

A complex circuit board with various electronic components like resistors, capacitors, and connectors, serving as the background for the report.

SAND2016-9388R

2016

HIGH PERFORMANCE COMPUTING
ANNUAL REPORT



Sandia National Laboratories

αΓΩΣ



SNLSimMagic®

AUGMENTED REALITY APP

Scanning this code with an iPhone or iPad will provide access to SNLSimMagic; an augmented reality iOS application that can be downloaded to the device. You can also download directly onto your device from the Apple App store. Readers with the application can use their mobile devices to scan images in this document that show the Augmented Reality icon, and an associated movie clip will be played on their device. SNLSimMagic was developed at Sandia National Laboratories.



CREDITS

Editor

Yasmin Dennig

Design

Stacey Long

Video

Regina Valenzuela

Photography

Lonnie Anderson

Dino Vournas

Randy Montoya

Danielle Oteri

Contributing Writers

Megan Davidson

Laura Sowko

SNLSimMagic

Trace Norris

Sandia Video Footage

Sandia Creative Services

Stock Video and Photography

Pond5.com stock footage

iStockphoto.com

thinkstockphotos.com

gettyimages.com

Special Video Attribution

1. Louisiana Salt Dome Composite Cross Section Diorama 20446_615b

Brian Lewis, Photographer

Courtesy of Louisiana State Exhibit Museum, Shreveport

(Article: Hexahedral Finite Element Mesh to Fit Complex Geological Geometries)

2. Blue Waters Computer Image

Courtesy of the National Center for Supercomputing Applications

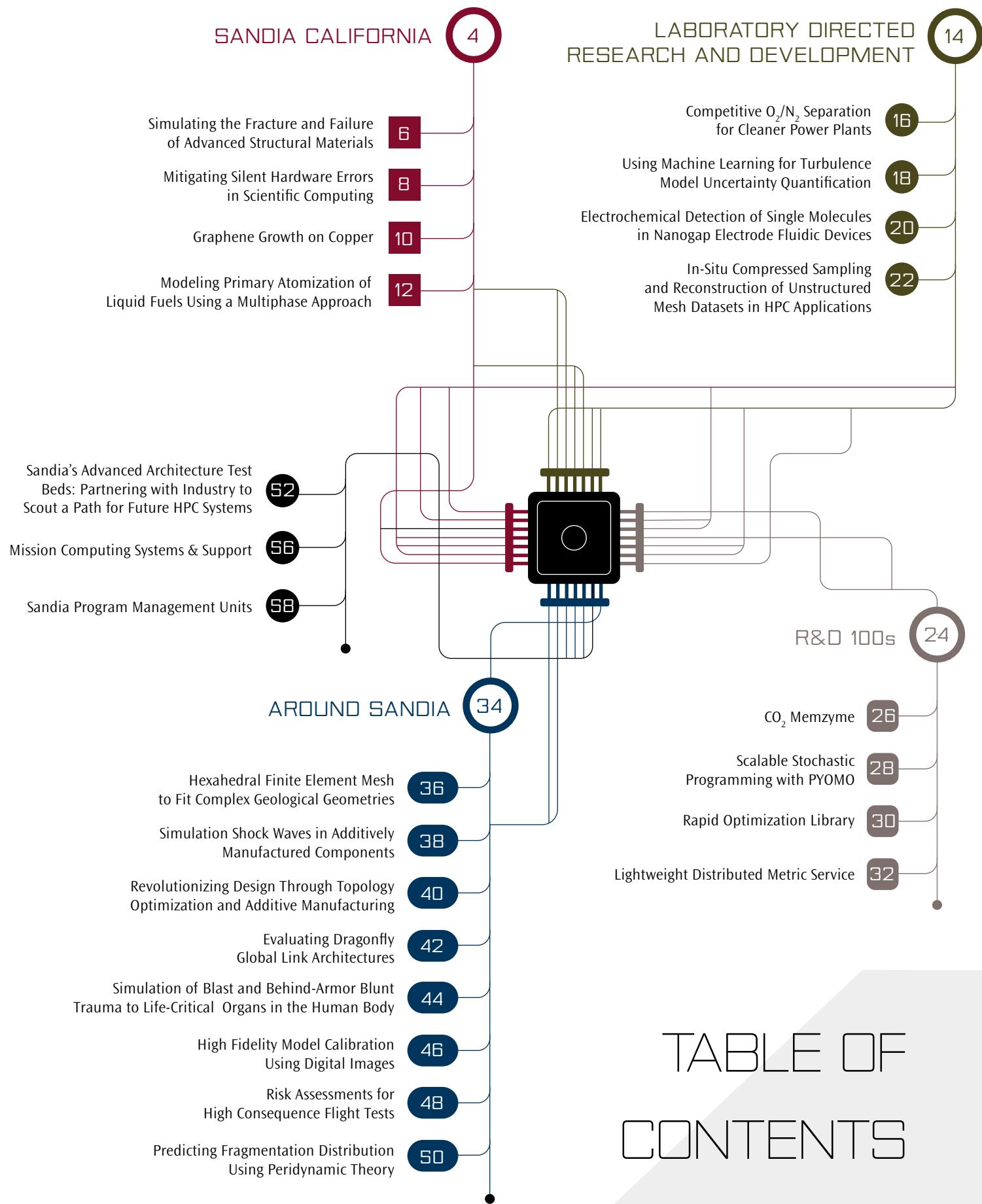
and the Board of Trustees of the University of Illinois

(Article: Lightweight Distributed Metric Service)

3. Antares and Space-X Launch Footage

Courtesy of NASA

(Article: Risk Assessments for High Consequence Flight Tests)



SANBIA NATIONAL LABORATORIES

**HIGH
PERFORMANCE
COMPUTING**



INVESTING IN EXPANDED COMPUTATIONAL CAPABILITIES

The projects highlighted in this edition of our High Performance Computing (HPC) Annual Report were enabled by Sandia National Laboratories' HPC capabilities. Computational simulation has become central to many areas of research and Sandia continues to invest in growing this critical capability, while establishing leadership as a nationally-recognized HPC facility.

