Computational Science and High Performance Computing at Lawrence Livermore National Laboratory

Teresa S. Bailey

LLNL-PRES-2002095

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

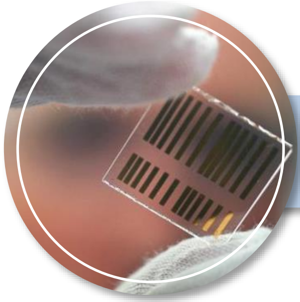
The Department of Energy

- Transform the nation’s energy system and secure U.S. leadership in energy technologies
- Maintain U.S. effort in science and engineering as a corner-stone of economic prosperity
- Enhance nuclear security and nuclear deterrence through defense, non-proliferation, and environmental efforts
- Carry out environmental clean-up from Cold War nuclear mission
- Manage the 17 National Labs

DOE’s Core Missions

Energy

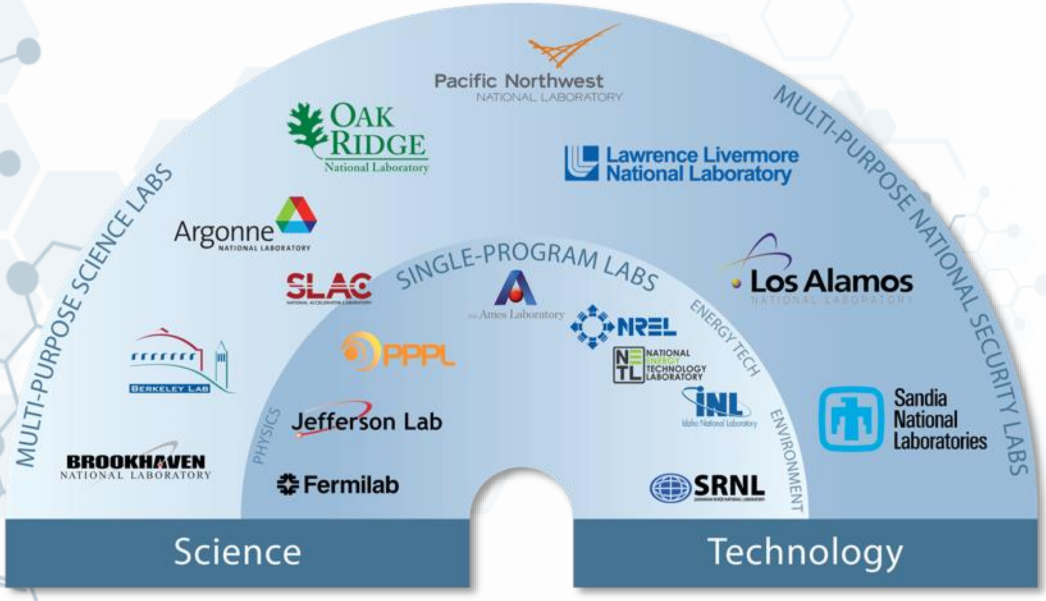
Nuclear Safety & Security



Science & Innovation

Management & Ops Excellence

LLNL is one of 17 Department of Energy national laboratories



- Major scientific user facilities; >77,000 employees
- Largest funder of basic research in the physical sciences and of energy research
- 16 of 17 are FFRDCs* and "GOCOs*"
- Extensive university and industry partnerships, and student programs

FFRDC = Federally Funded Research and Development Center
"GOCO" = Government Owned, Contractor Operated

The National Nuclear Security Administration

NNSA
National Nuclear Security Administration

- Established by Congress in 2000 as a semi-autonomous agency within the DOE
- Responsible for enhancing our national security through the military application of nuclear science
- Manages the three NNSA National Laboratories and the production complex

Maintaining the Stockpile

Counterterrorism

Nonproliferation

Powering the Nuclear Navy

Reduce global danger from WMD

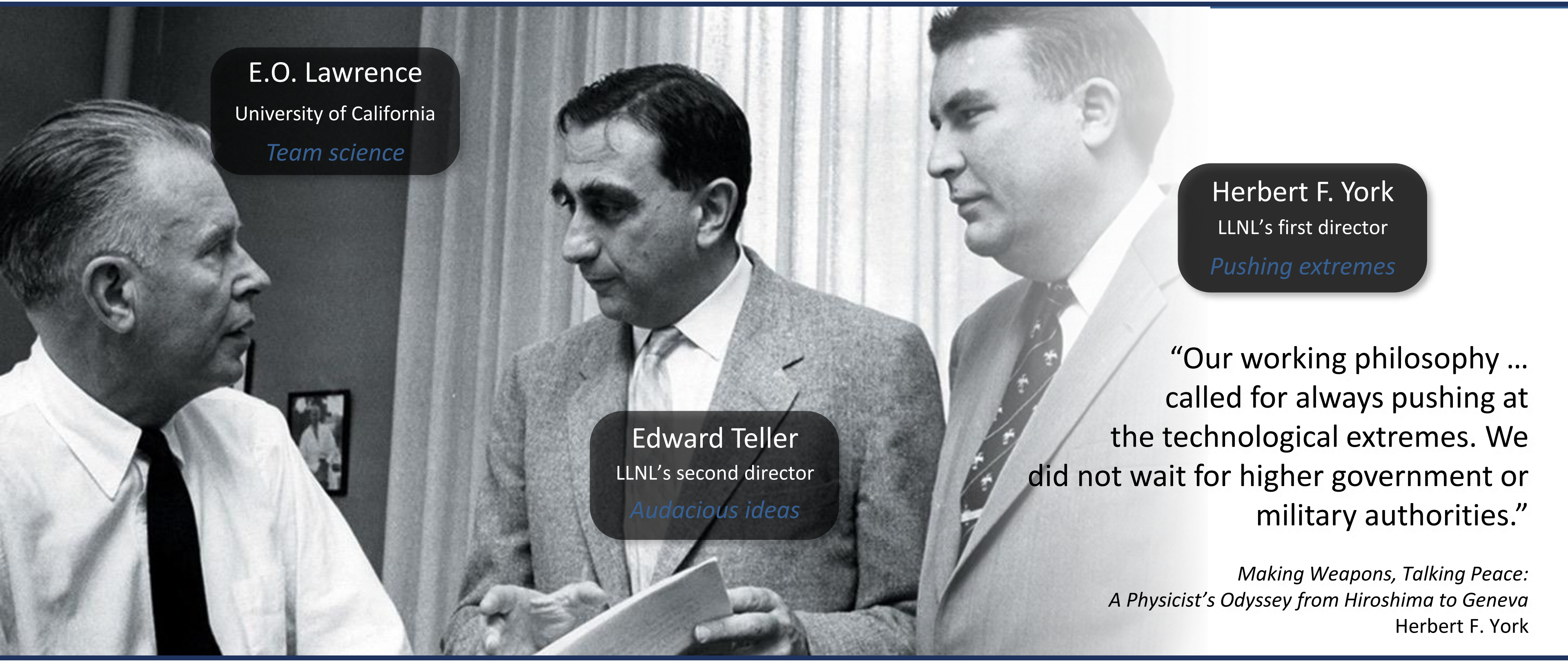
Supporting U.S. leadership in S&T

- LLNL established in 1952
- ~9,300 LLNS employees
- 1 square mile
  and 521 facilities
- Annual budget: ~$3.3B
- Operated by LLNS, LLC
  for U.S. Dept. of Energy

