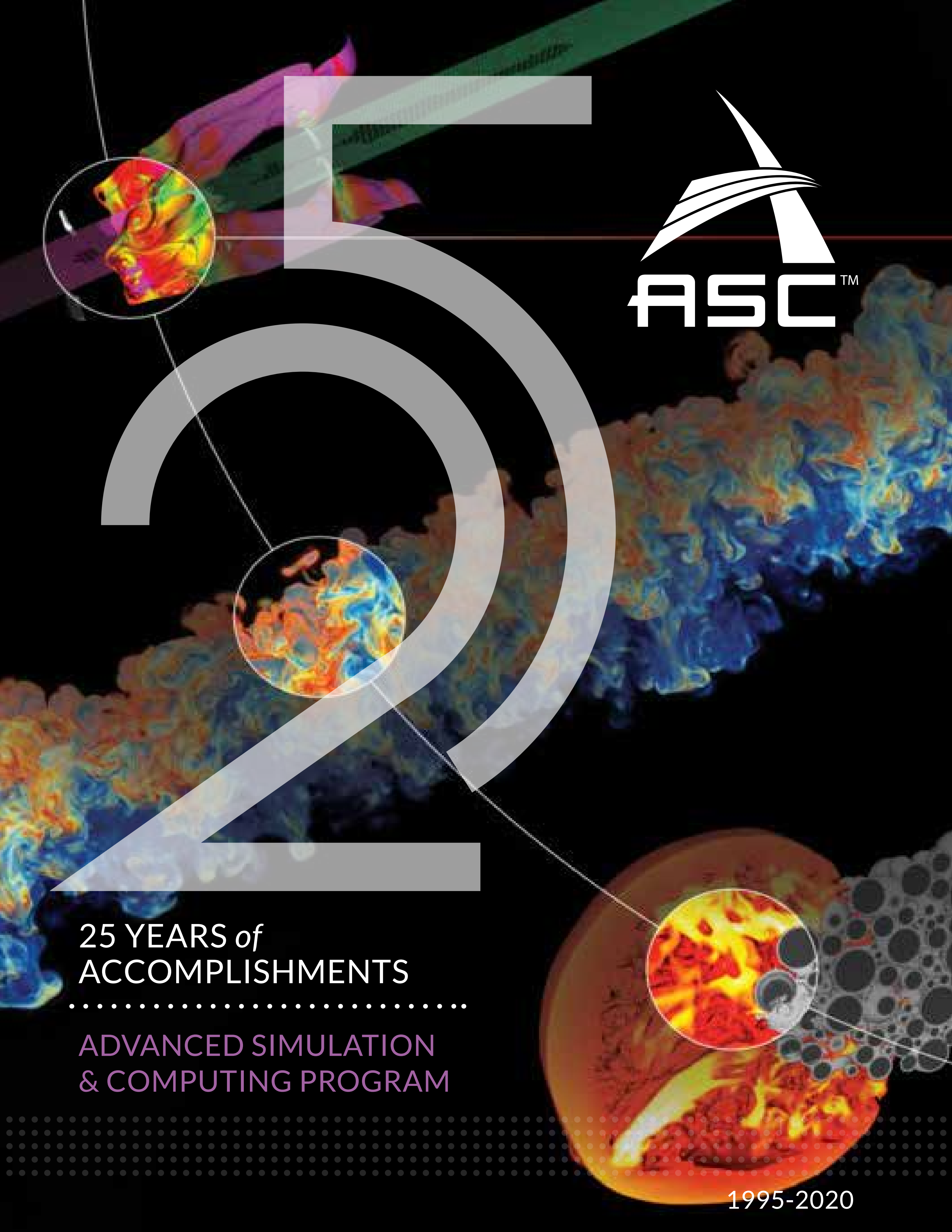


ASC™

25 YEARS *of* ACCOMPLISHMENTS

ADVANCED SIMULATION & COMPUTING PROGRAM

1995-2020

TABLE OF CONTENTS

ACKNOWLEDGMENTS

This report presents the 25-year record of accomplishments of the U.S. Department of Energy's Defense Programs Advanced Simulation and Computing (ASC) program.

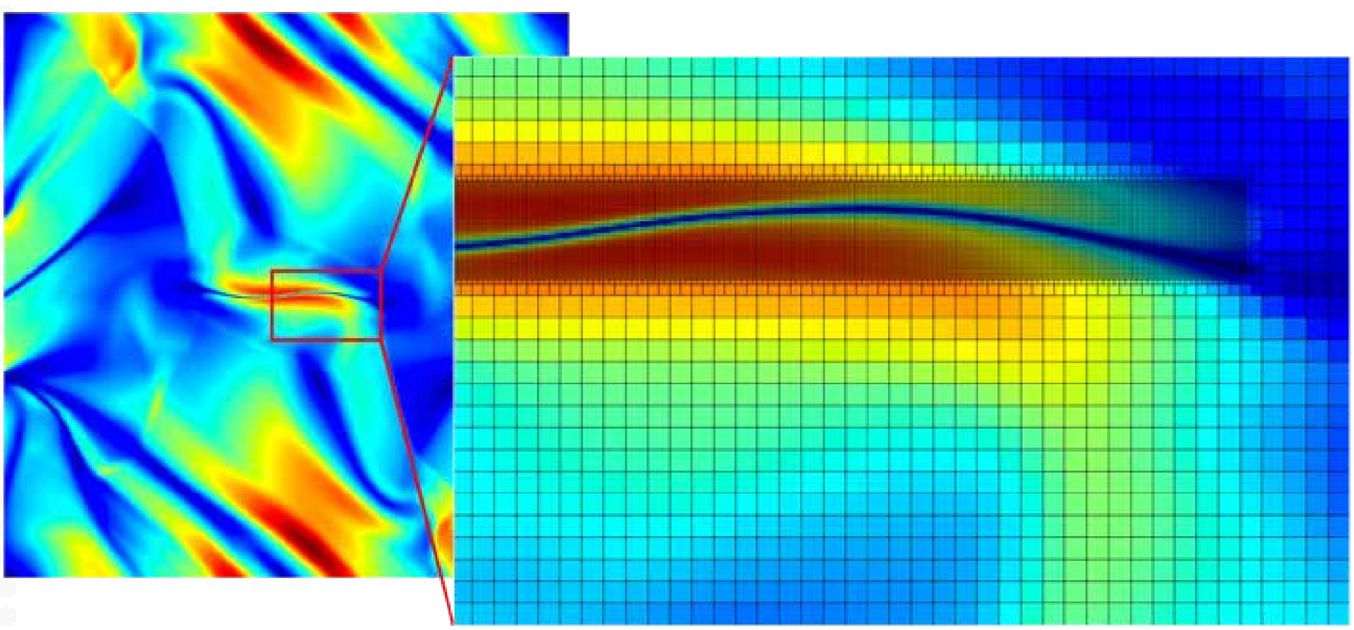
The authors of this document are indebted to Alex Larzelere's "Delivering Insight, The History of the Accelerated Strategic Computing Initiative (ASCI)." (Larzelere, 2009). That work provided a comprehensive view of ASCI from the perspective of 2009.

The authors also wish to thank the reviewers who provided in-depth expertise to each section of this report. This report would not be possible without them.

Approved for public release
SAND2022-11472 M.

Cover Photo: Three images of computer models. Top: Unstable intermixing of heavy (sulfur hexafluoride) and light fluid (air). Middle: Turbulence generated by unstable fluid flow. Bottom: Examining the effects of a one-megaton nuclear energy source detonated on the surface of an asteroid. (LANL)

- **Lead**: David Stevens (LLNL Detailee)
- **Committee members and authors:** Nicholas Lewis (LANL), Rob Neely (LLNL), and Sophia Lefantzi (SNL)
- **Editors:** Steve Scott (SNL), and Andrea Penner (SNL)
- **Designer:** Tara Renee Camacho-Lopez (SNL)
- **Reviewers:** Thuc Hoang (NNSA), Chris Clouse (LLNL), Samuel Scott Collis (SNL), Erik Strack (SNL), Jason Pruet (LANL), Shawn Burns (SNL), David Daniel (LANL), Kevin Elzie (NNSA), David Etim (NNSA), Aimee Hungerford (LANL), Mike Lang (NNSA), Anthony Lewis, (NNSA), Terri Quinn (LLNL), Robert Rieben (LLNL), and Becky Springmeyer (LLNL)



Magnetohydrodynamic (MHD) simulation with LANL xRAGE multi-physics code under the Eulerian Applications Project. MHD simulations are important to many inertial confinement fusion (ICF)/high-energy-density (HED) and stockpile applications. The simulations continue to be improved to support modeling experiments using laser-driven hohlraums.

FOREWORD

It is an immense pleasure for me to write the introduction for this ASC 25-year Accomplishments document. I have been working for the program since its beginning and can attest to the national ASC team's hard work in delivering accomplishments mentioned in this document.

The overall mission of the National Nuclear Security Administration (NNSA) Office of Defense Programs (DP) is to maintain a safe, secure, and effective nuclear weapons stockpile for the United States. Essential to this mission is the Advanced Simulation and Computing (ASC) program's high-performance simulation and computing capabilities that inform critical DP stockpile stewardship decisions. The ASC program's success relies on its ability to provide end users with validated simulation tools. These enable broad and deep knowledge of the individual processes involved in the detonation of a nuclear weapon and provide a comprehensive understanding of the complex interactions among these processes. By continually developing and deploying for NNSA the credible, science-based simulation tools to certify the current and future stockpile, ASC assures confidence in the nation's nuclear deterrent.

Although ASC has been successful in meeting its program responsibilities in the last 25 years, it cannot slow down or rest on its laurels. More challenges await in the near future. To ensure a strong nuclear deterrent in the face of a changing geopolitical climate, DP is required to maintain U.S. nuclear weapons systems and strengthen responsiveness of the larger nuclear enterprise. The ASC program will be called upon to contribute to the development of new weapons options, which could be very different from those in the current stockpile. To enable the nuclear security enterprise to field these options with agility, the ASC program also will be required to extend the application of its tools to improve the efficiency of the NNSA production complex. In the near future, the modeling of weapons systems beyond the stockpile's lifetime and the simulation of production processes to enable the production complex to be more efficient will represent a significant growth in the scope of the ASC program.

Regardless of the many obstacles awaiting us in the near future, whether budgetary or technical, I am certain that the ASC program, staffed by its very capable national laboratory and NNSA Headquarters team, will meet them head-on with high confidence of success.

Thuc T. Hoang
Director, Office of Advanced Simulation and Computing
& Institutional Research and Development Programs
NNSA Office of Defense Programs

NNSA
National Nuclear Security Administration
DEFENSEPROGRAMS

SUMMARY OF ACCOMPLISHMENTS

Today's Advanced Simulation and Computing (ASC) program has its origins in the aftermath of the Cold War. The program's predecessor, the Accelerated Strategic Computing Initiative (ASCI), was established in the mid-1990s as a national response to the need to preserve the Nation's nuclear deterrent with the cessation of underground nuclear testing. This need coincided with a period of great change in computing, and the ability to support stockpile stewardship through combined simulation and experiments alone was far from certain. ASCI's primary objective was to dramatically improve the performance of computing platforms and the accuracy of the models and codes that run on them. This document describes the high-level accomplishments of the resultant ASC program and how the program has become integral to the success of stockpile stewardship.

The organization established at the start of ASCI has performed admirably and has persisted largely intact in the ASC program. Current components retained from ASCI include the "One Program/Three Labs" principle (as represented by its logo), annual ASC Principal Investigator (PI) meetings, the use of technical program milestones to measure progress, and multi-year partnership with U.S. academic institutions and industry (ASCI Program Plan, 2000). Several technical concerns of the time are still relevant today, including effects of material aging on the stockpile, maintenance of critical skills, and preparations for unforeseen changes in the stockpile. These concerns are compounded by new challenges such as the simultaneous need to address evolving threat environments and stockpile modernization. Therefore, the solutions the program arrived at in the early part of the program are still relevant. These solutions have successfully enabled ASC to respond to rapid changes in technologies and deliver capabilities adapted to today's challenges, even though these new challenges are often quite different from those of 25 years ago. The program's long record of accomplishments in computing and the capabilities it enables will serve the Nation well in the rapidly evolving technological and threat environments of the 2020s and beyond.

ASC's initial 10-year goal was to achieve 100 teraFLOPS (one teraFLOPS is 10^{12} floating-point operations per second or FLOPS) of computing to demonstrate an initial capability to run 3D, stockpile-to-target-sequence simulation codes (ASCI Program Plan, 2000). The current goal is to achieve exascale computing, of more than one exaFLOPS (10^{18} FLOPS), which is a 10,000-fold increase in computing power over the original 100 teraFLOPS goal, by 2023 to support the DP stockpile stewardship and modernization missions. Both the multiphysics models and the computing platforms they run on have grown more sophisticated and capable, and can now perform calculations unfathomable by the original founders of ASCI. Using current ASC capabilities, calculations that were considered heroic or intractable at the start of the program are now commonplace. This document details these accomplishments in enough depth to illustrate how they were realized and provides examples of how ASC continues this record of achievement. More details of the program's rich history can be found in the documents listed in the bibliography.

Advanced Simulation & Computing™

The ASC program was created to support the DOE Science-Based Stockpile Stewardship (SBSS), now better known as the Stockpile Stewardship Program (SSP). It is now transitioning to the next-generation SSP with fundamentally modified or new weapons being incorporated into the stockpile. Notable accomplishments enabling this transition include the following:

- Simulation capabilities that are vital in negating the need for underground nuclear testing. Simulation results are a cornerstone of the Annual Assessment Reports (AARs) that detail the status of each weapon system.
- The ability of simulation to synthesize priceless data from underground nuclear testing, which informs the models used to fill gaps in the historical record.
- The use of three-dimensional (3D) simulations of the nuclear explosive packages as the basis of many assessments. 3D simulations are now regularly replacing earlier 2D simulations in assessments and eliminating many of the shortfalls associated with 2D simulation. The incoming ASC supercomputers, El Capitan and Crossroads, will enable routine 3D parameter studies with high resolution, which extends a core goal of the original initiative.
- The establishment of multi-disciplinary teams and multiphysics codes as standard practice with well-designed computer science infrastructure. This allows codes more maintainable, extensible, and better able to achieve higher fidelity and complexity.
- The multi-disciplinary code teams facilitated incorporation of new experimental data, thus ensuring steady growth in the breadth of physics considered and the material models implemented.
- A steady progression in the use of surveillance data, underground nuclear test data, and new experimental data to support simulation. This enables continued confidence in systems that are often more than 40 years old and well beyond their original projected lifetimes.
- Newly incorporated digital engineering efforts enable the development of the future stockpile. These efforts provide needed flexibility and resilience as the threat spectrum and peer competitor technologies evolve.

HIGHLIGHTS

ASC drove hardware and technology developments that transformed the high-performance computing (HPC) industry, achieving a series of global firsts and helping to maintain a fundamental U.S. industrial advantage over our competitors.

Arrival of the **first teraFLOPS (10^{12}) machine in ASCI Red,** occurring in the second year of ASCI (1996)

InfiniBand, the most common interconnect in today's supercomputers, initially funded by ASC for use in HPC systems (2000)

The development of **Lustre, the most commonly used file system** among major computer centers, through ASCI PathForward investments (2003)

The **Red Storm** system, which showed the value of balance between compute and communication and was the **blueprint for the highly successful Cray-XT line of supercomputers**, widely used across the HPC community (2004)

The **first 100 teraFLOPS machine in ASCI Purple**, achieving a key 10-year goal of ASCI (2005)

The **first 500 teraFLOPS machine in Blue Gene/L** in the first 10 years of the program, which helped the U.S. reclaim the global race for computing from Japan's Earth Simulator (2007)

The **first large-scale, accelerator-based system** in ASC **Roadrunner, the first petaFLOPS (10^{15}) machine** (2008)

Sequoia's 1.6 million central processing units (CPUs), which pushed the extremes of distributed-memory computing and lightweight low-power CPUs (2012)

Sierra and its Department of Energy (DOE) Office of Science sister machine, **Summit**, which reclaimed **top rankings for the U.S.** in the global race for the most capable supercomputers with graphics processing unit (GPU) accelerated systems (2018)

Astra, was the largest system with Arm Ltd. designed processors and was the **first Arm system to exceed one petaFLOPS** (2018)

El Capitan, will become the **first multi-exaFLOPS (10^{18}) system** fielded for the national security mission (2023)

EVOLUTION OF STOCKPILE STEWARDSHIP

A B61-12 test unit strikes a target during a complex forward ballistics test. The test, which mimicked a high-speed accident, allowed engineers to examine the weapon's safety features. (SNL)

Simulation has always been a component of nuclear weapon development. From the early 1950s to the end of the Cold War, nuclear weapon design codes were used to develop concepts and analyze trends that would then be supported by underground tests (UGTs). Often multiple UGTs were needed to evaluate all the operational conditions experienced by a weapon. The ability to rely on UGTs as confirmatory tests made it possible to use relatively simple forms of computation. As new computing technologies were introduced, major advances were made in the accuracy and analytical capabilities of the codes used to analyze the U.S. stockpile (Archer, 2017; Carr, 2013). However, prior to ASCI, the existing hardware and software would not support codes for anything beyond narrow weapon analyses.

The end of the Cold War in 1991 radically transformed the nuclear weapons establishment. One consequence was the end of full-scale, underground nuclear testing. A U.S. unilateral moratorium began in 1992, leading to the current Comprehensive Test Ban Treaty in 1996. This forced the NNSA laboratories to completely rethink their ability to ensure the performance and safety of the Nation's nuclear stockpile. It was not clear that a stockpile stewardship program could replace full-scale testing with simulation and experiments. The successes of stockpile stewardship as enabled by ASCI and its successor, the ASC program, have been central to the Nation's ability to manage a reliable and robust nuclear deterrent. Over the last 25 years, the program has consistently evolved to meet the needs of the time. This consisted of supporting refurbishment of weapon systems through modifications and Life Extension Programs (LEPs) throughout most of the program's history and is now transitioning to manufacturing new components and replacement of systems.

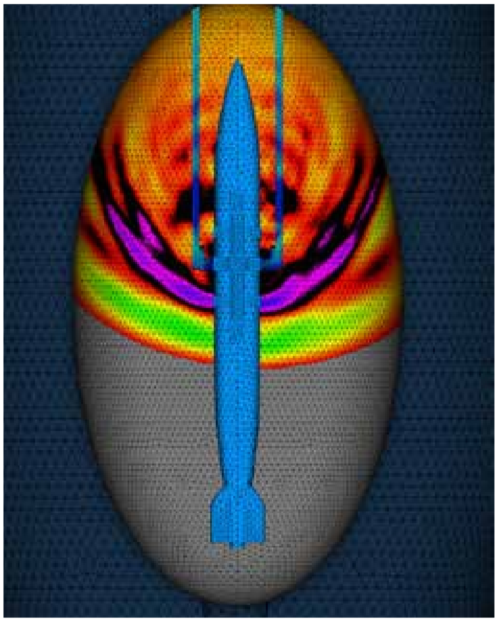
The current DP modernization effort presents challenges similar to those at the start of ASCI. Decades of computing breakthroughs and other ASC achievements place the program on a firm foundation for supporting NNSA in addressing these challenges. At the start of ASCI, the concept of SBSS replacing the use of underground nuclear testing was not widely accepted. The fact that SBSS and now SSP have enabled the current stockpile to serve the Nation without underground testing—in some cases with systems that are 40 years old—is an amazing accomplishment.

One current challenge is that replicating some production processes from the end of the Cold War is no longer cost-effective or even possible. This is due to many factors such as the loss of legacy methods, and/or safety and environmental restrictions. Another challenge is the development of more sophisticated threat vectors that might require new weapon system attributes to preserve the U.S. deterrent. Indeed, all the military platforms will see tremendous change or replacement. These include the new B-21 bomber, the Columbia submarine and the new Long Range Standoff (LRSO) and Ground Based Strategic Deterrent (GBSD) missiles. In the past, these platforms would have been accompanied by new warheads supported by their own set of UGTs. ASC is focused in the coming decades on providing simulation capabilities to meet a growing number of evolving needs. The next several pages illustrate important simulation capabilities that support the stockpile stewardship mission. Examples show the important roles simulation plays in supporting design, production, qualification, and secure transportation activities.

Currently, five weapon modifications and LEP efforts are underway - the most in NNSA history. The confidence NNSA has in taking on these endeavors was made possible, in part, by the success of realizing the earlier ASCI goal of using simulation to conduct stockpile stewardship. The B61-12 and W80-4 programs are highlighted below to illustrate the ASC program's comprehensive impacts on NNSA weapon activities.

B61 Life Extension Program

The B61 is a nuclear gravity bomb that first entered service 50 years ago. Numerous modifications have been made since that time to increase safety and reliability. The weapon is a core component of the nuclear deterrent and has been flown on a wide variety of aircraft types. It is also a key element of the nuclear deterrent umbrella that supports the North Atlantic Treaty Organization (NATO). The B61 LEP consolidates and replaces all the B61 variants into a single B61–12 weapon, with the exception of the B61–11. This has the benefit of extending the bomb's service life by at least 20 years and ensuring its continued safety, security, and effectiveness. Los Alamos National Laboratory's (LANL's) multiphysics codes enabled both the consolidation of these versions and the development of the modern weapon without underground nuclear testing. This weapon also requires a wide variety of experimental tests including impact, vibration, drops, extreme temperatures, and massive electromagnetic impulses. These physical tests are conducted to show that both individual components and the entire B61–12 weapon system will operate as intended. ASC multiphysics codes, in particular



Simulation using the structural dynamics (SD) component of the SIERRA multiphysics code, depicting structural response for the B61-12 to acoustic loading. (SNL)

the Sandia National Laboratories' (SNL) SIERRA code suite, enabled advanced computational analysis of these field tests. This ensures that the weapon performs as designed, while the components, systems, and warhead integration are safe and secure so as to prevent unplanned detonations or unauthorized use (Baker, 2019).

The B61–12 required close collaboration with the Air Force and other Department of Defense (DOD) organizations. To meet tight production schedules, the design work on both the F-35 and B61-12 had to be simultaneous (Baker, Virtual support, 2020). The accuracy and flexibility of the ASC codes were such that design work was able to proceed, ensuring successful deployment of the B61-12 on the F-35, well before the F-35 was available. Later flight tests demonstrated the accuracy of the simulations, marking a huge success for the LEP.

Several different tests and rigorous computational analysis are used to qualify the B61-12 system and components. (SNL)

An F-35A Lightning II opens its bomb bay doors and drops a mock B61-12 at SNL's Tonopah Test Range.

W80-4 Life Extension Program

The W80-1 warhead was first introduced in 1982 and deployed on an air-launched cruise missile (ALCM). Both the warhead and delivery systems are well past their planned lifetime, which led the DOD to start replacing the ALCM with the LRSO cruise missile. Simultaneously, NNSA is modernizing the W80-1 through the W80-4 LEP, managed by Lawrence Livermore National Laboratory (LLNL). This LEP requires flexible options for replacing aging components and materials, including new manufacturing methods that minimize costs, increase production throughput, and reduce the use of environmentally sensitive materials and processes. This work is all being done under considerable time pressure to meet the Air Force's operational needs, requiring strong relationships between LLNL, LANL, SNL, NNSA production plants, and the Air Force. ASC simulation codes and platforms have provided high-fidelity computational surrogates, allowing rapid design exploration to meet the requirement that the modified warhead be compatible with the LRSO missile, while also increasing safety, reducing cost, and extending the warhead's service life.

As with all modernization programs in the NNSA, a systems engineering approach on the W80-4 was taken to identify improvements in safety, reliability, and manufacturability. This systems approach requires the full complement of computational, experimental, and manufacturing capabilities.

Accurate simulations help reduce design iterations compared to building and testing prototypes. Scientists combine extensive experimental data with ASC computer simulations to understand and assess the safety and performance of refurbished and new components. For example, a family of increasingly accurate computer codes predicts the performance of new energetic material formulations, as validated with experiments. In all, LEP-affiliated simulations consume a large fraction of ASC computing resources. Sierra, known as an Advanced Technology system (ATS-2) at LLNL, plays an important role in assessing W80-4 performance and LEP certification. It also demonstrates the need for the increased computing power to be provided by NNSA's first exascale system, to be named El Capitan (ATS-4), which will be deployed at LLNL in 2023 (Heller, 2018).

W80-4 test unit.

A fundamental accomplishment of the ASC program in support of the W80-4 has been the ability to accurately model data from NNSA experimental facilities. Modern LEPs would be far more expensive without ASC simulations to reduce, and in some cases eliminate, the need for expensive and time-consuming experiments. Simulation and experimental facilities have an important interplay where advances in one drive innovation in the other, resulting in more accurate simulations and a much better understanding of observed system behavior.

LLNL's Scott McAllister, associate program director for hydrodynamic and subcritical experiments states: "We combine such [simulation] models with data from current experiments and past nuclear tests to verify we have confidence that our nuclear stockpile will perform. Codes are improving all the time, so diagnostic tools such as [the Flash X Ray] FXR must also evolve to maintain parity," (Parker, 2018).

The Flash X Ray (FXR) at the LLNL Contained Firing Facility (CFF) is an important diagnostic tool for hydrodynamic testing that enables scientists to capture images of an exploding device. (LLNL)

Transportation of Nuclear Weapons and Special Nuclear Materials

The National Nuclear Security Administration's Office of Secure Transportation (OST) has been safely and securely transporting nuclear weapons, weapon components, and special nuclear materials in purpose-built tractor-trailers, called SafeGuards Transporters (SGTs) along U.S. highways for decades. Each transportation system is required to ensure the safety of the public and the security of on-board assets under a wide range of conditions that include normal operation, severe or unusual accidents (e.g., highway crash, large fuel fires, lightning), and hostile environments.

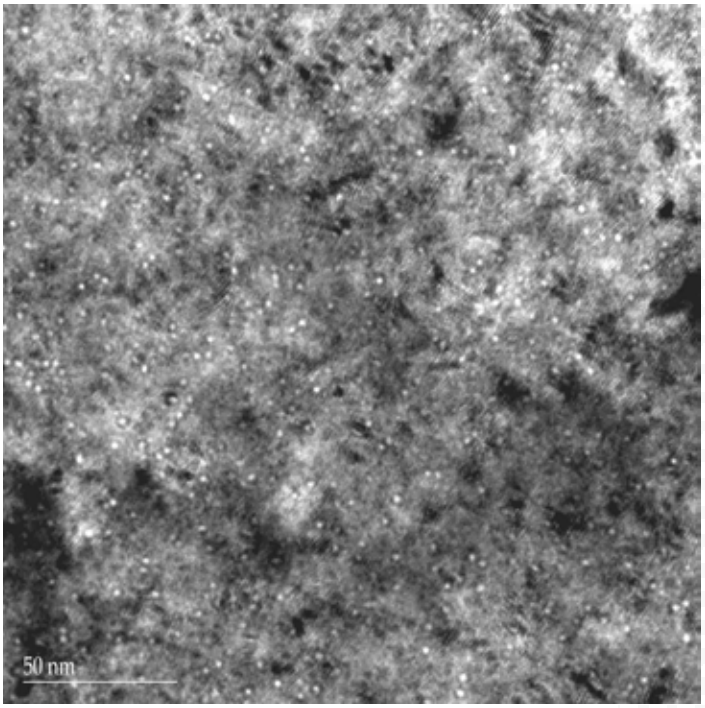
ASC capabilities (HPC platforms and the SNL SIERRA suite of codes) are being used in tandem with selective testing to help OST answer pressing

Side impact crash scenario investigated analytically using SNL's high-performance computing resources.

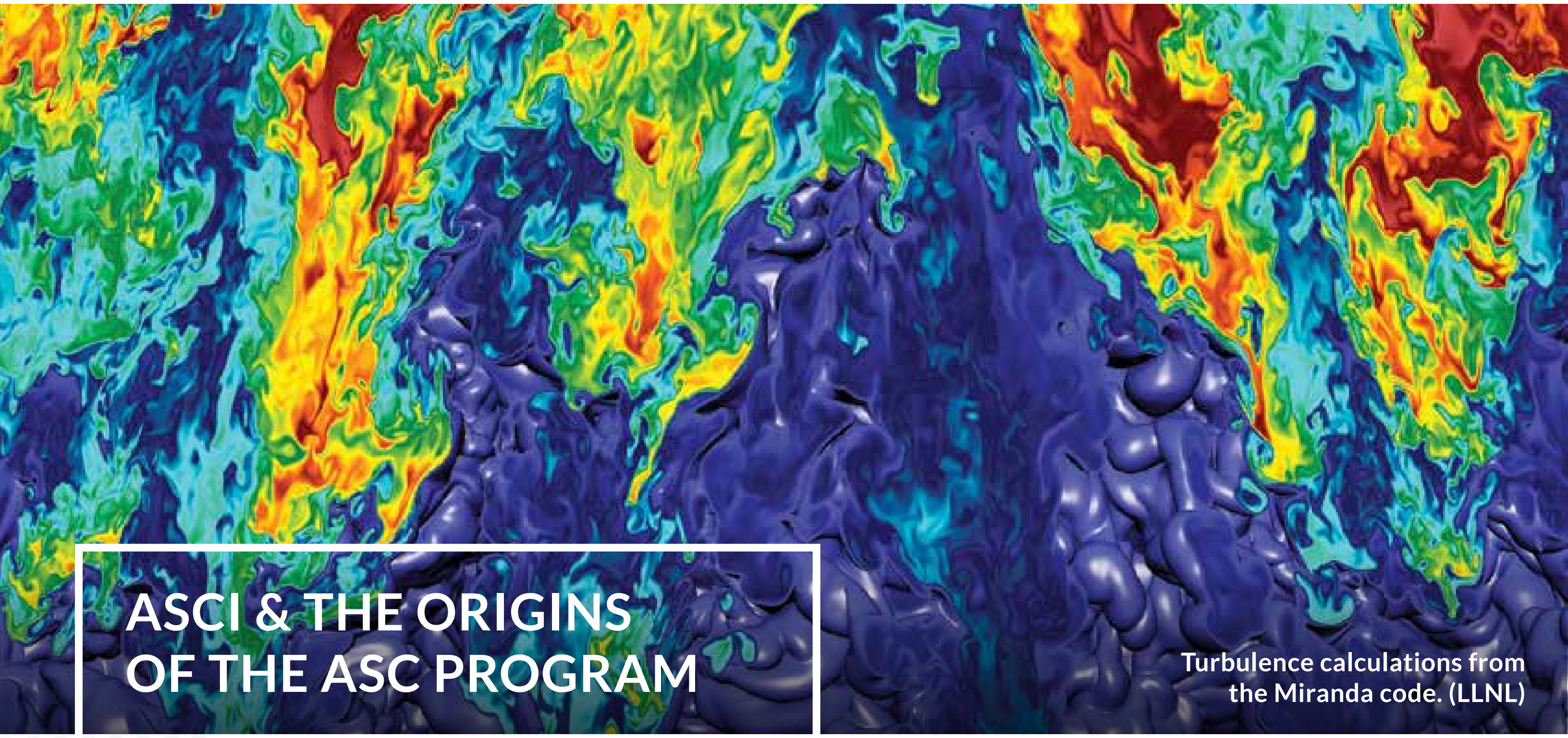
questions efficiently and effectively. Simulations are used to investigate various crash scenarios, including the side impact of the SGT by a second SGT, and results are used to inform planning for future tests. Data from the tests will not only demonstrate system performance but will also help to assess and improve the accuracy of the simulation models. These models will then be applied to further investigate the crash scenario space and arrive at an improved understanding of the underlying phenomena controlling the performance of the SGTs. Using these new ASC capabilities to more effectively and efficiently support the mission of OST ultimately ensures the safety and security of these critically important transportation systems.