Each year we work toward enhancing our high performance computing capabilities by investing in new systems and improving infrastructure as we

continue to acquire faster systems. This year, through the Advanced Simulation and Computing (ASC) program, we acquired the first petaflop machines to be sited at Sandia, Cayenne and Serrano. Through the Institutional HPC program, we also acquired new scientific computing systems to meet mission demand – “Solo,” a 400 teraflop system, for our External Collaboration Network, and “Dark Ghost,” an 800 teraflop system, to address sensitive computing needs. We also expanded our capabilities in data analytics by placing into operation two new platforms acquired at the end of last year: Ray (sited in New Mexico) and Kahuna (sited in California). Additionally, we’re investing in infrastructure changes to support liquid cooling to reduce operating energy costs and reduce noise hazards in our data center. In addition to SkyBridge (operational in fiscal year 2015), both of the new ASC machines (Cayenne and Serrano) are liquid cooled. The liquid cooled machines save 40% in infrastructure costs and about \$75,000 annually per machine in operating costs.

Our computing resources and expertise allow Sandia to continue to provide exceptional service in the national interest by using sophisticated computational modeling to provide solutions to national scientific challenges. As the importance of computational modeling grows, it is critical to continue expanding our computing capabilities to remain on the cutting edge of innovation and discovery.

— **Leonard M. Napolitano, Jr.**
*Acting Chief Information Officer and
Vice President for Information Technology Services*



SANDIA
CALIFORNIA



CELEBRATING 60 YEARS OF INNOVATION AT SANDIA CALIFORNIA

Sandia National Laboratories' California site is celebrating its 60th anniversary (1956 to 2016), and high performance computing has been a key enabler for its scientists and engineers throughout much of its history.

Since its founding, Sandia California has helped pioneer the use of HPC platforms including hosting Sandia's first Cray-1 supercomputer in the 1970s and supporting development of scalable cluster computing platforms to create a new paradigm for cost-effective supercomputing in the 1990s.

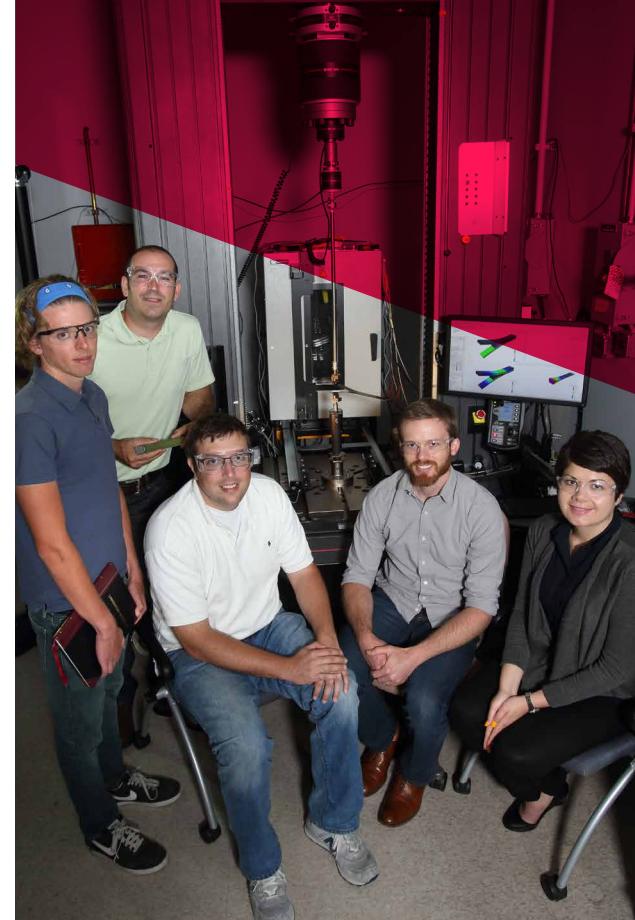
Recent decades of investment in creation of scalable application frameworks for scientific computing have also enabled new generations of modeling and simulation codes. These resources have facilitated computational analysis of complex phenomena in diverse applications spanning national defense, energy, and homeland security.

Today, Sandia California researchers work with partners in academia, industry, and national labs to evolve the state-of-the-art in HPC, modeling, and data analysis (including foundational capabilities for exascale computing platforms) and apply them in transformational ways. Research efforts include mitigating the effects of silent hardware failures that can jeopardize the results of large-scale computations, developing exascale-capable asynchronous task-parallel programming models and runtime systems, formulating new techniques to better explore and analyze extreme-scale data sets, and increasing our understanding of materials and chemical sciences which has applications spanning nuclear weapons stockpile stewardship to more efficient automobile engines.

The following section highlights some of these research and applications projects and further illustrates the breadth of our HPC capabilities.

— **Heidi R. Ammerlahn**
*Acting Director, Research & Development
Science & Engineering*

SIMULATING THE FRACTURE AND FAILURE OF ADVANCED STRUCTURAL MATERIALS



STACY NELSON (PI)

ALEXANDER HANSON

BRIAN WERNER

SHAWN ENGLISH (PI)

TIMOTHY BRIGGS

Without the computing resources currently available to Sandia's analysts, a model's predictions could not be fully verified and validated and, subsequently, a lower level of confidence would be held in the numerical tool's ability to formulate realistic simulations.