(LLNS: University of California,
Bechtel, BWXT, and Amentum)

Experimental Test Site
(11 square miles near Tracy, CA)

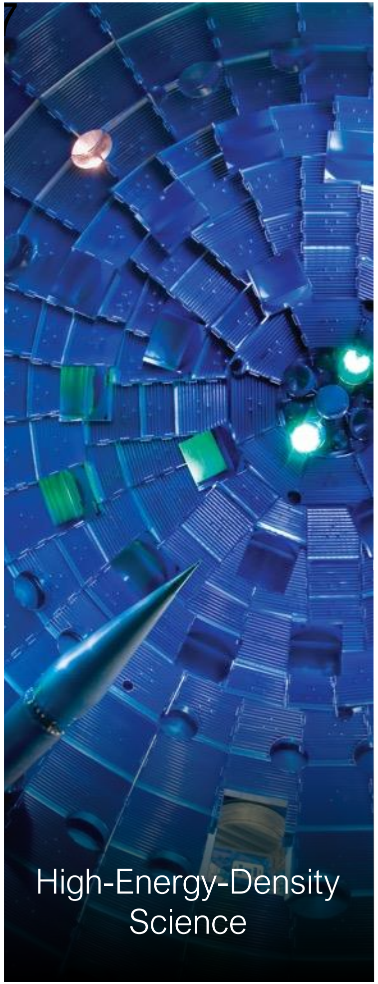
Our heritage: team science, audacious ideas, and pushing the extremes



"Our working philosophy … called for always pushing at the technological extremes. We did not wait for higher government or military authorities."

Making Weapons, Talking Peace: A Physicist's Odyssey from Hiroshima to Geneva
Herbert F. York

LLNL Core Competencies



Our approach is rooted in challenging predictions with experiments

High-Performance Computing

Livermore Computing

- 11 systems on Top500 list
- A total of ~218 petaflops + El Capitan (~2 exaflops)

Advanced Experimental Capabilities

National Ignition Facility

- World's most energetic laser
- Fusion ignition achieved Dec. '22
- Laser Energy increased to 2.2 MJ

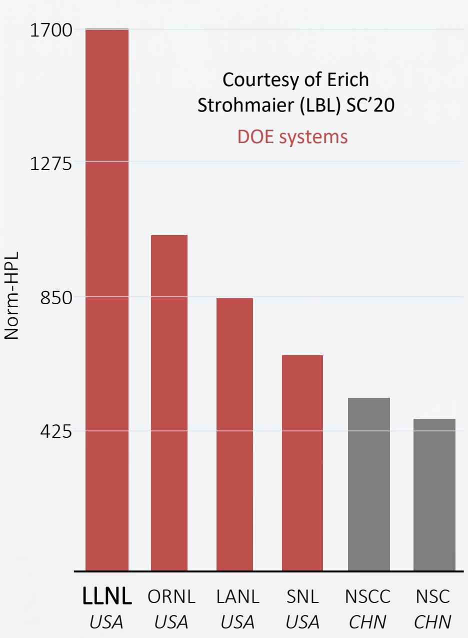
Accelerated Materials and Manufacturing

Advanced Manufacturing Lab

- Material development
- Systems design
- Specialized manufacturing

LLNL has defined the “bleeding edge” of high-performance computing

Top Computing Institutions
(past 25 years)



Sierra

- 125 petaflops (peak)
- On-line in 2018

El Capitan

- NNSA’s 1st exascale supercomputer
- Silicon based
- 1.74 exaflops (peak measured)
- Operational in 2024

