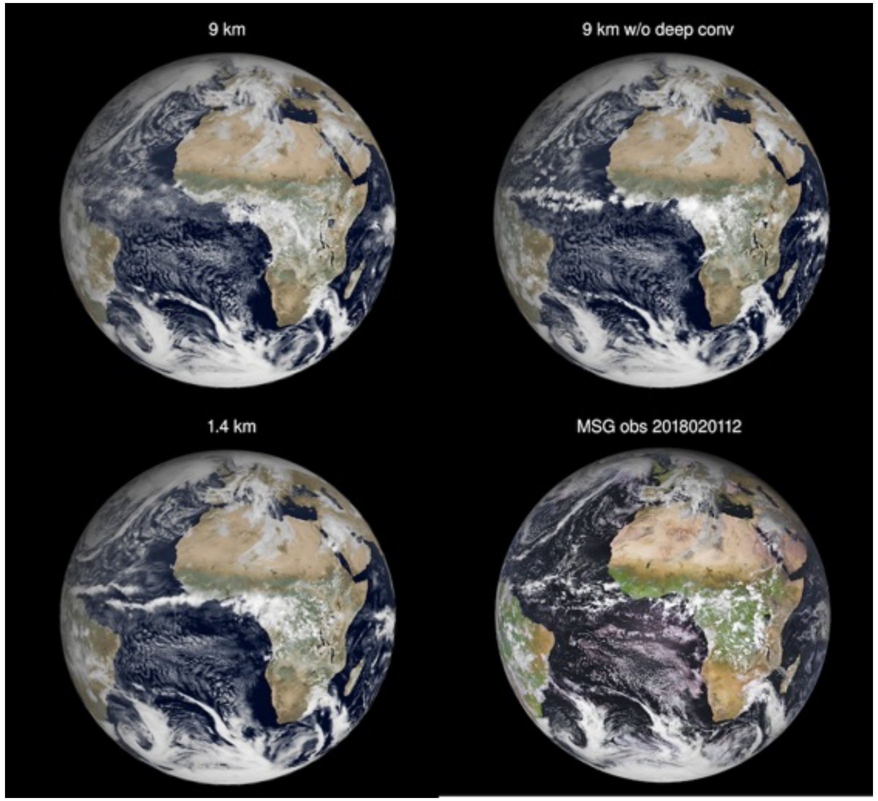
- "The last great unsolved problem in classical physics" (One of the 7 Millennium Problems)
- Werner Heisenberg assuredly never said: "When I meet God, I'm going to ask him two questions: why relativity? And why turbulence? I really believe he'll have an answer for the first."
- We remain far away from being able to resolve turbulent physics from the largest scales where it is generated (even in terrestrial settings) to the molecular dissipation scale.
- But, there are many places where turbulence is important where other physics arrests the impact of the turbulent cascade before it gets to the smallest scales.

Reaching New Heights in Weather Forecasting's Exascale Future

ECMWF and ORNL researchers use the power of Summit to simulate the Earth's atmosphere for a full season at 1-square-kilometer grid-spacing

The Science
Using Summit, a team of researchers from ECMWF and ORNL achieved a computational first: a global simulation of the Earth's atmosphere at a 1-square-kilometer average grid-spacing for a full 4-month season. Completed in June, the milestone marks a big improvement in resolution for the "European Model," which currently operates at 9-kilometer grid-spacing for routine weather forecast operations. It also serves as the first step in an effort to create multi-season atmospheric simulations at high resolution, pointing toward the future of weather forecasting—one powered by exascale supercomputers.

The Impact
The team has made the simulation's data available to the international science community. By eliminating some of the fundamental modelling assumptions prevalent in conventional simulations, the high-resolution data may help to improve model simulations at coarser resolutions.



These simulated satellite images of Earth show the improvement in resolution of the ECMWF Integrated Forecasting System from 9-kilometer grid-spacing with parametrized deep convection (top left), to 9-kilometer grid-spacing (top right), and 1-kilometer grid-spacing (bottom left). On the bottom right is a Meteosat Second Generation satellite image at the same verifying time. Image courtesy ECMWF.

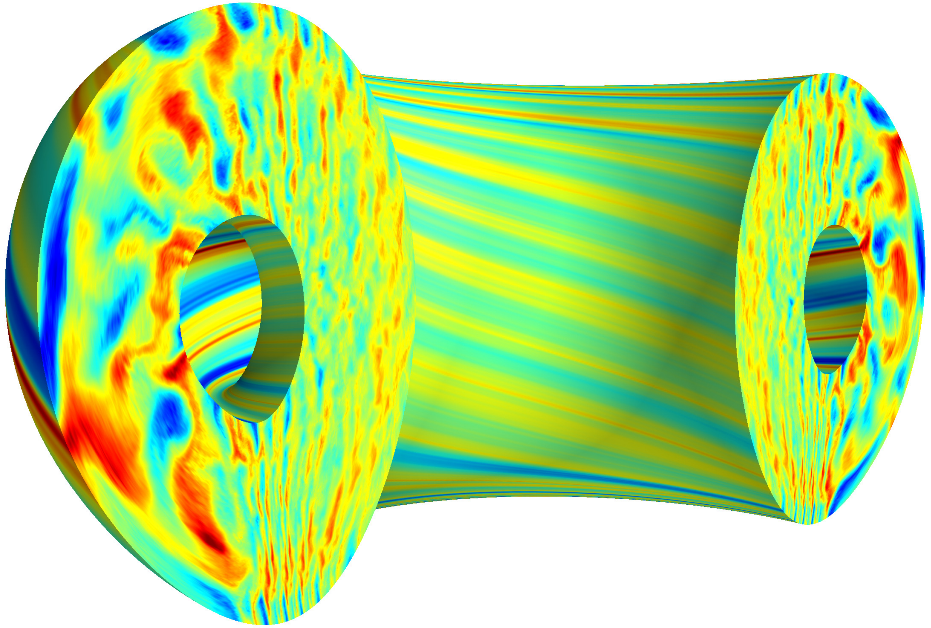
PI(s)/Facility Lead(s): Nils Wedi, ECMWF
ASCR Program/Facility: INCITE/OLCF
ASCR PM: Christine Chalk
Publication(s) related to this work: Wedi, N. P., et al. A baseline for global weather and climate simulations at 1 km resolution. *Journal of Advances in Modeling Earth Systems*, 12 (2020), e2020MS002192. doi: 10.1029/2020MS002192

Closing In on Fusion

A team modeled plasma turbulence on the nation's fastest supercomputer to better understand plasma behavior

The Science

The same process that fuels stars could one day be used to generate massive amounts of power here on Earth. Nuclear fusion—in which atomic nuclei fuse to form heavier nuclei and release energy in the process—promises to be a long-term, sustainable, and safe form of energy. But scientists are still trying to fine-tune the process of creating net fusion power. A team led by computational physicist Emily Belli of General Atomics has used Summit supercomputer at the Oak Ridge Leadership Computing Facility to simulate energy loss in fusion plasmas. The team used Summit to model plasma turbulence, the unsteady movement of plasma, in a nuclear fusion device called a tokamak. The team's simulations will help inform the design of next-generation tokamaks like ITER—the world's largest tokamak, which is being built in the south of France—with optimum confinement properties.