Scientists and engineers at Sandia are conducting collaborative research to develop the finite element analysis (FEA) tools necessary to make realistic and reliable predictions regarding the performance of composite materials and structures

The methodology under development utilizes a building block approach, in which the relevant phenomena exhibited by composite materials experiencing fracture are isolated and examined individually prior to investigating more complex combinations of fracture related behaviors. First, simulations of a composite's microstructure, via a representative volume element (RVE), are used to understand and define the composite's macroscopic material properties in preparation for global simulations of a composite structure. (See Figure 1)

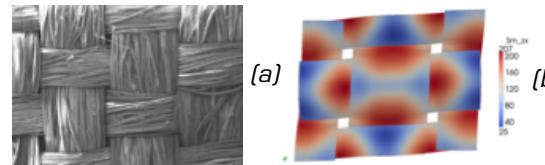


Figure 1. RVE of a fiber-reinforced composite sample
(a) micrograph and (b) simulated



The next step in the building block approach involves the examination of interlaminar fracture, or delamination, between the individual plies within a fiber-reinforced composite material. This behavior is well understood with experiments, such as the double cantilever beam (DCB) test. Therefore, delamination within a fiber-reinforced composite material is characterized with DCB experiments that are concurrently modeled in the process of developing the computational tools required to predict fracture. (See Figure 2)



Figure 2. Example double cantilever beam (a) experiment and (b) FEA model

The building block approach then considers the residual stresses that form within a composite structure due to the material's thermally-driven manufacturing process and differences in the rates of thermal expansion among the composite's constituents. This phenomenon is being experimentally characterized and computationally modeled through the examination of simple composite structures, such as flat plates composed of dissimilar composite materials, that demonstrate the presence of residual stresses with measurable deformations. (See Figure 3)

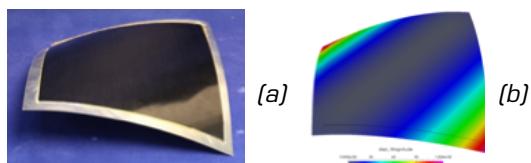


Figure 3. Deformation of a composite plate exhibiting residual stresses (a) experiment and (b) computational model

Finally, these capabilities are being applied to models of realistic test structures experiencing the coupled effects of thermal residual stresses and applied mechanical loads, such as the low velocity impact (LVI) of composite plates. LVI is a common concern within the aerospace industry, as highly probable events, such as tool drop, bird strike, and interactions with runway debris, can cause the barely visible impact damage (BVID) of composite structures. BVID must be assessed with either sophisticated non-destructive testing techniques or through simulation. Therefore, with the existing collaborative relationships, physical puncture and impact tests have been completed to experimentally validate the computational simulations concurrently under development. (See Figures 4 and 5)

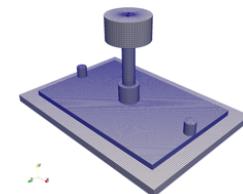


Figure 4. Computer model of the low-velocity impact experiment

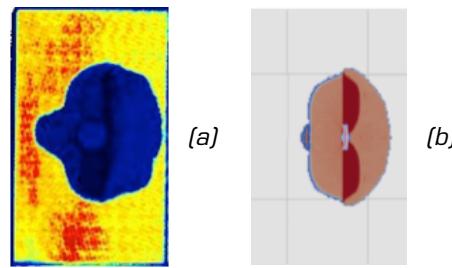


Figure 5. Composite panel tested under low velocity impact showing region of interlaminar delamination (a) ultrasonic scan and (b) modeled result

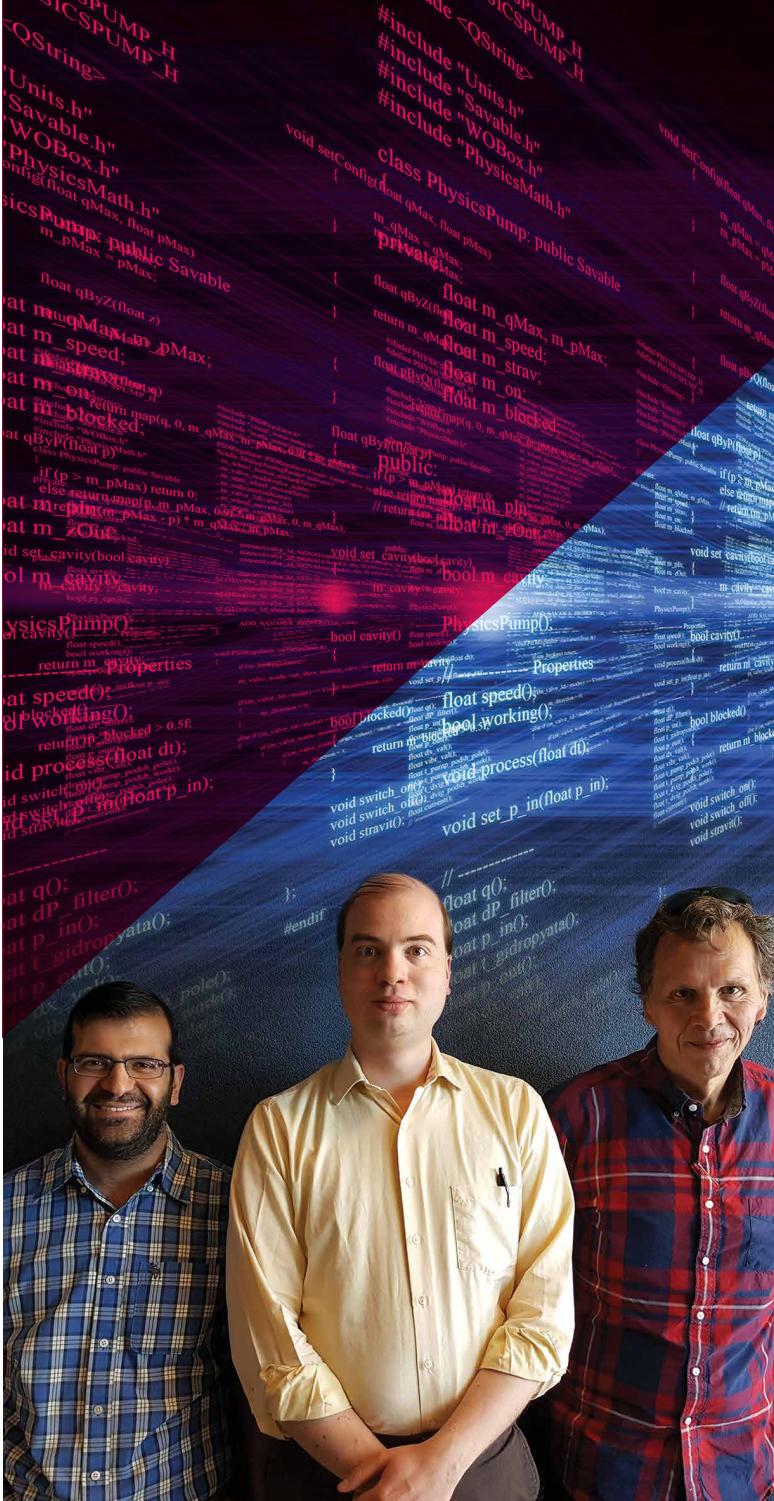
As the computational methods under development continue to be numerically verified and experimentally validated, these analysis techniques can be implemented with increasing confidence to support various aspects of Sandia's national security mission.