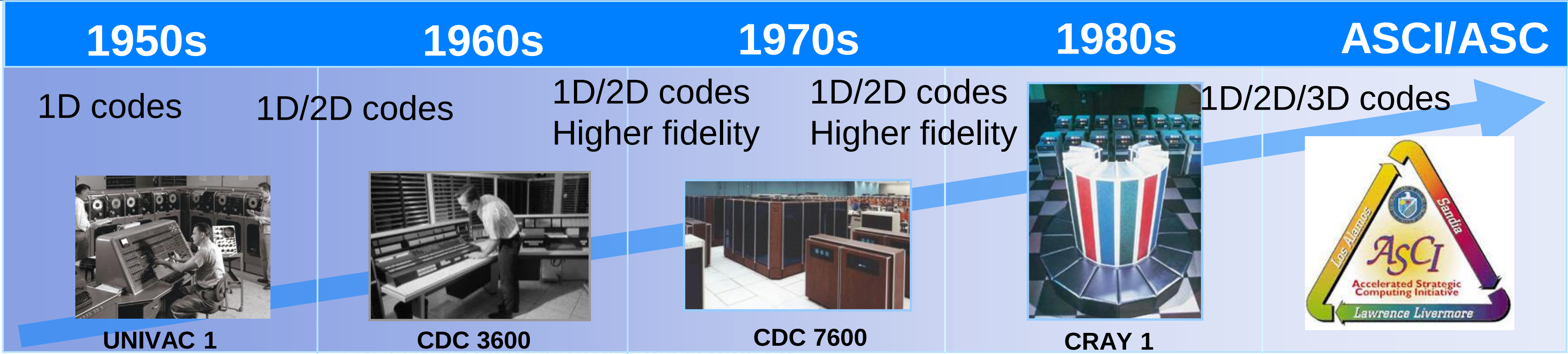
Beyond Moore’s Law

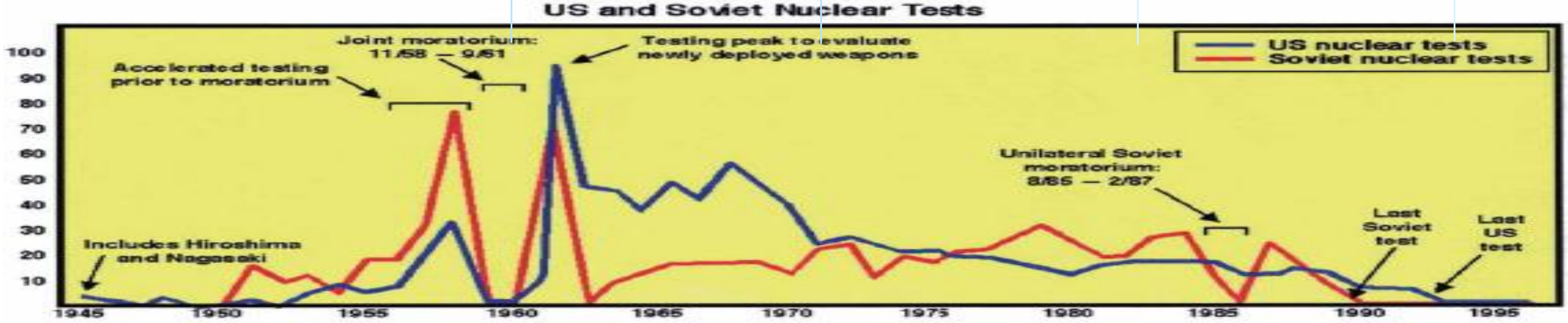
Neuromorphic

Quantum

Our High Performance Computing History

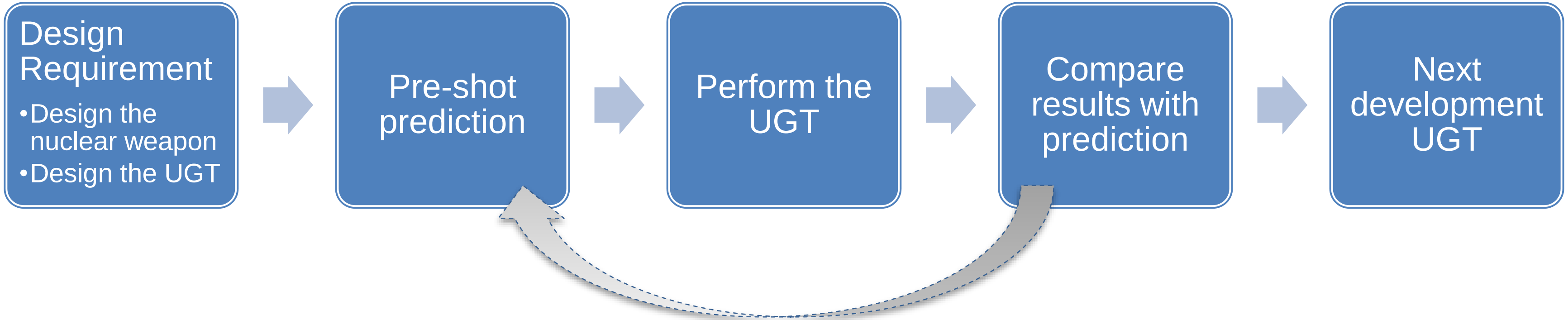
High Performance Computing and code development directly supported design efforts during the Cold War





Success in this era provided confidence that ASCI vision could be realized

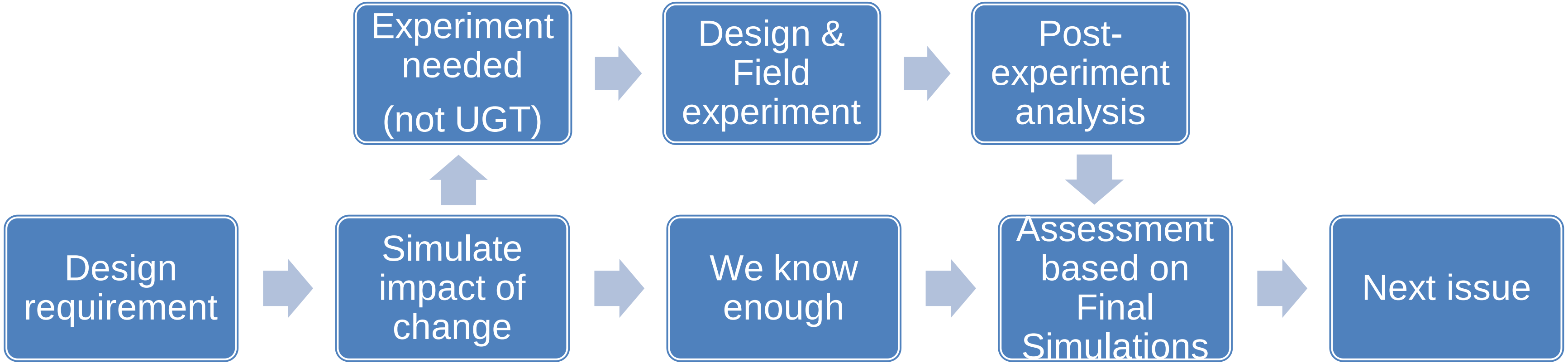
Historic LLNL code development responded directly to designer needs to develop the modern stockpile



- Code development was supported throughout the testing era, but the focus was on testing
  - Code developers and designers worked closely together in the same organization
- Codes were valued because they helped designers design better systems, faster
  - Insight into the physical processes in the weapons is a big part of the value codes provide
  - A few key surprises resulted in improvements to codes and to the modeling choices being made
  - New physics in codes enabled design flexibility
- Code teams were small, but were multi-disciplinary
  - Physicists, Mathematicians, and Computer Scientists were part of the code teams
  - Projects would focus on one part of the problem and link results together

LLNL's multi-disciplinary approach to code development is rooted in our history

Today’s workflow is fundamentally different



- In the absence of nuclear testing, multi-physics codes have become much more integral in assessing stockpile questions
- Continuously advancing predictive science is key. This enables us to say “we know enough” more often.
- We pursue multiple avenues to advance predictive science
  - Improved numerical methods
  - Improved material models
  - Improved mathematical models
  - Advanced computer architectures enabling higher-fidelity simulations
  - Clever experiments