Plutonium Science

Understanding the behavior of plutonium, uranium, and other fissile materials is central to stockpile stewardship. The ASC Physics and Engineering Models (PEM) subprogram's development of models that describe the behavior of these materials under extreme conditions has proven invaluable for the ongoing development of predictive models of nuclear weapon behavior. Better understanding of plutonium and how it ages is central to this model development. By the year 2030, several weapons in the stockpile will have plutonium parts that are over 50 years old and may be required to serve another 20-30 years. The presence of small amounts of helium decay products in these plutonium parts as well as other effects from plutonium decay have long been a cause for concern in ensuring that these parts can perform as intended. In 2006, JASON, an independent scientific group that is chartered through the MITRE Corporation to advise the U.S. government on sensitive science and technology areas, released the Pit Lifetime report which compiled several years of laboratory research on plutonium aging. This report concluded that the available research supports an initial claim that these parts have credible lifetimes that will allow for extended deployment and storage (Pit Lifetime, 2006). The report was guided by decades of work by scientists who developed detailed models of aged plutonium and its effect on weapons. This work formed the basis of the 2006 JASON report and multiple follow-on studies. This is instrumental in providing a much clearer picture for how these parts will perform as they age. The growing understanding of plutonium behavior is fundamental to providing NNSA flexible and science-based options for maintaining a sustainable deterrent.



Helium bubbles in a Transmission Electron Microscopy (TEM) image of an aged plutonium sample. These bubbles affect the strength, compressibility, and other properties of the material. (LLNL)



ASCI & THE ORIGINS OF THE ASC PROGRAM

Turbulence calculations from the Miranda code. (LLNL)

Before the end of the Cold War, LLNL and LANL relied on underground nuclear testing as the primary means of certifying the U.S. nuclear stockpile. The end of the Cold War created substantial debate on how to maintain a nuclear deterrent, while accounting for the reality of no new confirmatory nuclear testing. After a long career with the DOD and Science Applications International Corporation (SAIC), Vic Reis became the Clinton administration's Assistant Secretary for Defense Programs (DP) in the U.S. Department of Energy from 1993 to 1999. He was one of the first to recognize the need for a formal program to maintain the U.S. stockpile in response to the proposed nuclear test ban. Vic Reis, together with Gil Weigand from SNL, led what became ASCI in 1995. ASCI was part of a new technical DP portfolio established by the 1994 National Defense Authorization Act (NDAA). The NDAA established SBSS as the strategy for assessing performance and reliability of the U.S. nuclear stockpile without underground nuclear testing. Reis and Weigand realized that a science-based program was needed to develop simulation and experimental technologies accurate enough to help certify nuclear weapons without full-scale testing. They also determined that the program should be a broad collaboration between the laboratories, industry and universities to persuade skeptics that simulation could essentially replace UGTs as a credible virtual test environment.

Several foundations of the resulting ASCI Initiative are still with ASC today:

- Support of national nuclear deterrent via SSP, relying on supercomputer simulations and small-scale, non-nuclear experiments
- An integrated program, coordinating deployment of HPC systems with development of modeling and simulation codes
- A cross-disciplinary approach to develop detailed and accurate computational simulations
- Cooperation among the NNSA laboratories to attain goals beyond the scope of a single laboratory

» Close collaboration with industry to foster the HPC ecosystem
» Close collaboration with academia to foster intellectual exchanges and build a workforce pipeline

These foundations were accompanied by a willingness to act boldly in response to a rapidly changing technological landscape. This led to an impressive string of achievements in the first 10 years of the program and, in turn, made stockpile stewardship a success (Crawford, 2021).

ASCI tri-lab executives were concerned with several drivers that still have impacts on the current stockpile. The major nuclear weapon systems had typically been replaced with new systems at the end of their initial design lifetimes. There was little expectation for these systems to outlive their initial design lifetimes. Excellent craftsmanship has allowed these still-fielded systems to age as well as they have. By the early 1990s, many of the scientists and engineers who maintained these weapon systems were also nearing retirement and they did not know how these systems would behave as they aged. The ASCI approach of embracing collaboration and change has helped enable the U.S. to surpass all initial expectations in preserving its nuclear deterrent. The discussion that follows will dive into how this was accomplished.

The 1994 NDAA gave the Department of Energy (DOE) authority to form ASCI as a joint effort among LANL, LLNL, and SNL. The three NNSA labs, in close collaboration with industry and academia, were directed to aggressively pursue development of the needed modeling and simulation capabilities. An initial goal called for an approximate 10,000-fold performance increase over about 10 years to cross a 100 teraFLOPS threshold that the labs had calculated would be needed to begin the transition from 2D to 3D simulations. This would start with a first phase of relatively coarse 3D simulations between 1996 and 2000, moving to a second phase of finer-resolution 2D and 3D simulations coupled with improved physics models between 2000 and 2006. Over the much longer third phase, the program would transition to physics models fully based on fundamental understanding of underlying processes and increase simulation resolution to enable predictive capability.

An Enduring Foundation

Vic Reis, Gil Weigand, and the ASCI executive team (led by Alex Larzelere) created a program structure that endures to this day as a way to manage the three NNSA laboratories charged with executing a very ambitious technical initiative. (Larzelere, 2009) (ASCI Program Plan, 2000). This structure includes:

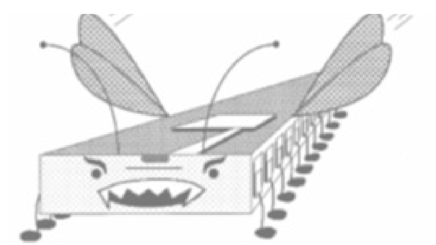
» Multi-disciplinary code teams to encourage tight collaborations between and within the laboratories themselves with the necessary skills and experience to incorporate new physics, material models, and algorithms as they are developed
» Milestone-driven implementation to ensure progress in all ASC subprograms to deliver simulation capabilities that are verified and validated
» Well-defined strategies and program plans to provide the labs with clear tasking and guidance on the scope of their work
» ASC PI meetings to foster engagement among laboratories and with external partners, assess technical progress, share recent program accomplishments, and position the program at the forefront of technological innovation
» Academic alliances to foster and ensure a steady flow of new staff and ideas into the NNSA laboratories
» Support for research organizations like the LLNL Center for Applied Scientific Computing (CASC) and SNL's Computer Science Research Institute (CSRI) to enable the exchange of technical, innovative ideas with appropriate external collaborators

Industrial Partnerships

While ASCI was being formulated, the supercomputer industry had entered a state of disarray after the Cold War ended. A reduction in government purchases of large computer systems, along with the rapid growth of personal computer (PC) and workstation sales, significantly reduced the profits and market share of mainframes and supercomputers. Vendors for massively parallel computers had found very few takers for their products before the market began to shift and were hit particularly hard by the lack of interest in supercomputing.Many of these companies, such as Thinking Machines Corporation, ultimately declared bankruptcy. Longer-established supercomputer and mainframe companies were also not spared. Of the top 15 computer companies in 1975, only four survived into the early 2000s. One such industry leader in the mid-1970s, Control Data Corporation, was not among them. IBM was no longer the behemoth it had once been. Another well-known company, Cray Research, was purchased by Silicon Graphics in 1996. PC, workstation, and commodity semiconductor companies had become the new giants of the computer market. In this altered marketplace, the DOE/DP laboratories found themselves with diminished influence and weakened ability to drive platform improvements by the vendors. This raised a well-founded concern that HPC machines would not be powerful enough to perform the needed simulations.

These factors meant that ASCI had to be inventive with the initiative's limited funding. While the rapid growth of microprocessors negatively impacted the development of conventional supercomputers, the consumer-grade technology also provided a solution to reverse the stagnation afflicting high performance computing. The spirit of the time was captured by Eugene Brooks in his 1989 presentation, "Attack of the Killer Micros," (Brooks, 1989). He and others realized in the late 1980s that, while commodity microprocessors were substantially inferior to supercomputing

The innovative Connection Machine, CM-5, was the first parallel computer used for stockpile related purposes at LANL. It was built by the Thinking Machines Corporation. (LANL)



"

The Enemy, the Killer Micro. Some are quite vicious, eating their parents within a year of birth.

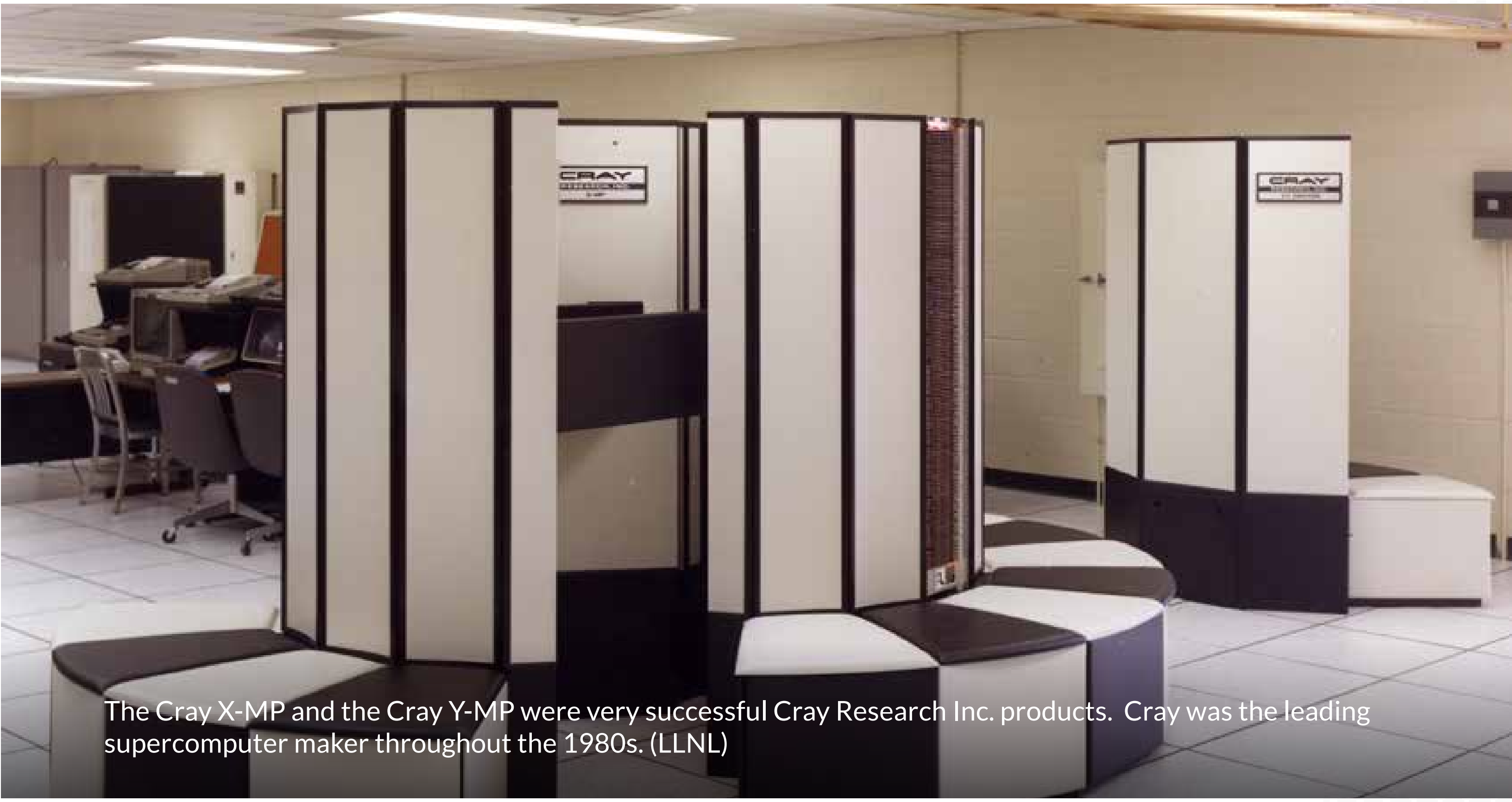
-Eugene Brooks
"The Attack of the Killer Micros"
Super Computing 1989.

"

processors, their computational growth curve was substantially faster. In around a decade, commodity microprocessors would eclipse the processing speed of the more powerful specialized processors that were then being used in supercomputers. The industry realized that very powerful supercomputers could effectively be created through chaining together thousands of commodity processors to form a massively parallel system.

The founders of ASCI, as well as experts in other areas, recognized that the exploration and maturation of these massively parallel systems could be the best path with the greatest long-term potential. This effort would need a considerable amount of research and development (R&D) investment to progress beyond the experimental systems that existed in the early 1980s. There was a substantial fear that without adequate inducement from the U.S. government to encourage continued R&D investments, the U.S. HPC industry might lag behind other nations in supercomputer development. Therefore, establishing a long-term market for this strategic capability was essential and required strong partnerships with industry.

ASCI would use proven collaboration models with its industrial partnerships, based on experience with the Technology Transfer program and the CRADA (Cooperative Research and Development Agreement) framework. One example is the High Performance Storage System (HPSS) collaboration between IBM and several DOE national laboratories. Marking its 30th anniversary in 2022, the effort enables long-term storage for huge volumes of archived simulation data. HPSS produces storage system technologies, used in many supercomputing centers, that have scaled from terabytes at the project's start to more than an exabyte in 2021. ASC continues to support HPSS, which has grown

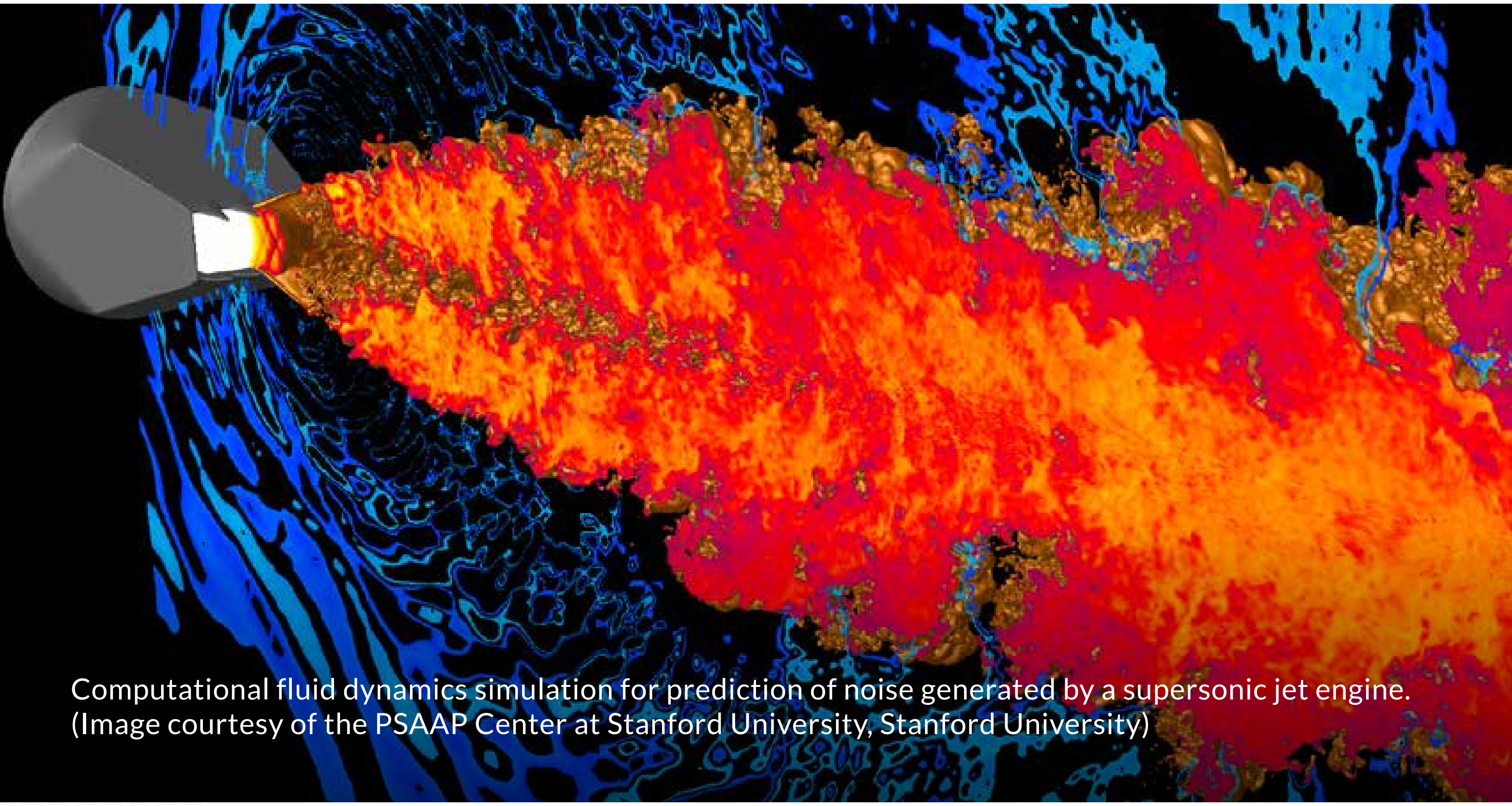


The Cray X-MP and the Cray Y-MP were very successful Cray Research Inc. products. Cray was the leading supercomputer maker throughout the 1980s. (LLNL) to include many other collaborators in its mission to preserve high-value data for governmental, academic and commercial organizations. (Heer, 2021).

The goal of supporting hardware and software solutions with viable business models was central to the solutions that were proposed. These solutions would use existing commodity products, and focused on cluster computers, which had enormous potential for delivering rapid performance increases. These cluster computers were composed of many computational nodes, each containing one or more commodity processors, which when combined would allow simulations to be performed on many thousands of processors simultaneously. Clusters could also exploit commodity components, networks, control software, and programming aids. The cluster computing paradigm required significant R&D to move forward, but its success continues to drive its position as the dominant HPC paradigm used today.

Academic Alliance Program

Universities would be a key part of the effort to demonstrate that simulation could serve as virtual proxy for large-scale experiments, such as UGTs, that cannot be performed. The original alliance of universities, the Academic Strategic Alliance Program (ASAP), was formed in 1997. This program focused on high-fidelity modeling and simulation on the initial parallel architectures that ASCI procured. Five ASAP centers were selected at leading universities to develop large multiphysics applications as academic analogs to the ASCI integrated codes that focus on stockpile stewardship. This was a time when the physics and mathematical algorithms that are now commonly used in massively parallel ASC applications were still being formulated. At the time, the Message

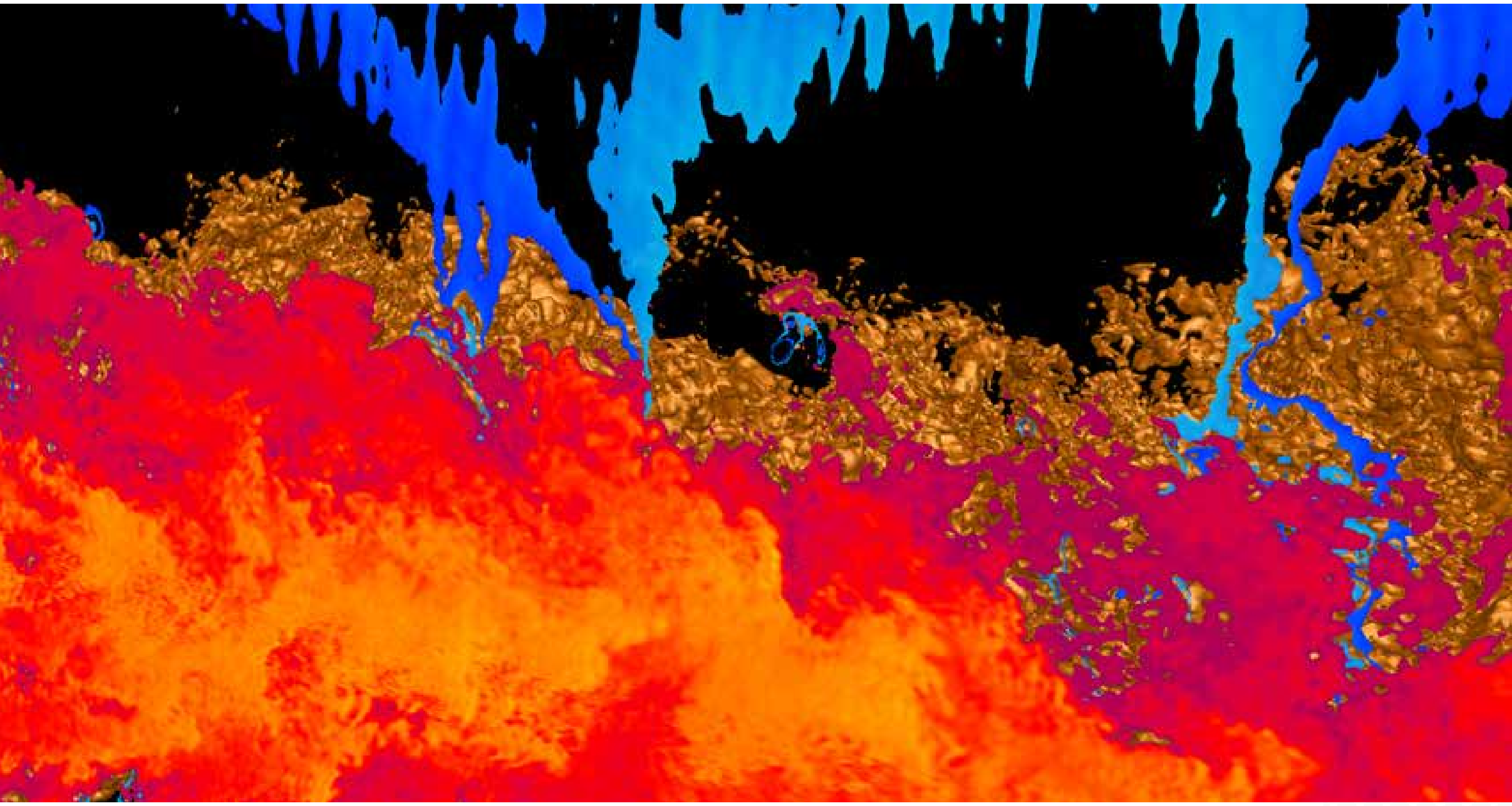


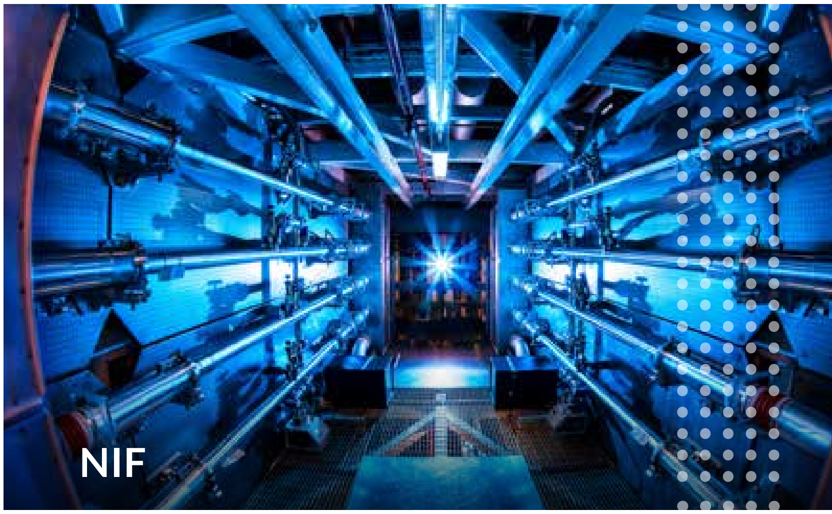
Computational fluid dynamics simulation for prediction of noise generated by a supersonic jet engine. (Image courtesy of the PSAAP Center at Stanford University, Stanford University)

Passing Interface (MPI) was not yet the dominant method for exploiting parallelism that it is today, and universities were instrumental in this development.

As ASC applications matured, an emphasis on predictive science was formally recognized in the program. In 2008, the Predictive Science Academic Alliance Program I (PSAAP I) replaced ASAP by adding an emphasis on validation, verification, and uncertainty quantification. The follow-on program in 2014, PSAAP II, added a focus on extreme-scale computing. The PSAAP III portfolio, initiated in 2020, was expanded to include topics in hypersonics and artificial intelligence/machine learning (AI/ML). Each of the nine centers in PSAAP III is categorized as either a Multidisciplinary Simulation Center (MSC), a Single-Discipline Center (SDC), or a Focused Investigatory Center (FIC). This reflects the growing needs of the ASC program for expertise in both integrated applications and individual disciplinary focus areas.

Overall, the engagement with academia has evolved over time. While initially the goal was to do R&D on complex predictive simulation capabilities, the engagement grew into a more comprehensive and synergistic relationship. Many of the algorithms and physics modules that would become standard to ASC applications had their origins at universities, eliminating the need for the laboratories to develop them in-house and bringing in a wider diversity of ideas (Advancing Simulation Science, 2011). The academic collaborations continue to serve as an important pipeline to bring both new staff and ideas into the NNSA laboratories.





Partnerships with Experimental Facilities

Stockpile stewardship intensified the challenge of understanding the extreme conditions associated with nuclear weapons. Previously, the only way to probe the physics of these conditions was via nuclear testing itself. To create effective models without that testing, designers needed new sources of experimental data. This need drove the establishment of several new NNSA facilities including the National Ignition Facility (NIF) at LLNL, the LANL Neutron Science Center (LANSCE), and the SNL Z machine. Together, these facilities continue to yield vital data related to high-energy density (HED) conditions, nuclear forensics, material analysis and other weapons-related purposes. The data help scientists and engineers properly account for the relevant physical processes for the integrated codes, as well as analyzing and improving the performance of model subcomponents.

Integrated experimental measurements were also needed. Several facilities were constructed after the initiation of SBSS that provide integrated weapons data of various forms. Although none of these facilities could completely replace underground testing, they each provide unique and highly accurate experimental data on desired components. A partial list of these facilities includes the Contained Firing Facility (CFF) at LLNL, the Proton Radiography Capability (PRAD) sited with LANSCE at LANL, the Dual-Axis Radiographic Hydrodynamic Test facility (DARHT) at LANL, the Annular Core Research Reactor (ACRR) at SNL, and facilities at the Nevada National Security Site (NNSS). For a more comprehensive list, see the Fiscal Year 2021 Stockpile Stewardship and

Management Plan listed in the References. A fundamental impact these facilities make on current weapons assessment is through the integrating ability of the ASC multiphysics codes. Experimental data for physics models affects analysis in the form of improved physical models and through providing integrated weapons data for comparisons and other purposes.

Government-wide Approach to HPC

In 1991, Congress passed the High-Performance Computing and Communications Act to enact recommendations from a 1987 report by the White House Office of Science and Technology Policy (A Research and Development Strategy for High Performance Computing, 1987). This led to the creation of the DOE Office of Science (SC) High Performance Computing and Communications (HPCC) Program in 1994, which eventually led to the formation of the DOE Office of Advanced Scientific Computing Research (ASCR) Program in 1999. DOE SC/ASCR currently supports multiple computing sites including Oak Ridge National Laboratory (ORNL), Argonne National Laboratory (ANL), and the National Energy Research Science Center (NERSC) at Lawrence Berkeley National Laboratory (LBNL). LBNL, ORNL, and ANL host very large supercomputing centers, managed as user facilities to provide academic and industrial users access to machines of the same caliber as those sited at the three NNSA laboratories (ASCR@40, 2020). DOE SC/ASCR and NNSA/ASC have together accumulated over two decades of R&D collaboration on HPC platforms, software, and algorithm investments. The capstones of this inter-program collaboration (as of 2021) are the ongoing Exascale Computing Project (ECP) and joint HPC procurement programs involving both NNSA and DOE SC laboratories, which are projected to field three exascale platforms between 2021 and 2023.

The computing and industrial drivers for HPC behind the formation of ASCI affected other parts of government outside of DOE. The National Science Foundation (NSF) and National Aeronautical and Space Administration (NASA) also had longstanding HPC efforts prior to ASCI. Many of these efforts have influenced the hardware and software developments of ASC. NASA developed cutting-edge meshing and computational fluid dynamics algorithms at its computing centers. NSF developed five supercomputing centers to directly support universities. The National Science Foundation Network (NSFNET) was created in 1985 to link these centers together through the communication protocols developed for the Advanced Research Projects Agency Network (ARPANET). The commercialized version of NSFNET would go on to become the main backbone of what is now the internet. The parallel development of Hypertext Markup Language (HTML) and the initial Mosaic web browser from the NSF-sponsored National Center for Supercomputing Applications (NCSA) would launch a revolution in how the world uses computers today.