MITIGATING SILENT HARDWARE ERRORS IN SCIENTIFIC COMPUTING

JACKSON R. MAYO (PI)

MAHER SALLOUM

ROBERT C. ARMSTRONG



HPC can be programmed to simulate virtually any physical or digital system and thus inherits the dynamical complexity of such systems. This makes the design and analysis of HPC both difficult and powerful: HPC is not only a tool for science but also an exemplar. Understanding the resilience of HPC is synergistic with understanding the resilience of the complex systems that HPC can simulate.



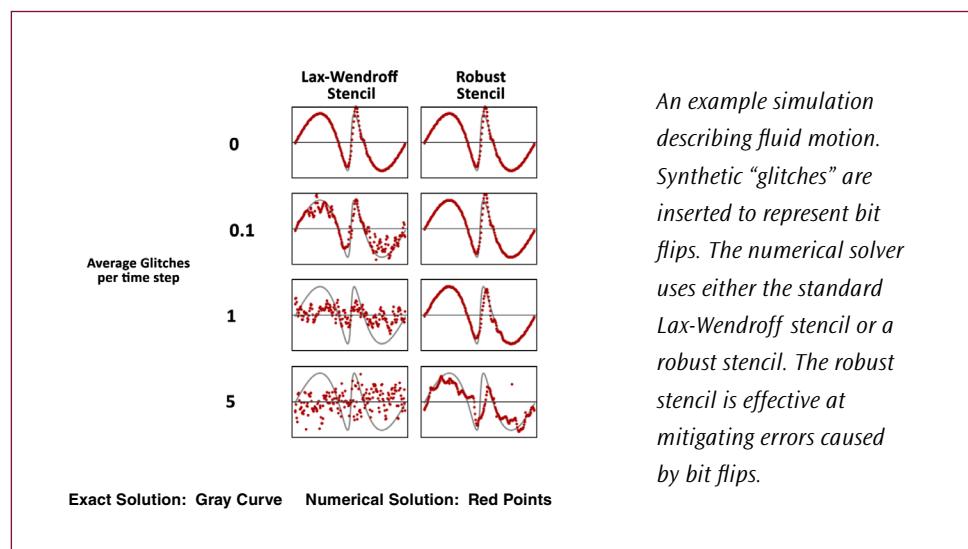
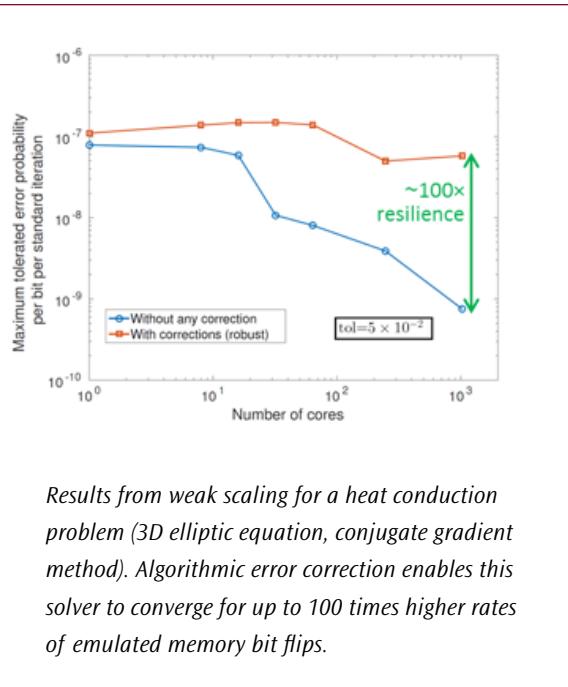
Computing hardware cannot be treated as perfectly reliable because failures that happen rarely in any given compute node will be seen much more frequently in a system of thousands of nodes. The goal of HPC resilience is to enable the design of hardware and software so that useful computational results are nevertheless obtained. Traditionally, programmers addressed directly observable failures such as nodes that “crash” and stop responding.

A different type of failure involves hardware that silently deviates from intended behavior. Silent data corruption (SDC), such as bit flips, can lead to wrong numerical results in a computation that otherwise appears normal, jeopardizing the correctness of large-scale scientific application runs.