Development during the ASC program has been an exciting time

	1995-2001	2002-2008	2009-2015	2016-2024	Today
Users	Legacy Codes	Adoption of ASC codes	Reliance on ASC codes	Reliance on ASC codes	Next-Gen ASC codes
Code Advances and Highlights	Original ASCI milestones demonstrate: • Physics capability • 3D capability • Massively parallel capability Invest in research	2D focus • Enables fast turnaround • Enables studying new physics 3D for special cases Invest in research	2D ensembles for UQ • ~80,000 2D • ~100,000 1D 3D in everyday use • Demonstrated on Sequoia Invest in research	Sierra enables same day turnaround for 3D • Can be used as a standard design tool • 3D ensembles for UQ are possible Invest in research	Widespread use of 3D for more applications Better numerical methods Machine Learning and AI New mathematical models? Invest in research
Code Teams	Multi-disciplinary team started	Teams established, building skill in CS and Physics areas	Teams established, building skill in CS and Physics areas	Modular Physics and CS teams established, Agile teams developed	ML and AI experts become part of teams, Agile teams established
Code Infrastructure	Exploring parallel algorithms	Parallel algorithms mature, physics modularity explored	Physics modularity adopted, development of proxy apps	CS abstraction and modularity, reliance on proxy apps	CS abstraction standardized, proxy apps become "predictive"
Hardware	 ASCI Blue Pacific	 BG/L  ASCI Purple	 BG/P  Sequoia BG/Q	 Sierra	 El Capitan

Today's culture of code development reflects our history

- The mission and the users are the focus of LLNL's code development
  - We establish strong partnership with the user communities – our code developers become SMEs
  - We encourage our staff to learn about the application space
- Multi-disciplinary code teams are a key to our success
  - We are experimenting with flexible teams of code developers who move between code projects
- We invest in multiple ways to improve overall capability: improving hardware, numerical methods, underlying physics, theoretical understanding, modeling methodologies, etc.
- Research investment to improve our capability is pursued continuously
  - Engage in university and vendor partnerships
- An evolutionary approach to building capability generally leads to success
  - We have created Computer Science focused projects that are peers to physics
  - We have embraced modularity
- We teach new code developers our history and try to learn from it

The future is now

How do we get ready for the world's faster computer?

NNSA Tri-lab resource

Step 1: Facilities

ECFM scales utility capacity of B-453 to 85MW and 28,000 tons cooling to prepare for Exascale

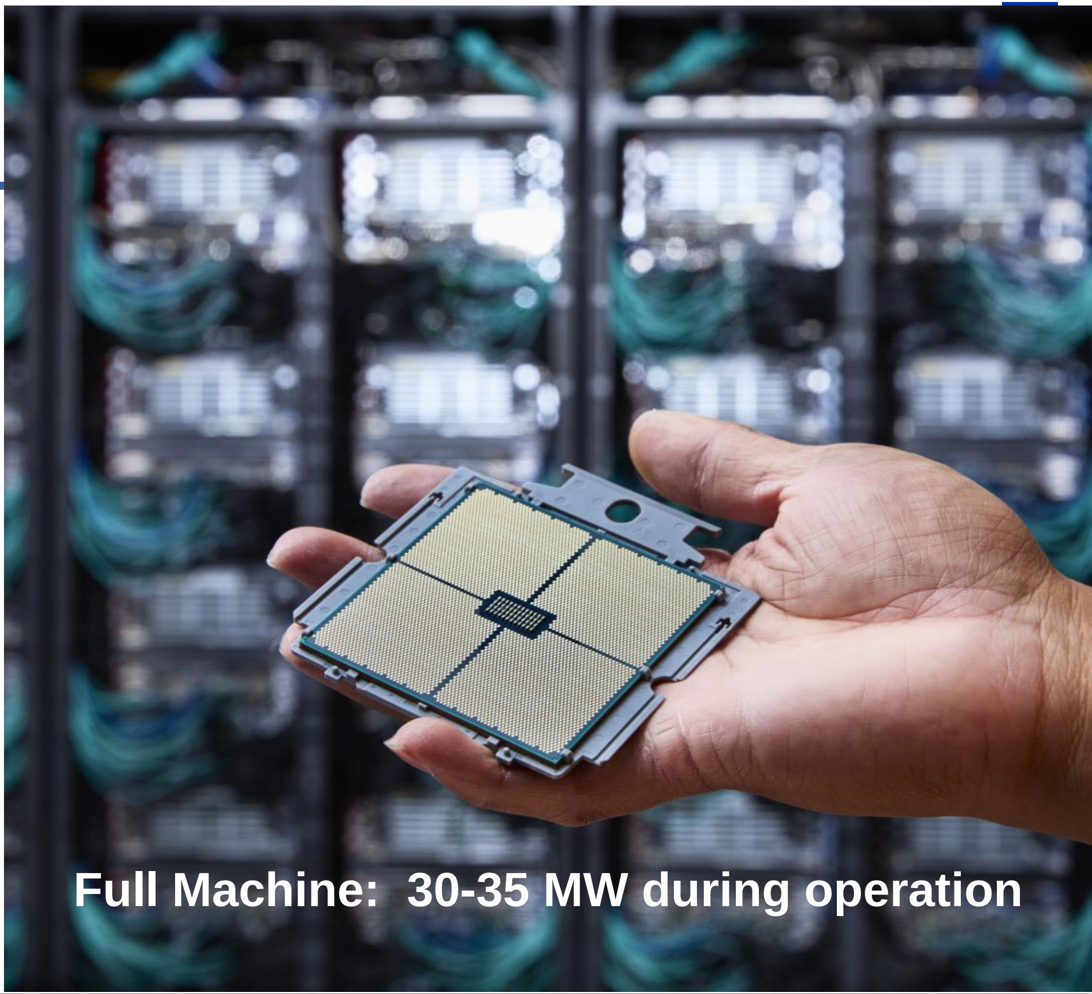
Step 2: Hardware

Sustained performance of 1.742 exaFLOPs → First exascale supercomputer for the NNSA

Peak speed of 2.79 exaFLOPs → Built by Hewlett Packard Enterprise with Cray EX architecture

AMD Instinct MI300A APUs optimized for AI workloads → HPE Slingshot interconnect and near node local storage technology

Unprecedented modeling & simulation for national security → Unclassified sibling system Tuolumne will be used for open science



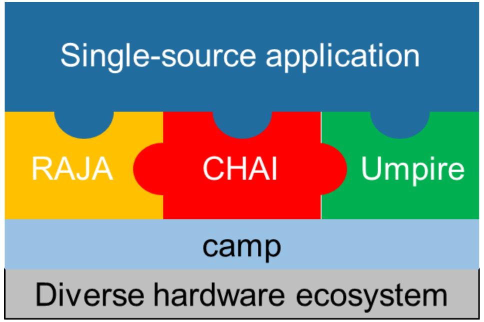
El Capitan (the hardware) on top of El Capitan (the mountain)