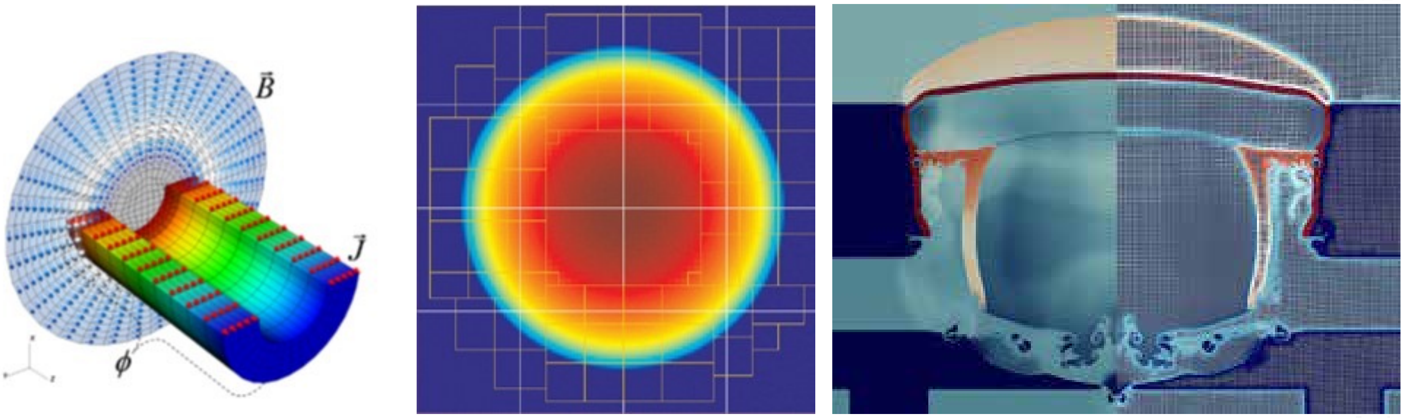
MULTIPHYSICS APPLICATIONS ENABLE DESIGN & ANALYSIS

Spent Fuel Fire Simulation. (SNL)

A fundamental operating principle of ASCI was to avoid the tendency toward defining any problem too narrowly and thereby hindering insight and innovation. One formalized approach to this lay in considering nuclear weapons as "systems," as opposed to individual components. This resulted in the creation of an Integrated Codes (IC) subprogram. Prior to ASCI, the small individual weapon codes were often maintained by only a few people. But simulating a full nuclear weapon system called for much larger codes, incorporating multiple types of high-fidelity physics models. This in turn required larger, integrated teams of subject matter experts (SMEs) with enhanced training for expertise in more than one field.

This new coupling of physical processes was much more complex to verify and validate than the previous simplified set of physics. It also created new demands on the underlying data that was used to support these models. Before ASCI, models had primarily been used to guide analysis and design. After the formation of ASCI, models became a key element in the stockpile certification process. To inform these new models, new experimental facilities such as those described earlier provided underlying data on various aspects of weapons behavior that had not been fully understood during the testing era.

Developing multiphysics applications was a daunting endeavor. The applications were required to simulate many complex behaviors, from the passive aging of materials over decades, to the workings of electromechanical devices, and behavior of the nuclear materials exposed to extreme conditions (such as those found inside a star). Each laboratory pursued applications that focused on their mission areas. SNL focused on engineering challenges and the associated physics needed to qualify complete systems. This resulted in the development of the SNL SIERRA code suite for engineering mechanics and the RAMSES code suite for radiation and electrical systems. LLNL and LANL focused on the extreme hydrodynamic conditions associated with explosively driven

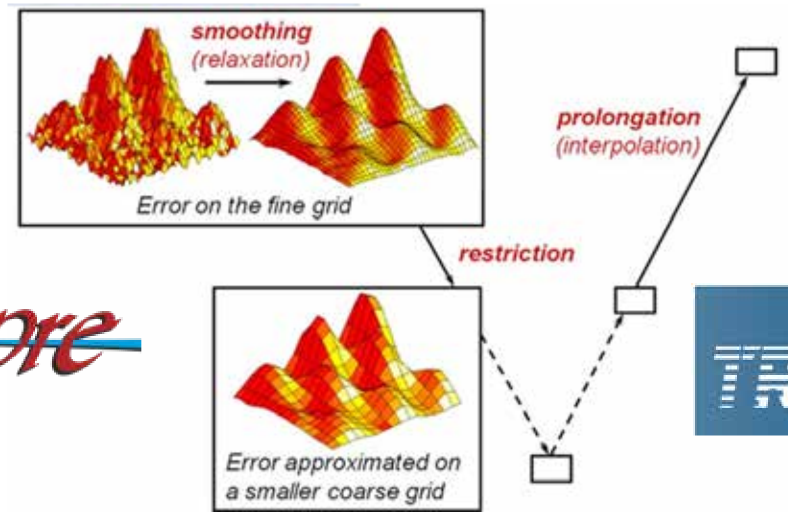


ASC multiphysics applications integrate multiple computer models of isolated physical processes to better understand how those processes interact in certain situations. The figure at left, for example, illustrates motion generated in a cylinder via electric and magnetic fields. The center image illustrates the radiative transfer of energy by high-energy photons in an experiment related to inertial confinement fusion. The simulation at right depicts a detonation wave passing through multiple layers of material.

materials and the HED physics encountered in nuclear weapons.

These new applications would couple together complex physics models, such as the coupling of photon radiation transport and magnetohydrodynamics with strength and other material behaviors. It was not possible for each code project to support all of the physics it would need to incorporate. This resulted in physics models being implemented as distinct packages or software libraries that could be shared among multiple multi-disciplinary codes. The code efforts behind these physics models would eventually grow to rival their supported integrated codes in terms of programming effort and staff commitment. Aggregating these resources allowed sufficient flexibility for the physics packages to be predictive. The ASCI founders realized that the ability to excessively "tune" parameters would have severely limited the utility of codes to simulate new, potentially threatening situations for the stockpile, by saddling new codes with the calibrated physics derived from the testing era. An emphasis on developing models based on fundamental understanding of the underlying physical processes would eliminate this problem of over-reliance on empiricism and calibration.

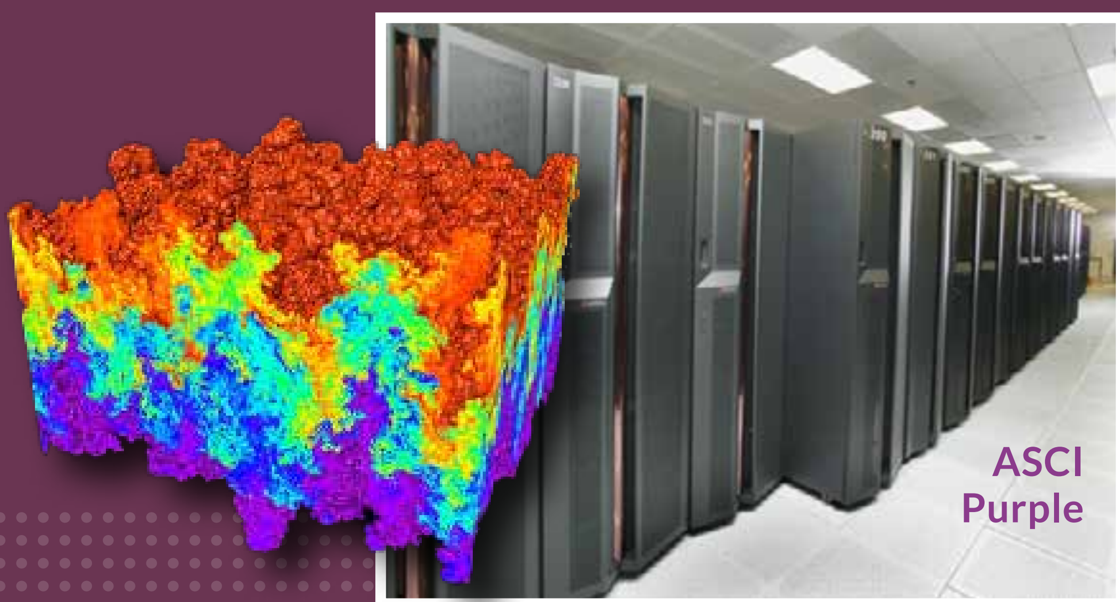
The leadership team also realized that teams of dedicated computer scientists and mathematicians were needed to implement integrated codes on advanced platforms. At the start of ASCI, the programming models needed to implement the physics on large cluster computers were still being formulated. It was also realized that surrounding the integrated codes was an ecosystem of problem setup, analysis, and visualization tools. The high level of numerical resolution required the invention of new mathematical algorithms that would scale to large numbers of processors and would not suffer from stability and efficiency issues as they were pushed to new extremes.



The solver packages Trilinos (SNL) and Hypre (LLNL) make it possible to efficiently solve the large systems of linear and non-linear equations used in multiphysics applications. Both packages enable use of the multi-grid method, combining multiple solution techniques to arrive at models containing high-resolution representations of relevant features.

The evolution of compute capabilities for simulating turbulent flows of thermonuclear fusing materials.

The ASCI Purple machine and a planar turbulent simulation from this platform. ASCI Purple was the first machine capable of true multiphysics simulations of turbulent thermonuclear fluids due to its peak performance estimated at 100-teraFLOPS. (Barney, 2010)



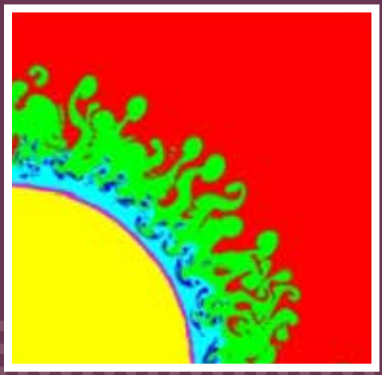
Thermonuclear Burn Initiative (TBI)

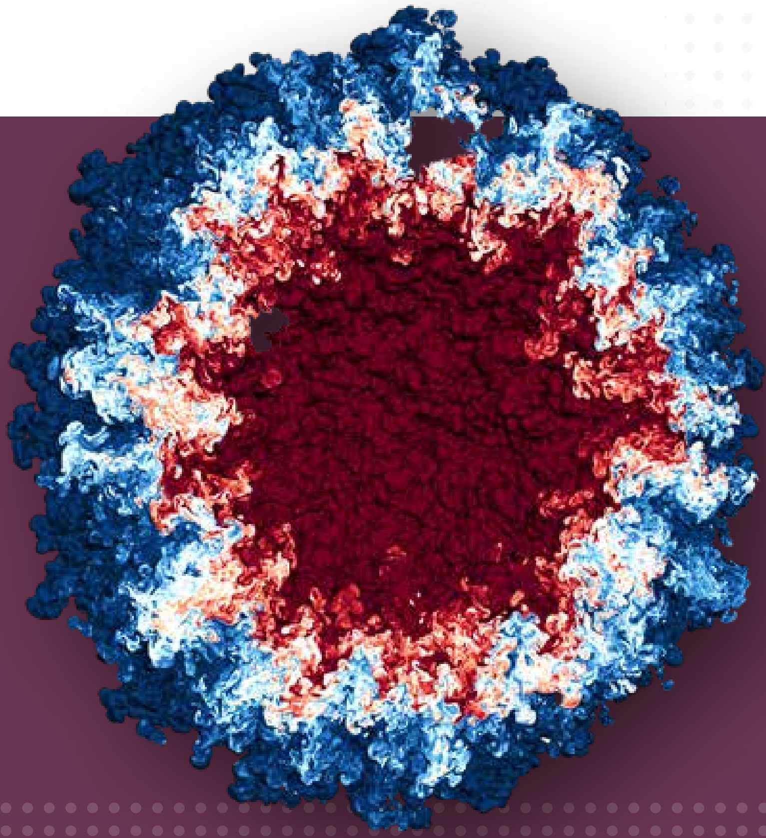
Fusion or thermonuclear burn is an example of the need for large multiphysics codes. Early, high-resolution simulations on the ASCI Purple system provided tantalizing insights, which resulted in the creation of the Thermonuclear Burn Initiative (TBI).

Much of the simulation difficulty stems from needing to better understand fusion in a turbulent environment. Turbulence is notoriously challenging to simulate at the resolutions required in even the most benign of fluids. The addition of reactive physics required for thermonuclear burn makes this problem even more difficult. The history of this problem parallels that of ASCI and the ASC program itself. Initially, ultra-high-resolution 2D calculations provided insights that then needed to be followed up with much more involved 3D simulations requiring even more sophisticated multiphysics capabilities. Several initial successes occurred with the initial ASCI machines up through the ASCI Purple machine, a 93 teraFLOPS machine. (Barney, 2010). The resulting TBI focused on refining these simulations and increasing the understanding of these types of turbulent reacting flows. Between 2000 and 2020, an exponential increase in computing power allowed unprecedented resolution to be applied.

The images below illustrate what a typical simulated turbulent flow looked like in 2000 compared to that possible in 2018. The capabilities of the Sierra platform enabled an unprecedented (98 billion cell) Inertial Confinement Fusion (ICF) motivated simulation of two-fluid mixing in a spherical geometry. The simulation uses a resolution unthinkable at the beginning of the ASCI Program in 1995.

The bottom left figure shows the resolution that was capable of being simulated in 2000 (Accelerated Strategic Computing Initiative (ASCI) Program Plan, 2000). The right figure shows a similar problem executed in 2018 on the Sierra platform, demonstrating more than a 10,000 factor resolution increase.





The SIERRA Code Suite

ASC™

The SIERRA code, developed at SNL starting in 1995, was ASCI's first code suite to be based on a common core infrastructure to support multiple applications.

SIERRA contains a set of application codes for several of SNL's diverse mission areas. SNL realized that maintenance could be minimized and overall functionality increased if its ASC application codes used a common core infrastructure. This enabled SIERRA developers to accomplish several key objectives:

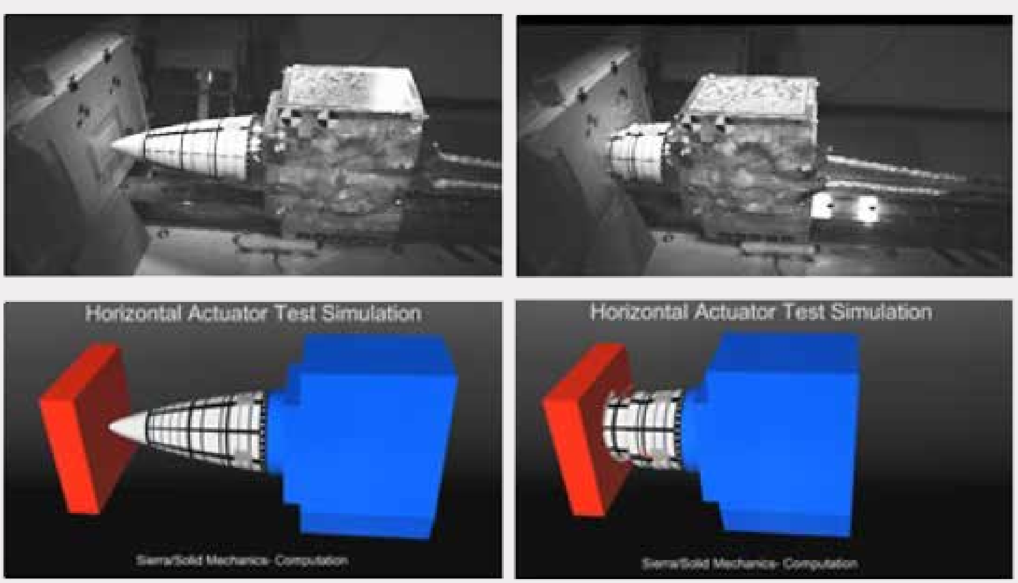
- Interoperability of the various physics modules (thermal, fluids, aerodynamics, solid mechanics, and structural dynamics)
- Built-in three-dimensional parallel scalability
- An integrated set of workflow components (geometry and meshing, boundary condition generation, and input management and processing)
- Focus and priority on software quality assurance (SQA), testing, verification, and validation
- Software development tools and infrastructure supporting developer productivity
- Application-aware development (tailored code features driven by use cases)
- Advanced and highly flexible C++ software infrastructure shared by the physics modules

A key early success for the codes suite came when engineers used SIERRA to help them understand several unexpected physical test results involving a safety component in the W76-1 weapon system. A part within the component was failing unexpectedly, and engineers struggled to design experiments that would help them fully understand what was happening. A finely detailed simulation created with the SIERRA solid mechanics module was able to reveal why and under what conditions the part could fail. This understanding, coupled with a large number of qualification simulations and experiments, allowed final certification of the mechanism. This was an early example of the insight that predictive simulation can provide for a complicated 3D problem. In the years since, the SIERRA code suite has been applied successfully to a multitude of important problems, including identification of previously unknown thermal safety failure mechanisms, early predictions of the environment inside the F-35 bomb bay before the new fighter was available, and quick evaluation of requirement changes to address production limitations of the

THE SIERRA CODE SUITE

nose-bomb subassembly for the B61 LEP. The flexibility and interoperability originally built into the SIERRA code suite continues to have benefits. Focus areas for current and future development are:

- Next-generation simulation targeting use of the suite in engineering design with an eye towards qualification
- Implementation of highly scalable algorithms needed to run effectively on next-generation HPC systems like Sierra, Crossroads, and El Capitan
- Increased collaboration with PEM and Verification and Validation (V&V) subprograms towards improved credibility
- Prediction of complicated combined environments, including re-entry for the upcoming W87-1 and future weapon systems



Comparison of experimental results with a mechanical computational model using SIERRA solid mechanics in support of weapons qualification.

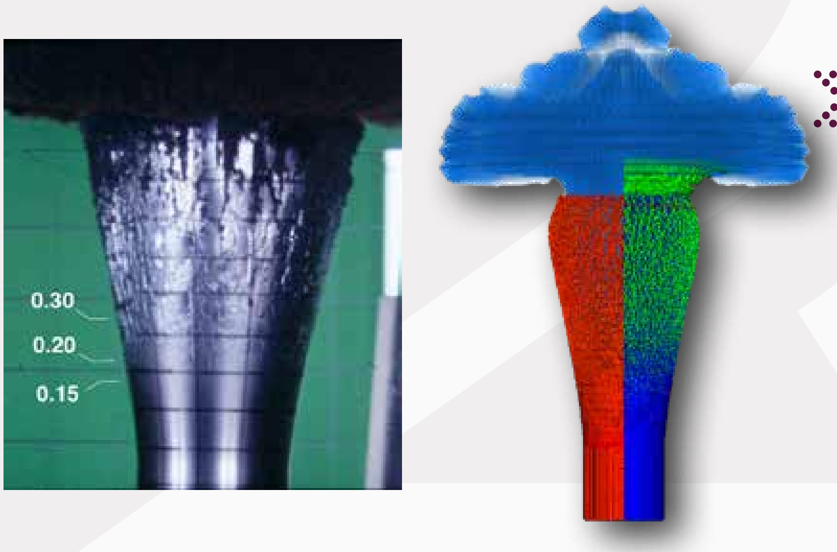
Sandia National Laboratories

Hydrodynamic Algorithms

HED physics are extremely challenging to simulate due to the wide range of densities and energies involved.

This required the development of a hybrid form of hydrodynamics to simulate the evolution of an experimental object from its initial room temperature solid state where strength is important to a regime that is potentially several times its initial density and tens of millions of degrees in temperature. The codes developed at LANL and LLNL often used Lagrangian methods, by which an object was discretized into a mesh of discrete elements that deformed when put under stress according to its material properties. These methods, which track the evolution of solid bodies typically associated with engineering solid mechanics, were then combined with Eulerian methods, where the material is able to flow through the mesh as if it were a fluid, such as water. The combined Arbitrary-Lagrangian-Eulerian (ALE) method allows complex multiphysics regimes associated with a weapon to be routinely performed in a seamless fashion. ALE methods have evolved substantially over the 25 years of the program, as seen by the figure of the ALE shaped charged jet simulation from the LLNL MARBL code. The images depict the evolution of the jet from its initial static geometry to its final fluid state. Here, both the inner liner and explosive are flowing through the simulated mesh, which is still evolving (Rieben, 2020).

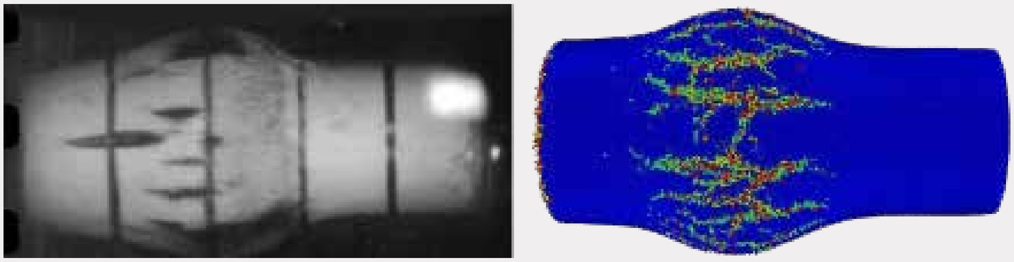
Another form of material motion that ASC is pursuing involves particle-



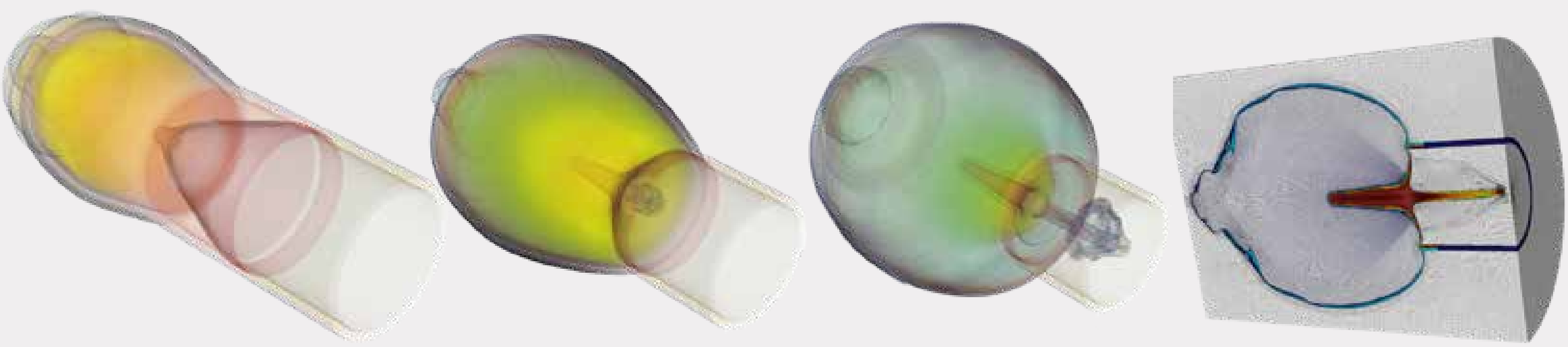
Simulation showing the explosively driven fragmentation of an AerMet steel cylinder at 25 microsec, with an image of the actual experiment for comparison. In the experiment, the high explosive is lighted at the top of the cylinder, and the burn front travels downward, expanding the explosive outward and fracturing the cylinder. In the simulation, the left side (red) shows the mass density of the intact material, while the right (green) shows the material damage.

HYDRODYNAMIC ALGORITHMS

based methods. The figure (right) shows an explosively driven cylinder that is fragmenting along with a simulated version of that cylinder. The simulation matches the fragmentation of the cylinder via a collection of particles that can accurately track the mass and velocities of the cylinder fragments. This form of simulation was impractical at the start of ASCI, due to the immature methods employed and the restricted ability of the platforms to simulate enough particles to be meaningful. Advances in methods such as smoothed particle hydrodynamics (SPH) and access to petascale platforms such as Sierra at LLNL have made it possible to routinely perform these simulations and develop good verification datasets (Owen, 2018).



The comparison between an image of an explosively driven cylinder and a simulation using the SPH method.



The evolution of a shaped charge jet as simulated by MARBL.

Lawrence Livermore National Laboratory

Turbulence Modeling

ASC™

Turbulence is another field of study for ASC. Many materials behave as turbulent fluids when placed under extreme conditions.

The three laboratories have developed sophisticated methods for simulating turbulence. These methods use accurate mathematics coupled with subgrid turbulence models to account for unresolved motion below the scale that the computational mesh can resolve. Often, they employ Adaptive Mesh Refinement (AMR) to concentrate numerical resolution in areas where needed to resolve important features such as material interfaces and boundary layers.

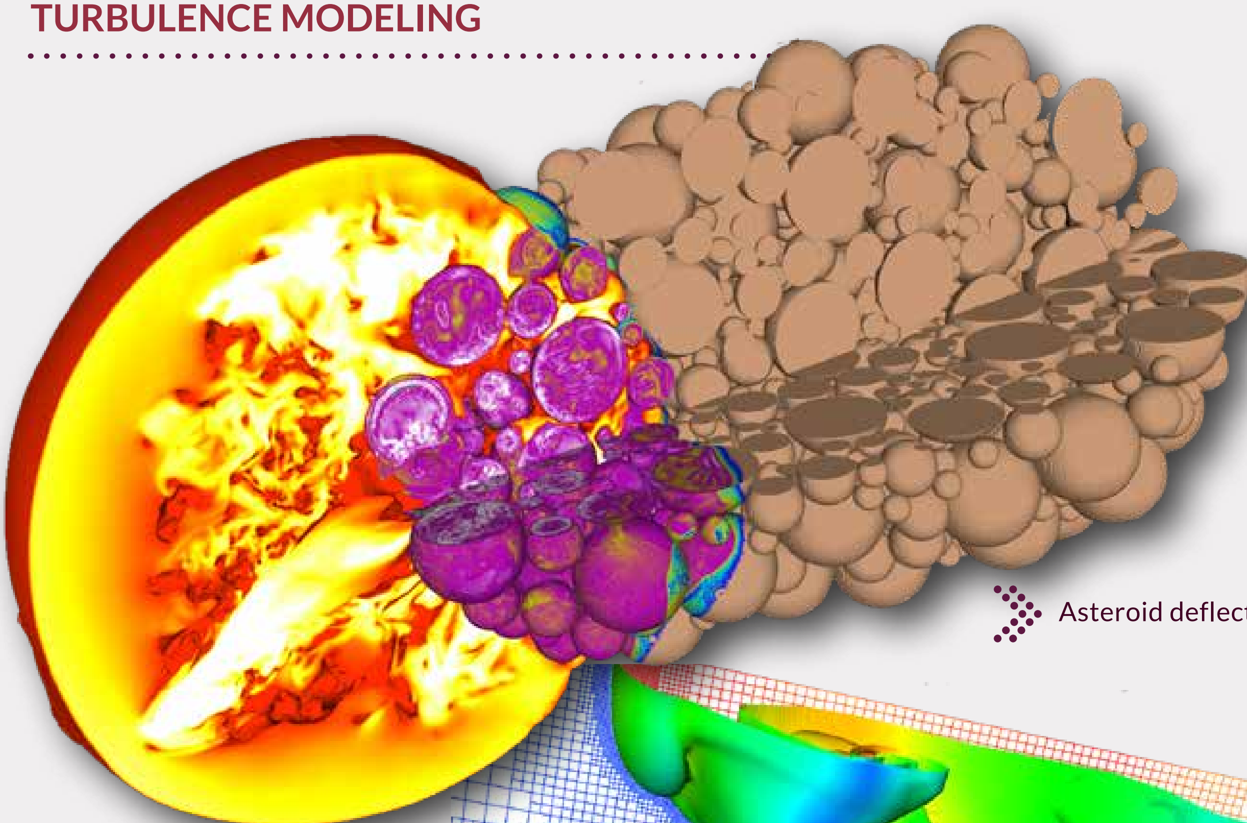
The figure (on page 32) shows a 3D single-mode Richtmyer-Meshkov instability growing on the interface between two gases as this interface is first shocked and then re-shocked by a reflected shock. The sides of the figure show the AMR structure of the underlying 3D mesh and the interface surface is colored by velocity. This is from an xRAGE problem run on LANL's first ASC platform, the Blue Mountain machine circa 1999 (Weaver, 2002).

The extremely fast growth in simulation capabilities and computational performance is shown on page 32 (top figure). This plot shows a 3D xRAGE simulation of an asteroid deflection using the blast and thermal radiation from an explosive energy source to impart an impulse to the asteroid. The simulated asteroid consists of a large rock pile intended to represent the structure of the real Earth-crossing Apollo asteroid 25143 Itokawa (Weaver & Dearholt, 2011). The image shows the propagation of the pressure wave launched through the rock pile by the burst. This simulation was performed on the Cielo machine circa 2011.