A visualization of deuterium-tritium density fluctuations in a tokamak driven by turbulence. Areas of red are representative of high density and areas of blue are representative of low density. Image Credit: Emily Belli, General Atomics

The Impact

Until now, almost all fusion simulations have only included only deuterium or tritium isotopes, but Summit enabled the team to include both as two separate species, model the full dimensions of the problem, and resolve it at different time and spatial scales. The results provided estimates for the particle and heat losses to be expected in future tokamaks and will help scientists and engineers understand how to achieve the best operating scenarios in real-life tokamaks.

PI(s)/Facility Lead(s): Emily Belli
ASCR Program/Facility: ALCC and INCITE / OLCF
ASCR PM: Christine Chalk
Publication(s) for this work: Emily A. Belli and Jeff Candy, "Asymmetry between Deuterium and Tritium Turbulent Particle Flows," *Physics of Plasmas* 28, no. 6 (2021), doi:10.1063/5.0048620.

GE Spins up Supercomputer Models to Zero in on Energy Loss in Turbines

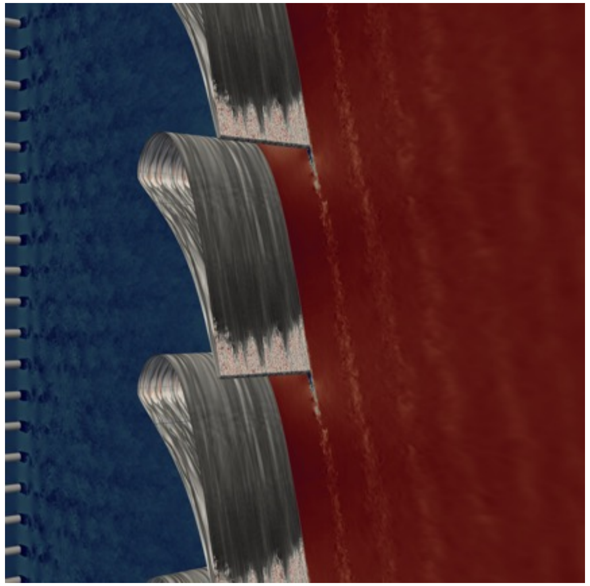
A team at GE Aviation and the University of Melbourne is studying turbulent flows on the Summit supercomputer for better engines

The Science

High-pressure turbines are vital components of gas turbines used to propel jet engines. The more efficient these jet engines are, the better they are for the aircraft industry and their customers. But these large, dynamic systems are difficult to study via experiments and physical testing. A team led by scientists at General Electric (GE) Aviation and the University of Melbourne used the Summit supercomputer to run for the first time real-engine cases capturing the largest eddies, or circular fluid movements, down to those that were tens of microns away from the turbine blade surface. From the simulations, the researchers determined which regions near a turbine blade experience a greater loss of energy. For the case with the highest Mach number, which describes the flow's velocity compared with the speed of sound, they discovered an extra loss of energy resulting from strong shock waves, or violent changes in pressure, that interact with the edge and wake of the flow to cause a massive amount of turbulence.

The Impact

More accurate prediction of real-engine conditions will lead to more efficient engines that consume less fuel and other positive derivative effects. A 1 percent reduction in fuel consumption across a fleet of engines is equal to about 1 billion dollars a year in fuel cost savings. Reduced fuel consumption also translates into reduced emissions—a 1 percent reduction in fuel burn reduces CO_2 emissions by roughly 1.5 percent.



A row of upstream bars produces highly turbulent flow that gets accelerated through a high-pressure turbine blade row and interacts with the blade surface, causing significant temperature variations. Image Credit: Richard Sandberg, University of Melbourne

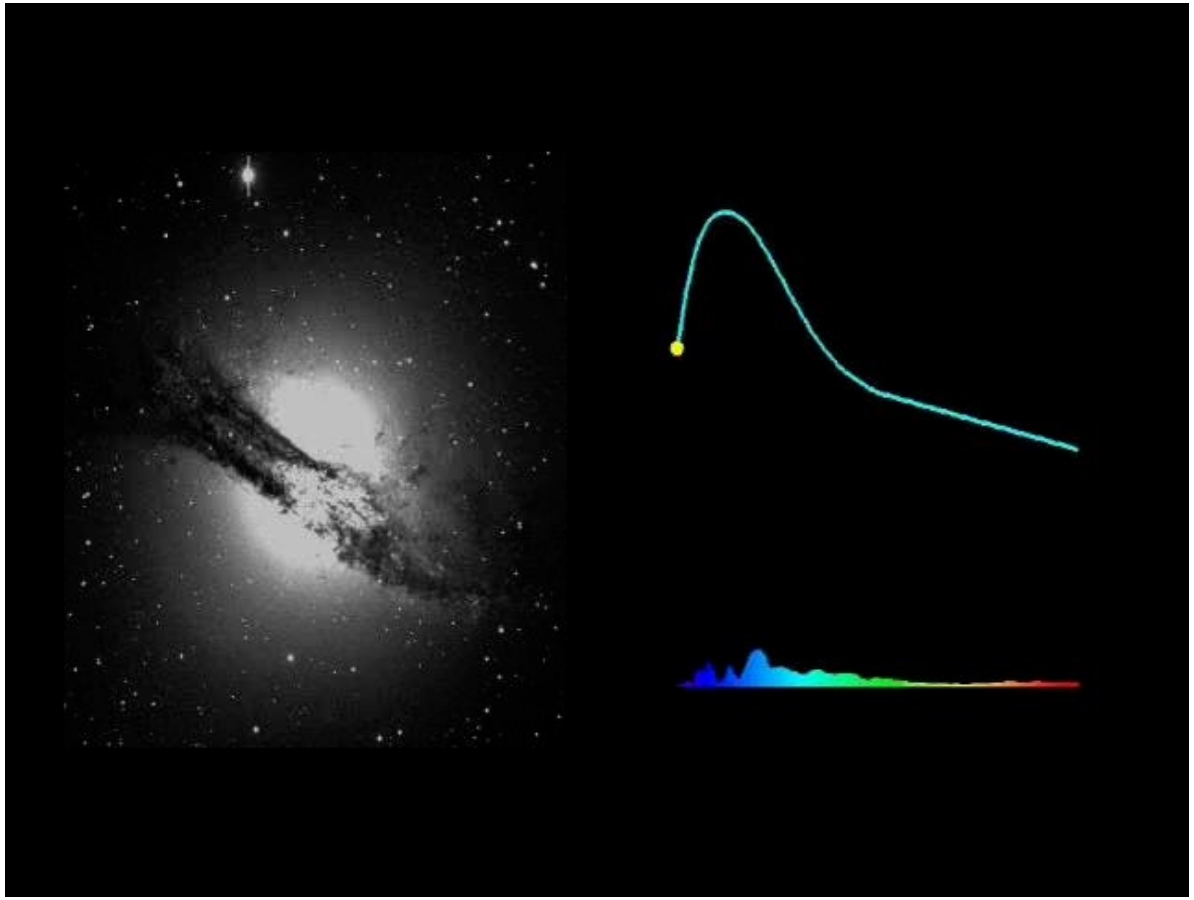
PI(s)/Facility Lead(s): Richard Sandberg, Univ. Of Melbourne; Sriram Shankaran, GE Aviation
ASCR Program/Facility: INCITE/OLCF
ASCR PM: Christine Chalk
Publication(s) for this work: Y. Zhao and R. D. Sandberg, "High-Fidelity Simulations of a High-Pressure Turbine Vane," *Journal of Turbomachinery* 143, no. 9 (2021).