Research on application-level resilience aims to enable tolerating the increased rates of these errors at extreme scale. This introduces the possibility that HPC systems could purposely use less reliable but more efficient hardware in the future. Ongoing Sandia work under the Advanced Simulation and Computing program is investigating algorithmic techniques for detecting and correcting SDC in physics simulations at low cost.

Prototype robust solvers have been implemented and tested for a scenario of random bit flips in memory. The mitigation works by comparing data points with their neighbors to identify outliers such as a simulated temperature value that is far off from its surroundings. These possibly corrupted points are replaced with an interpolation that is likely closer to the correct value.

We evaluated the solvers using software-injected bit flips as a synthetic source of errors. Example equations describing fluid motion and heat conduction are solved accurately by the robust methods in the presence of bit flip rates that are orders of magnitude higher than those tolerated by standard solvers, with runtime overhead as low as a few percent. Ultimately, this research can contribute to a practical resilience toolbox for production codes, adapting to the error characteristics of future HPC hardware and supporting both hardware and software design choices to achieve correctness most efficiently.



Scan this image with your SNLSimMagic App to see an associated movie clip.

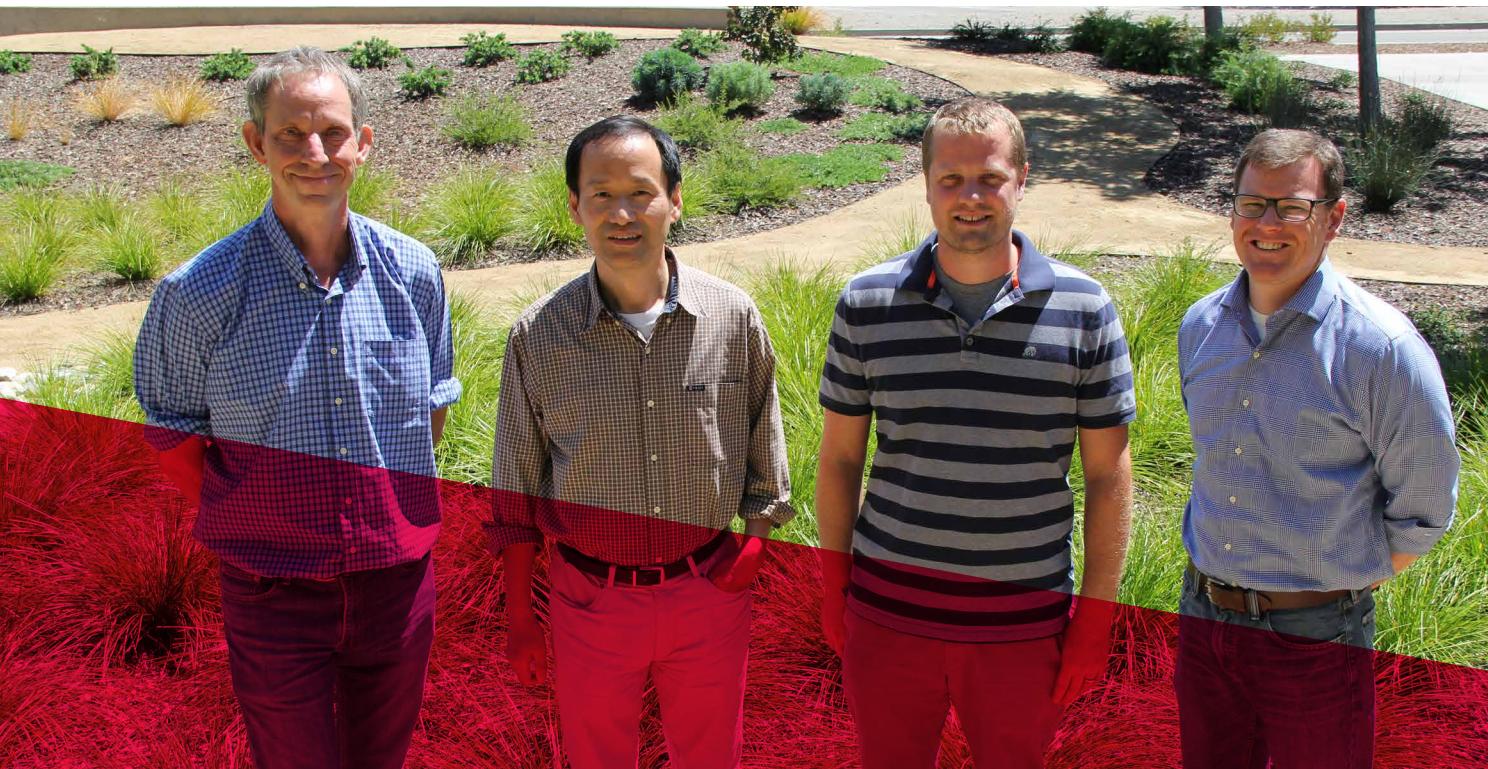
GRAPHENE GROWTH ON COPPER

NORMAN C. BARTELT (PI)

XIAOWANG W. ZHOU

DONALD K. WARD

MICHAEL E. FOSTER



HPC is essential for us to visually explore the atomic assembly mechanisms of graphene growth. Such visualization of atomic assembly mechanisms is inaccessible directly through any existing experimental techniques, but it can help explain experimental observations and guide the development of new graphene growth techniques with reduced cost and improved graphene products.

Graphene is a remarkable material with unique electrical, optical, thermal, and mechanical properties. While these properties provide transformational opportunities for many national security applications, an emergence of new graphene-based technologies will not occur unless a low-cost process can be used to manufacture large graphene sheets with low defect contents.