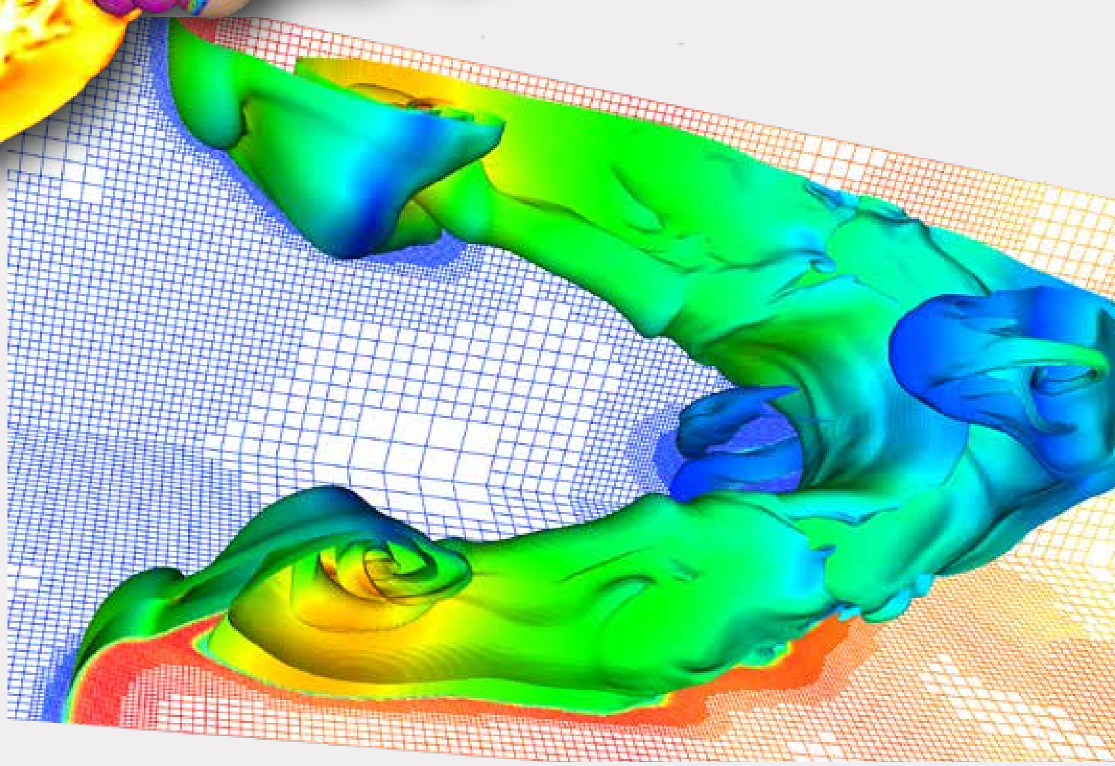
Both simulations illustrate the complexities of fluid flows encountered by the ASC codes. These range from understanding instabilities in flows that start as laminar flows and turn turbulent, as well as flows resulting from initially solid objects undergoing extreme deformations.

Los Alamos NATIONAL LABORATORY EST.1943

TURBULENCE MODELING



Asteroid deflection simulation



Simulated Richtmyer-Meshkov instability

Visualization

The end products of the integrated codes lie in the analysis of their results and generated technical insight.

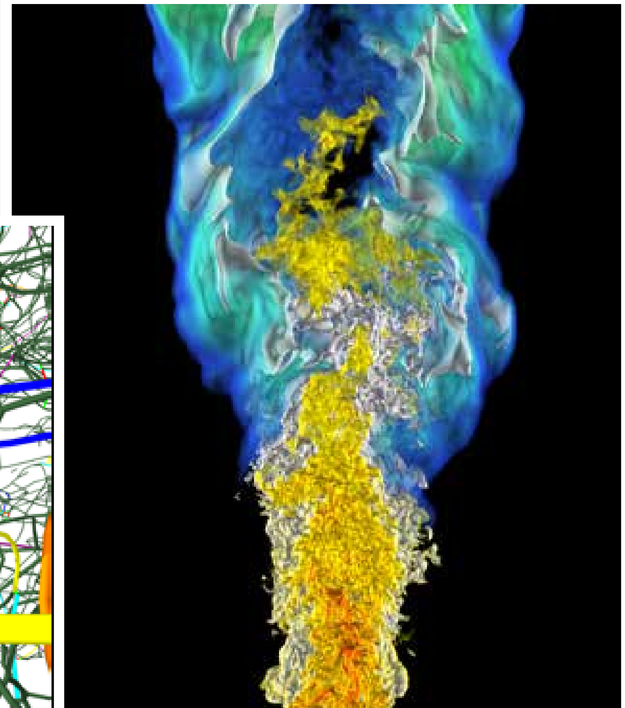
The ASCI founders realized that multiple efforts were required to build a robust analysis workflow around the integrated codes. This took the form of VisIt at LLNL, ParaView at SNL and LANL, other LANL-developed tools, plus the commercial product Ansys EnSight. These products accounted for the unique needs of each laboratory, while providing scalable visualization tools that could be used outside of the laboratory as well. Visualization is worth noting—although it is separate from the supercomputer application of an integrated code, it still requires interaction with that code. Integrated code users are often sitting at a personal desktop far removed from the platform that is deploying thousands of processors for their simulations. This leads to the need for common ways of communicating graphical and analysis data across networks and between different types of machines.

There have been several successful technology spinoffs from ASC visualization efforts. One was the widespread adoption and use of common binary data formats, such as Hierarchical Data Format (HDF). Another was standardization of the formats used to communicate graphical data. The open-source Visualization Toolkit (VTK) format, now supported by Kitware, was widely adopted. Another advancement in supercomputer visualization by ASC was the development of scalable visualization algorithms, which made it possible to explore highly parallel datasets. Scalable visualization algorithms greatly increased the interactivity of supercomputing applications, which had previously been highly serialized and cumbersome.

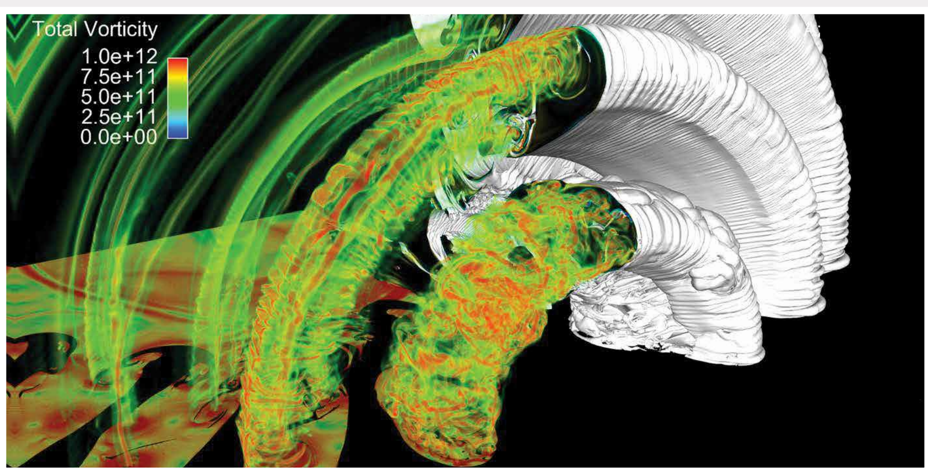
Los Alamos NATIONAL LABORATORY EST.1943 — Lawrence Livermore National Laboratory — Sandia National Laboratories

VISUALIZATION

Rendering of a dislocation multijunction discovered in a ParaDiS dislocation dynamics simulation. Multijunctions are new topological dislocation configurations discovered through simulation that are key to understanding the hardening of body centered cubic alloys. (LLNL)



Direct numerical simulation reveals turbulent combustion of a lifted ethylene/air jet flame in a heated coflow of air. (SNL)



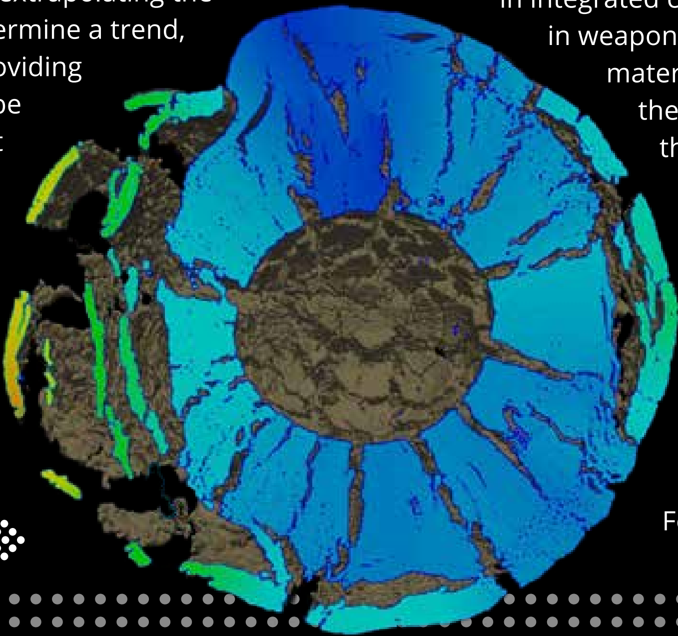
One billion cell 3D xRage simulation showing the turbulent development of vortex tubes in an ICF capsule. (LANL)

PREDICTIVE CAPABILITIES FROM MATERIAL MODELS

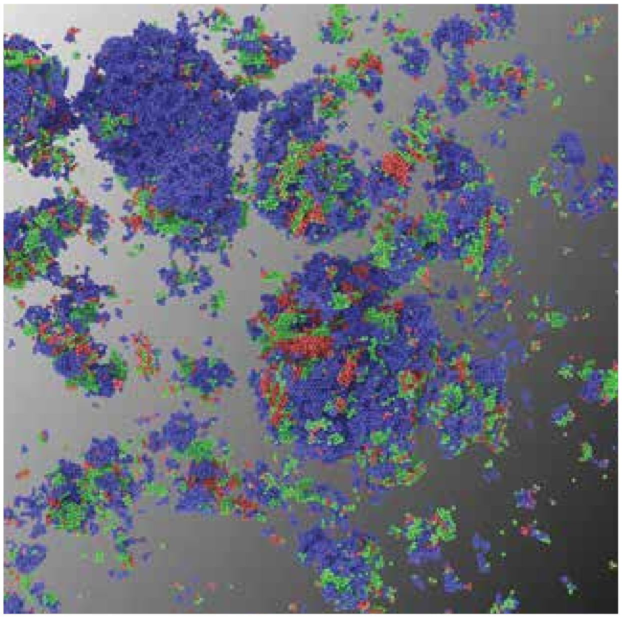
Accurately simulating the behavior of different materials in complex and extreme conditions has been integral to the success of ASC. The need for better material models led to creation of the Physics and Materials Modeling subprogram of ASCI, later renamed as Physics and Engineering Models (PEM). At the start of ASCI, understanding the fundamental behaviors of materials was often achieved through empirical methods, and the simulation of those materials required tuning or "fudge" factors that calibrated models to given experiments. This was useful for extrapolating the behavior of an experiment to determine a trend, but was wholly inadequate for providing a predictive capability that could be applied to conditions that had not already been observed.

Simulations require knowledge of a wide variety of material behaviors. Materials exhibit several different phases. For example, a solid material may rapidly melt and recrystallize into a different solid. Or a material might become ionized, requiring it to be modeled as a plasma. The strength of materials can also change substantially, depending on how they are manufactured, independent of their material composition. Depending on the application, the compressibility of the material may also be important, requiring an understanding of the equation of state. These strength and compressibility properties are often tabulated using PEM codes to explore a range of pressure and temperature states and released as tables for use in integrated codes. Material behavior in weapons depends on how the materials are manufactured, how they age, and how they change through different insults, such as being dropped or heated in a fire. All these factors place additional requirements on material models.

The behavior of a material extends beyond its mechanical response. For applications in which



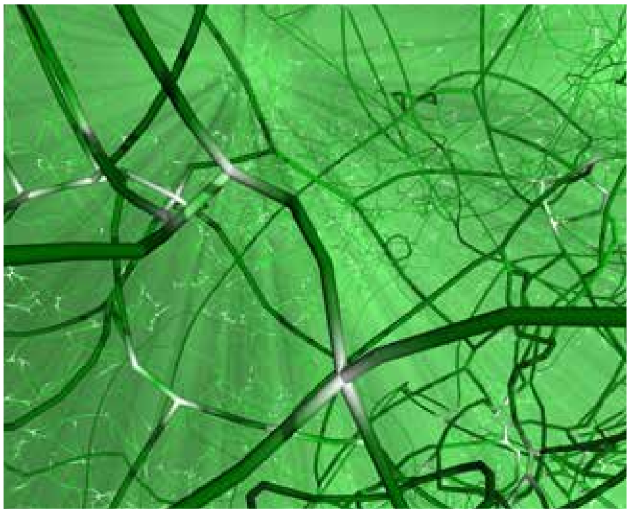
Golevka Asteroid Explosion Simulation (SNL)



Molecular dynamics models simulate collections of atoms to expose phase transitions, such as the freezing of liquid metals in extreme conditions. (LLNL)

radiation is an important factor, the modeling of opacities is needed to estimate the ability of radiation energy to flow from one location to another. Nuclear cross-section data is also needed for several applications. Nuclear data is provided to the outside community through databases such as the Evaluated Nuclear Data File (ENDF) from LANL and the Evaluated Nuclear Data Library (ENDL) from LLNL. A new international standard for nuclear data, the Generalized Nuclear Database Structure (GNDS), has been born out of PEM efforts. These databases provide an invaluable data source used by the nuclear power industry, as well as other industries.

ASC multiphysics codes also require the ability to simulate chemical effects. High explosives (HE) operate on scales ranging from hours in the case of thermal cookoff experiments to microseconds for a high-order detonation. The Cheetah kinetics model from LLNL and SESAME data tables from LANL are widely used outside of DOE. The DOD widely uses ASC models for HE in numerous conventional munitions applications.



A dislocation multijunction discovered in a ParaDiS (LLNL) dislocation dynamics simulation. These simulations provide insight into how materials fail when put under stress.

VERIFICATION & VALIDATION ESTABLISH CONFIDENCE

Simulated threat to coastlines from an asteroid ocean impact. (LANL)

Trusted analysis capabilities are needed to enable stockpile stewardship in the absence of underground testing. This requires methodologies for systematic analysis of the stockpile. One concern is that analysis is only as good as the connection to the underlying reality. This led to the development of new experimental facilities, like those described earlier, that directly measure the physical properties used in the physics models. Additional experimental facilities arose from testing weapon components or sub-assemblies in isolation, which improved confidence in combined multiphysics simulations.

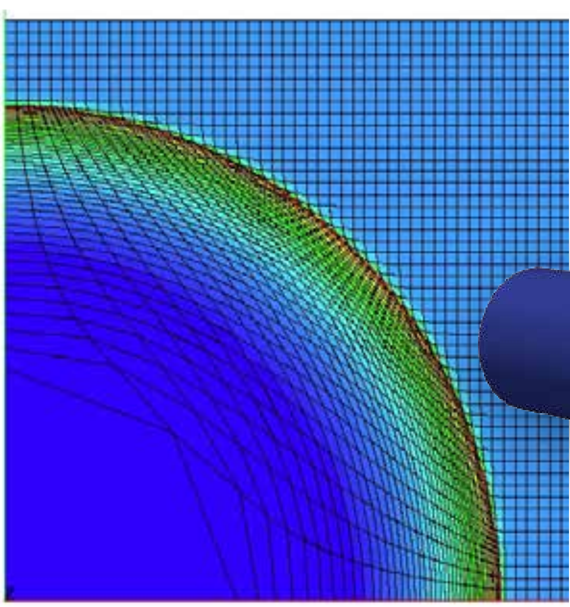
The developers at each laboratory designed methods for establishing confidence in the predictive abilities of their code through implementation of standard verification and validation (V&V) methodologies. Verification is often easier to pursue because it uses analytic mathematical solutions to check individual algorithms in a model for accuracy and correct implementation. Often, it is even possible to verify correctness of an algorithm from analytic estimates of how fast the algorithm converges as resolution is increased.

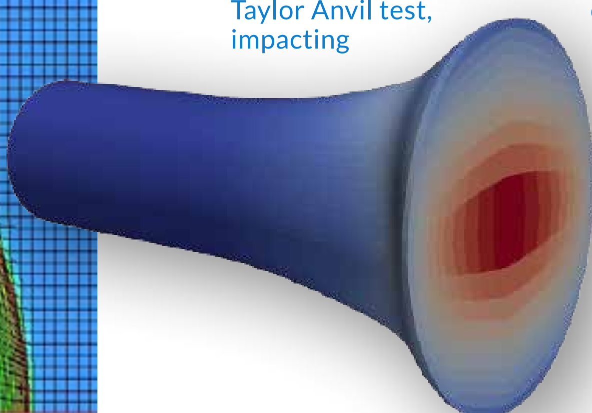
Validation is more subjective as one measures how relevant the implemented algorithm is to simulating reality. Validation is typically conducted by measuring the difference between physical experiments and their simulated results. Throughout the history of ASC, major milestones have been achieved to improve integrated code validation. A fundamental tool of validation is the use of common model frameworks. These are suites of simulations targeted at collections of experiments. By simulating these suites with a fixed model configuration, the overall accuracy of the model in predicting the results without tuning for each experiment allows understanding to be applied across multiple problems of interest. This fixed configuration establishes the validity of the model for predicting outcomes of interest without bias from a single experiment. It also ensures applicability over a wide range of conditions.

Uncertainty Quantification (UQ) is the science of understanding simulation uncertainties that are caused either by unpredictable conditions in the environment or by approximations in the models themselves. UQ is the natural complement to V&V, in that UQ is required to establish confidence in the ability to apply results from multiphysics codes to weapons-relevant problems. V&V establishes the precision of what was discretized, while ensuring that it was the correct thing to discretize. By itself, V&V does not fully quantify how well a model is related to the real world. Every model is an imperfect analog. Successful UQ determines the level to which models can be trusted and used in a predictive fashion (Bishop A., 2009) (Assessing the Reliability of Complex Models, 2013). The ASC program has demonstrated successes in applying V&V and UQ methodologies to nuclear weapons and has enabled the ongoing evolution of stockpile stewardship from refurbishment of legacy systems to their replacement as appropriate with modernized systems.

Reanalysis of the UGT Database

The U.S. conducted more than 1,000 UGTs during roughly 40 years of nuclear testing. The database of test results is a priceless asset that is still yielding new insights, despite nuclear testing having ended almost 30 years ago. By the end of nuclear testing, a single test could cost several hundred million dollars to complete, due in part to the wide variety of diagnostics that the highly skilled experimentalists were able to field. Much of this direct expertise has now been lost through personnel retirement and aging out of legacy technologies. The ASC multiphysics codes play a unique role in interpreting these tests and helping to preserve this heritage. Because the models are mathematically self-consistent, they often uncover misunderstandings in the historical record and explain phenomena that were observed, but not previously understood. This is often done using rediscovered and reanalyzed UGT data that laboratory experimentalists continue to mine from the historical record.





Based on an outgoing spherical shockwave, the Sedov problem (left), is a common test for model validation. Comparing models against Sedov's analytic solution can identify errors that point to defects in a simulation (Sedov, 1993). Another form of model validation, the Taylor Anvil test, consists of a moving cylinder impacting a rigid plane (right). Comparing deformation in the simulated cylinder against results of a live experiment enables validation of the hydrodynamic algorithms used in the simulation (Hungerford, 2021; Zecevic, 2021).

A temporary modular tower designed to hold the diagnostic rack for the Icecap underground test is a current landmark at the Nevada National Security Site (left). The last U.S. underground test was the LANL Divider event on September 23, 1992 (right).

Today, multiphysics models underpin their UQ analysis with this testing database. Linking these models to the stockpile test basis has been a guiding principle to keep model development focused and relevant. The large diversity of nuclear tests also enables these models to be valid across a wide variety of scenarios, which provides confidence in their analysis products. The increasing sophistication and power of multiphysics codes, in terms of their physics and their ability to incorporate high-resolution 3D features, has brought new insights from the testing database. Due to their improved physics, modern models can now be compared against more advanced experimental diagnostics, thereby explaining even more of the historical record. The investment in ASC multiphysics codes has been extraordinarily effective at maximizing the value of the UGT record.

ASCI Red

EARLY SUCCESS WITH MASSIVELY PARALLEL COMPUTING

At the start of ASCI, the concept of “massive” HPC platforms was easier to propose than implement. The technologies involved—collections of commodity processors linked by commodity networks, running general purpose operating systems and analysis software—had never been assembled at such scale and robustness. To minimize risk and maximize the chances for a successful parallel supercomputer, ASCI pursued three separate projects, each with its own architecture, vendor, and laboratory team. These were ASCI Red, sited at SNL in 1996, with Intel as the system integrator; ASCI Blue Pacific, sited at LLNL in 1998, working with IBM; and ASCI Blue Mountain, sited at LANL in 1998, working with SGI. These platforms proved integral in developing the applications needed to support each laboratory’s stockpile stewardship mission. These three machines would go on to win multiple Gordon Bell awards, as described in the Recognition & Awards section of this document. They were key in providing the computing environment to produce the first results from the ASCI integrated codes.

PathForward Program

To help develop the follow-on machines, ASCI implemented the PathForward program which funded R&D partnerships with industry to develop critical technologies needed to make massively parallel clusters possible. Technologies were chosen for this program in areas where it was either not economically feasible or prioritized for private sector commercialization. It was desirable that the chosen technologies would be in areas that industry might be able to turn into future products and markets with additional government investments. This would enable an enduring supply chain of technology upgrades based on commercially viable components. An ambitious program that started in 1997, PathForward, focused in 2000 on the development of technologies that would be required for 30+ teraFLOPS systems, when the highest computer performance was around three teraFLOPS (Larzelere, 2009).

One successful collaboration was the Lustre parallel file system, developed in the early 2000s, and still in continuous development and use today. As a major advancement, Lustre made it possible for thousands of processors to read and write to a unified file system.

Another critical technology was the interconnect, the communications network between the nodes of a massively parallel supercomputer. A persistent challenge with these machines lies in finding an appropriate balance between memory, Input/Output (I/O), and processor speed. If the interconnect is too slow relative to the processor, then the advanced capabilities of the processor will be wasted due to time lost in communication. An early successful collaboration was the Quadrics interconnect, which became a widely used technology for building low-cost UNIX clusters. PathForward contributed to many of the technologies behind open-source, Linux-based clusters, which made the cluster-computing paradigm very appealing and affordable to HPC users outside the DOE lab system, such as universities and other research institutions. Although halted in the mid-2000s, PathForward was later reborn as a joint collaboration between ASCR and ASC that would make important contributions to the development of exascale computing.

ASCI Red

An extraordinary leap in performance in very little time, ASCI Red was the first ASCI system delivered (1996) and the first system to break the one teraFLOPS barrier—a seven-fold increase in performance in only two years—three years earlier than expected. This advance demonstrated ASCI's ability to promote rapid platform advancement, even in a radically altered computing market (Christon, 1997).

ASCI achieved this feat through a contract with Intel, using the Intel Paragon platform and more than 9,000 off-the-shelf Pentium Pro microprocessors, each with its own memory. The contract was really an R&D partnership, with Intel taking considerable technical risk and producing a system that competed with many of its existing customers. Despite skepticism that such a large collection of hardware could ever be made reliable, ASCI Red proved to be remarkably dependable, helping to assuage fears about the feasibility of massively parallel commodity-based systems, and reservations about ASCI's ability to meet extremely ambitious performance goals.

ASCI Red was an extraordinarily successful platform in operation from 1996 to 2006, almost double the lifespan of most supercomputers. The machine's architecture was so successful that when ASCI Red reached the end of its operational life, it was replaced with Red Storm, a similar system that used much of SNL's experience with fielding ASCI Red. The new machine would be built by a reconstituted Cray Inc. and utilized AMD Opteron processors instead of the original Intel Pentium Pros. Being x86 compatible, the Opteron processors allowed SNL to transition its applications easily to the new machine. In Red Storm, the Cray XT3 architecture was introduced and it would comprise a line of supercomputers that would continue up through 2021. Red Storm was upgraded in 2006, resulting in a #2 ranking in the TOP500. The machine was further upgraded to XT4 in 2008, doubling its computational speed. As a follow-on architecture, a Cray XT5 at ORNL, Jaguar, would become the TOP500 #1 machine in 2009. Jaguar itself would be upgraded in place to the Titan machine in 2012 using CPUs accelerated by GPUs. This GPU-accelerated architecture is the dominant form of computing architecture today. Experience gained from ASCI Red contributed to more than a decade of U.S. leadership in supercomputing.

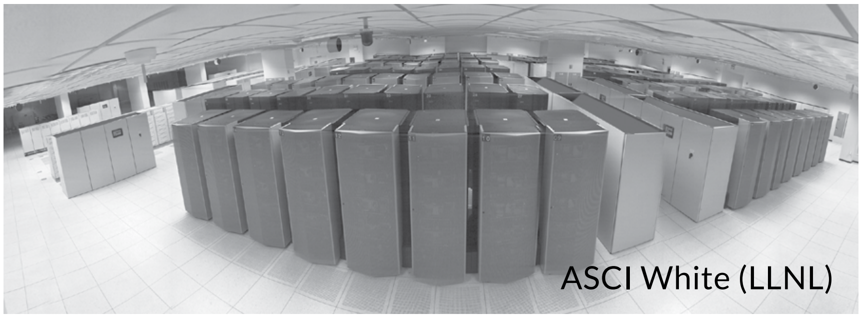
The first ASCI machine, ASCI Red, was the first ASCI computer delivered and the first to break the one teraFLOPS barrier. Its robust architecture allowed it to remain in use for 10 years, roughly double the usual lifespan for a supercomputer.

ASCI Red (SNL)

ASCI Blue Mountain was another early ASCI machine, along with ASCI Red and ASCI Blue Pacific. These machines used thousands of commodity Symmetric Multi-Processing (SMP) nodes to achieve calculation sizes not previously possible.

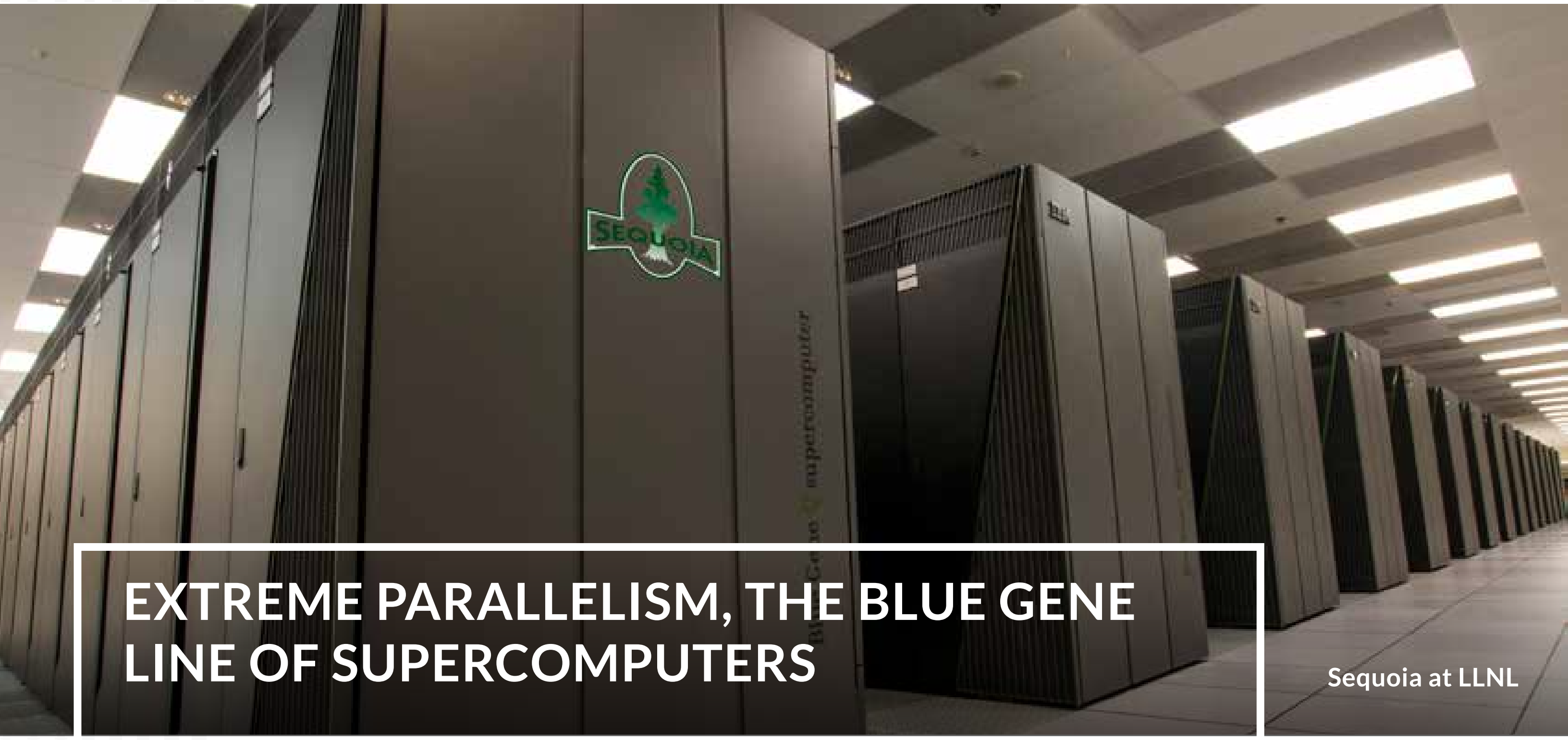
ASCI Blue Mountain (LANL)

The second wave of commodity massively parallel machines included ASCI White, Q, and Purple. More complex and capable than their predecessors, these were the machines that reached the initiative's ambitious 10-year goals, including Purple's achievement of 100 teraFLOPS (Barney 2010).



ASCI White (LLNL)

EXTREME PARALLELISM, THE BLUE GENE LINE OF SUPERCOMPUTERS



Sequoia at LLNL

Although the initial ASC platforms led the world for several years, they were beginning to be eclipsed by machines from external organizations, such as the Japanese NEC Earth Simulator in the early 2000s. This competition pushed ASC and its partners to find new ways to compete. IBM realized that one way of continuing the extremely fast growth in supercomputing was through using hundreds of thousands of medium-performance processors, rather than several thousand high-performance processors, which was the current practice. This is analogous to the revolution that occurred in the mid-1990s, when several thousand commodity microprocessors were able to outcompete custom-built specialty processors.