Y. Zhao and R. D. Sandberg, "Using a New Entropy Loss Analysis," *Journal of Turbomachinery* 142, no. 8 (2020): 081008

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY
THE UNIVERSITY OF MELBOURNE
GE Aviation
U.S. DEPARTMENT OF ENERGY INCITE LEADERSHIP COMPUTING

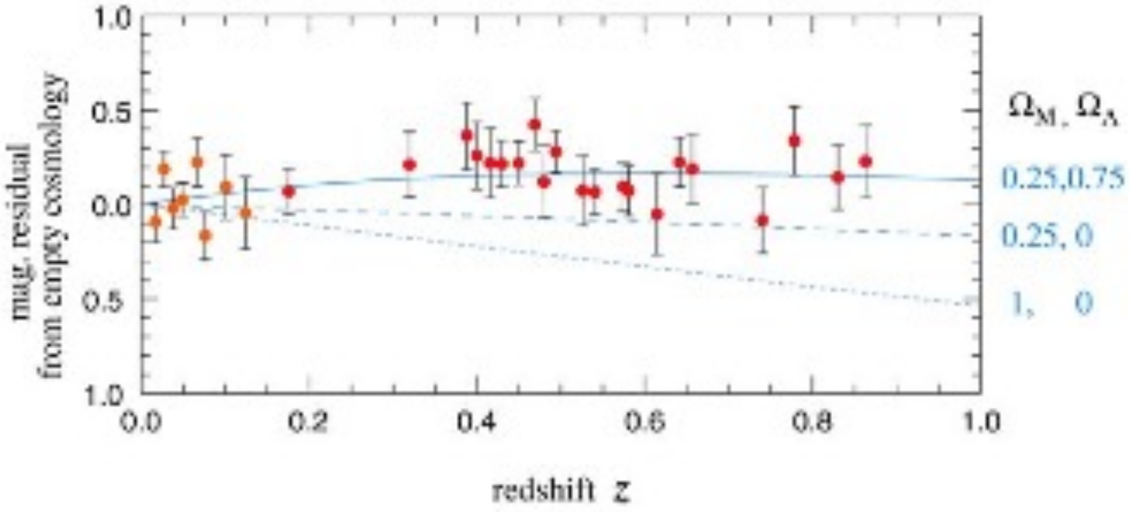
Type Ia supernovae

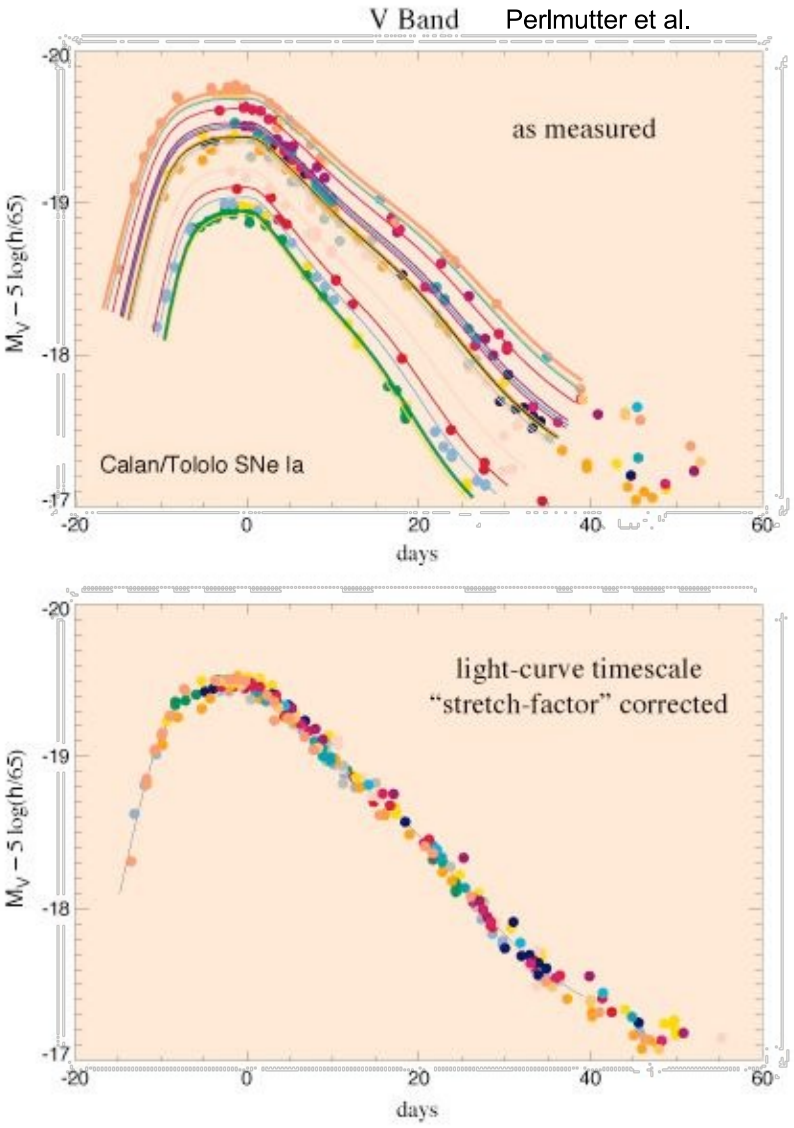
- Brightness rivals that of the host galaxy (L ~ 10^{43} erg/s)
- Larger amounts of radioactive ^{56}Ni produced than in CCSNe
- Radioactivity powers the light curve ("Arnett's Law")
- Not associated with star-forming regions (unlike CCSNe)
- No compact remnant - star is completely disrupted
- Likely event – the accretion-induced **thermonuclear explosion** of a white dwarf (Pankey 1962)



Type Ia supernova cosmology

- SNe Ia are ‘standardizable’ candles
  - Robust lightcurve - variations can be corrected with a single-parameter function (Phillips relation)
- Distant Ia’s appear dimmer than expected in a Universe without a ‘dark energy’ component.