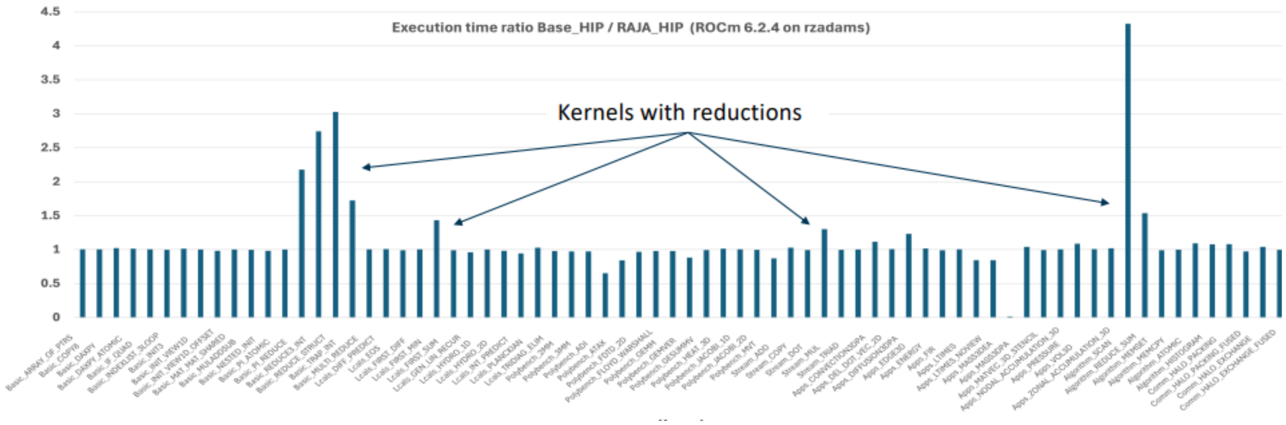
Step 3: We have the machine, now what?

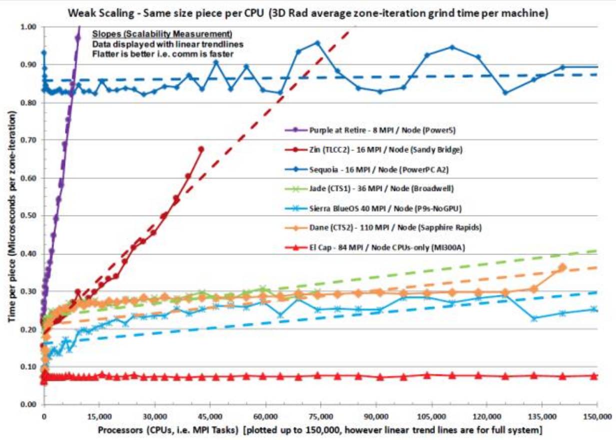
- Milestones we have met in delivering El Capitan capability to NNSA users
  - Procurement
  - Early Access Hardware
  - Acceptance
  - Early Access and Open Science Phase
- Main considerations:
  - Computational Performance
  - Machine Stability
  - Special features
- Procurement: Benchmarks/Proxy Apps established
  - Focus on key, simple algorithms
  - Benchmarks are open source, and represent Tri-labs
- Early Access Hardware – e.g. MI250 GPU machines
  - Test compilers, benchmarks, and key applications at small scale
  - Test new scheduling hardware
- Acceptance – Benchmarks/Proxy Apps are evaluated, Linpack benchmark is run for Top500 list
- Early Access and Open Science Phase
  - Focus on scaling, demonstrating machine stability at all scales
  - Use Open Science efforts to challenge the machine and get new science results at the same time
  - Key phase for collaborating with the vendors
  - Explore new workflows
- Then on to next steps – focusing on mission and delivering to users

We first test our Computer Science abstractions and simple scaling

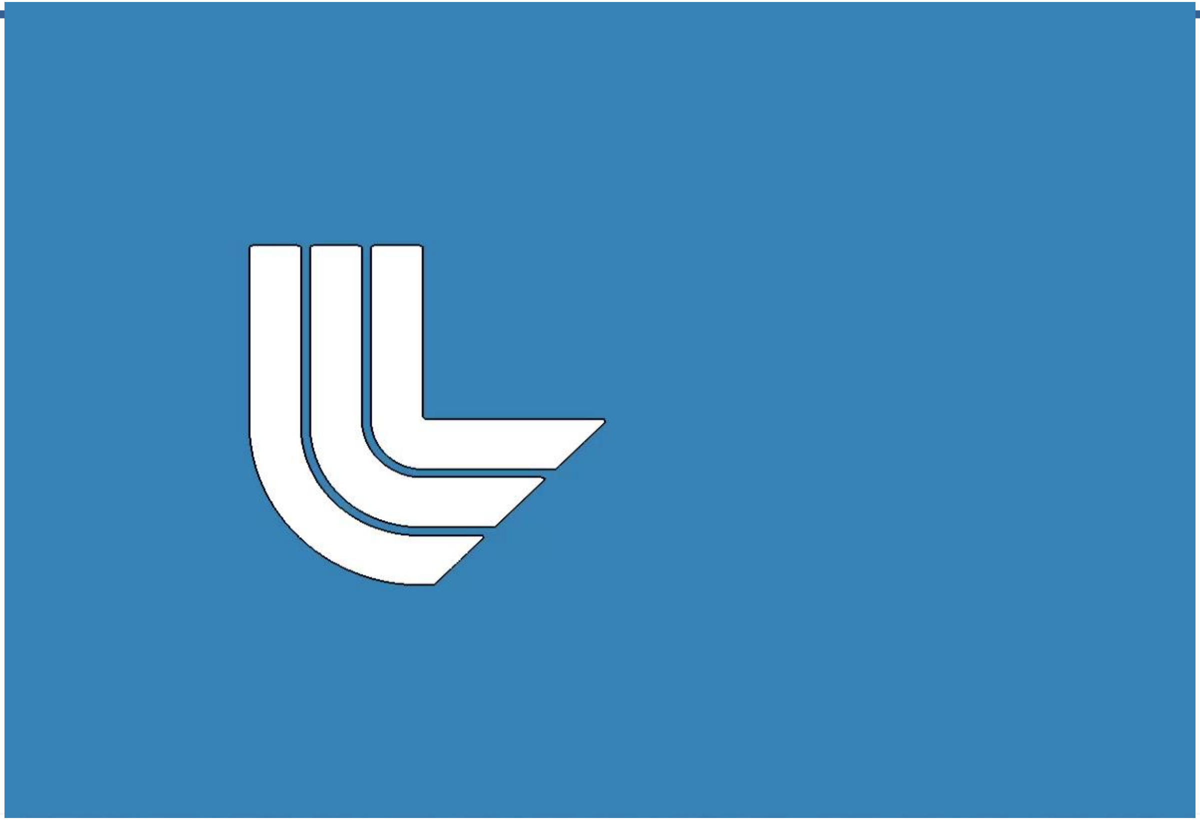
- Computer Science abstraction
  - This modularity is a relatively new advance, and is a big part of our Sierra success
  - Additional sharing has added additional burden for coordination between major projects
  - RAJA: 1 means as good as HIP, > 1 means outperforming HIP