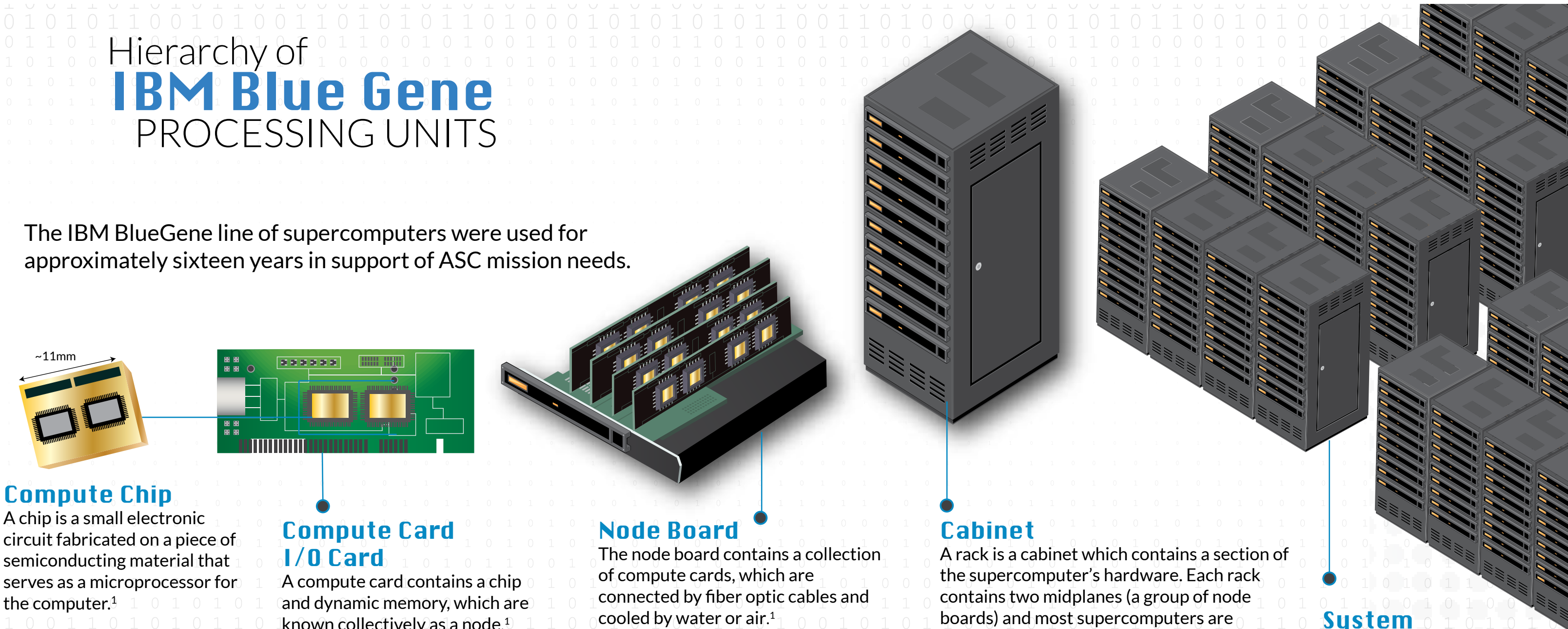
The result of these developments was BlueGene/L, a machine developed through collaboration between LLNL and IBM that shifted from modest numbers of increasingly power-hungry processors to a massive number of low-power PowerPC processors— building on the tremendous gains in distributed-memory computing over the previous decade. The multipart 3D torus network topology enabled a very fast interconnect that was both reliable and could reliably incorporate more than a million processors. This more than overcame the performance tradeoffs of using slower processors. This enabled LLNL's BlueGene/L to remain the #1 system on the TOP500 list for nearly four years after successive expansions, plus become the first architecture to break the 500 teraFLOPS barrier. It also became the first system to achieve a sustained 100+ teraFLOPS on real-world applications (Supinski, B.R., & Louis, S., 2005). BlueGene/L achieved this while taking up only a fraction of the floorspace and power of similarly capable systems. This system was also a success outside of LLNL. BlueGene/L's low-cost and low-power-consumption model made smaller versions of the system attractive to universities, and to government and commercial research centers. This increase in machine acquisitions also helped distribute the cost of development and fostered a larger user base. This system was finally overtaken by the LANL Roadrunner system, described in the next section.

Blue Gene/L at LLNL

This line of computers resulted in the LLNL Sequoia BG/Q platform, which regained the top spot in the TOP500 list from the Japanese K computer. Also of note was its low power consumption, requiring only 7.9 megawatts of power versus the K computer requiring more than 12 megawatts. Sequoia was preceded at LLNL by a prototype BG/P machine called Dawn, which was able to obtain the 9th position on the TOP500, resulting in three LLNL BlueGene-based systems breaking the top 10 list over almost a decade. Sequoia was remarkable in that it reliably fielded more than 1.5 million cores on almost 100,000 nodes. Sequoia was decommissioned in early 2020. This ended an approximately 16-year span, beginning with BlueGene/L, in which a BlueGene computer worked in service of ASC mission workloads.

Hierarchy of **IBM Blue Gene** PROCESSING UNITS

The IBM BlueGene line of supercomputers were used for approximately sixteen years in support of ASC mission needs.



Compute Chip
A chip is a small electronic circuit fabricated on a piece of semiconducting material that serves as a microprocessor for the computer.[1]

Compute Card I/O Card
A compute card contains a chip and dynamic memory, which are known collectively as a node.[1]

Node Board
The node board contains a collection of compute cards, which are connected by fiber optic cables and cooled by water or air.[1]

Cabinet
A rack is a cabinet which contains a section of the supercomputer's hardware. Each rack contains two midplanes (a group of node boards) and most supercomputers are composed of multiple racks.[1]

System
A series of 64 cabinets composed of multiple racks.

[1] https://www.energy.gov/articles/infographic-everything-you-need-know-about-supercomputers-0

Roadrunner at LANL

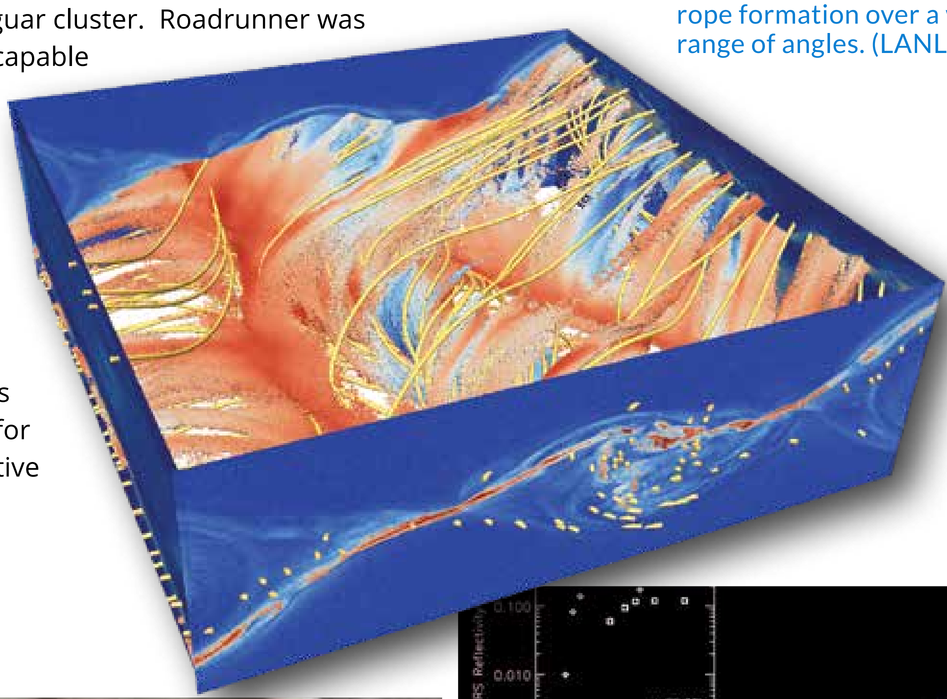
ROADRUNNER & THE INITIAL USE OF GRAPHICS PROCESSORS

Another approach to scalability is to put more parallelism into each processor on a node. This was the approach taken by the LANL Roadrunner system, which pioneered the use of a central processor in a supercomputing node paired with several floating-point accelerators. This approach avoids the reliability issues that may accompany very large machines with millions of independent processing cores. This heterogeneous approach now accounts for a large fraction of the systems in the TOP500 list, including the Summit system (acquired through the DOE Office of Science ASCR program), and the NNSA Sierra system (acquired through the ASC program).

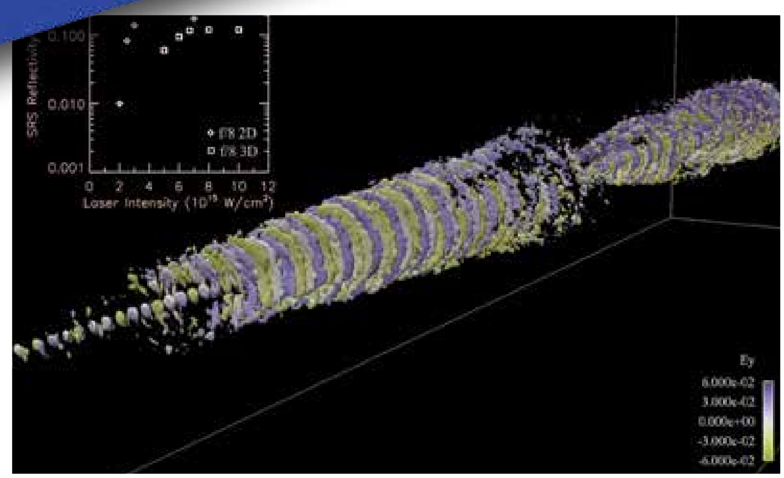
IBM Roadrunner was a pioneer that matched a large number of mathematics acceleration chips with a CPU to provide the bulk of its computing power (Patterson, 2013). The IBM PowerXCell-8i chips on this machine were similar to Cell chips found on the Sony PlayStation 3, but were modified to offer double-precision arithmetic. This system, installed at LANL in 2008, was the first large-scale heterogeneous cluster, using a combination of about 6,500 conventional AMD Opteron processors and about 13,000 IBM Cell chips, which provided the bulk of Roadrunner's processing power.

Utilizing Roadrunner was the start of the difficult but necessary journey for LANL and the ASC program to address the challenges associated with hybrid systems, particularly the challenge of programming for heterogeneous architectures. Hybrid systems offered the benefit of custom-built computational hardware, in which a single chip could support hundreds or thousands of simultaneously executing threads. This approach is effectively the Cray vectorization approach from the 1970s taken to an extreme. However, hybrid systems require special techniques for managing the movement of data to and from the accelerator chips, greatly complicating model development.

Partnering with IBM provided LANL access to the company's Cell-chip line at a time when GPUs were insufficiently mature for scientific computing uses. Breaking the petaFLOPS barrier in 2008, Roadrunner assumed the number one spot on the TOP500 list, and the highest spot for a large computer on the Green500 list. It did this while using less than one-third of the power per petaFLOPS of Oak Ridge's similarly powerful, but conventional, Jaguar cluster. Roadrunner was the first system capable of routinely performing finely resolved 3D calculations for full-scale weapons simulations. This level of performance was a key milestone for achieving predictive 3D simulations.



Open boundary simulations on Roadrunner in a guide field geometry feature highly elongated electron current layers that are unstable to flux rope formation over a wide range of angles. (LANL)



The Open Science period on Roadrunner (2009) demonstrated that the initial, difficult reprogramming effort rewarded scientists, as with this record-breaking 400-billion particle simulation of laser-light scattering. (LANL)

Roadrunner in the Nicholas C. Metropolis Center for Modeling and Simulation, 2010. (LANL)

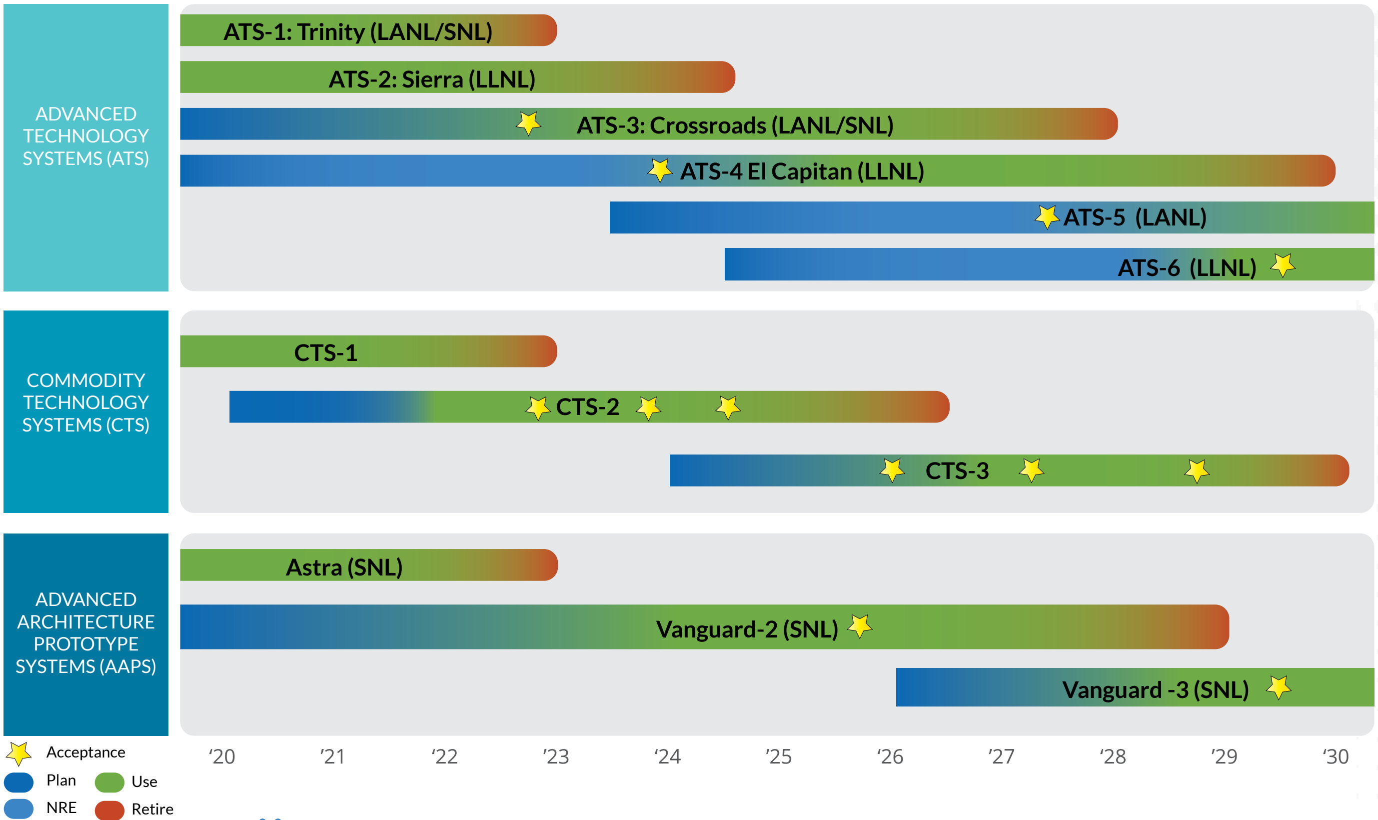
CHANGES IN PLATFORM ACQUISITION

Magma (CTS-1) at LLNL

The success of parallel clusters, commodity processors such as the x86 architecture, and the ubiquity of Linux as a stable operating system, allowed for the deployment of relatively inexpensive, large-scale Linux clusters. This enabled ASC to provide additional computational resources for the increasing demands of the design community. Coupled with the continued need to push the boundaries of computing at the high end, ASC formulated a platform strategy in the early 2010's built around two classes of systems (ASC Computing Strategy, 2013).

The first type of system is the Advanced Technology System (ATS). These are machines at the leading edge of computing and are aquired to solve the largest and most challenging problems. These capability computing machines are first-of-a-kind platforms that take on more technological risks and require much more engagement with the system and component vendors. Often, these machines would be in non-recurring engineering (NRE) development for several years before being deployed at either LANL or LLNL for the very first time.

The second type of system is the Commodity Technology System (CTS). These systems provide a large amount of computing via an umbrella, open contract with a 4-year window, by which multiple CTS machines can be procured by the three NNSA labs and other sites. The CTS line was formalized in the ASC platform strategy after two successful cycles of the Tri-Lab Linux Capacity Cluster (TLCC procurements initiated in the mid-2000's, in which the labs first explored and deployed large Linux-based systems with the goal of having similar hardware and software environments. Like TLCC before them, CTS platforms are selected and fielded und are selected and fielded under a fielded under a single tri-lab procurement model. The CTS machines use more commercially available computing technologies than the ATS machines that don't require NRE collaboration with the vendors. These platforms can be scaled up in multiple scalable units and interconnected to build



The March 2022 plan for future ASC procurements. A third class of machines, Advanced Architecture Prototype Systems (AAPS), appeared after 2013, allowing experimental technologies to be matured and evaluated for feasibility before being deployed as an ATS or CTS platform.

Fire/ICE (CTS-1) at LANL

Quartz (CTS-1) at LLNL

CTS-1 was awarded to Penguin Computing in 2015 for the procurement of scalable units (SUs) that can be combined into powerful multi-SU systems to meet the capacity computing demands of the ASC program. Several multi-SU systems have achieved top 100 rankings in the TOP500 list.

computing environment that's consistent across the three labs and serves as a disaster-recovery option for the program. The CTS machines are lower-risk and less-expensive clusters compared to the ATSs, meant to be deployed quickly, and can run problems with more modest computational requirements. Together, ATS and CTS platforms allow ASC to continue to push the possible boundaries of HPC, while supporting the full range of simulations needed for stockpile stewardship and other mission needs.

Later in 2018, ASC authorized SNL to initiate its Vanguard program to manage the procurement of Advanced Architecture Prototype Systems (AAPSs) by which advanced technologies can be hardened further and assessed for possible inclusion into future ATSs or CTSs. This bridges the gap between lower and higher technology readiness concepts. One success of this program is the SNL Astra platform which was the first petaFLOPS computer using Arm-based processors.

The Astra supercomputer at SNL, which runs on ARM processors, is the first result of the Sandia Vanguard Program, intended to explore emerging technology for supercomputing. (SNL)

Technicians build the Attaway (CTS-1) supercomputer from Penguin Computing at SNL.

JOINT PROCUREMENT- A DRAMATIC SUCCESS

Cielo at LANL

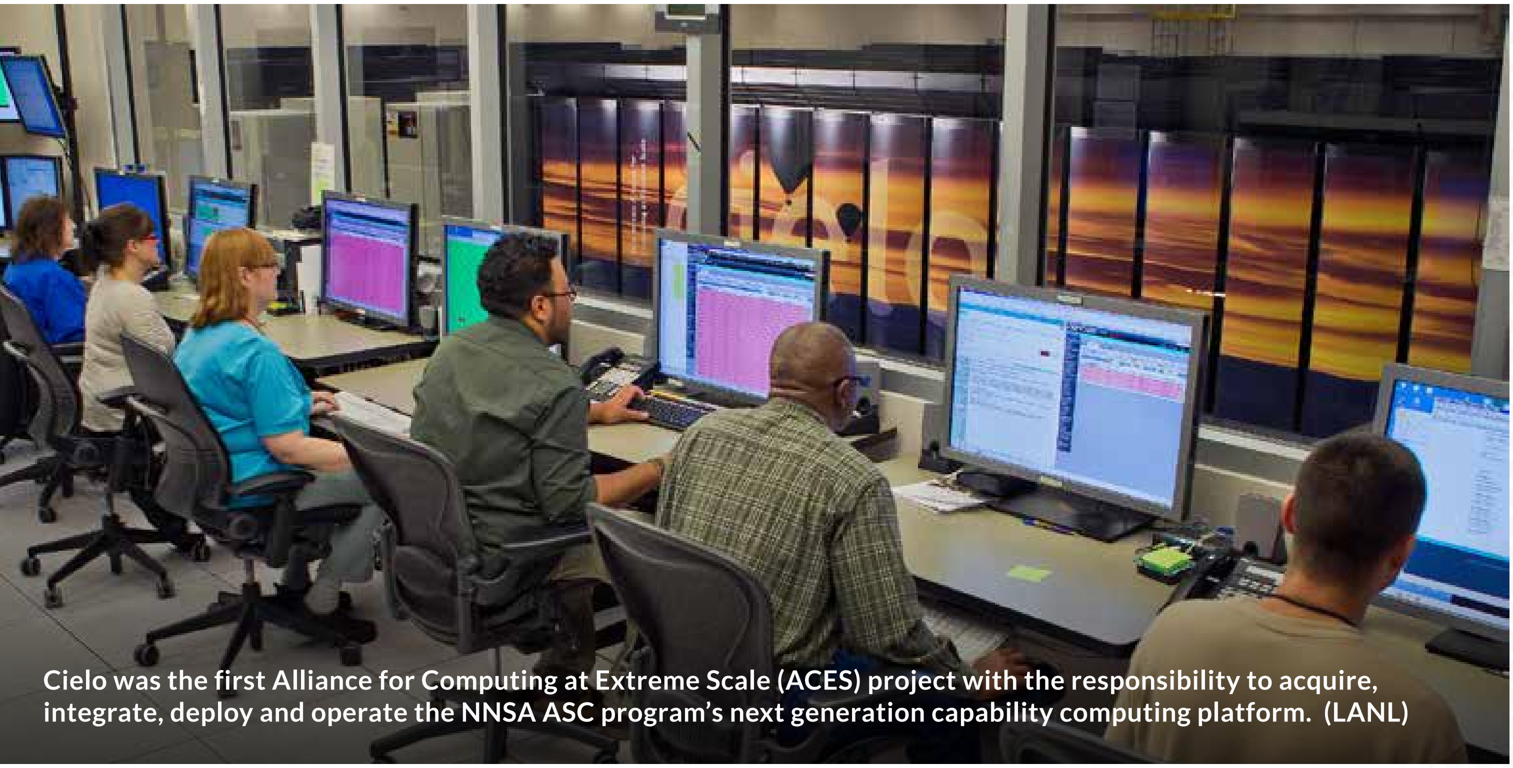
One way to manage the development and deployment risks associated with large platforms is through joint procurements, as pioneered by ASCI and ASC. By pooling expertise, the partnering organizations can issue a single request for proposal (RFP), thus making it easier (and more attractive) for vendors to respond to, while stimulating competition for a successful bid. The organizations would also have a much more rigorous process for developing the technical specifications of the RFP and evaluating the resultant bids. Finally, bigger platforms with longer lead times could be purchased through pooling resources. Platform vendors can more viably compete and their experience is further developed by selling multiple platforms of a given or similar type. Even if only a single platform is purchased, risk is also reduced.

After Roadrunner broke the petaFLOPS barrier, it soon became apparent that the exaFLOPS (10^{18} FLOPS) barrier would be the next national priority for computing performance, amounting to a 1,000-fold increase over the 1 petaFLOPS Roadrunner. To reach this new strategic performance goal more quickly, the DOE SC and NNSA DP organizations signed a memorandum of understanding (MOU) to formalize their collaboration (Approval of the DOE Exascale Initiative MOU). This MOU commited NNSA and DOE SC to work together in several areas, including hardware technology, algorithm research, applications, and system software. The MOU created a holistic approach to developing and producing the technologies most needed to achieve exascale computing. This was especially apparent in the APEX and CORAL system procurement efforts (described below) that are jointly coordinated activities between NNSA ASC and DOE SC ASCR programs.

The Alliance for Application Performance at Extreme Scale (APEX)

In 2008, a MOU between LANL and SNL established the NNSA New Mexico Alliance for Computing at Extreme Scale (ACES). The ACES partnership recognized the benefit of pooling intellectual capabilities and resources towards the goal of exercising continued leadership in HPC. Following ASC Purple, Cielo from Cray, Inc. was the first ASC platform to emerge from the ACES partnership (Ang, J. D., 2010). Deployed in 2011, Cielo used more than 140,000 processing cores to exceed 1 petaFLOPS, a 10-fold performance increase over ASC Purple. An XE6 machine, Cielo's upgraded Gemini interconnect set it apart from similar XT6 systems, which were direct descendants of the XT5 systems previously discussed with ASC Red Storm. The Gemini interconnect would become a common component of many of Cray's products going forward. Cielo's success demonstrated that this type of partnership could succeed at delivering a leadership-class machine, while sustaining a viable business model for Cray and other vendors.

The ACES partnership was joined later by LBNL NERSC computing facility, which procured Hopper (NERSC-7), an XT6 machine similar to the ORNL Jaguar system. The partnership with NERSC would be extended in 2015 through the formation of the Alliance for Application Performance at Extreme Scale (APEX). The APEX collaboration is responsible for the NERSC Cori (NERSC-8) and LANL Trinity (ATS-1) computers, both of which are Cray XC40 computers. The APEX collaboration was involved in the procurement phase of the LANL procurement of Crossroads (ATS-3) and the NERSC procurement of Perlmutter (NERSC-9).

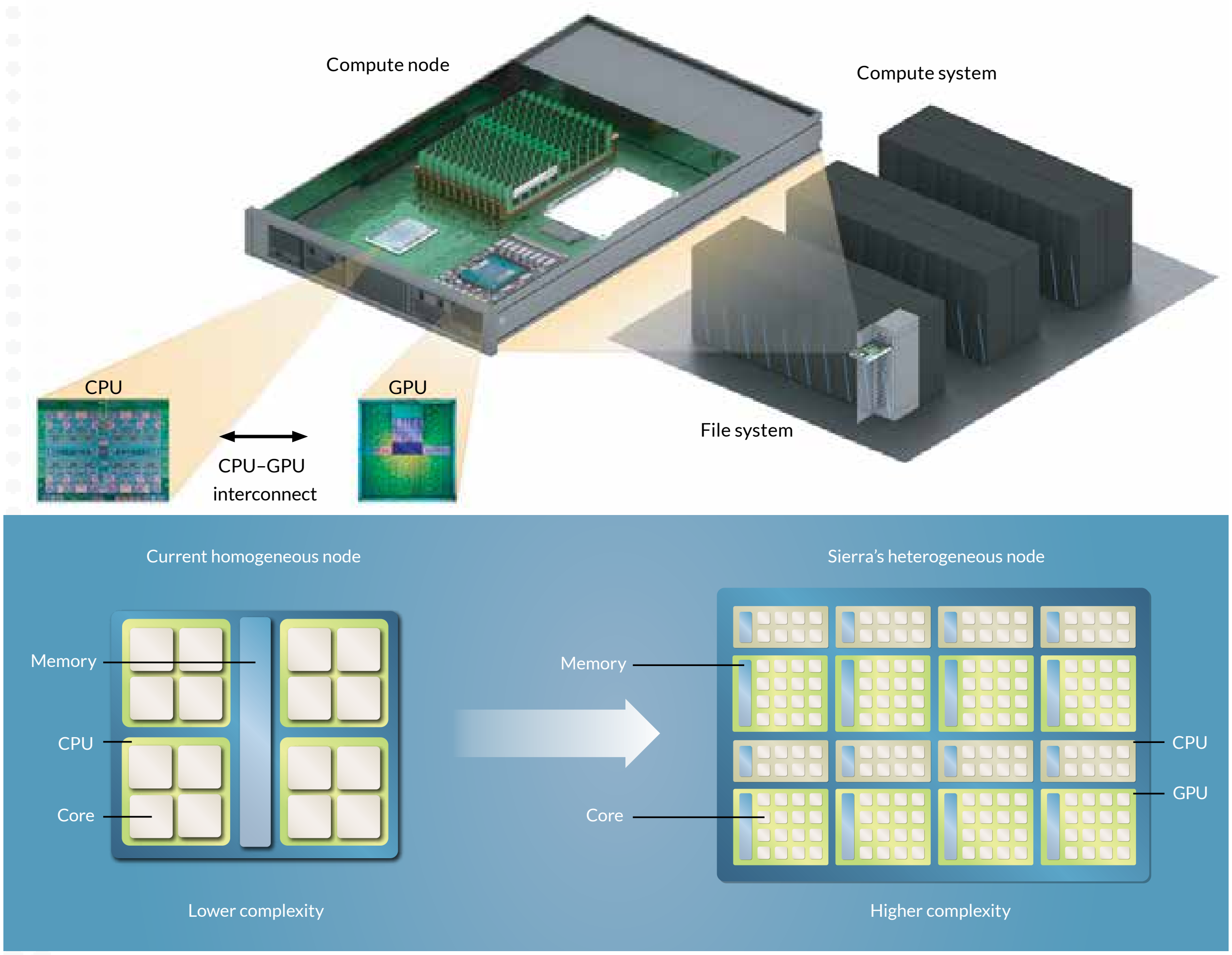


Cielo was the first Alliance for Computing at Extreme Scale (ACES) project with the responsibility to acquire, integrate, deploy and operate the NNSA ASC program's next generation capability computing platform. (LANL)

The CORAL Partnership

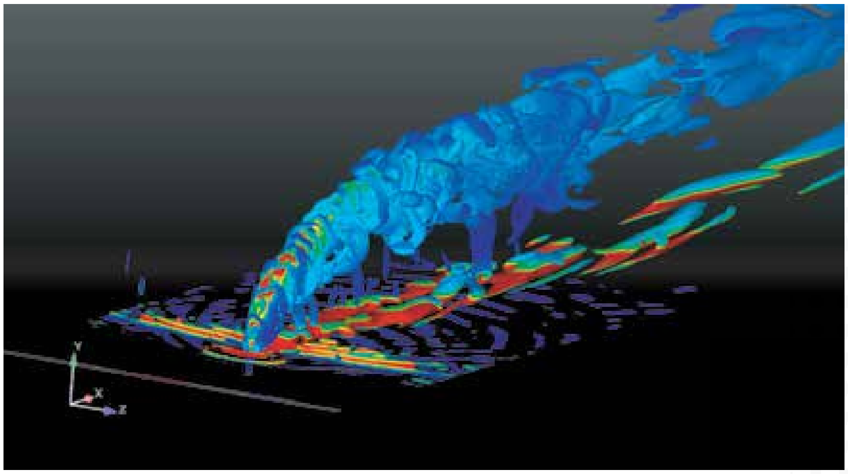
Other DOE national laboratories responded to difficulties inherent to supercomputing by forming partnerships similar to APEX. In 2014, the Collaboration of Oak Ridge, Argonne, and LLNL (CORAL) partnership was formed as a joint procurement activity (Vazhkudai, 2018). From the CORAL-1 procurement in 2014, ORNL and LLNL selected a similar, hybrid architecture based on several thousand nodes consisting of IBM Power9 CPUs accelerated by NVIDIA Volta GPU processors. LLNL purchased the Sierra (ATS-2) platform and Oak Ridge procured Summit. Argonne ultimately delayed the selection of its system, Aurora, now scheduled to launch in 2022 as an exascale platform. Aurora was recast as a heterogeneous architecture equipped with Intel CPUs and GPUs, following the success of the Sierra and Summit heterogeneous systems. Intended to build on the accomplishments of CORAL, the CORAL-2 procurement was launched in 2018, resulting in two exascale-class computers, Oak Ridge's Frontier in 2021, and LLNL's El Capitan (ATS-4) in 2023.