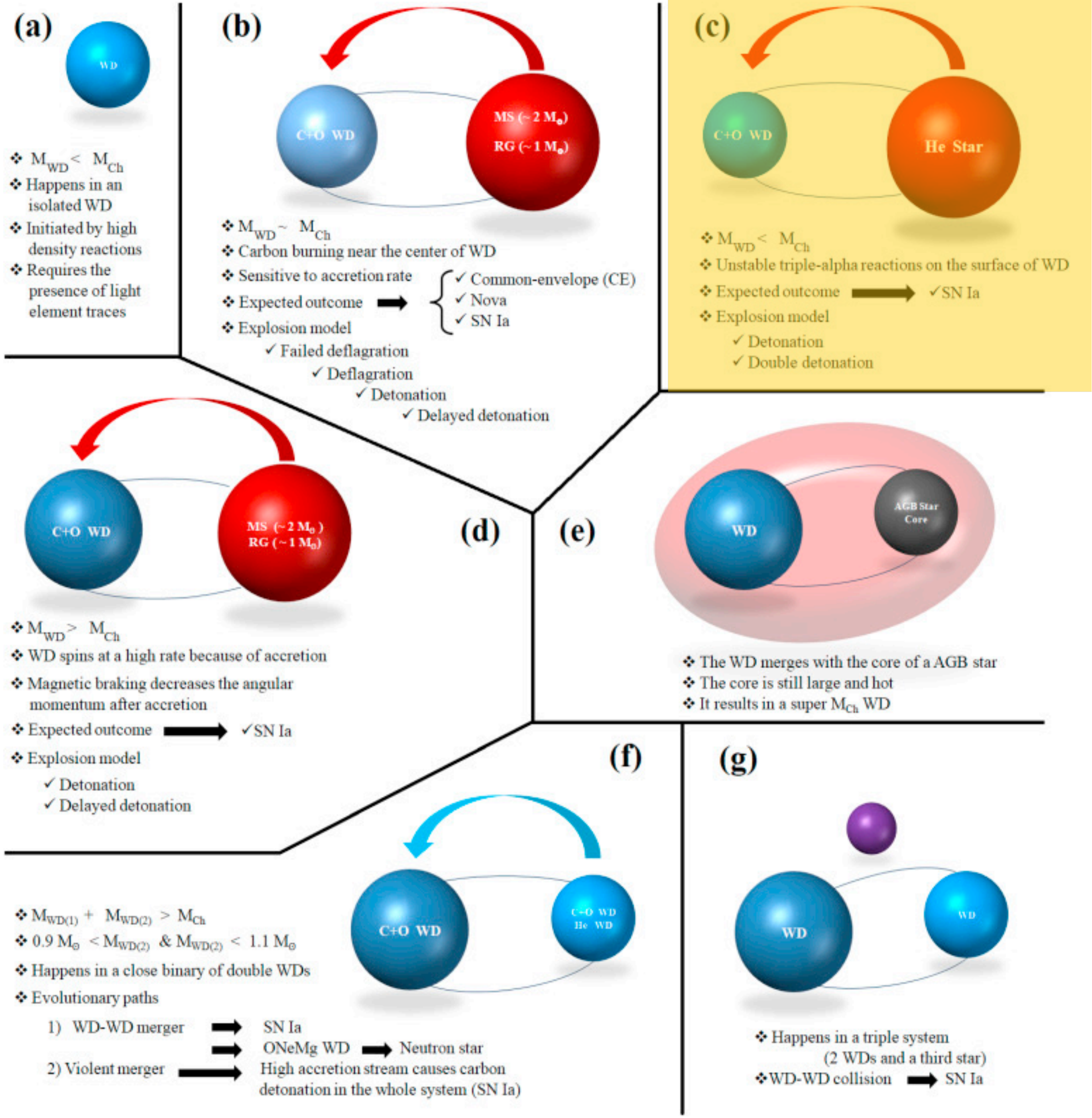


2011 Nobel Prize (Perlmutter, Schmidt, & Riess)

Several possible scenarios

- One discriminator: progenitor
  - Chandrasekhar-mass WD
  - sub-Chandrasekhar mass WD

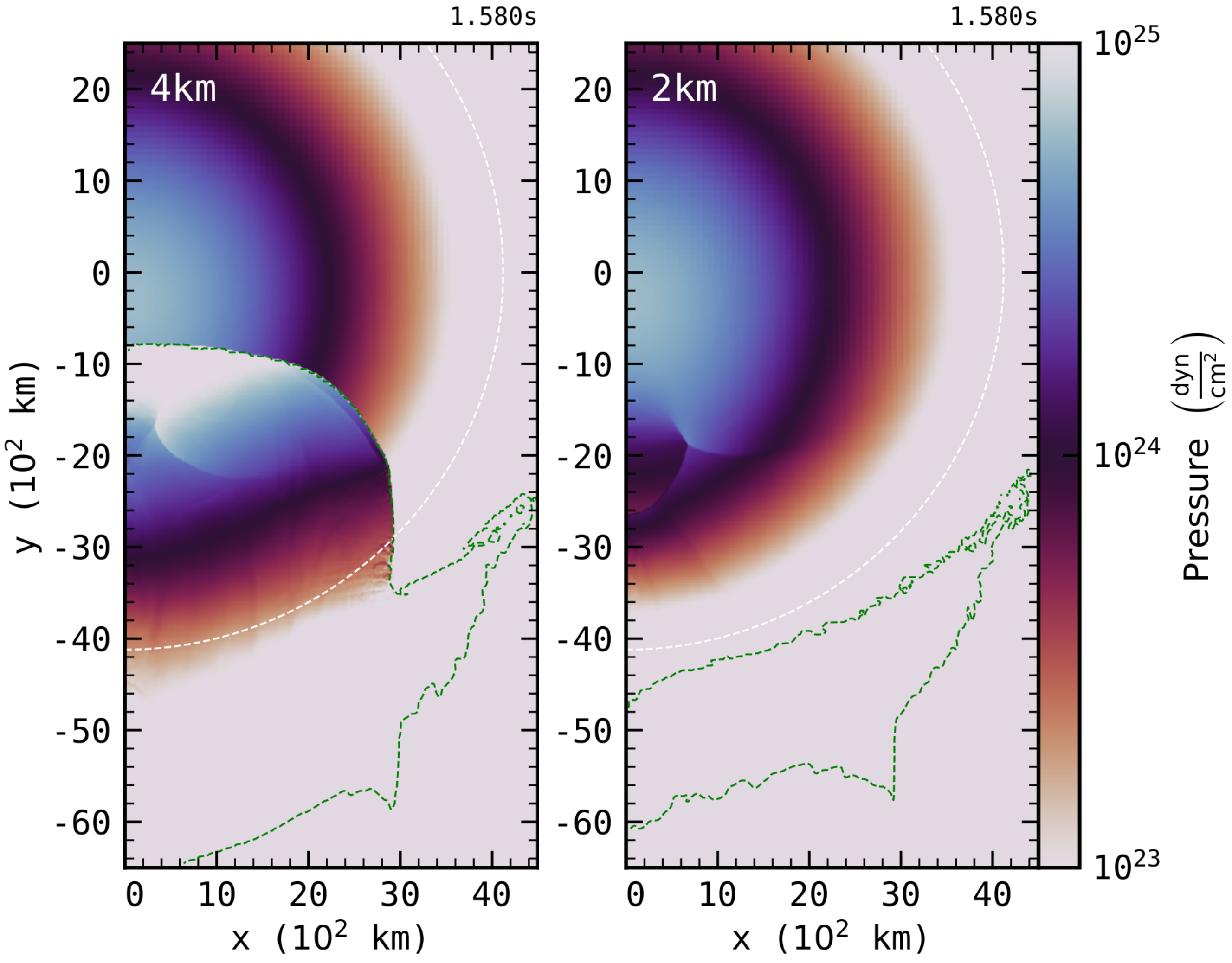
- Scenario C – “double detonation”) has increasing observational support (e.g. Shen & Moore 2014)