A group of Sandia scientists have developed an analytical bond order potential for carbon. This potential enables molecular dynamics (MD) simulations of crystalline growth of graphene on copper as shown in Figure 1. In such simulations, carbon atoms are randomly injected to a copper surface and positions of atoms are then solved from Newton's equation of motion without any assumptions. As a result, the simulations can reveal realistic growth mechanisms. However, graphene is composed of one atomic layer of carbon atoms constructed on a "honeycomb" lattice. Capturing the formation of such a unique structure from random atom clusters had been an elusive goal in the field of molecular dynamics simulations.

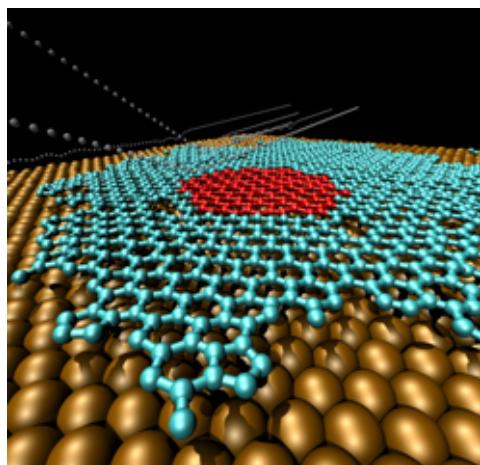


Figure 1. Simulated graphene growth on copper.

- big bronze balls: copper
- small red balls: carbon in pre-existing graphene island
- small blue balls: carbon in new graphene created during the growth simulation

Conventionally, defects in graphene have been thought to form at the boundaries between graphene islands when these islands impinge. Time-resolved atomic images of growing graphene obtained from our MD simulations are shown in Figure 2, where defects can be identified as the 5- or 7-member rings rather than the perfect 6-member rings, and are marked by white and green circles. These images surprisingly indicate that rather than the impingement mechanism, defects are continuously formed along the edge of graphene islands throughout the growth process. This finding suggests that improved growth should not focus only on the island impingement sites as done previously, but on the early stage of the growth as well. This new understanding can guide experiments to improve the growth of graphene.

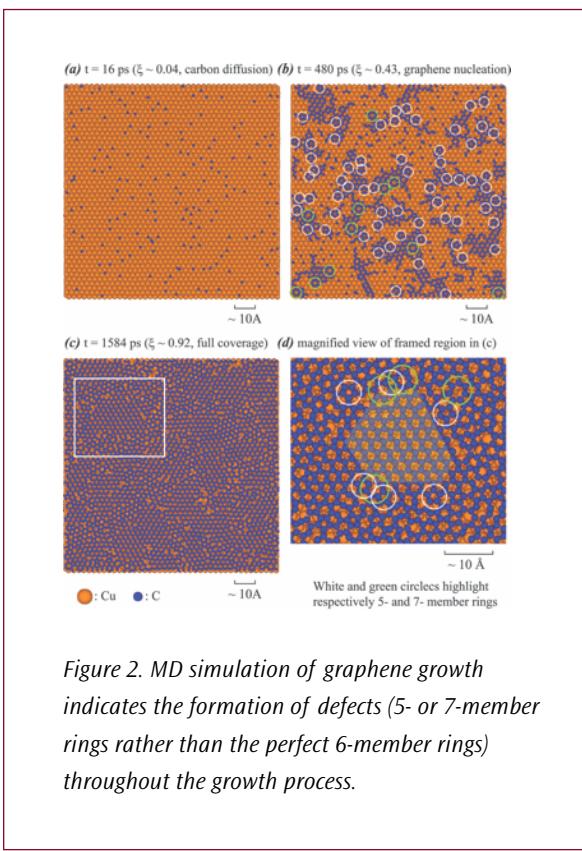
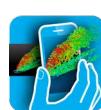
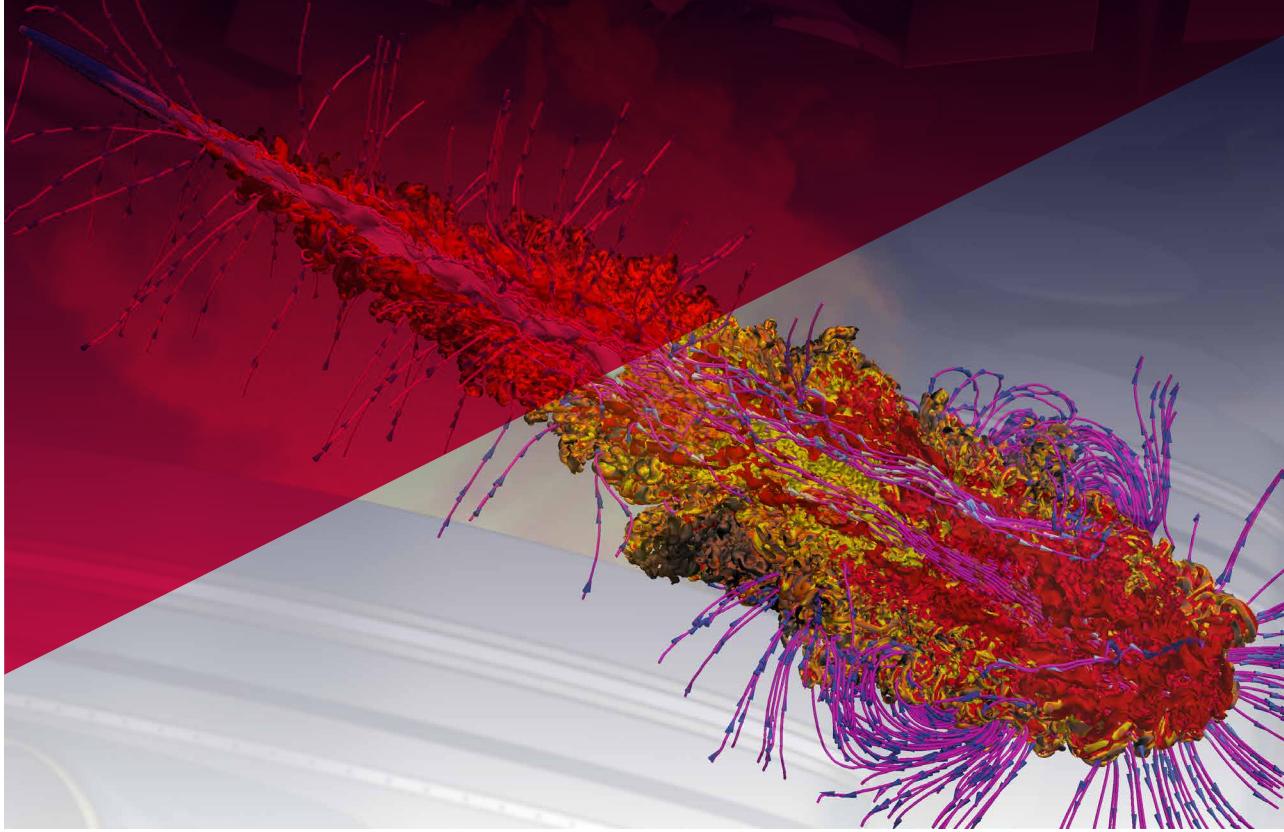