Sierra is designed to maximize speed and minimize energy consumption to provide cost-effective modeling, simulation, and big data analytics. Sierra nodes combine several types of processing units: central processing units (CPUs) and graphics-processing units (GPUs) which result in greater parallelism, faster results, and energy savings.

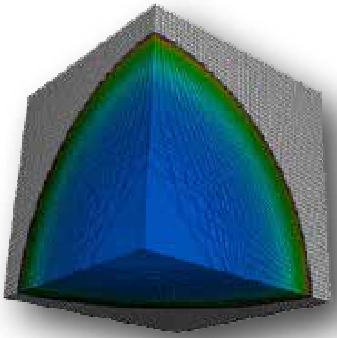


DesignForward and FastForward

The earlier ASCI PathForward program was very successful in developing needed HPC technologies and transitioning them to viable product lines. The 2011 MOU between DOE SC and NNSA DP made it clear that similar programs were needed to create viable exascale computers. This led to a series of collaborations between DOE and industry between 2011 and 2015 called DesignForward and FastForward, jointly funded by DOE SC ASCR and NNSA ASC. The collaborations used the co-design concept to improve the effectiveness of two-way feedback loop of the application, hardware and software requirements, and their design. An extremely efficient means of collaboration, co-design leveraged the expertise of vendors, hardware architects, software developers, computer scientists, and others to ensure that future HPC architectures were well-suited to DOE applications. An example of this two-way interaction involves proxy apps that serve as models for one or more features of the parent applications. Proxy apps can be easily optimized to vendor hardware, yet are reflective of target computational workloads. The FastForward 1 and 2 programs focused on accelerating innovation in processors and memory systems, while DesignForward 1 and 2 focused on interconnect networks and system design and integration. Together, these four programs facilitated the development of many technologies that will be integral to the exascale systems arriving in the 2021-2023 timeframe (Ang J., 2015).



Iso-surfaces of Q-Criterion colored by eddy-viscosity from a Hybrid RANS-LES simulation of a jet-in-crossflow flow field. The image shows the complex interacting vortex structures in the flow field. The algorithms were influenced by information on new hardware changes. (ASC Computing Strategy, 2013).



Proxy apps or mini-apps are simplified versions of applications and major physics packages that retain key features of the original. Proxy apps illustrate the principles of co-design. These apps can be easily shared with and studied by vendors and academic researchers to evaluate new architectures. The figure illustrates one of the earliest proxy apps, the LLNL Unstructured Lagrangian Explicit Shock Hydrodynamics code (LULESH). This proxy app simulates the Sedov explosion problem and can be easily optimized by vendors to test their products on an algorithm relevant to many ASC integrated codes (Karlin, 2013).

THE EXASCALE COMPUTING INITIATIVE (ECI)

The need for coordination to continue the growth of HPC extended beyond NNSA and DOE SC. In 2015, a presidential Executive Order launched the National Strategic Computing Initiative (NSCI). This order states, “in order to maximize the benefits of HPC for economic competitiveness and scientific discovery, the United States Government must create a coordinated Federal strategy in HPC research, development, and deployment,” (Creating a National Strategic Computing Initiative, 2015). The Executive Order formally launched a government-wide effort to develop exascale computing. The effort recognized that advancements in all aspects of computing were needed to reach the goal of exascale computing. In short, the founding principles of ASCI in 1995 were restated and updated to focus on the requirements of exascale computing.

The lead organization to implement the push for exascale computing, specified in the NSCI Strategic Objective 1, would be DOE via a joint partnership through the DOE SC and NNSA DP. DOE would launch the Exascale Computing Initiative (ECI) to execute exascale computing with the Exascale Computing Project (ECP) as a subset of more closely coordinated, unclassified R&D activities funded by ASC and ASCR programs. The execution of ECP was facilitated by an updated MOU in 2016 (Approval of Memorandum, 2016). This MOU establishes both the ECI and the ECP. It builds upon the 2011 MOU cited previously, which laid the organizational foundations that would be used to manage the ECP mission.

This mission is summarized by the ECP Director, Doug Kothe (2019):

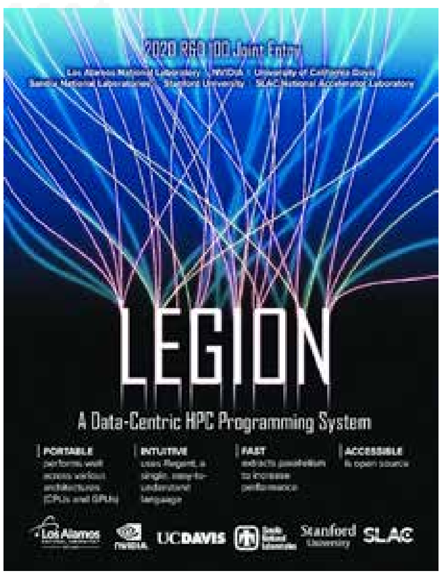
1. Deliver exascale-ready DOE applications and solutions that address currently intractable problems of strategic importance and national interest.
2. Create and deploy an expanded and vertically integrated software stack on DOE HPC exascale and pre-exascale systems, defining an enduring U.S. exascale ecosystem.
3. Leverage U.S. HPC vendor R&D activities and

products into DOE HPC exascale systems. The ECP mission directly aligns with the spirit of ASCI's founding principles. The project's goal is to create a comprehensive ecosystem where the relevant multi-disciplinary applications are tightly integrated and work directly with HPC vendors in delivering the needed new capabilities. This seven-year project (ending in FY23) will create an enduring set of scientific application codes, software packages, and hardware technologies that perform well at exascale speeds and to provide users with computing platforms to execute their tasks.

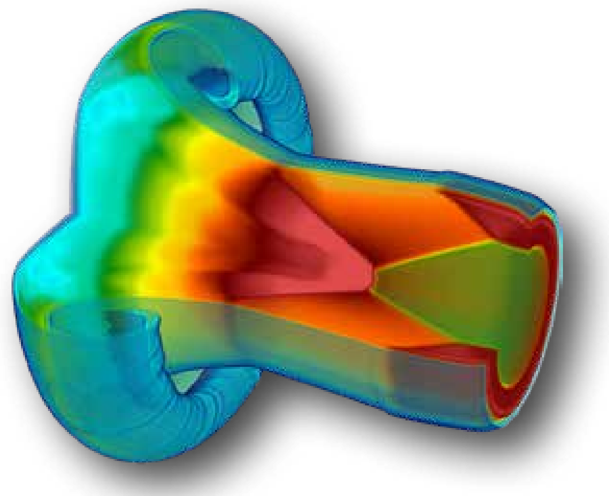
The particular focus on exascale applications recognized that exascale-class machines would require the ability of codes to run on multiple types of computer architectures, many of which would require the development of new programming models. This put extreme demands on the ASC integrated codes which had now grown into large applications relied upon for the assessment and certification of the Nation's nuclear stockpile. This was equivalent to changing the transmission on a car when it is rolling down the road filled with passengers. NNSA created the Advanced Technology Development and Mitigation (ATDM) subprogram to manage this challenging endeavor. There are two key components of ATDM. Code Development and Applications (CDA) builds the next generation of integrated codes. Architecture and Software Development (ASD) funds the ECP PathForward program and develops compilers, libraries, and programming models to support the next-generation platforms and codes. LLNL, LANL, and SNL have each contributed uniquely to ECI through ATDM as discussed in the ATDM section. Each laboratory would develop novel solutions that would allow their integrated codes to adapt to exascale without losing any of their predictive capabilities developed over the decades since 1995.

Renewed Connections with Industry

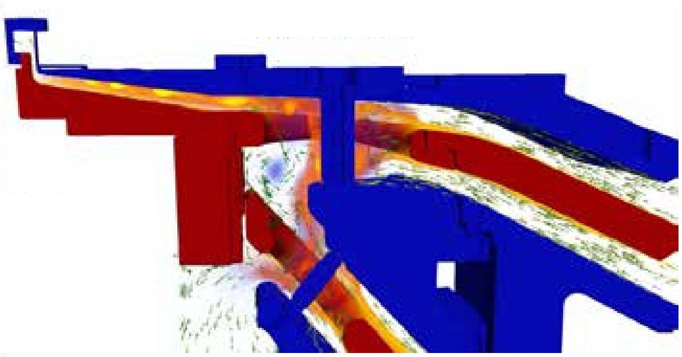
ECP built upon prior work done to prepare U.S. industry for exascale computing. The ECP PathForward program, started in 2016, is the successor to the earlier ASC/ASCR-funded DesignForward and FastForward programs. The goal of the 2016 program was to build a sustainable industrial base for exascale computing, with capability increased by a factor of more than 33,000 over the original ASCI PathForward goal of 30 teraFLOPS. With ECP PathForward focusing on node and system design, a co-design approach was central to the engagement with industry—with vendors focusing on conceptual system and node designs, analysis of potential performance improvements, and technology demonstrators that quantify performance gains. The DOE SC and NNSA labs were fully engaged in evaluating the vendor proposals, participating in project reviews, and evaluating deliverables relevant to the performance of actual ECP applications.



ECP supports the development of abstraction libraries that enable large multiphysics codes to quickly use complex and heterogeneous supercomputers with only minor modifications. This significantly reduces the cost and risk of adapting existing codes to new exascale machines. The Legion abstraction library at LANL won a 2020 R&D 100 Award for a single language approach to automating task scheduling and data movement.



MARBL simulation of a shape charge jet. MARBL uses the RAJA library at LLNL to easily support a wide variety of architectures like those used in ATS machines such as Sierra and the forthcoming El Capitan.



EMPIRE simulation of Sandia Z-accelerator experiment showing electron density in the magnetically insulated transmission line using Kokkos as the performance portability layer.

Trinity at LANL

TRINITY & THE BEGINNINGS OF ATS

Deployed in 2015, Trinity (ATS-1) was the first ATS machine. The ATS machines were discussed in the 2013 ASC Computing Strategy, with ATS machines for ground-breaking capability and CTS machines for capacity computing. Trinity was built with the twin goals of achieving a balance between usability of current simulation codes and enabling adaptation to new computing technologies and programming methodologies (Hemmert, 2016).

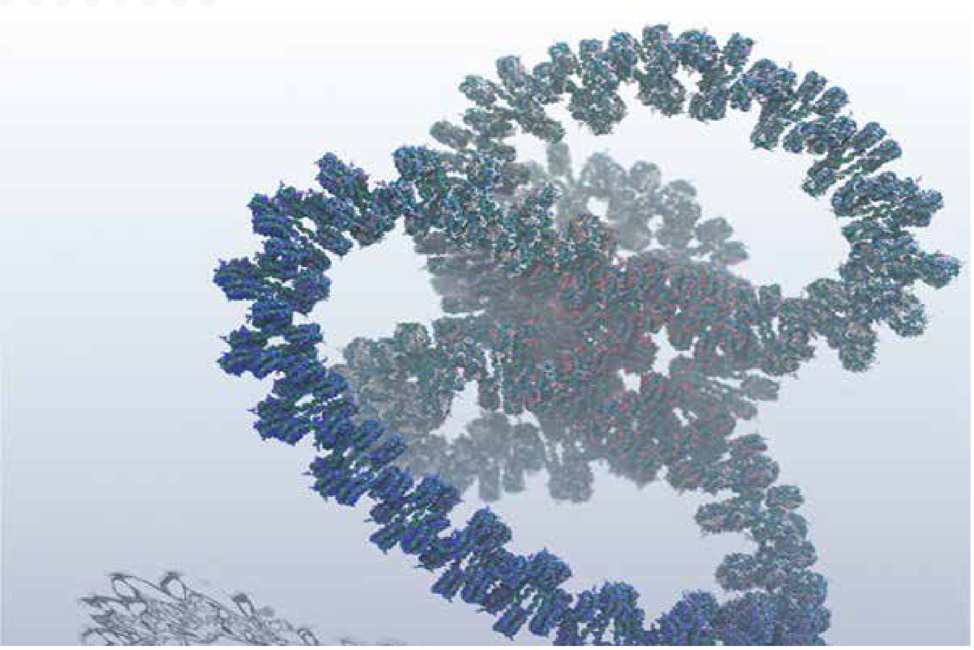
The Trinity supercomputer, a product of Cray, Inc., was based on the XC40 platform architecture. Trinity was built in two partitions with traditional Intel Xeon (Haswell) CPUs in one partition and Intel Xeon Phi (Knights Landing) processors in the other. This made Trinity a heterogeneous architecture that supported both existing programming models while providing a bridge to more challenging accelerated computing approaches. The Haswell partition provided a computing resource for legacy codes that were accustomed to CPU-only environments. The Knights Landing partition required ASC code teams to employ multiple memory spaces and higher levels of thread- and vector-level parallelism than had been needed before. To help facilitate this transition, the Trinity Center of Excellence was established, with staff from the three ASC laboratories, Cray, and Intel.

Future exascale systems will place strong demands on parallel file systems. If not managed properly, time spent merely reading in and out simulation data could easily overwhelm the time spent performing the simulation. Trinity introduced tightly integrated nonvolatile

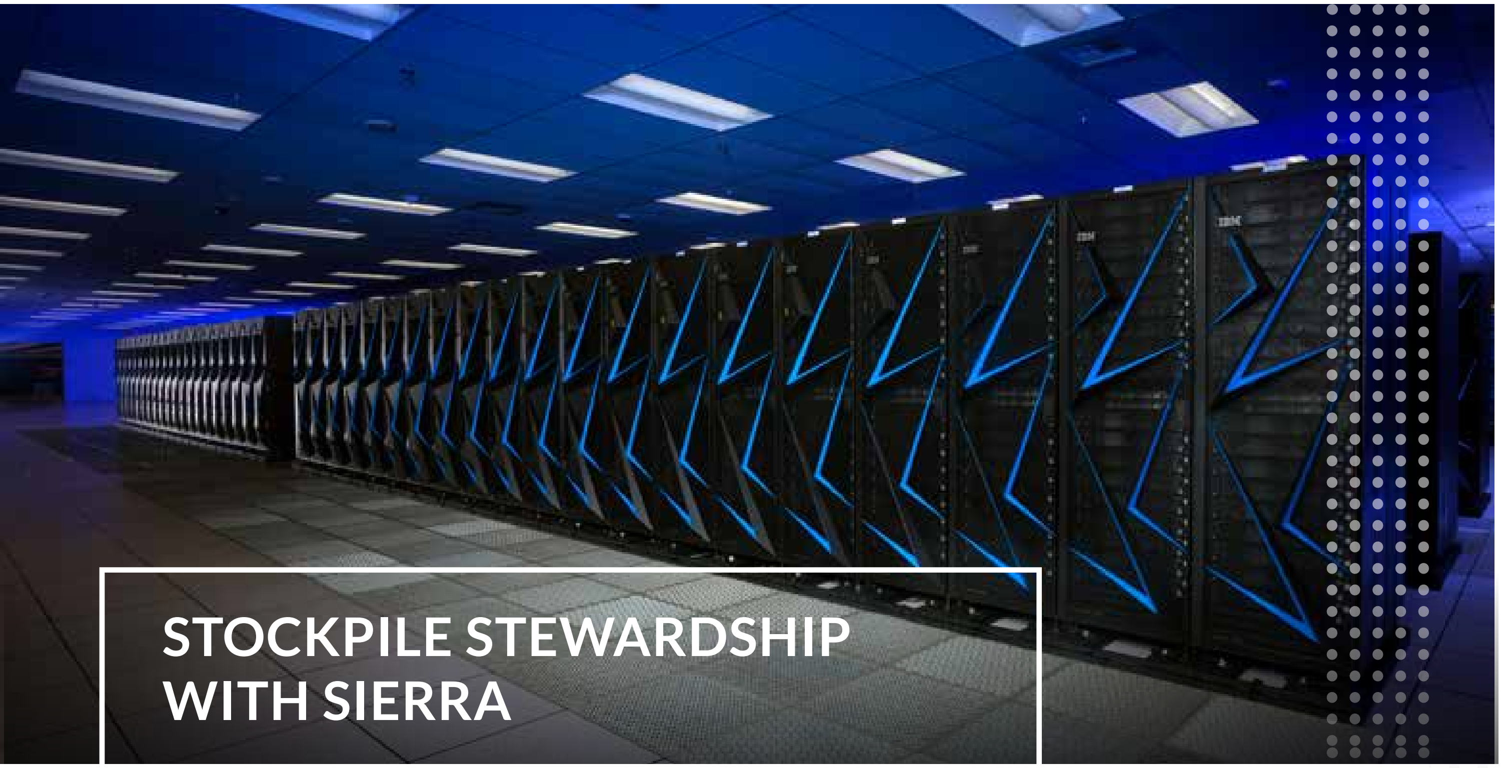
The Trinity supercomputer was manufactured by Cray, Inc. and was based on its XC40 platform architecture. Trinity was a mixture of Intel Xeon and Xeon Phi processors. (LANL)

"burst buffer" storage capabilities. Embedded within the high-speed fabric were nodes with attached solid-state disk drives. The burst buffer capability allowed for accelerated checkpoint/restart performance and relieved much of the pressure normally loaded on the back-end storage arrays. In addition, the burst buffer supported novel new workload management strategies such as in-situ analysis, which opened a new space where projects can manage their overall workflows.

Trinity also introduced advanced power management functionality that allowed monitoring and control of power consumption at the system, application, and component levels. Although advanced power management was not needed for the existing power and operational budget, its functionality was used to gain a better understanding for future system requirements and features. This power management capability is needed for the upcoming Crossroads (ATS-3) and El Capitan (ATS-4) systems which will use 3-4 times the power of Trinity.



The largest simulation to date of an entire gene of DNA, a feat that required one billion atoms to model. (LANL)



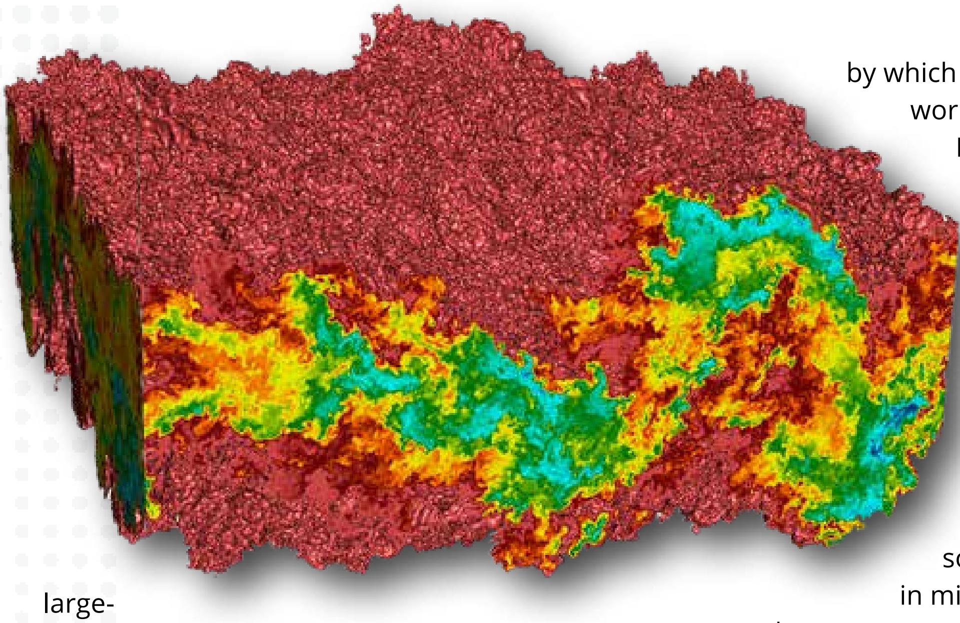
STOCKPILE STEWARDSHIP WITH SIERRA

The Sierra machine (ATS-2) is the latest of the ATS supercomputers as of April, 2022. This machine is making substantial contributions to stockpile stewardship in the form of high-resolution 3D simulations and other weapon analyses. (LLNL)

Acquisition of the Sierra (ATS-2) computer was the next step on the path to exascale after Trinity. It was purchased as part of the CORAL partnership and is the latest of the ATS line of computers deployed. This machine was deployed in 2018 and has been among the top three supercomputers in the TOP500 list from June 2018 until 2022. This was the first ASC machine to have a 125-petaFLOPS peak performance, which is roughly a factor of 3 greater than Trinity. Sierra is quite energy efficient, using around half the power while being more performant than the former TOP500 #1 machine it displaced, the Chinese Taihu Light platform. This represents a roughly six-fold increase in computing power from the previous ASC machine at LLNL, Sequoia (Thomas, 2018).

Sierra is a heterogeneous architecture that combines IBM's Power 9 processors and NVIDIA's V100 Volta graphics processing units (GPUs). It comprises 4,320 nodes, with each node consisting of two IBM Power 9 CPUs, four NVIDIA V100 GPUs and a Mellanox EDR InfiniBand interconnect. Programming on this machine requires the developer to consider which portions of the application are best run on the GPUs versus CPUs, with optimization for GPUs often requiring a new way of formulating algorithms to best take advantage of the massive parallelism available. In addition, previous machines such as Cielo and Sequoia had only one type of memory that was uniformly available to all the processing cores. Sierra's heterogeneous design has multiple memory spaces, a small but fast memory near the GPU, a larger memory near the CPU, and a third tier of

Turbulent mixing of two fluids simulated using more than 6 billion cells on Sierra.



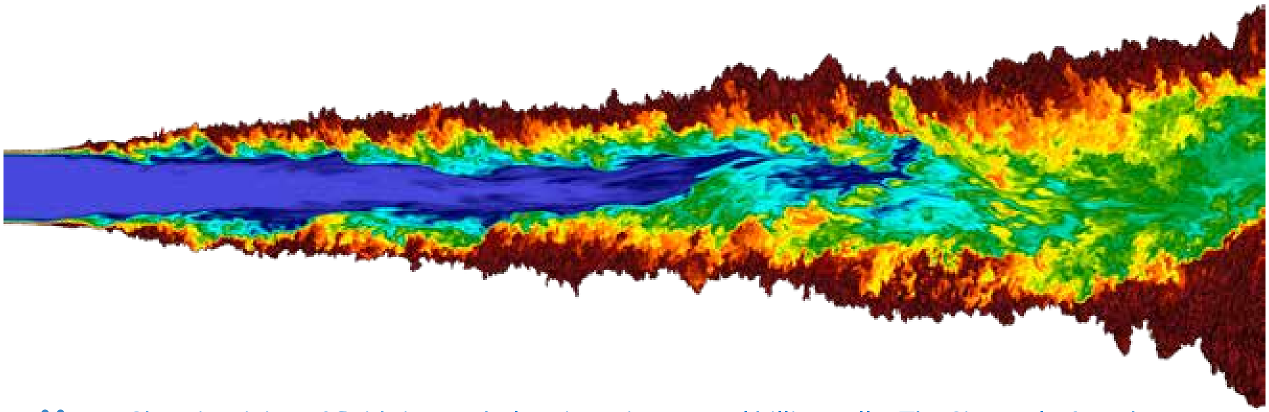
large-capacity memory made up of non-volatile random access memory (NVRAM) that built on the burst buffer concept pioneered with Trinity. Although optimal performance requires that applications explicitly manage these multiple spaces, Sierra's nodes feature a CPU-to-GPU connection via the NVIDIA NVLink interconnect that allow for the appearance of a single coherent memory space.

Sierra presented a considerable technical challenge in that over 90% of the total computational resources are on the GPUs. To use these resources, one would have to substantially modify the integrated codes to run on these machines. LLNL partnered with IBM and NVIDIA to rapidly develop software solutions that would effectively optimize the CPU-GPU architecture. The mini-apps mentioned earlier were a fundamental tool of this effort. The mini-apps provided relevant, computationally intensive kernels that represent the most challenging portions of the multiphysics codes. This provided industry many opportunities to optimize their technologies and minimize hurdles with Sierra being widely accepted.

As part of the co-design strategy, a "Center of Excellence" for the Sierra NRE effort was developed. This created a joint organization by which IBM and NVIDIA personnel working closely with the three NNSA laboratories, both on-site and remotely, were able to collaborate on achieving maximum performance through code development and restructuring. This enabled lab personnel to directly provide feedback on system design and the software stack to the vendor. This direct feedback allowed quick iteration on arriving at software solutions, which was crucial in mitigating the potential risk that these heterogeneous systems initially represented.

The GPUs on Sierra provide unique machine learning attributes. Their ability to provide extremely large amounts of lower-precision computing makes them very well suited to employ machine learning and artificial intelligence methods. This capability enables the use of these methods for such applications as the use of deep learning to accelerate time-to-solution of physics codes. Simulations, with acceleration coming from the use of artificial intelligence technology, will be increasingly employed over the coming decade.

The performance of Sierra has made the early preparation of the ASC codes worthwhile. A large number of scientific publications have resulted from simulations run on this machine, ranging from the human heartbeat to plasma physics research. Sierra has produced several groundbreaking stockpile stewardship simulations, not least of which is the ability to use high-resolution, 3D simulations for LLNL's weapon baseline assessments. The ability to perform such high-resolution 3D simulations was one of the stretch goals of ASCI. With Sierra, this 3D goal has now become common as a powerful tool for the stockpile stewardship community.



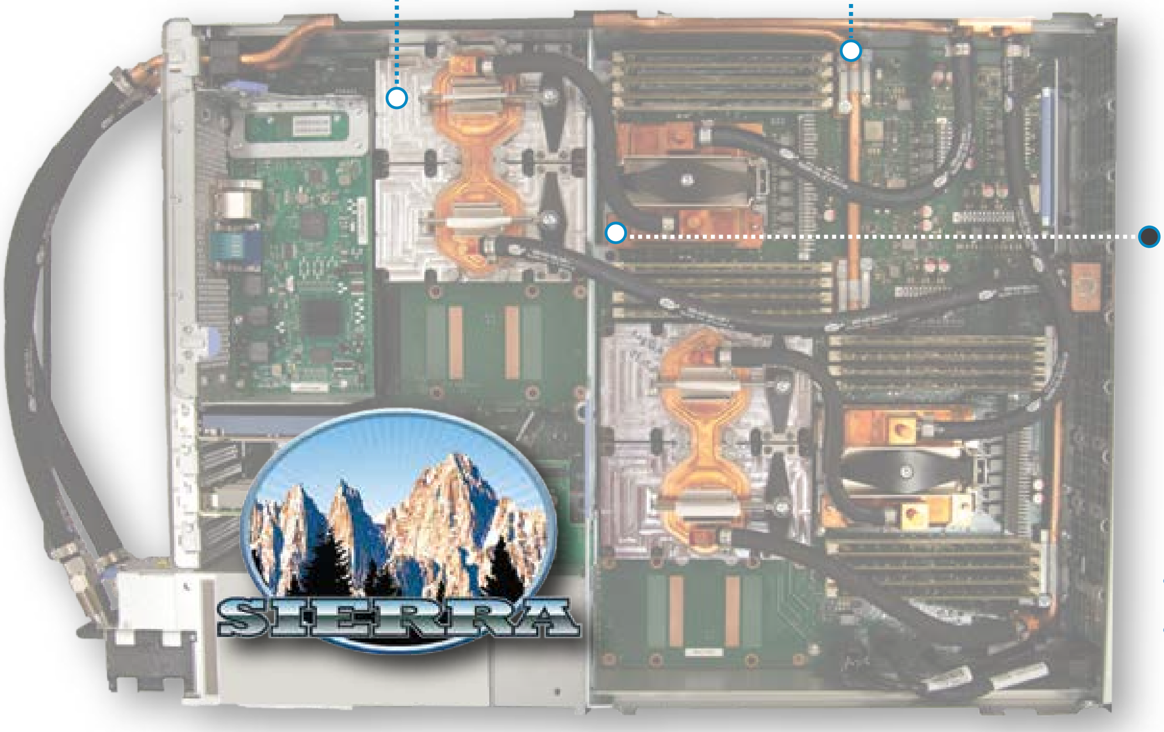
Chaotic mixing of fluids in a turbulent jet using several billion cells. The Sierra platform has reduced the time required for these high-resolution simulations from more than a month to the span of two or three days. Rather than one-off calculations, these simulations can now be run as part of routine studies to better understand the complex flow features of turbulent mixing and validate engineering models.