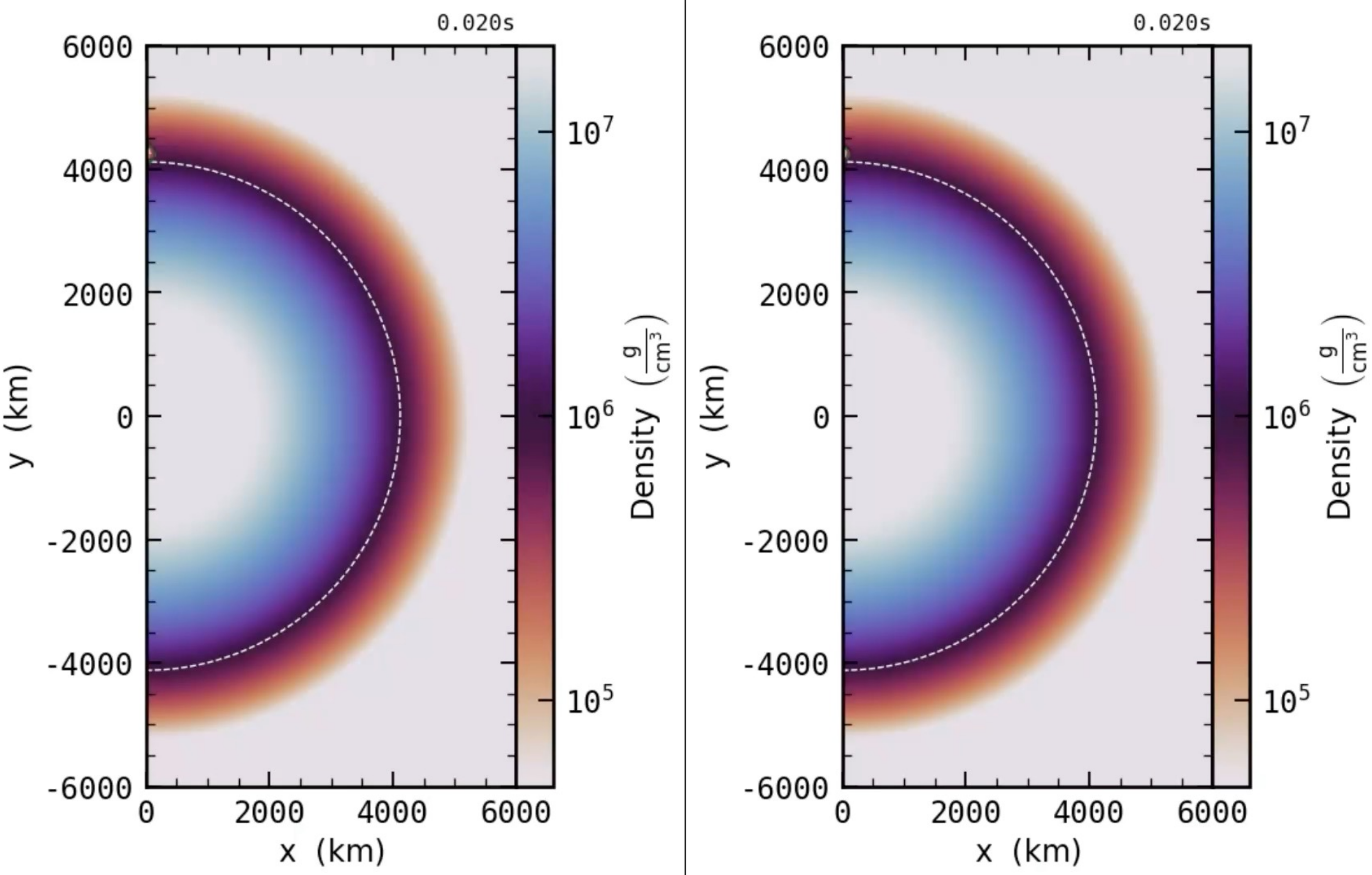
Nouri+ (2019)

Double-detonation models (Rivas+ 2022)

- SD, sub-M_{Ch} scenario: detonation in accreted He layer leads to compressional, off-center carbon detonation
- Details of outcome depend on resolution
- Large-network technology (feasible on GPUs) and adaptive mesh refinement (ECP Flash-X) enable these and future simulations



Rivas, Harris, Hix, & Messer 2022 *ApJ* **937** 2


0.020s
6000
4000
2000
0
-2000
-4000
-6000
y (km)
0
2000
4000
6000
x (km)
10^7
10^6
10^5
Density (g/cm^3)
0.020s
6000
4000
2000
0
-2000
-4000
-6000
y (km)
0
2000
4000
6000
x (km)
10^7
10^6
10^5
Density (g/cm^3)

Turbulent mixing is the source of the difference: Sets a minimum necessary resolution to obtain reliable results