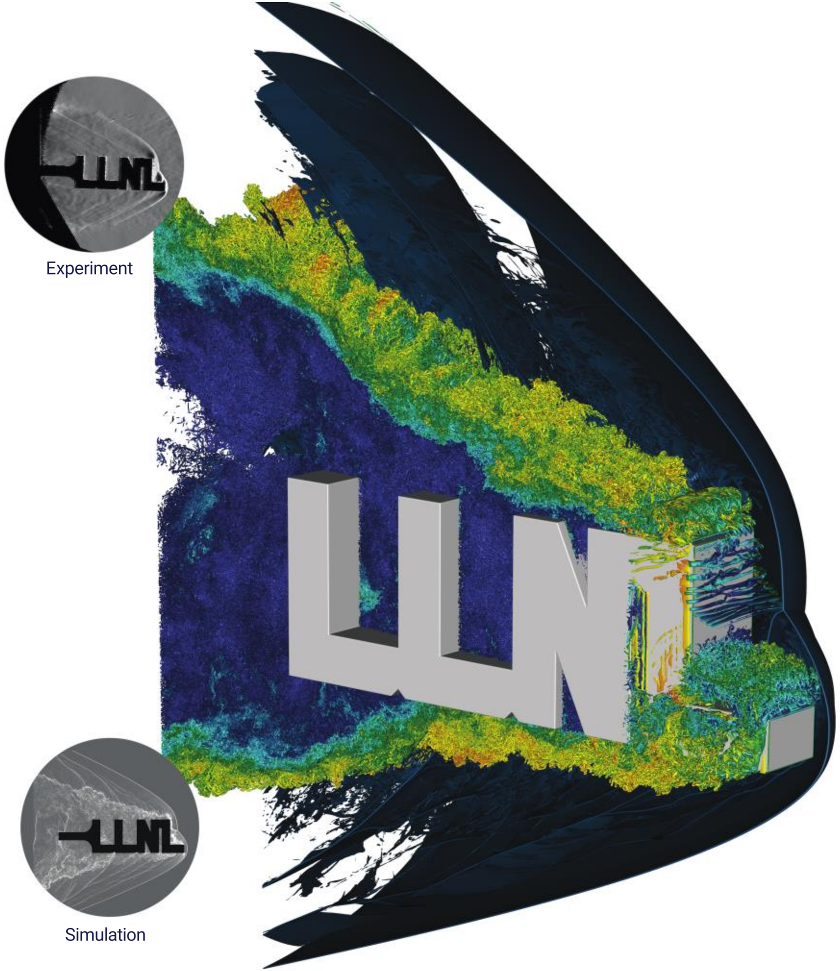


2D simulations allow us to test physics at small scales, and make sure multi-physics codes can compile and run



MARBL 2D simulation using Miranda hydro package and an immersed boundary method of Mach 3.0 flow past the LLNL logo on 16 nodes of Tuolomne

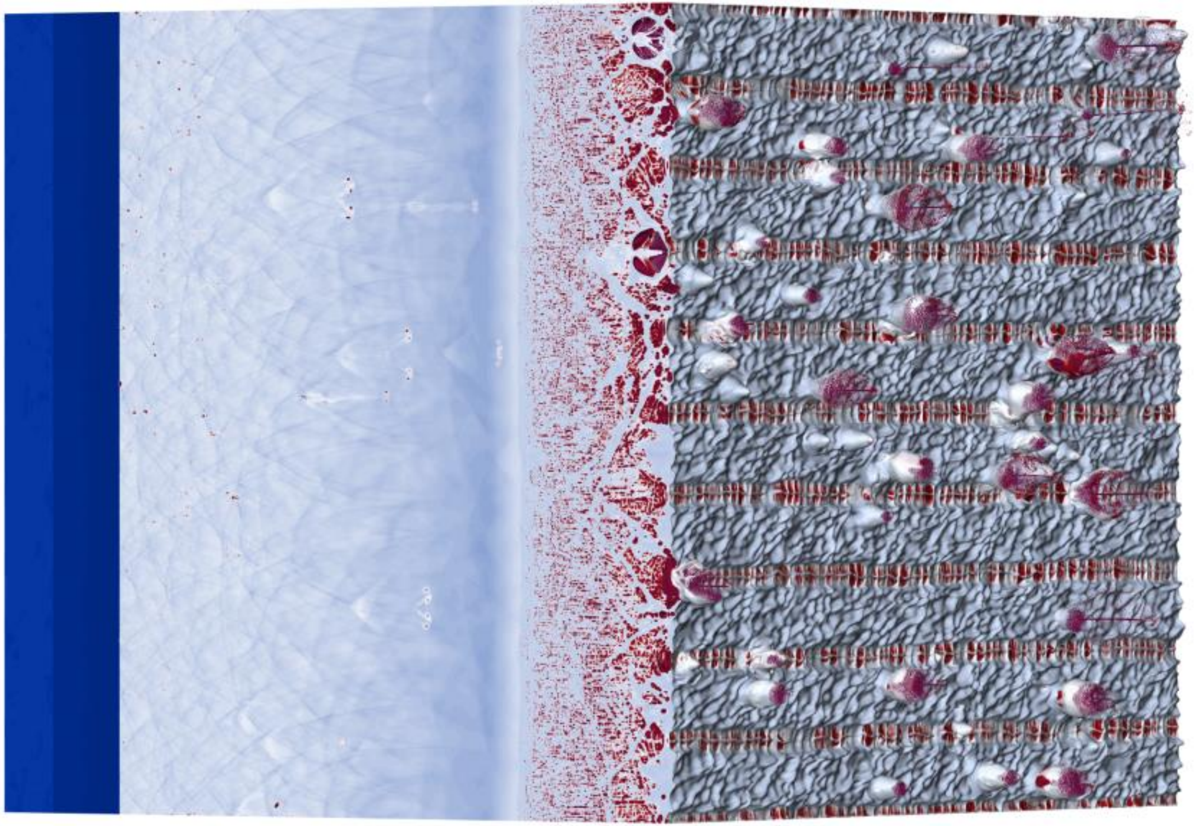
Scaling up results in new science and challenges the machine



Visualization of Mach 3.0 flow past a complex geometry and comparison to experimental data from the EMIT-LLNL wind tunnel. High resolution simulation is required to capture the turbulent shedding and the calculation above used ~1/3 of El Capitan (4096 nodes) for 24 hours.