GPUs

Provide most of the computing power in a Sierra Node

NVLink

Enables high speed communication between CPUs and GPUs and simplifies programming requirements



Power9 CPUs

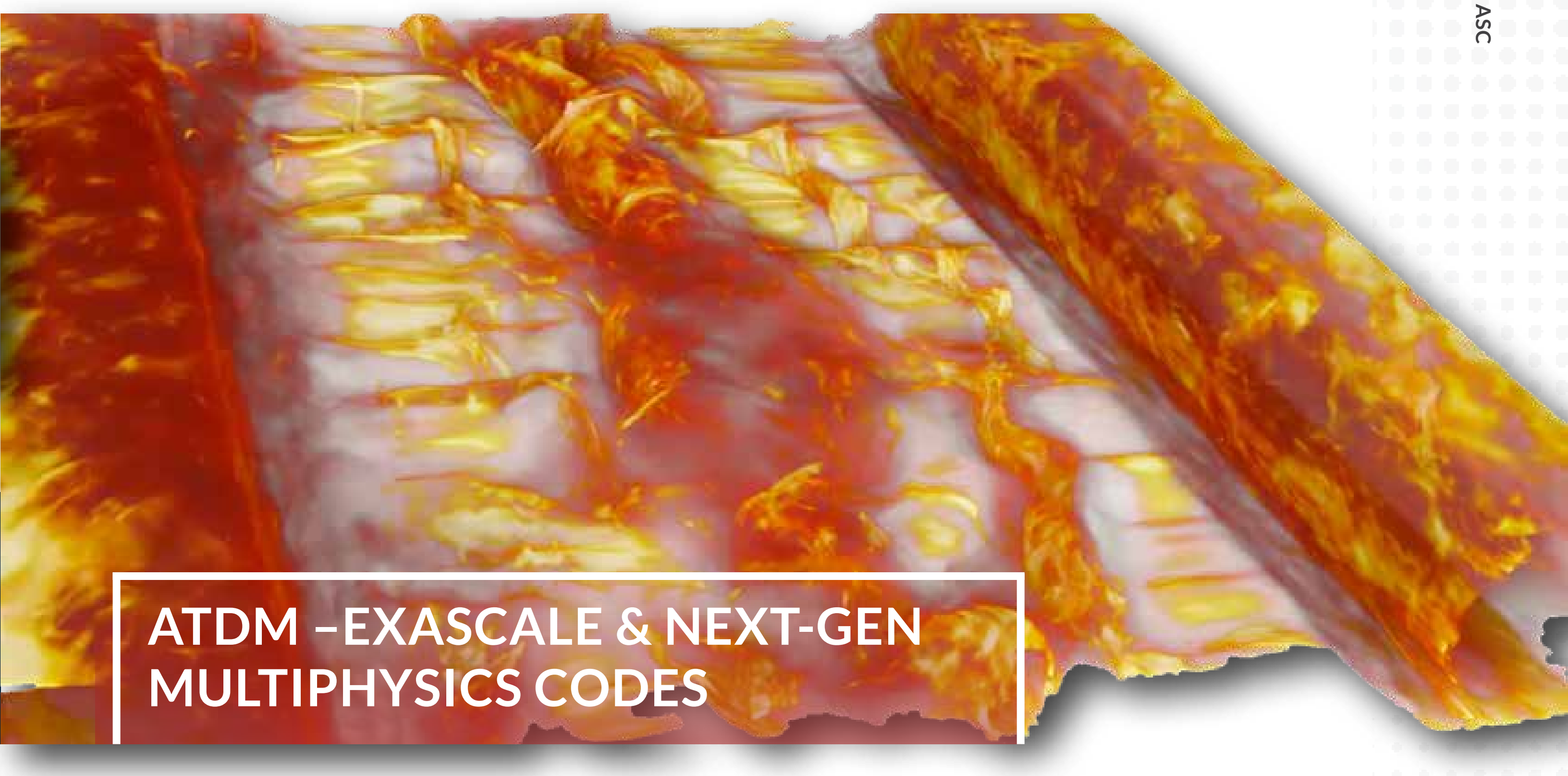
Handles serial and other tasks enabling the GPUs to focus on numerical tasks

Sierra's optimized use of CPUs, GPUs, and high-speed connections enables ground-breaking stockpile stewardship calculations.

CROSSROADS & EL CAPITAN

The ATS line of supercomputers continues the march towards the goal of exascale computing. Crossroads (ATS-3), at LANL, is the next computer in line after Sierra (ATS-2) and will feature several technologies, such as new memory technologies, that will be integral to increasing the efficiency of applications on HPC systems. This computer is slated to be deployed in 2022 and will help bridge the gap from Sierra to exascale computing.

When deployed in 2023 at LLNL, El Capitan (ATS-4), will be the first ASC exascale machine with at least 2 exaFLOPS peak performance. This machine will be the second platform of the CORAL-2 collaboration after the ORNL Frontier platform deployment in 2021. Both machines have a GPU-based architecture similar to Sierra, but with AMD-produced CPUs and GPUs.
The deployment of El Capitan will be the capstone of the NNSA exascale computing journey, addressing several national security challenges that were previously not tractable.



ATDM –EXASCALE & NEXT-GEN MULTIPHYSICS CODES

Magnetic reconnection, the continuous breaking and rearrangement of magnetic field lines in a plasma, is a fundamental process in physics. Understanding reconnection phenomena has broad implications and may eventually help protect astronauts, communications satellites, and electrical power grids. This 3D simulation shows how instabilities in the reconnection layer lead to multiple flux rope structures and turbulent magnetic fields. (LANL)

The extreme challenges of achieving exascale first came into broad awareness through a seminal Defense Advanced Research Projects Agency (DARPA) report (Kogge, 2008). Within several years, a series of DOE-sponsored workshops had added a clear mission need for the U.S. to reach that exascale capability — particularly in national security and stockpile stewardship. By 2013, it was apparent that a major effort was needed to transition the ASC multiphysics codes into the exascale era. Most of the codes had outlived the original ASCI platforms by more than a decade, and it was time to make some fundamental architecture changes. However, at that time it was still largely speculative what those exascale systems would look like —other than they would certainly not resemble a simple continuation of the homogeneous CPU-based architectures that defined ASCI and ASC platforms to date. There was thus great uncertainty as to whether the existing production codes could make the leap to exascale, and a clean-slate approach to developing a next-generation set of codes would either be necessary or at least a critical risk mitigation. Predating the NSCI and ECI, the Advanced Technology Development and Mitigation (ATDM) subprogram was launched to afford the three NNSA Laboratories the necessary opportunity and personnel to focus on next-generation technology development. Each laboratory employed its ATDM funding to investigate a unique approach to developing next-generation application codes and system software. All set out to create codes built on lessons learned from two decades of ASC research. These experienced teams would construct codes to fully exploit new exascale machines and deliver unprecedented stockpile stewardship insight. All three labs developed and deployed performance portability software abstractions. The RAJA and Kokkos packages were developed in-house by LLNL and SNL, respectively, to insulate application codes from the plethora of emerging programming models aimed at exploiting the massive parallelism on node. LANL teamed up with Stanford and co-developed the Legion framework to exploit asynchronous task models that use a sophisticated runtime to orchestrate work across the system. These abstractions enable the laboratories to centralize the optimization effort needed to get optimal code performance on potentially any future platform. Thus, by using software abstractions one could achieve excellent performance on disparate systems using a single base of source code, or by modifying only a very small portion of a very large application code that spans millions of lines.

The pre-exascale multiphysics codes often developed their supporting software, such as visualization and analysis, as needs arose. As the next-generation ATDM codes were being designed, it was realized that much could be standardized, and substantial savings could be realized, if this supporting software was built upon the same modules and underlying computer science utilities. Each lab used its own approach for their new codes. The approaches taken are described in depth on the following pages, illustrating the substantial impact the new codes will have when fully deployed.

ATDM was always envisioned as a limited-duration project. As it will conclude by end of FY 2023 along with ECP, the ATDM application and software efforts are being integrated back into the core ASC program elements where they will live on as production capabilities to be exercised on the El Capitan exascale system and subsequent ASC HPC systems.

Finally, within the last decade it was recognized that the supercomputing endeavor had become a global race. The size of the supercomputing community had grown from just a few government-sponsored HPC programs in the U.S., China, Europe, and Japan to a globally broad and diverse community. At the start of ASCI, much of the software had to be developed in-house as it did not exist anywhere else. The community beyond ASC has now created a wide range of new techniques and algorithms. ASC and its sister HPC program, ASCR in the DOE Office of Science, were developing world-class utilities for a wide range of high-performance computing tasks. A large benefit to the U.S. and international HPC communities, and a legacy of all three lab ATDM efforts, is that many of the unclassified technologies are being released as open-source software. In turn, by harnessing the intellectual contribution of the broader scientific user community, the ASC codes and software packages can be sustained for the next 25 years and beyond.

Open-source products developed by NNSA labs with commercial support

LUSTRE

ASCI PathForward developed parallel file systems. Lustre is the most commonly used file system in the Top500 machines.

SLURM

Open-source resource manager developed through an LLNL led consortium deployed on a majority of the Top500 machines.

OPEN FABRICS ALLIANCE

Open-source ecosystem for high speed data transport, initially funded through ASC, that is spurring the deployment of InfiniBand (most common interconnect in HPC systems).

VISIT

Open-source visualization tool developed at LLNL, widely used across DOE, DOD, and at a few international sites.

SPACK

Software deployment tool utilized by the ECP software stack and most supercomputing centers.

Commercial spin-offs & vendor supported

ARCHIVAL STORAGE

HPSS is the first hierarchical storage management capability, initially developed as a DOE collaboration and now used by 40+ HPC centers and research sites.

BLUEGENE

First advanced technology platform: A successful commercial IBM system co-designed over three generations at LLNL.

LINUX

TOSS (Tri-lab Operating System Stack) enhances RedHat distribution optimized for HPC linux clusters.

ROADRUNNER

First heterogeneous accelerator based Supercomputer. Developed for LANL by IBM. Accelerator based systems now dominate the Top500.

REDSTORM

Stabilizes Cray as a premier HPC provider, establishing a foundation for partnerships with multiple national labs (e.g., ORNL).

Money spent by the ASC program has impact beyond the Nuclear Deterrence Mission.

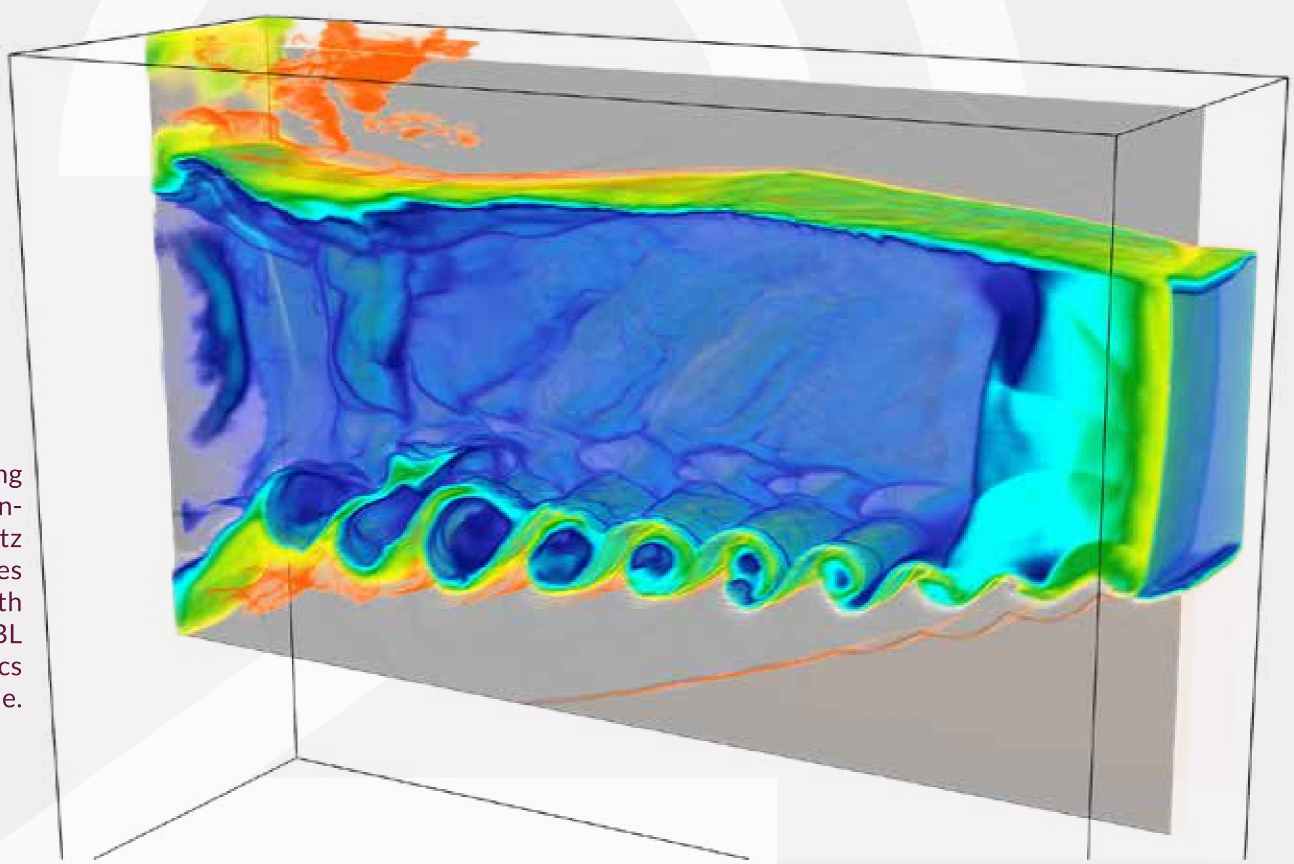
ASC open-source software has a strong history of providing a broad benefit to the HPC community. ASC software is a world leader in numerics, linear algebra, visualization, and containerization of applications. The Spack and SLURM utilities are the most widely used utilities for managing software package builds and job scheduling, respectively. The Lustre and HPSS file storage solutions are used for a wide class of supercomputers.

Multiphysics on Advanced Project Platforms (MAPP)

LLNL's ATDM application efforts are developed under the Multiphysics on Advanced Platforms Project (MAPP).

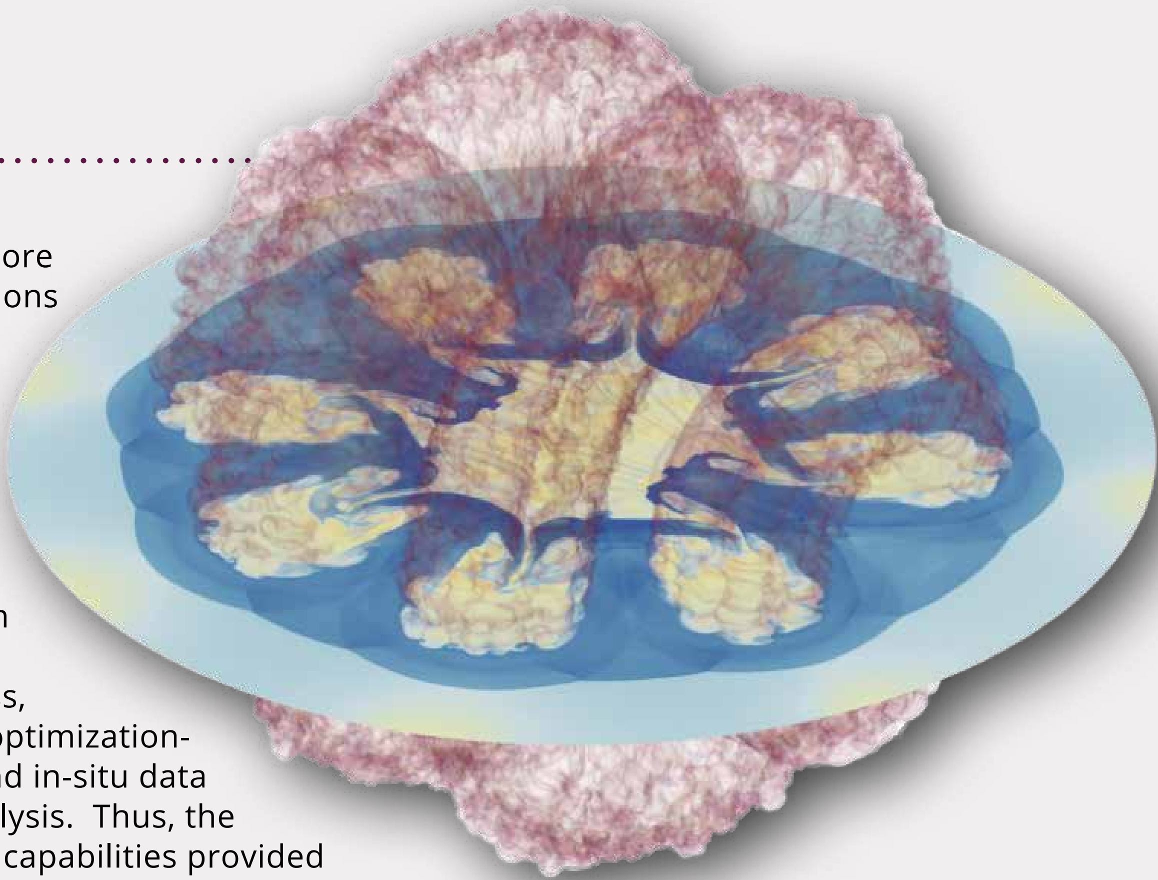
Applications developed within MAPP are based on high-order methods and a highly modular approach using reusable computer science and physics components to assemble the final application (Rieben, 2021). MAPP includes the MARBL code designed to address the modeling needs of the high-energy-density physics (HEDP) and ICF communities as part of NNSA stockpile stewardship activities.

The high-order numerical discretizations that are a distinguishing feature of MARBL use advanced, interchangeable hydrodynamics packages supporting both high-order finite element ALE and high-order finite difference Eulerian methods. High-order numerical methods were chosen due to their proven increases in accuracy and their ability to play to the strengths of exascale architectures by taking advantage of specialized hardware for small, dense-matrix operations, as



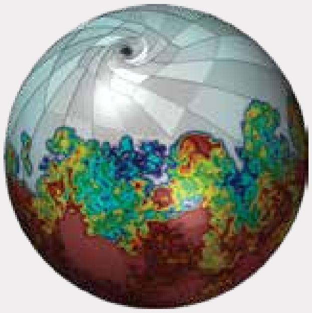
Radiating Kelvin-Helmholtz instabilities models with the MARBL multiphysics code.

MARBL



well as performing more floating-point operations for each piece of data retrieved from memory. Another key goal for MARBL is enhanced end-user productivity, including improved workflow for problem setup and meshing, simulation robustness, support for UQ and optimization-driven ensembles, and in-situ data visualization and analysis. Thus, the advanced simulation capabilities provided by MARBL improve user throughput and productivity along two axes: faster turnaround for multiphysics simulations on advanced architectures and less manual user intervention.

The LLNL MARBL project uses packages in a new flexible manner. Shown is a high-order ALE simulation using MARBL of an idealized ICF implosion to illustrate the Rayleigh-Taylor (RT) and Richtmyer-Meshkov (RM) fluid instabilities in a spherically converging geometry.



Lawrence Livermore National Laboratory

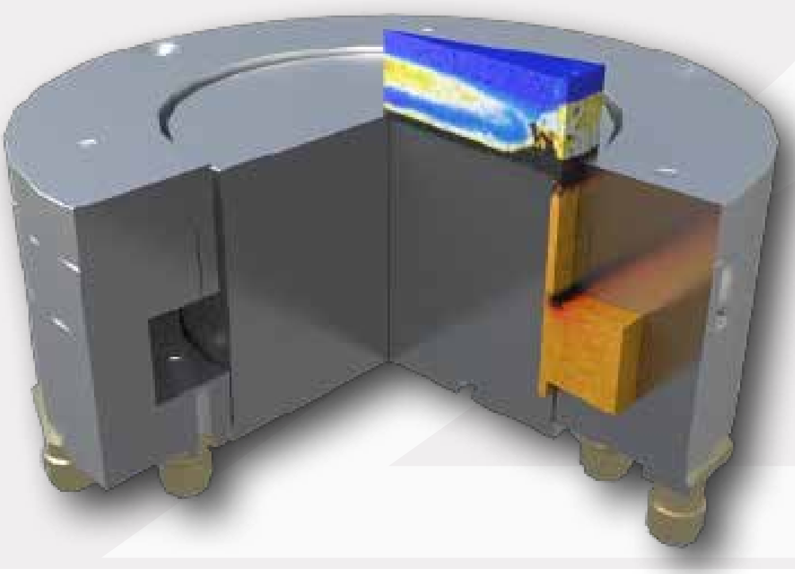
HIGHLIGHT

SPARC & EMPIRE Codes

Sandia implemented two codes for their ATDM effort.

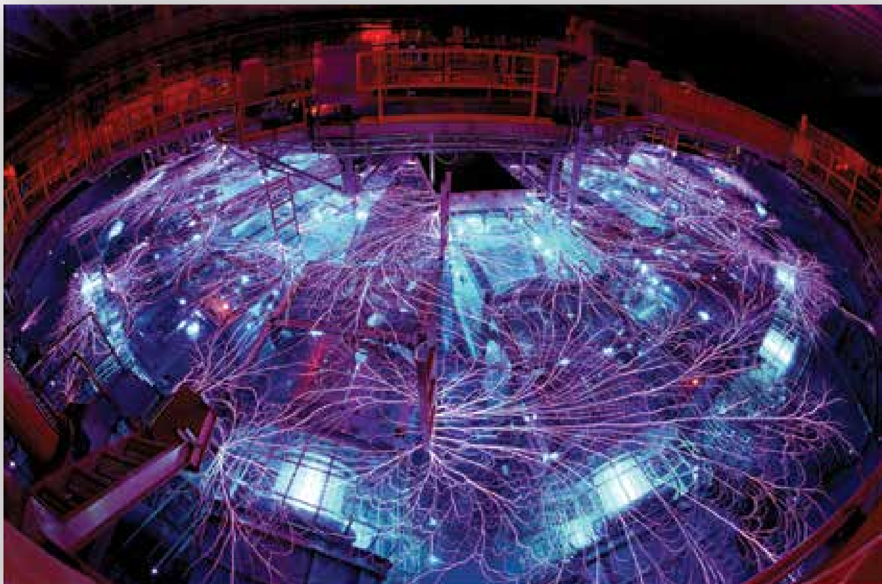
The SNL Sandia Parallel Aerodynamics and Reentry Code (SPARC) code enables a revolutionary hypersonic reentry simulation capability that captures the random vibration and thermal environments created by reentry of a vehicle into the earth's atmosphere. SPARC incorporates innovative approaches on several fronts, including effective harnessing of advanced and heterogeneous computing architectures using Kokkos and exascale-ready parallel scalability. SPARC enables its users through advanced uncertainty quantification, implementation of state-of-the-art reentry physics and multiscale models, use of advanced verification and validation methods, and improved workflows (Howard, 2017; Dennig, 2021).

The ElectroMagnetic Plasma In Realistic Environments (EMPIRE) application code complements SPARC by delivering an advanced radiation, electromagnetic, and plasma physics simulation capability that is well optimized for next-generation hardware architectures. Target applications include electromagnetic pulse (EMP) environments and related simulations of pulsed power facilities (e.g., Z, Saturn, and HERMES), high-voltage components, and high-power microwave sources. EMPIRE complements SPARC in that, without nuclear explosive testing, EMP environments must be extrapolated from what can be realized with above-ground test facilities. Hence, a validated computational simulation tool is critical to ensure the integrity and safety of these important nuclear weapon



Experimental CAD geometry as fielded on the Z pulsed power X-ray facility with an overlay of a plasma generated in an EMPIRE simulation. The extreme-scale computational resources of Trinity and Sierra are necessary to model the complex dynamics generated by the X-ray generation from the SNL Z facility.

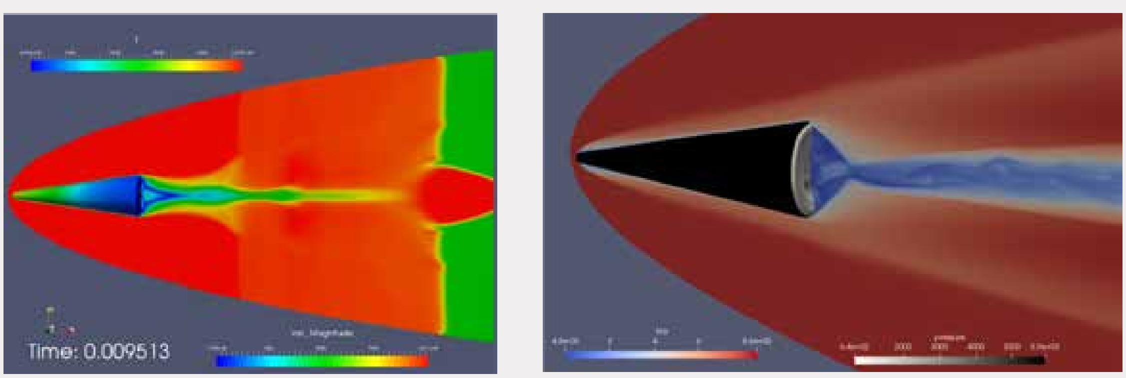
SPARC & EMPIRE CODES



Sandia's Z Machine oncentrates electrical energy and turns it into short, enormously powerful pulses to generate X-rays and gamma rays.

delivery systems. These problems are numerically challenging to simulate and can span vast length and time scales, making them highly suitable for exascale computation. To broaden the range of plasma conditions that can be efficiently simulated, EMPIRE also includes both kinetic (particle) and fluid (continuum) plasma representations—a first for this class of codes (Dennig, 2021).

This approach fosters a hierarchical code development ecosystem, where new application codes may be deployed with heightened efficiency due to reuse of shared software components, as well as an environment that facilitates the progression from prototype to production software deployment (Hungerford 2021).



SPARC simulation of warhead re-entry into the atmosphere showing the bow shock and wake. (SNL)

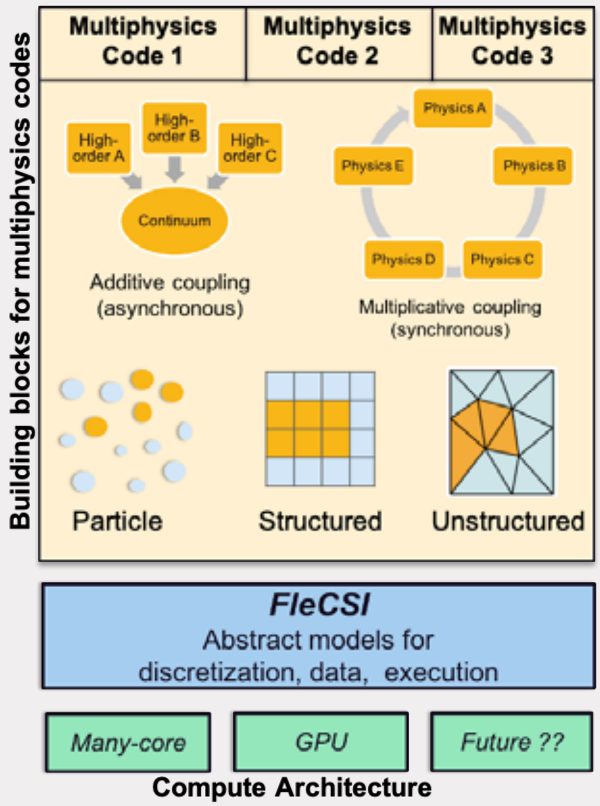
HIGHLIGHT

RISTRA Project

LANL has built an ambitious new approach to multi-physics code development based on FleCSI, a data-centric programming system that uses the Legion task-based parallel runtime.

The Flexible Computational Science Infrastructure (FleCSI) project defines a new software architecture for multiphysics code development that promotes code agility through the separation of concerns between foundational computer science and the expression of physics algorithms. This is achieved through an application-customizable interface that defines the mesh and particle needs of the code together with models for defining and accessing physics data as shown to the right. The new software architecture has been used to field a diverse set of physics application codes and is now being integrated into LANL's plans for its next generation of integrated design codes.

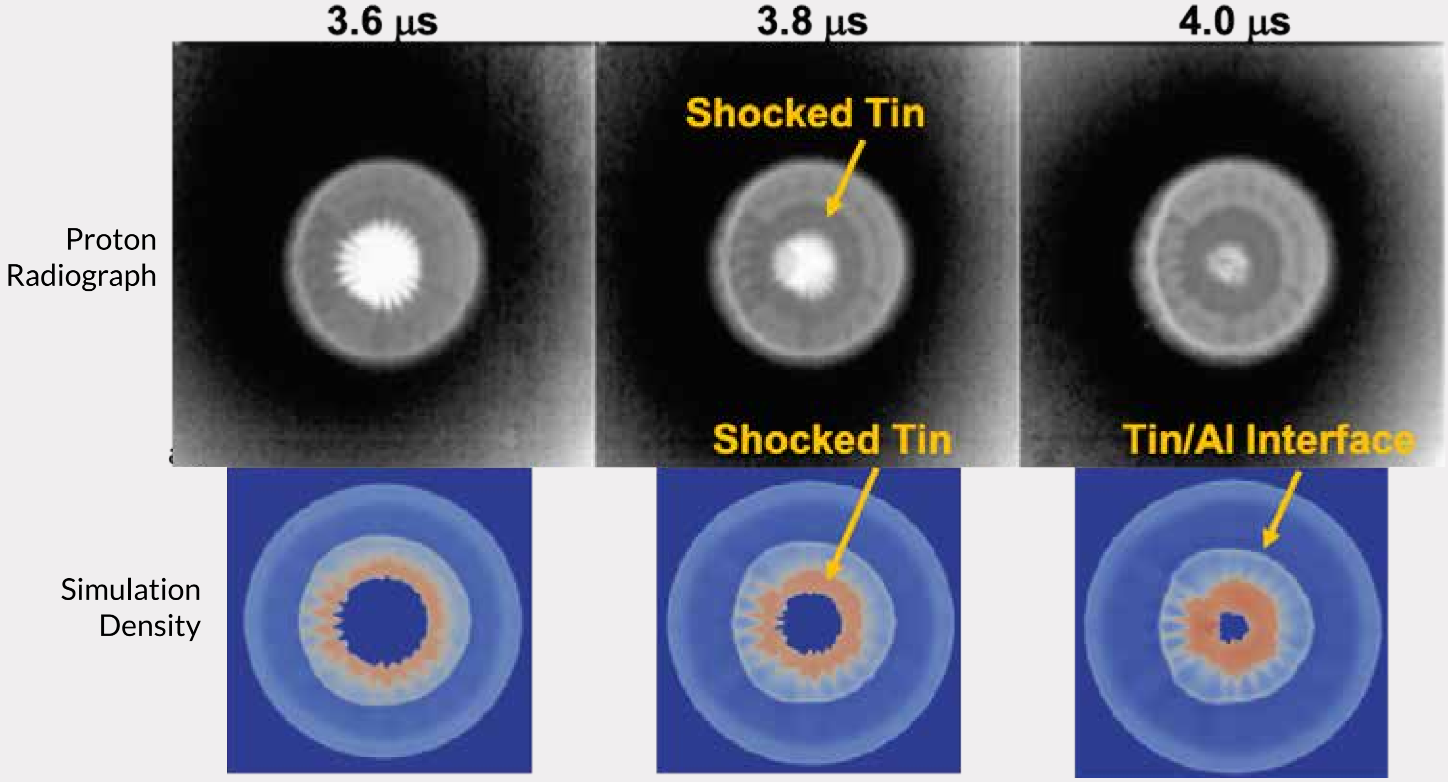
As part of the ATDM initiative, the FleCSI-based set of physics codes was developed to inform and validate the development environment, including mesh- and particle-based codes.