Figure 2. MD simulation of graphene growth indicates the formation of defects (5- or 7-member rings rather than the perfect 6-member rings) throughout the growth process.



Scan this image with your SNLSimMagic App to see an associated movie clip.



MODELING PRIMARY ATOMIZATION OF LIQUID FUELS USING A MULTIPHASE APPROACH

MARCO ARIENTI (PI)

JOE DEFELEIN

FRANÇOIS DOISNEAU

HPC is becoming an increasingly indispensable tool to study direct fuel injection in automotive engines because of the multiplicity of scales and physics that are entangled in this problem. As these scales cannot yet be directly observed with laboratory diagnostics, we must resort to increasingly high-resolution simulations in the attempt to approximate first-principles physics more closely.

As part of a Laboratory Directed Research and Development project, we are developing a modeling and simulation capability to study direct fuel injection in automotive engines. Predicting mixing and combustion at realistic conditions remains a challenging objective of energy science. Its relevance to many flows in defense and climate make it a research priority in Sandia's mission-critical area of energy security. High performance computing applied to this non-linear, multi-scale problem is key to engine calculations with increased scientific reliability.

Our team has developed a two-fold computational strategy to assess with unprecedented detail how fuel sprays are injected, atomize, and burn in engines. First, a state-of-the-art multiphase sharp-interface formalism with adaptive mesh refinement can track fuel injection on a time-scale of nanoseconds, capturing fragmenting liquid interfaces with micrometer resolution near the injector. This capability is unique since it does not require pre-existing knowledge

of spray characteristics. Second, the sequence leading to combustion is explored within the large eddy simulation (LES) framework. Major advances underlie a state-of-the-art Euler-Euler solver describing the dynamics of sprays strongly coupled to multicomponent gases. Multiple sizes of droplets interact, as described by a multi-fluid (MF) model, a sectional/high-order moment method with advanced velocity closures. The accurate numerics we developed for MF-LES achieve unprecedented robustness and parallel efficiency on target injection cases.

The knowledge from direct simulation of atomization at engine conditions as shown in Figure 1 is summarized and input in the high-fidelity MF-LES as shown in Figure 2. The high computational resolution is key: finely discretized physical and phase spaces alleviate spray and turbulent sub-model uncertainty. Thanks to low-dissipation numerics, the novel approach crisply renders injection and auto-ignition despite large density and temperature fluctuations.

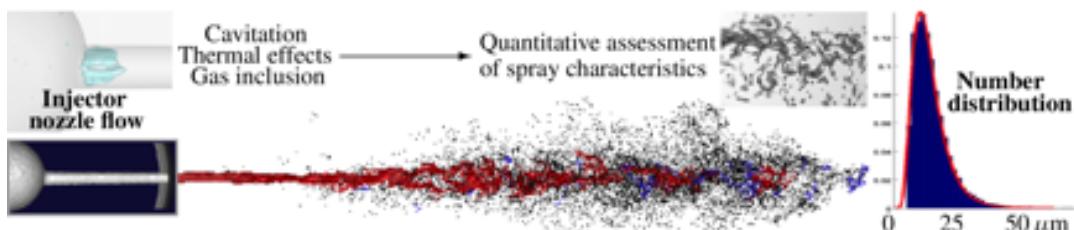


Figure 1. Direct numerical simulation of fuel injection with commanded needle opening at $t = 40 \mu\text{s}$.

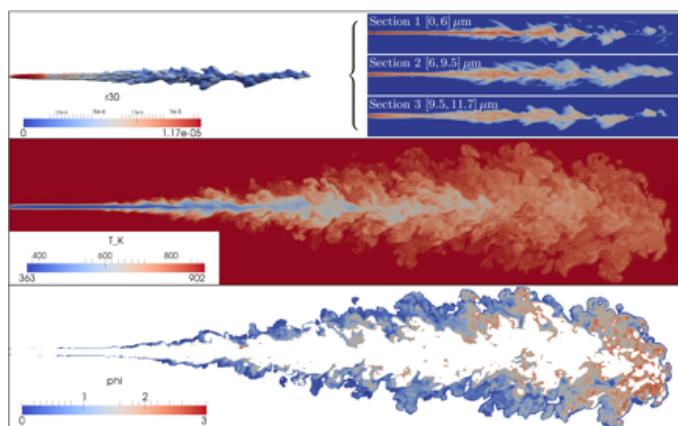
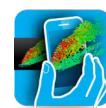


Figure 2. Large eddy simulation of polydisperse spray flame at $t = 400 \mu\text{s}$. Liquid fuel contour (top), gas temperature (middle) and equivalence ratio ϕ in warm regions ($800 < T_g < 899 \text{ K}$) (bottom).