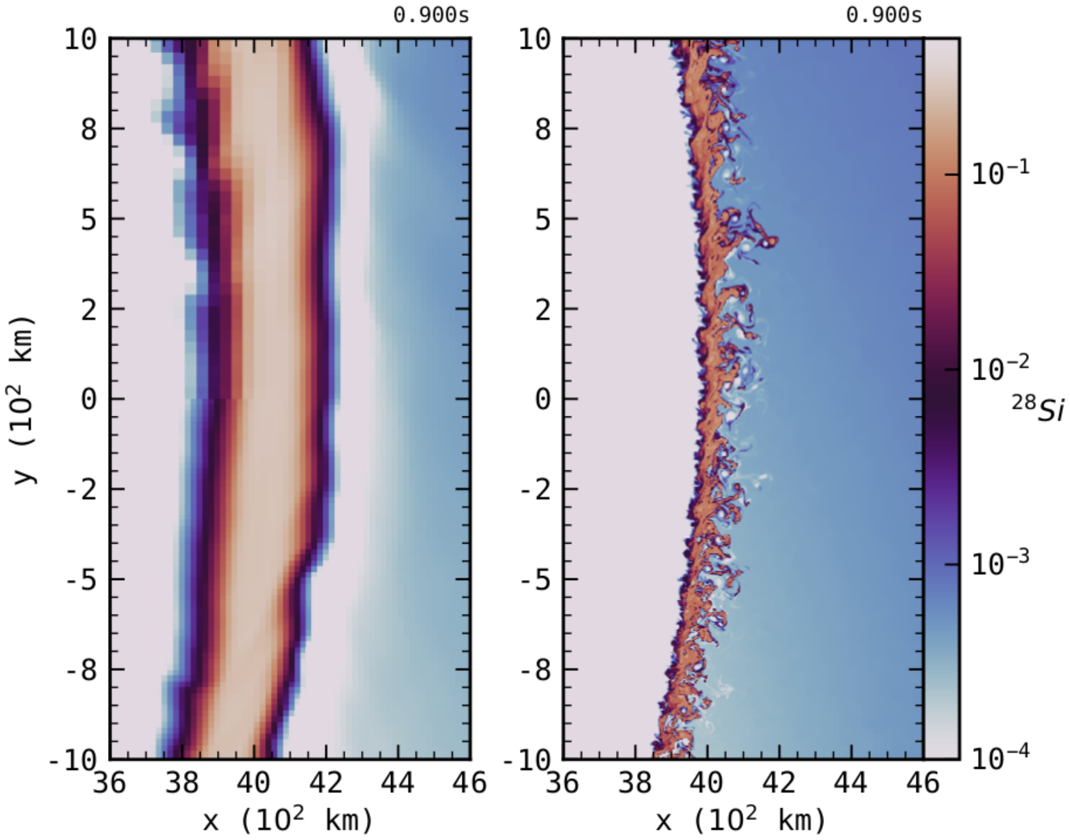
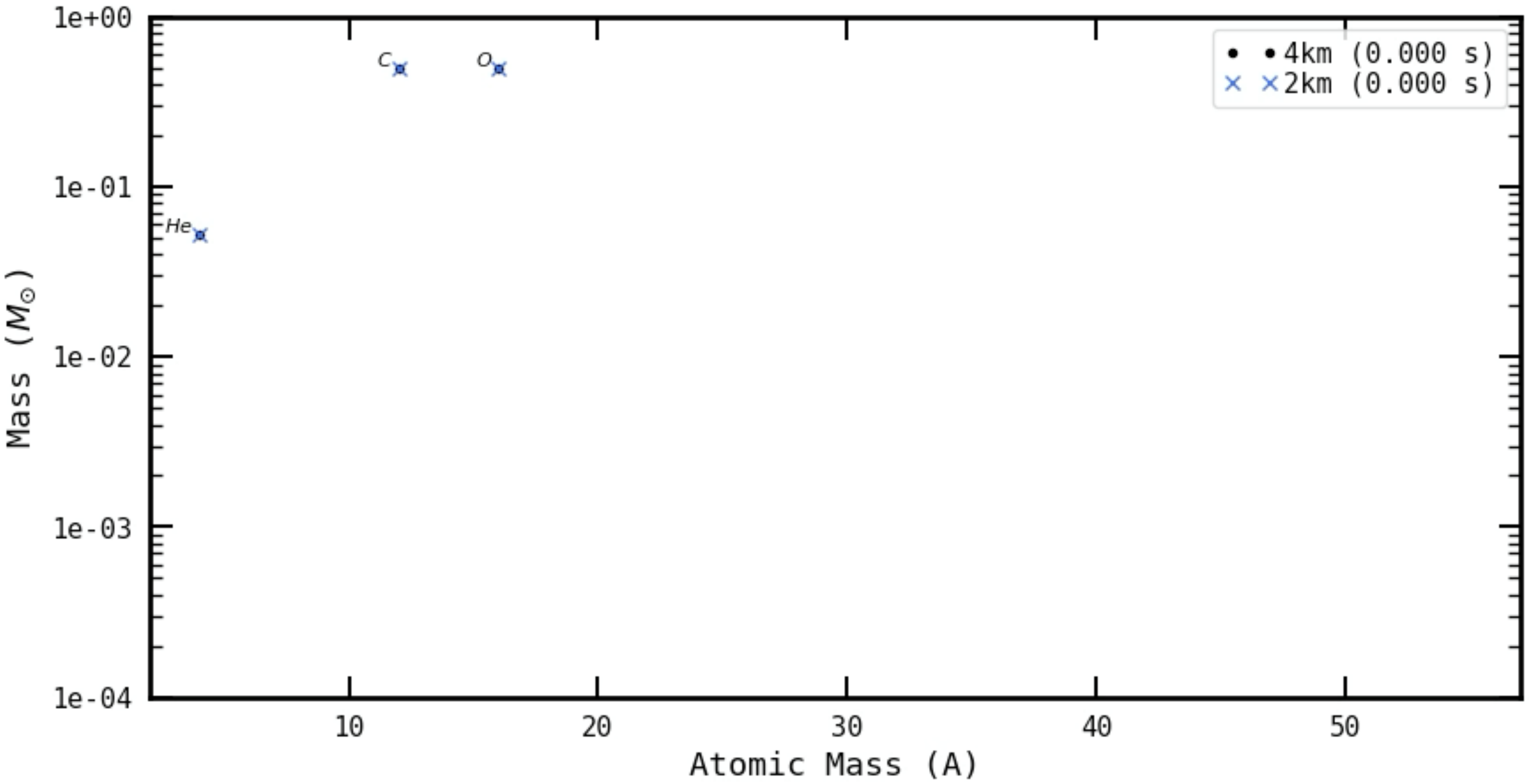


Figure 4. Outer shell burning remnants: equatorial slice of mass fraction of silicon-28 at the WD limb after shell burning traverses the whole cutout. At 16 km resolution (left), the thickness of the expanding band is 4 times larger than in the highest resolution case (0.5 km, right). Additionally, mixing