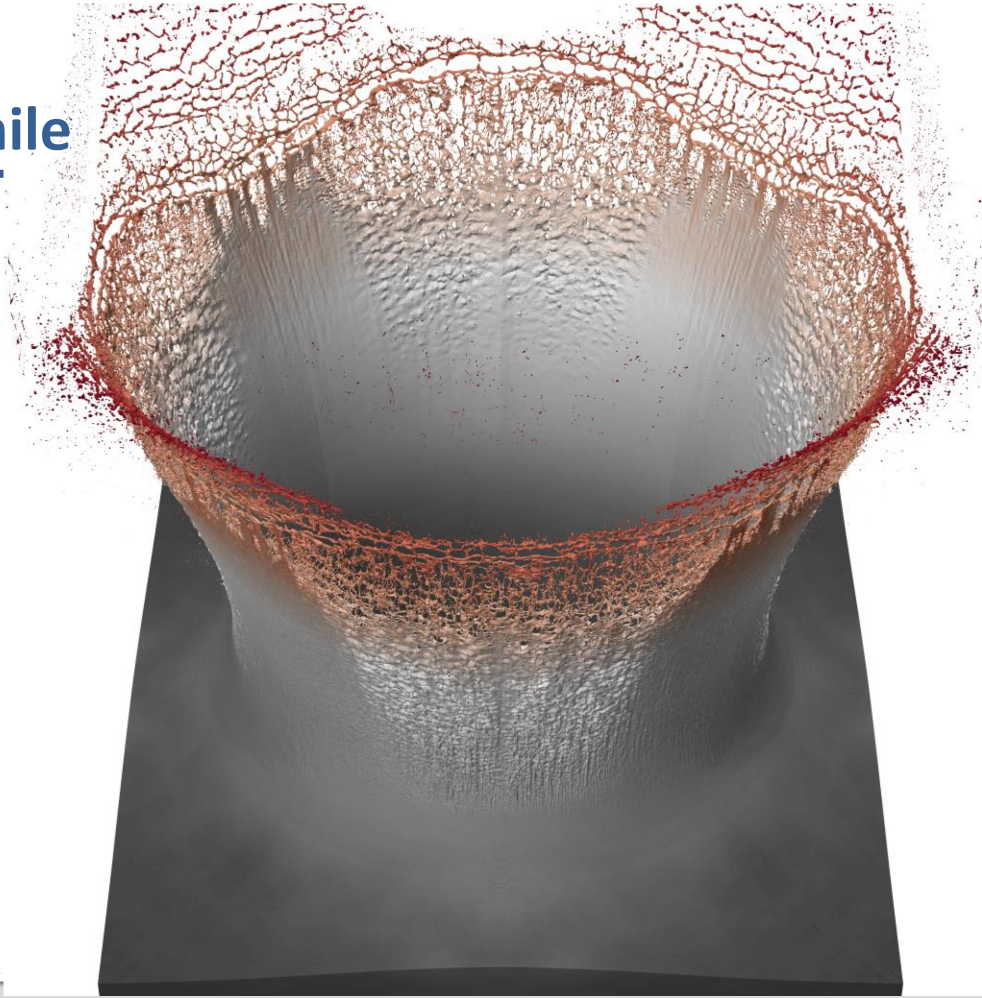
Multiple applications challenge the machine in different ways

- A density plot of a 139 Bz (2048 node) simulation of a shocked Tin surface undergoing Shallow Bubble Collapse (Jason Burmark, Kyle Mackay)



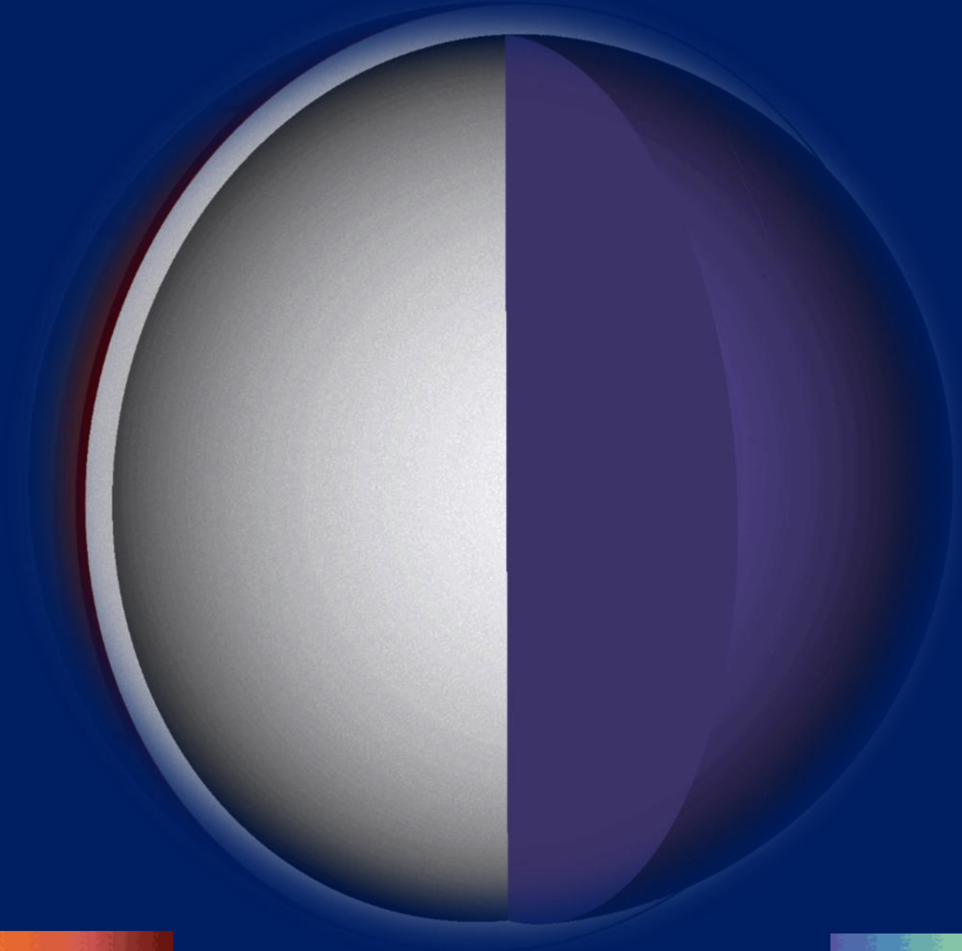
We are learning new things about physics models at these scales, while challenging the machine