This notional design of LANL's ATDM software stack includes an abstraction layer (FleCSI) separating computer architectures from numerical physics implementations. This provides agility in adapting to different parallel runtimes, and enables exploration of new numerical methods for the physics relevant to complex national security mission.

Los Alamos NATIONAL LABORATORY EST. 1943

RISTRA PROJECT



Simulations employing a diverse set of physics disciplines were carried out to ensure that this software framework supports the necessary physics complexity for LANL's national security mission spaces, as well as the agility to facilitate numerical methods exploration. An example from this test suite is pictured here. A FleCSI-based hydrodynamics code with a treatment for solids was used to model a Tin pulsed power experiment at SNL's Z-pinch facility. The simulated results (bottom) are shown at comparable times to experimental proton radiography results (top).

The key goal of accelerating mission-relevant, multi-physics code development was the focus of a 2020 milestone where multi-scale methods for both radiation hydrodynamics and for grain-structure-aware continuum dynamics were successfully demonstrated.

The figure above shows an example of experiment-to-simulation comparison for a Ristra hydrodynamics code with a treatment for solids.

MACHINE LEARNING & ARTIFICIAL INTELLIGENCE

The artificial intelligence accelerator from SambaNova Systems integrated into the Corona cluster. (LLNL)

New data-science needs are driving development in the multiphysics codes. Machine learning (ML) and artificial intelligence (AI), in their infancy when ASCI began, are now revolutionizing a wide variety of fields. ASC is investing in AI/ML across the stockpile stewardship mission space through several comprehensive and collaborative efforts.

AI/ML methods are automating and improving the effectiveness of engineering and physics analyses. These methods enable advanced data-driven analysis that can incorporate a wide range of surveillance and other diagnostics. For example, AI-driven image analysis techniques have been applied to the analysis of experimental images, such as radiographs, reducing the time of analysis and increasing reproducibility. These methods also can be applied to making seamless many of the workflows that characterize simulations in weapons design, production, qualification, and certification activities. For example, LANL is exploring machine learning to impact efficiency in pit production. Early results show promise for relaxing tolerances in manufacturing processes, as well as removing costly steps like chemistry analyses outside of gloveboxes. Optimization of process steps like these brings the potential for significant production cost and schedule savings.

These methods are also improving the predictive capabilities of the multiphysics codes. The laboratories are using machine learning to reduce the time required to develop simulation models, automate the selection of numerical solvers, and even automate the selection of computational resources that are applied to run the models. SNL is investigating use of machine learning to create data-driven materials models for solid mechanics, learn device radiation models, segment computed tomography (CT) scanned parts, classify aging of components, and predict microstructure damage. At LLNL, AI-driven simulation controls for ALE hydrodynamics methods are increasing robustness and decreasing user intervention for complex problems. AI-driven surrogate modeling at all three laboratories has been shown to effectively produce useful surrogate results for hydrodynamics simulations at ~1000-fold speedups over the full simulation. This technique allows an extremely fast exploration of design space to focus higher fidelity calculations where they can have the largest impact for scientific discovery.

AI and ML are also impacting decisions on the types of hardware acquired. Integrating AI-enabled hardware with more traditional HPC systems provides an approximately 40-fold increase in throughput over GPUs for atomic physics calculations. This work demonstrates the continued scaling gains that can be realized as AI solutions grow in scientific computing applications. Continuing and expanding work in AI can revolutionize the ASC program by providing new solutions to long-standing challenges. The impact of these solutions will be felt in both hardware and software areas.

Cerebras System CS-1 AI computer after integration with Lassen. (LLNL)

RECOGNITION & AWARDS

Supercomputing development has always involved competition. This competition helps propel dramatic growth in computing and, as a consequence, in scientific endeavor generally. Not only does it fuel the human desire to excel, but it also brings in peer review and outside perspectives.

Important benchmarking measures include the GREEN500 list, which compares the energy efficiency of machines running problems at large scales, and the TOP500, which tests for the speed at which a computer can perform floating-point operations. A third, widely used measure is the High Performance Conjugate Gradients (HPCG) benchmark, developed at SNL in partnership with university collaborators (Dongarra, J., & Heroux, M.A., 2013). The HPCG benchmark was designed to mimic the data access and usage patterns associated with a physics-based code. Together, the TOP500, GREEN500, and the HPCG provide relatively simple metrics to compare different aspects of HPC system performance.

The Association for Computing Machinery (ACM) Gordon Bell prize is a prestigious award focused on capturing the impact of supercomputing from groundbreaking simulations (Bell, 2017). This annual award focuses on the teams of scientists who perform unprecedented simulations that push the boundaries of supercomputing in a variety of fields, including turbulence, material research, and cutting-edge biomedicine. The machines required by these simulations tend to be in the top three of the TOP500 list each year they are awarded, due to the large amount of computing resources required. Because this prize measures the impact of computing, the number of prizes awarded every year varies from just one in the peak performance category to multiple prizes in this category, or even the creation of new categories to reward groundbreaking impacts. It is a capstone achievement for laboratory scientists to become finalists for the awards.

Teams from all three NNSA laboratories are frequent finalists or prize winners, due to the world-leading machines that ASC has deployed. Over time, LLNL, LANL, and SNL have participated in more than 20 prize-winning or finalist teams.

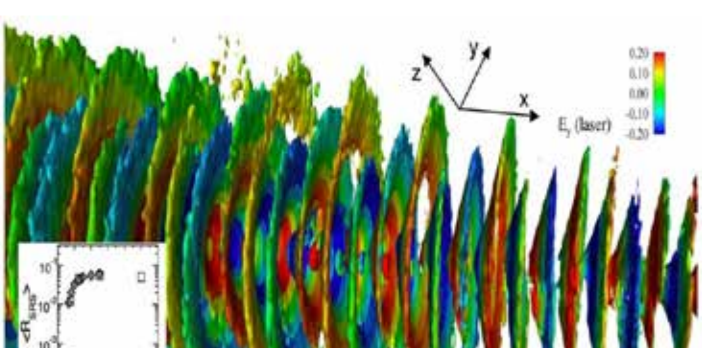
The list of accomplishments below demonstrates the range of impacts of ASC machines on the field of computing:

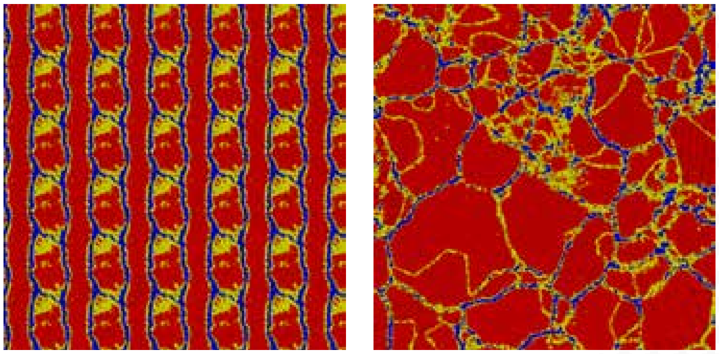
- 1997, Peak Performance, "Simulating the Motion of 322,000,000 Self-gravitating Particles": 430 gigaFLOPS on ASCI Red using 4,096 processors.
- 1999, Peak Performance, "Very High-resolution Simulation of Fluid Turbulence in Compressible Flows": 1.04 teraFLOPS

sustained on 3840 CPUs of the ASCI Blue Pacific machine.

- 2002, Special Award, "Salinas: A Scalable Software for High Performance Structural and Solid Mechanics Simulation." The Salinas code from the Sierra structural and solid mechanics code suite achieved a sustained 1.16 teraFLOPS performance on 3,375 ASCI White processors.
- 2005, Peak Performance, "100+ teraFLOPS Solidification Simulations on BlueGene/L." The team achieved a sustained rate of 101.7 teraFLOPS over a seven-hour run on the IBM BlueGene/L's 131,072 processors. LLNL researchers and the Blue Gene/L machine would go on to win the 2006 and 2007 Peak Performance awards.
- 2008, Peak Performance finalist, "0.374 petaFLOPS Trillion-particle Particle-in-cell Modeling of Laser Plasma Interactions on Roadrunner." The team used the Vector Particle-In-Cell (VPIC) code on Roadrunner to study details of laser-plasma instabilities.
- 2013, High Performance, "11 petaFLOPS simulations of cloud cavitation collapse," high throughput simulations of cloud cavitation collapse on 1.6 million cores of Sequoia reaching 55% of its nominal peak performance.
- 2015, Sequoia, Gordon Bell prize winner.
- 2020, Gordon Bell Special Prize finalist for High Performance Computing-Based COVID-19 Research, "Enabling Rapid COVID-19 Small Molecule Drug Design Through Scalable Deep Learning of Generative Models." The research team's computing approach, using tensor cores, was applied successfully to all of Sierra, achieving 97.7 percent efficiency. The team trained the machine learning model on an unprecedented 1.6 billion small molecule compounds and one million additional promising compounds for COVID-19.

The left image is from the 2005 Gordon Bell for Peak Performance won by LLNL on BlueGene/L. The right image is from the 2008 Peak Performance finalist team whose Vector Particle-in-Cell code ran on ASC Roadrunner from LANL.





THE FUTURE OF ASC™

Attaway system with glass floor tiles allows for data center facilities team members to easily verify the operational status of the system. (SNL)

The coming years hold several challenges for the ASC program. The largest supercomputers are running into limits of the technologies that enabled the forefront of computing to be relentlessly pushed (Future High Performance Computing Capabilities, Summary Report of the Advanced Scientific Computing Advisory Committee [ASCAC] Subcommittee, 2019). Over the entire span of the ASC program, increases in supercomputer speed have outpaced Moore's Law, roughly doubling at intervals of less than two years by harnessing design innovations in parallelism and micro-electronics. This impressive accomplishment represents the fruition of ASCI's "accelerated" computing approach.

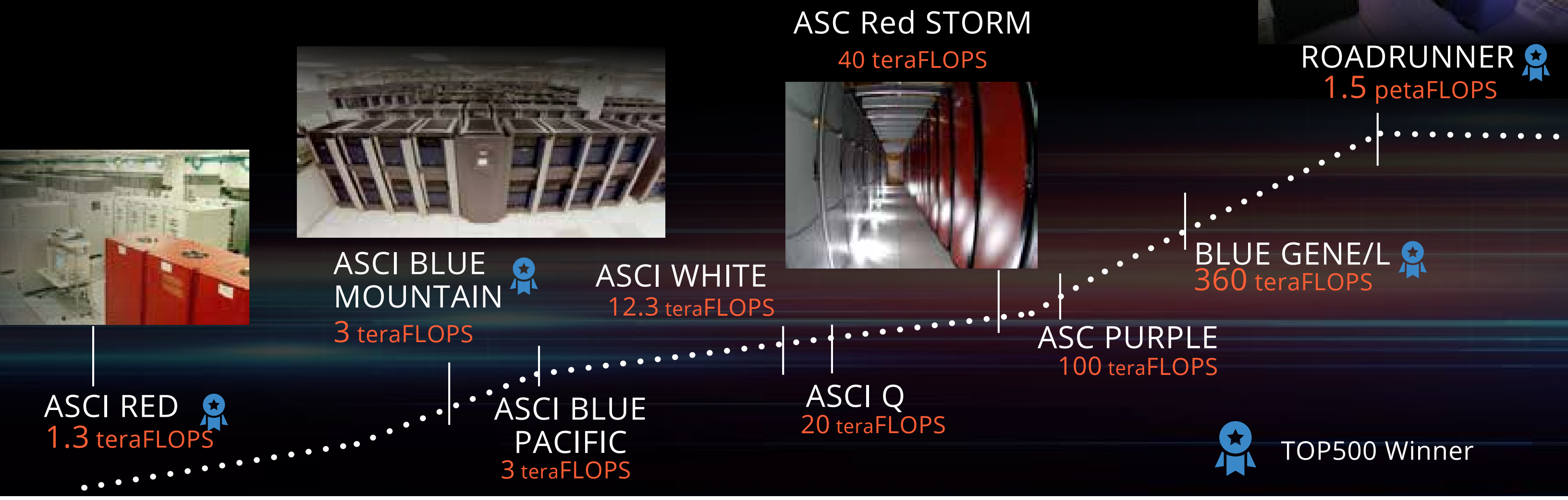
However, many of these innovations are becoming less effective at increasing supercomputer speed and there are hard limits that could soon be encountered. Perhaps the most fundamental limits are physical. Moore's Law, which is based on increased transistor density, has quantum mechanical limits on how far circuits can continue to shrink. Another limitation is reliability—once machines grow over 1 million processors, engineering-in reliability becomes increasingly difficult. While supercomputers, per FLOPS, are becoming much more power efficient, they still consume more power at the high end than ever before. An exascale system rated at 40 megawatts peak power could potentially see operating costs that are prohibitively expensive, which could substantially lower the attractiveness of procuring the system.

Fortunately, while traditional approaches to achieving performance are potentially slowing down, other areas remain to be explored. One new approach will continue advancing the relationship between CPUs and GPUs to simplify and improve heterogeneous computing even further.

Perhaps the most important tools in navigating the technical headwinds facing ASC are the collaborative mindset and the people that contribute to ASC and NNSA's national security mission. The collaborative framework embodied in the ASCI principle of "One Program/Three Labs" encourages innovation, cooperation, and collaboration to embrace the ongoing dramatic pace of change. Many of the technologies that have enabled computing to achieve petascale and exascale rates had not even been imagined at the start of ASCI. The next 25 years will likely deliver a similar new set of technologies that should dramatically increase NNSA's computing abilities. The ASC program, given its past record of accomplishment and its current culture, is well positioned to take advantage of these exciting opportunities.

Page intentionally left blank

A QUARTER CENTURY OF ADVANCED SIMULATION & COMPUTING PROGRAM (ASC) ACCOMPLISHMENTS



1995

- The Accelerated Strategic Computing Initiative (ASCI) is created immediately after the end of the Cold War. The Initiative funds simulation and modeling efforts that provide high-fidelity computer simulations of weapon systems and help maintain the credibility of the U.S. nuclear deterrent.
- SNL SIERRA becomes the first ASCI code suite to support multiple applications, which are built with a common core infrastructure.

1996

- ASCI Red at SNL is the first major ASCI system to break the 1-teraFLOPS barrier and remains the #1 supercomputer until 2000. Experience gained from ASCI Red contributes to over a decade of U.S. leadership in supercomputing.

1998

- LANL deploys its first ASCI platform, the Blue Mountain machine.

1999

- ASCI Blue Pacific at LLNL wins a Gordon Bell Prize for Peak Performance with turbulence simulation achieving 1.04 teraFLOPS sustained on 3840 CPUs.

2000

- ASCI White at LLNL starts the second wave of commodity massively parallel machines that will include ASCI Q and Purple. This machine remains the #1 supercomputer until 2002.

2002

- The SNL Salinas code for structural and solid mechanics wins a Gordon Bell Special Award. The code achieved a sustained 1.16 teraFLOPS performance on 3,375 White processors.

2003

- Lustre parallel file system is released for general usage. Lustre makes it possible for potentially thousands of processors to read and write to a unified file system.
- ASCI Q at LANL becomes the top ASCI supercomputer and remains in the top 3 supercomputers until 2004.

2005

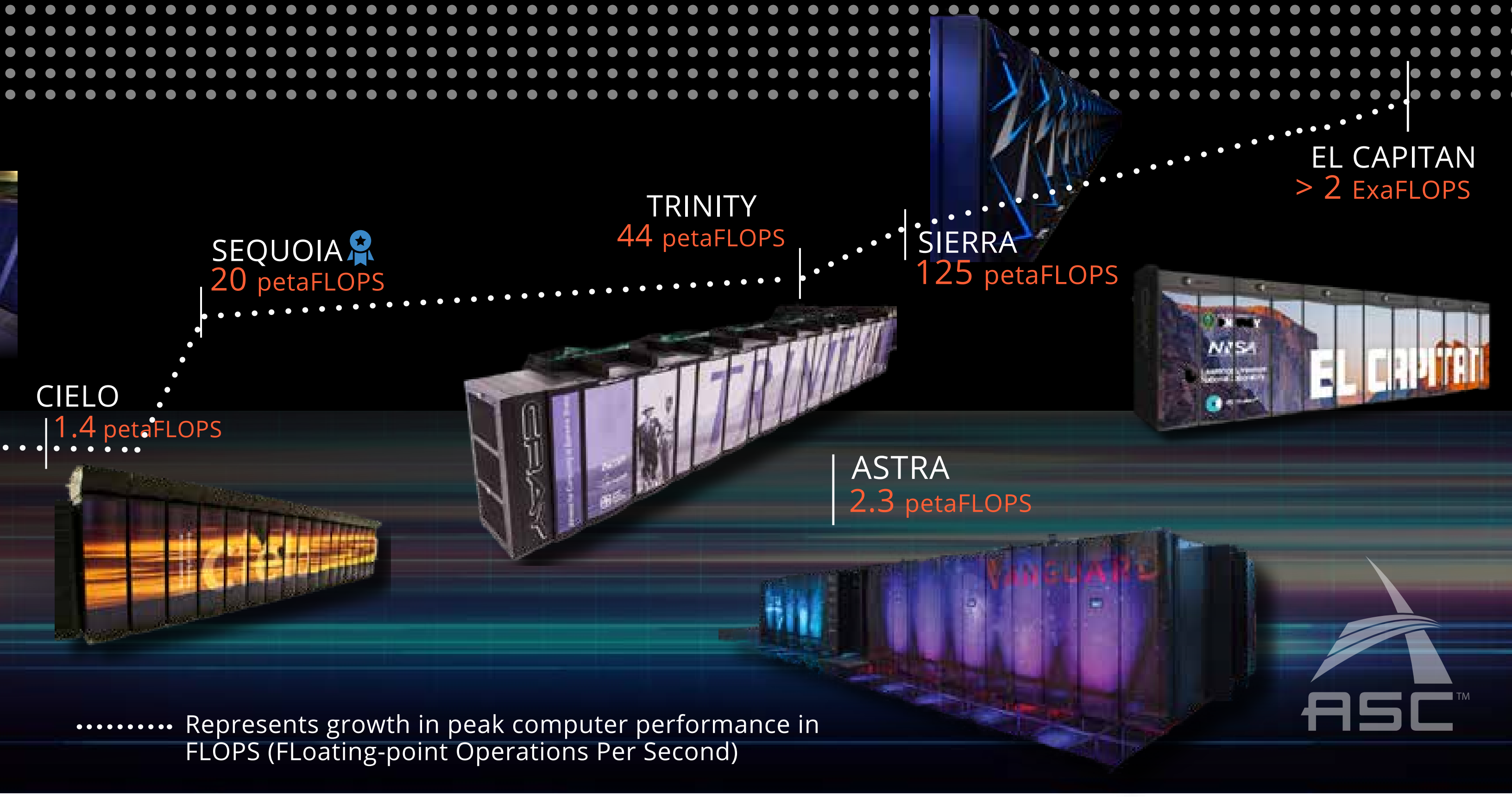
- Blue Gene/L at LLNL wins a Gordon Bell Prize for Peak Performance with simulations achieving a sustained rate of 101.7 teraFLOPS on 131,072 processors. Blue Gene/L simulations would go on to win Gordon Bell prizes in 2006 and 2007.
- ASCI Purple accomplishes the ambitious 10-year goal of fielding a 100-teraFLOPS machine capable of high-resolution 3D stockpile simulations.

2006

- Red Storm is upgraded resulting in a #2 ranking in the TOP500. The machine is further upgraded to XT4 in 2008, doubling its computational speed.

2008

- LANL Roadrunner system pioneers the use of a central processor in a supercomputing node paired with several floatingpoint accelerators
- Roadrunner breaks the petaFLOPS barrier and assumes the number one spot on the TOP500. Roadrunner is a Gordon Bell Prize for Peak Performance finalist with laser plasma simulations



achieving a speed of 0.37 petaFLOPS.

» The NNSA New Mexico Alliance for Computing at Extreme Scale (ACES) is created by LANL and SNL.

2011

» Cielo machine at LANL reaches over 140,000 processing cores, exceeding 1 petaFLOPS.

» DesignForward and FastForward, a series of collaborations between DOE and industry, are established by ASCR and ASC.

2013

» Sequoia at LLNL wins the Gordon Bell Prize for High Performance with simulations of cloud cavitation collapse achieving 11 petaFLOPS on simulations of 1.6 million cores.

2014

» The Collaboration of Oak Ridge, Argonne, and Livermore (CORAL) partnership forms as a joint procurement activity.

2015

» Launch of the National Strategic Computing Initiative (NSCI), a government-wide effort to develop exascale computing.

» Exascale Computing Initiative (ECI), a joint program through the Office of Science and NNSA, is created.

» Trinity (ATS-1) is deployed at LANL as the first ATS machine, a ground-breaking capability.

» Sequoia wins its second Gordon Bell prize for highperformance computing.

2016

» ECP PathForward program starts with the goal of a sustainable industrial base for exascale computing.

2018

» The Vanguard program starts developing experimental computing systems with the potential to mature into the next ATS platform.

» The Astra supercomputer at SNL developed through Vanguard is the first Arm petascale machine to use Arm processors.

» Sierra (ATS-2) at LLNL launches and stays in the top three supercomputers of the TOP500 list from 2018 to 2021.

2020

» Sierra at LLNL is a finalist for the Gordon Bell Special Prize for HPC-based COVID-19 research with work that used scalable deep learning to enable rapid small molecule drug design.

2021-2023

» Early Access EAS-2 (2021) and EAS-3 (2022) systems prepare ASC for the exascale El Capitan (ATS-4) system.

» El Capitan at LLNL, when deployed in 2023, is projected to be the first ASC exascale machine with at least 2 exaFLOPS peak performance.

» Crossroads Phase 1 will deploy in 2022 and feature memory and other technologies needed for exascale computing.

REFERENCES

A Research and Development Strategy for High Performance Computing. (1987). Washington, DC: Office of Science and Technology Policy.

Accelerated Strategic Computing Initiative (ASCI) Program Plan. (2000). U.S. Department of Energy.

Advanced Simulation and Computing Strategy. (2013). SNL, SAND 2013-3196P.

Advancing Simulation Science: The Legacy of ASC Academic Strategic Alliance Program. (2011). National Nuclear Security Administration Office of Advanced Simulation and Computing.

Ang, J. (2015). *DOE's FastForward and DesignForward R&D Projects: Influence Exascale Hardware.* SNL SAND2015-0626PE.

Ang, J. D. (2010). The Alliance for Computing at the Extreme Scale. *Cray User Group 2010 Proceeedings*, (p. 8).

Approval of Memorandum of Understanding (MOU) for the Coordination of Exascale Computing Project (ECP). (2016). DOE, eDOCS #EXEC-2016-006582/NNSA-2016-000332.

Approval of the DOE Exascale Initiative MOU between DOE Office of Science and NNSA Office of Defense Programs. (2011). DOE, eDOCS # NNSA-2010-002269.

Archer, B. J. (2017). *Seventy Years of Computing in the Nuclear Weapons Program.* LANL, LA-UR-17-24329.

ASCR@40 Highlights and Impacts of ASCR's Programs. (2020). U.S. DOE Office of Advanced Scientific Computing Research.

Baker, M. J. (2019). *B61-12 Team Reaches Milestones in Nuclear Deterrence Mission.* SNL, SAND2019-0080 L: Sandia Lab News.

Baker, M. J. (2020). *Virtual Support from Sandia Enabled B61-12 Flight Test on F-35A Jet Fighter.* SNL, SAND2020-8544 E: Sandia National Laboratories news release website.

Barney, B. A. (2010). *100 teraFLOPS Dedicated to Capability Computing.* LLNL, UCRL-WEB-207397.

Bell, G. B. (2017). A Look Back on 30 Years of the Gordon Bell Prize. *The International Journal of High Performance Computing Applications, 31*(6), 469-484.

Bishop A. (2009). *Scientific Grand Challenges for National Security: The Role of Computing at the Extreme Scale.* Washington, DC: PNNL, PNNL-19379.

Brooks, E. (1989, November 12-17). Attack of the Killer Micros. *1989 ACM/IEEE conference on Super Computing*. Reno, Nevada, United States.

Carr, A. B. (2013). *Punchcards to PetaFLOPS.* LANL, National Security Science.

Christon, M. A. (1997). *ASCI Red: Experiences and Lessons Learned with a Massively Parallel TeraFLOP Supercomputer.* SNL, SAND-97-0463C.

Crawford, D. (2021). The Causes and Origins of ASCI. *Salishan Conference on High-Speed CComputing, 40th Anniversary Lecture Series*.

Creating a National Strategic Computing Initiative. (2015). Executive Order (EO) 13702.

Dennig, Y. (2021). *2020 A New Decade of High Performance Computing.* SNL, SAND2021-2936M.

Dongarra, J., J., & Heroux, M. A. (2013). *Toward a New Metric for Ranking High Performance Computing Systems.* Sandia report, SAND2013-4744, 312.

Fiscal Year 2021 Stockpile Stewardship and Management Plan: Biennial Plan Summary. (2020). Report to Congress.

Future High Performance Computing Capabilities, Summary Report of the Advanced Scientific Computing Advisory Committee (ASCAC) Subcommittee. (2019).

Heer, T. (2021). *DOE High Performance Storage Systems (HPSS) Eclipse One Exabyte in Stored Data.* LLNL, LLNL-ABS-825363.

Heller, A. (2018, October). Extending the Life of a Workhorse Warhead. *Science and Technology Review*.

Hemmert, K.S. (2016). Trinity: Architecture and Early Experience. *Cray User Group 2016 Proceedings.*

Howard, M. (2017). Performance Portability in SPARC: Sandia'a Hypersonic CFD Code for Next-Generation Platforms. *DOE COE Performance Portability Meeting.* SNL.

Hungerford, A. (2021). A Look at Trends in Computational Methods: The Tie Between Machines and Mission Science. *Salishan Conference on High-Speed Computing.* LANL, LA-UR-21-23988.

Karlin, I. B. (2013). Exploring Traditional and Emerging Parallel Programming Models Using a Proxy Application. *Proceedings of the 27th IEEE International Parallel & Distributed Processing Symposium* (pp. 1-14). Boston, MA: LLNL-CONF-586774.

Kogge, P. (2008). *ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems.* DARPA IPTO, AFRL Contract #FA8650-07-C-7724.

Kothe, D. L. (2019). Exascale Computing in the United States. *Computing in Science and Engineering, 21*(1).

Larzelere, A. (2009). *Delivering Insight: The History of the Accelerated Strategic Computing Initiative (ASCI).* LLNL, UCRL-TR-231286.

Owen, M. J. (2018). *SPHERAL.* LLNL, LLNL-WEB-656052.

Parker, A. (2018, July). X-Ray Machine Does Double Duty. *Science and Technology Review*.

Patterson, E. (2013, April). Roadrunner: On the Road to Trinity. *National Security Challenge*, 7-13.

Pit Lifetime. (2006). The MITRE Corporation, JSR-06-335.

Rieben, R. N. (2020). The MARBL Multi-physics Code. *Exascale Computing Project Annual Meeting.* LLNL, LLNL-POST-802190.

Rieben, R. N. (2021). Mission Driven Innovation: A Case Study in Multiphysics Code Development. *Salishan Conference on High-Speed Computing.* LLNL, LLNL-PRES-821617.

Sedov, L. I. (1993). *Similarity and Dimensional Methods in Mechanics* (10th ed.). Boca Raton: CRC Press.

Supinski, B.R., & Louis, S. (2005). Initial BlueGene/L Application Performance Results. *Salishan Conference on High-Speed Computing, LLNL, UCRL-PRES-211410*.

Thomas, J. (2018). *Lawrence Livermore Unveils NNSA's Sierra, World's Third Fastest Supercomputer.* LLNL, LLNL-WEB-458451.

Vazhkudai, S. S. (2018). The Design, Deployment, and Evaluation of the CORAL Pre-Exascale Systems. *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis.* Dallas, TX.

Weaver, R. (2002). *Massively Parallel Simulations with DOE's ASCI Supercomputers: An Overview of the Los Alamos Crestone Project.* LANL, LA-UR-02-6424.

Weaver, R. P., & Dearholt, W. R. (2011). *A Comparison of 2D and 3D Hydro Simulations of Asteroid Mitigation by a Strong Surface Explosion.* LANL, LA-UR-11-06824.

Zecevic, M.E. (2021). Viscoplastic Self-consistent Formulation as Generalized Material Model for Solid Mechanics Applications. *Applications in Engineering Science, 6*, 1-16.

Back Cover: Using a computational fluid dynamics code, a supercomputer can produce a model that simulates turbulence when materials (such as gases, liquids, or metals) mix and change states. For example, when the cold water of Antarctica mixes with the warm Gulf Stream, the interaction produces turbulence, which is seen as vortices and curves and offers valuable information about material interactions. *National Security Science Magazine*, Winter 2000. (LANL)

25 YEARS *of*
ACCOMPLISHMENTS

ADVANCED SIMULATION & COMPUTING PROGRAM

1995-2020