is far more complex and evolved even at scales which do not represent the actual length scales of burning.

Difference in evolution leads to difference in yields



Significant differences in intermediate mass yields and velocity distributions

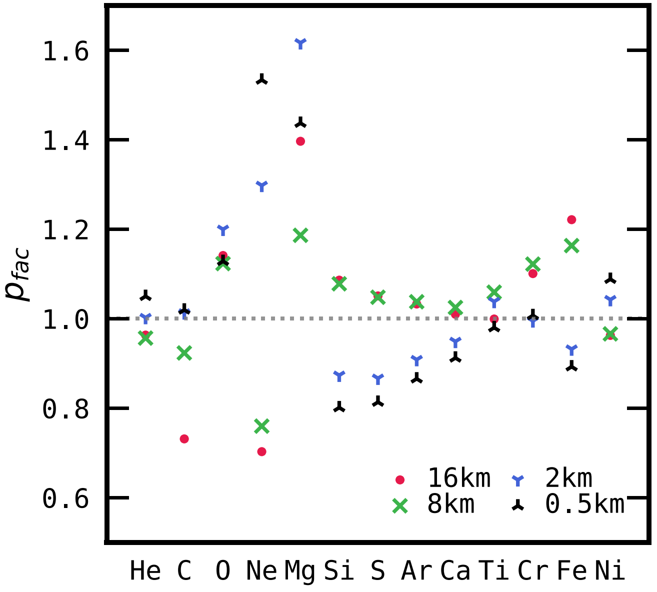
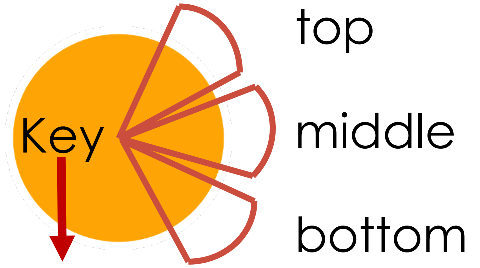
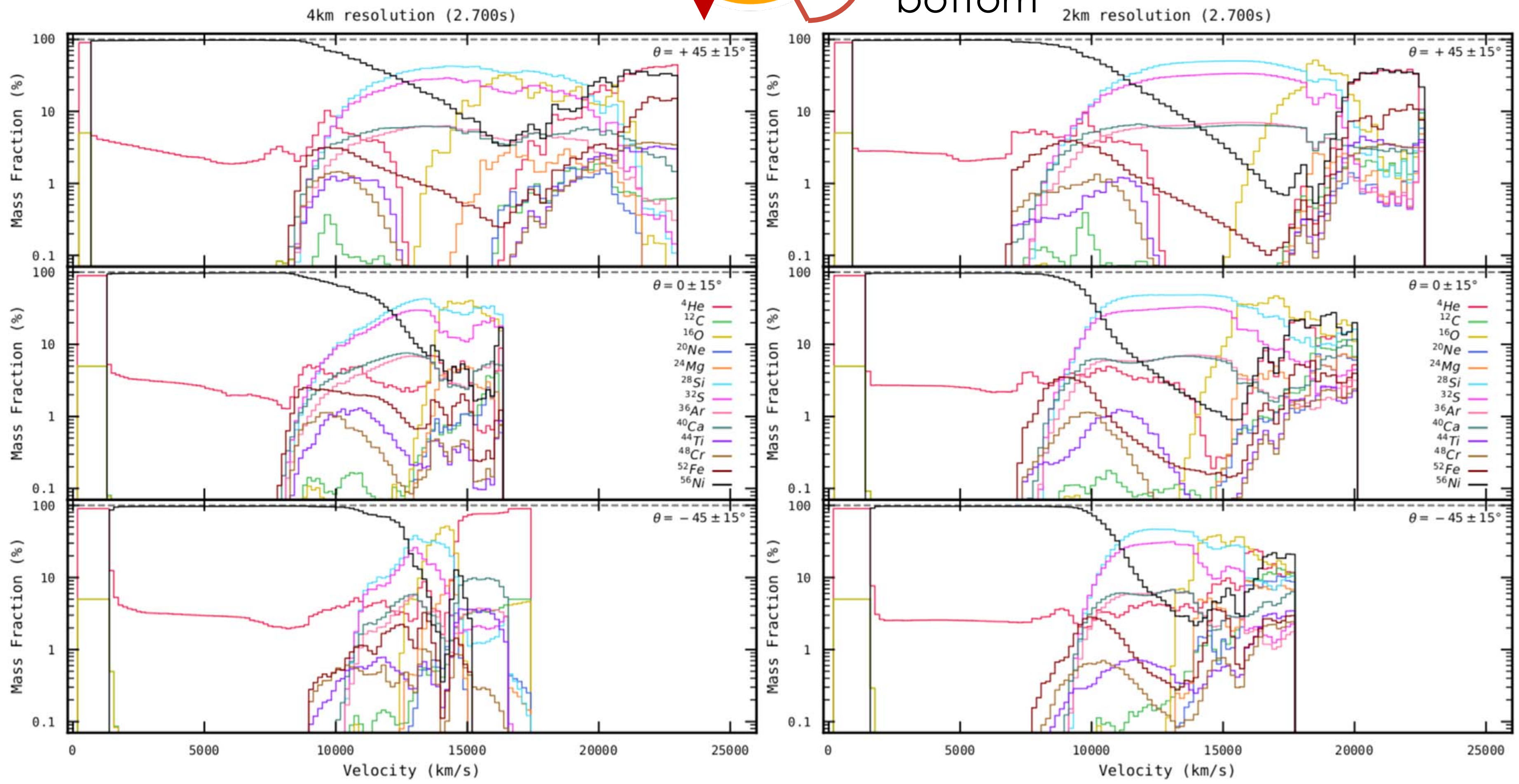