- A density plot of a 250 Bz (2048 node) simulation of a hypersonic impact of a Tin particle on a Tin surface including the effects of surface tension (Jason Burmark, Fady Najjar)



3D MARBL simulation of NIF Shot N210808, Run on 64 Nodes of Tuolumne

Early Access/Open Science



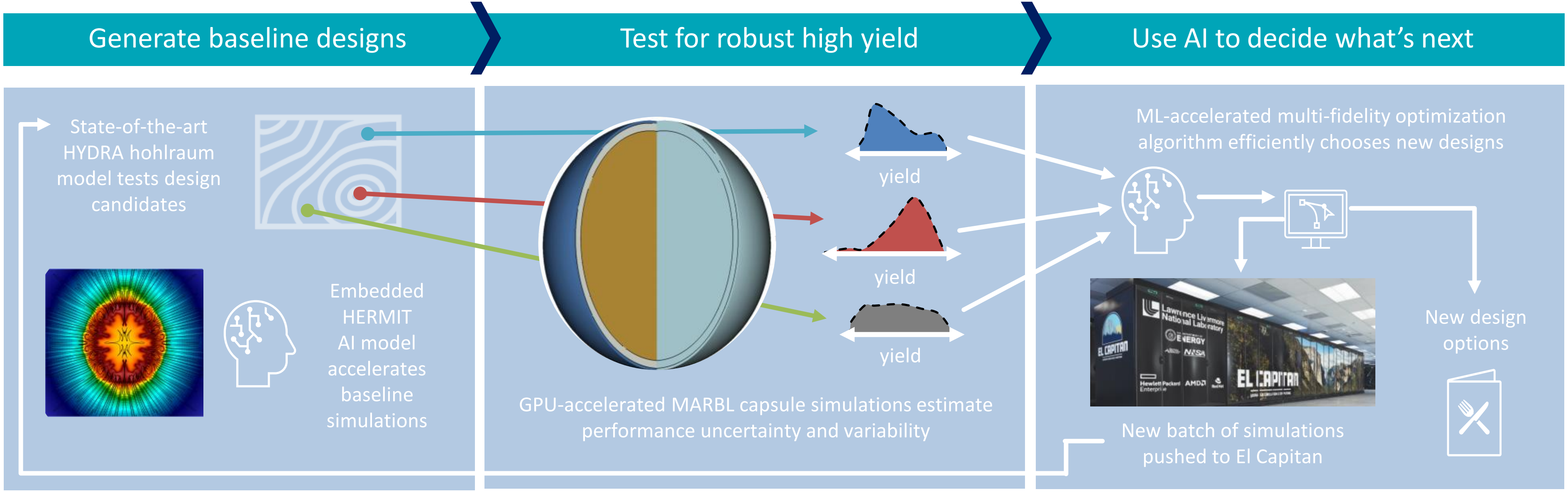
We are preparing design tools for Open Science applications and ML/AI workflows

Density

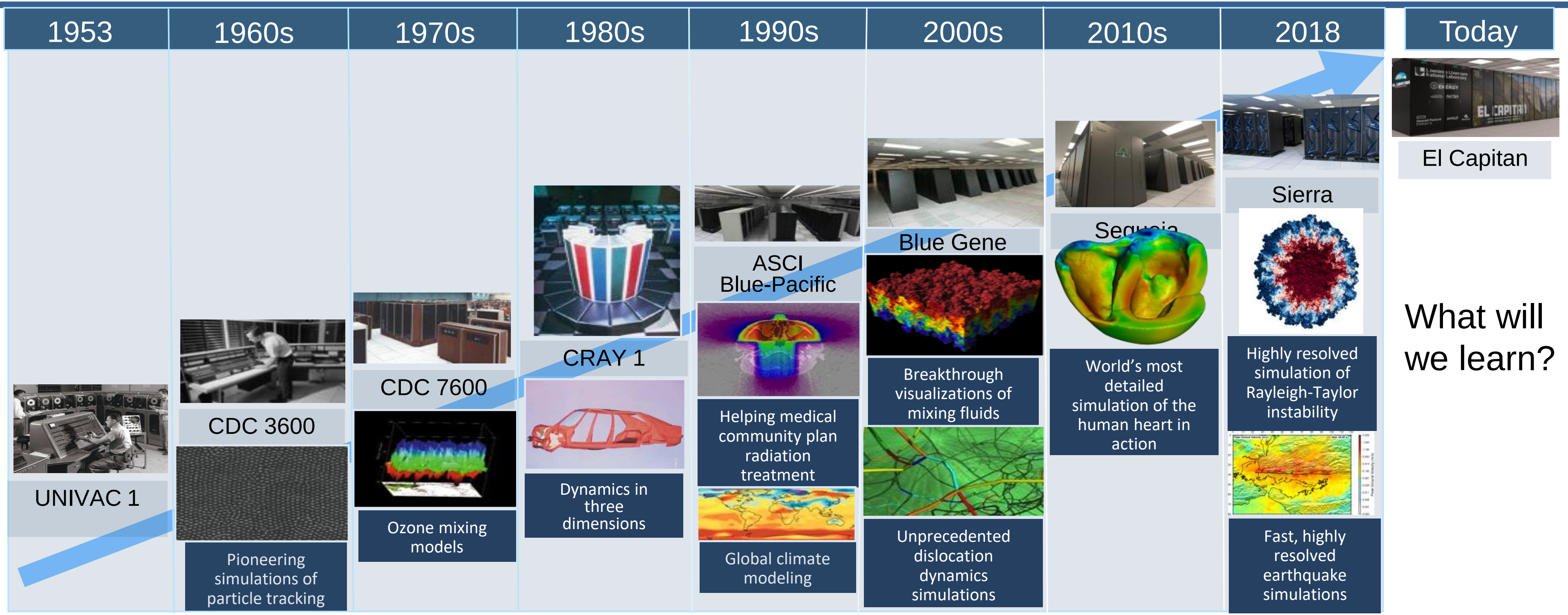
Temperature

El Capitan is the right machine at the right time to study high fidelity mod-sim, coupled with ML/AI workflows

→ Project ICECap aims to leverage exascale computing and AI to automate millions of simulations and discover new robust ignition designs
→ The first runs and designs are expected on El Capitan in FY25



LLNL has used its computing resources throughout its history to contribute to computational advances in many additional fields



LLNL programs benefit greatly from being part of the international computational physics community