Figure 8. Relative yields for simulations compared to the 4 km case.





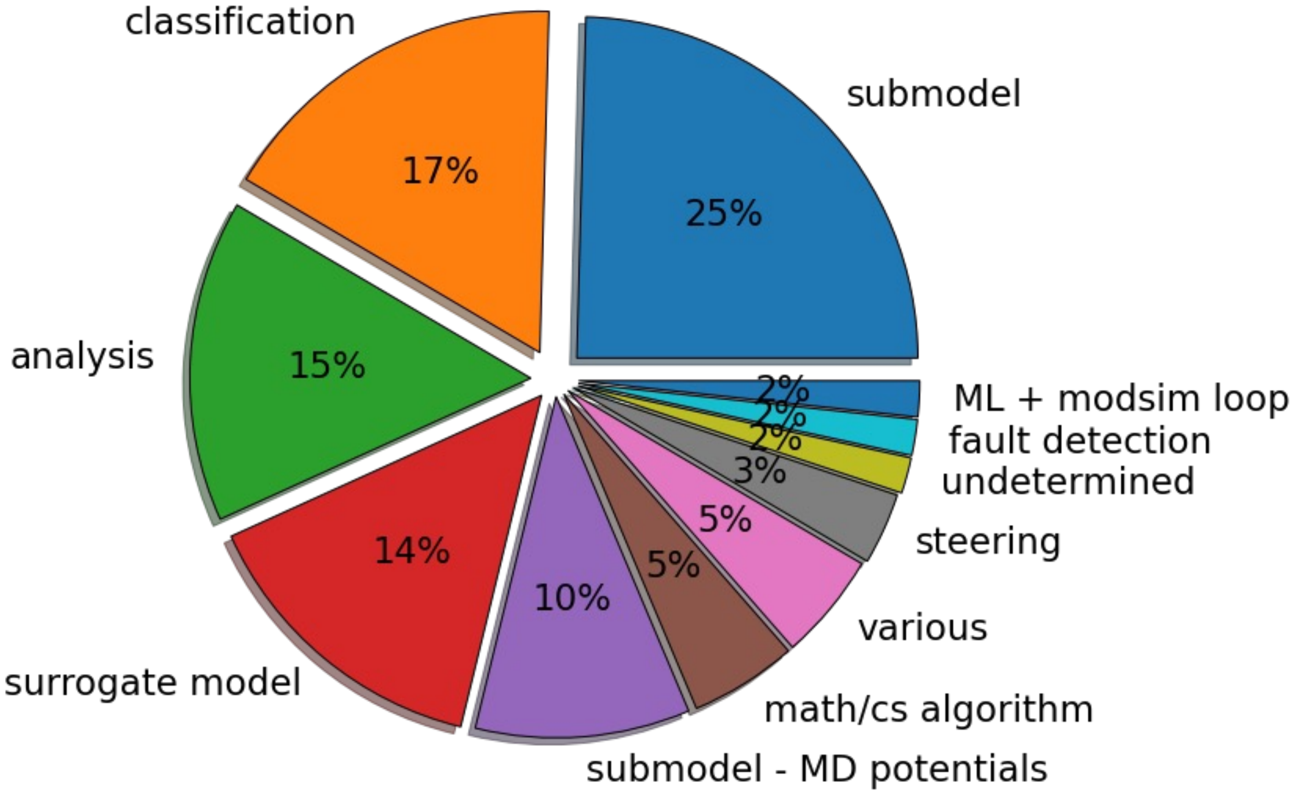
Conclusions

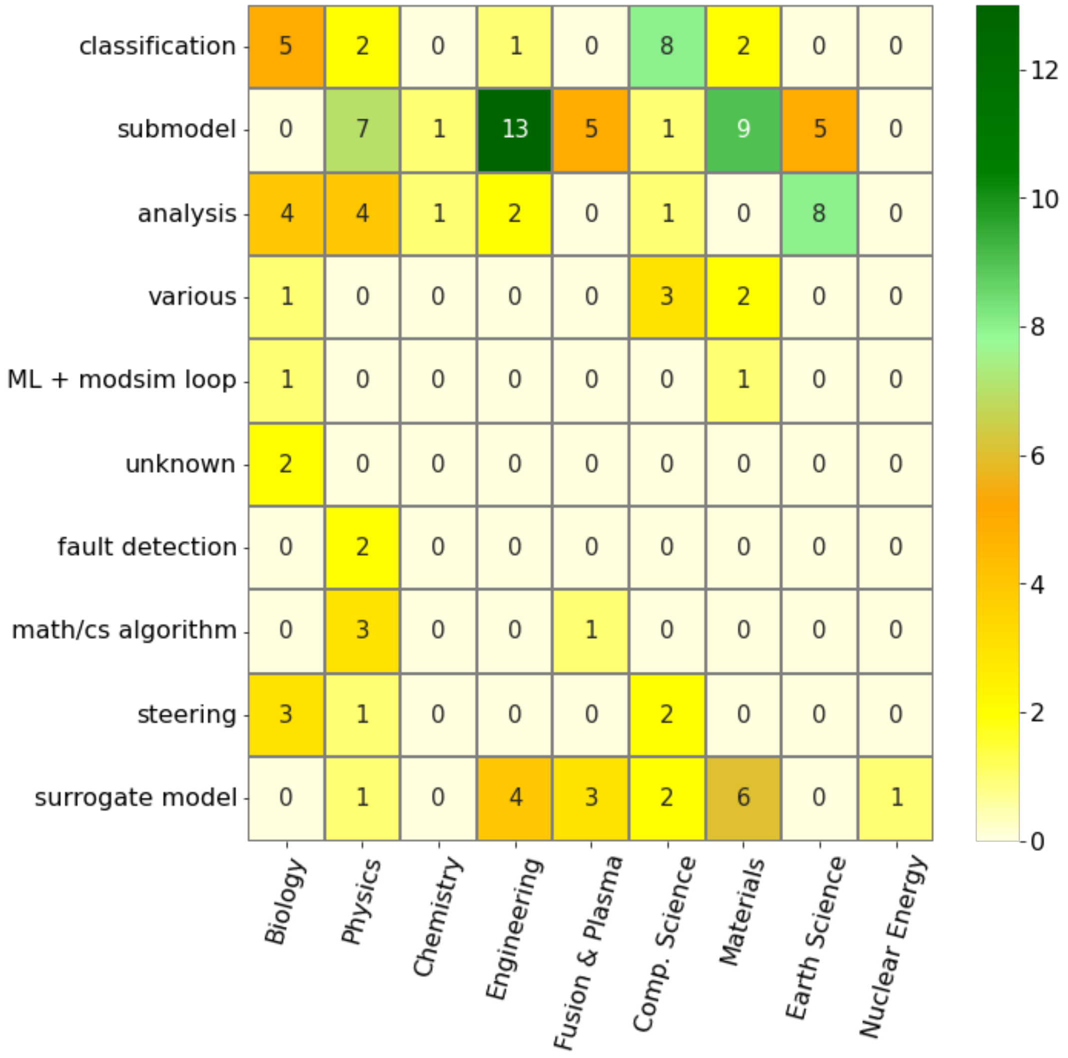
- Building Frontier was an adventure made all the more “interesting” by the pandemic.
- Leadership-scale supercomputers are unique scientific instruments, like JWST and LHC, but have catholic applicability in science.
- Turbulence might be one physical phenomenon where exascale computing is particularly suited to advance our understanding.
- Simulations of double-detonation Type Ia SNe require high resolution to obtain qualitatively meaningful results because of turbulence. They likely require other pieces of higher-fidelity physics to robustly confront observations.

A note about the elephant in the room…

- Hyperscalers are spending far more than the budget for OLCF on machine learning. The ”unreasonable success” of ML has generated a lot of excitement.
- Machine learning is not a competitor for simulation: It is already becoming widespread in many computational workflows.
- I predict (editorial, YMMV) that ML will continue to grow in usage in scientific computing, but it will not “replace” simulation in any sense. The most impressive uses will be in the design of experiments, in data analysis, and in classification of *simulation* results.

AI/ML are found throughout the OLCF workload today





"Learning to Scale the Summit: AI for Science on a Leadership Supercomputer." Wayne Joubert, Bronson Messer, Philip C. Roth, Antigoni Georgiadou, Justin Lietz, Markus Eisenbach and Junqi Yin, Accepted, ExSAIS 2022: Workshop on Extreme Scaling of AI for Science (IPDPS 2022)

Questions?


FRØNTIER
OAK RIDGE
National Laboratory
U.S. DEPARTMENT OF
ENERGY
DEPARTMENT OF ENERGY
UNITED STATES OF AMERICA

Acting on the data?

nature astronomy

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature › nature astronomy › comment › article

Comment | Published: 10 September 2020

The ecological impact of high-performance computing in astrophysics

Simon Portegies Zwart

Nature Astronomy **4**, 819–822 (2020) | Cite this article

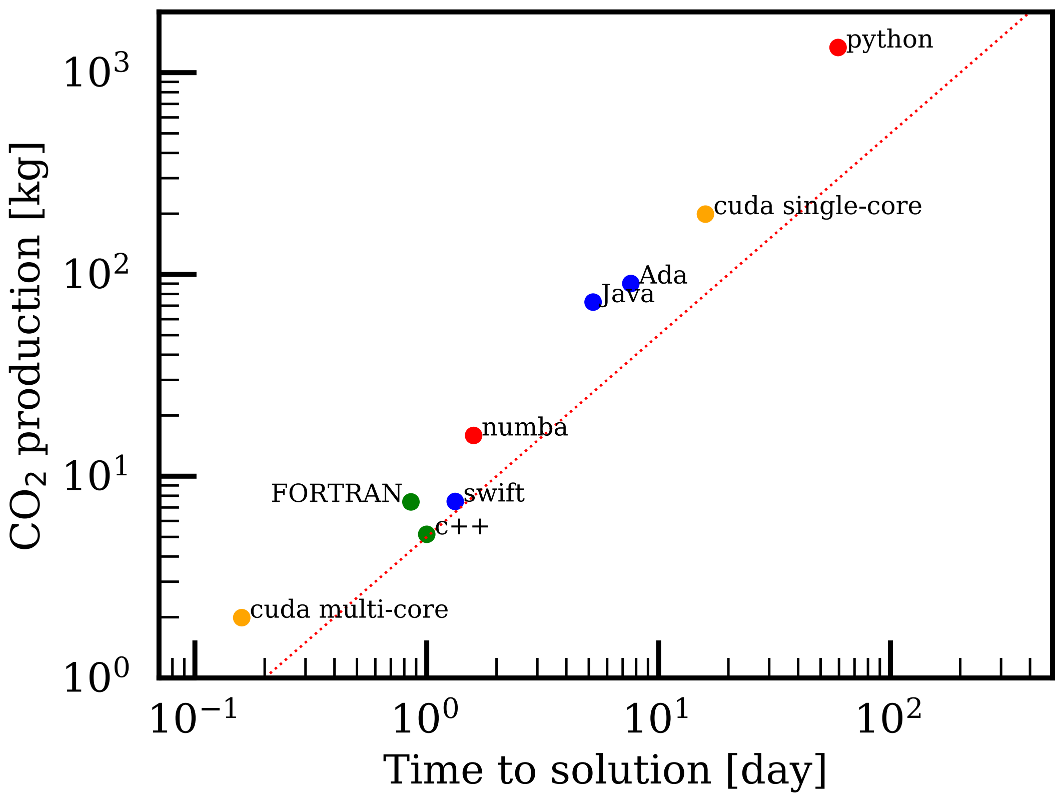


Figure 3: **Here we used the direct N-body code from [23] to measure execution speed and the relative energy efficiency for each programming language from table 3 of [22]. The dotted red curve gives a linear relation between the time-to-solution and carbon footprint (~ 5 kg CO_2/day). The calculations were performed on a 2.7GHz Intel Xeon E-2176M CPU and NVIDIA Tesla P100 GPU.**