Scan this image with your SNLSimMagic App to see an associated movie clip.



LABORATORY DIRECTED RESEARCH & DEVELOPMENT



25 YEARS OF LDRD—
CATALYST OF INNOVATION

The Laboratory Directed Research and Development (LDRD) Program, authorized by Congress in 1991, functions as a catalyst for the genesis of innovative science and applied advancements in engineering and technology that serve the Department of Energy and other national security missions. Now in its 25th year, the LDRD program—Sandia's only discretionary R&D to anticipate, innovate, and deliver solutions to future problems—continues to help shape Sandia's HPC capabilities and enable the research featured in this report.

LDRD proposals outlining R&D distinct from existing programs are reviewed and selected for funding by team members in the following mission relevant investment areas: Materials Science; Computing and Information Sciences; Engineering Sciences; Radiation Effects and High Energy Density Science; Nanodevices and Microsystems; Bioscience; Geoscience; Defense Systems and Assessment; Energy, Climate and Infrastructure Security; International, Homeland and Nuclear Security; Nuclear Weapons; Grand Challenges; and Research Challenges. Significant investments in cyber security are managed throughout the program and are distributed across these investment areas according to mission need. Strategic partnerships also support the professional development of graduate students and new staff at Sandia through LDRD projects.

Research activities throughout the LDRD program are high-risk but driven by anticipated mission needs. High performance computing capabilities and computational science expertise are critical to the program's and Sandia's success. Throughout the program, HPC algorithms and expertise are being developed and deployed to a variety of scientifically challenging problems. The LDRD projects featured within exemplify the important role Sandia's HPC resources play in developing the technologies and capabilities that support our nuclear weapons and national security missions.

— **Rob Leland**

Vice President, Chief Technology Officer

COMPETITIVE O_2/N_2 SEPARATION FOR CLEANER POWER PLANTS



TINA NENOFF (PI)

JEFFERY GREATHOUSE

MARIE PARKES

DORINA SAVA GALLIS

AIMD simulations are quite computer-intensive, requiring almost 4,000 processor hours to run a single 28.5-ps molecular simulation. Our collection of 36 different simulations required almost 6,000 processor days, a staggering amount that is simply not feasible on smaller systems. This ambitious project could only be completed with the 50,000-plus processors available on Redsky and Skybridge.

Increased demand for energy is accompanied by a disconcerting increase in emissions of carbon dioxide (CO₂), contributing to climate change. Energy generation accounted for 37% of the nation's total CO₂ emissions in 2015, with coal-burning power plants responsible for over 70% of those emissions. By increasing the efficiency of coal-burning power plants, the United States' CO₂ emissions can be decreased dramatically.

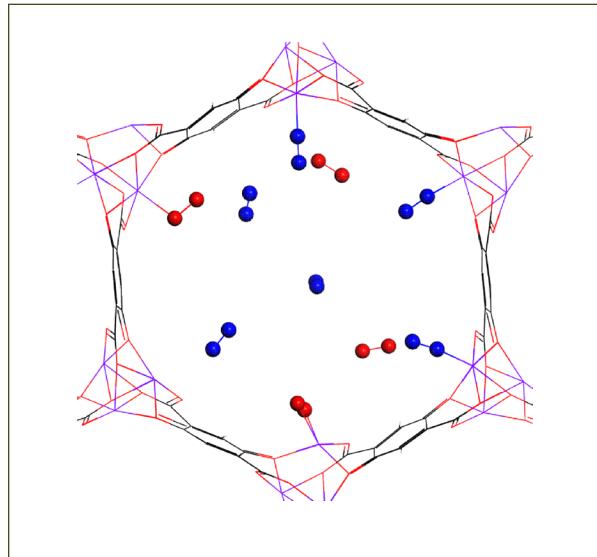
Coal is currently burned in air—composed primarily of N₂ with only 21% combustible O₂. Increasing the O₂ component of the combustion gas would increase the thermal efficiency of burning coal, decrease emissions of pollutants, reduce flue gas volume, and facilitate post-combustion CO₂ capture. Current O₂ purification is accomplished by cryogenic distillation, an energy-intensive process that cannot reasonably be carried out on the scale necessary for power plants. O₂-selective metal-organic frameworks (MOFs) are being studied as an alternative solution for the production of high-purity O₂. Success in this energy-efficient process will correspond to increases in the efficiency and cleanliness of coal-burning power plants.

A team of researchers at Sandia used quantum density functional theory (DFT) modeling to determine the binding energy of both O₂ and N₂ in two common MOFs, Fe₂(dobdc) and Cr₃(btc)₂. Similar calculations were completed on 14 variations of each MOF, with the Fe or Cr metal atoms substituted by other first-row transition metals, two second-row transition metals, and two alkaline earth metals. These calculations showed that early transition metal MOFs bind O₂ more strongly than late transition metal MOFs.

The last phase of this project utilized *ab initio* molecular dynamics (AIMD) modeling simulations to examine the pure-gas and competitive gas adsorptions of O₂ versus N₂ in the M₂(dobdc) (M = Cr, Mn, Fe) MOF series.

AIMD simulations, which combine traditional molecular dynamics with electronic structure methods, make use of DFT electronic optimizations at each time step, allowing temperature effects to be included in the simulations. These AIMD simulations allowed researchers to study in detail the subtle effects of gas-metal interactions and temperature on competitive adsorption in MOFs.

The results of these simulations have been used to direct experimental MOF development, with the goal of producing a MOF that is able to cheaply and efficiently separate O₂ from air for oxyfuel combustion power plants. Success in this LDRD project has led to four related publications, numerous national and international presentations, one provisional US patent, and ongoing industrial collaboration interest.

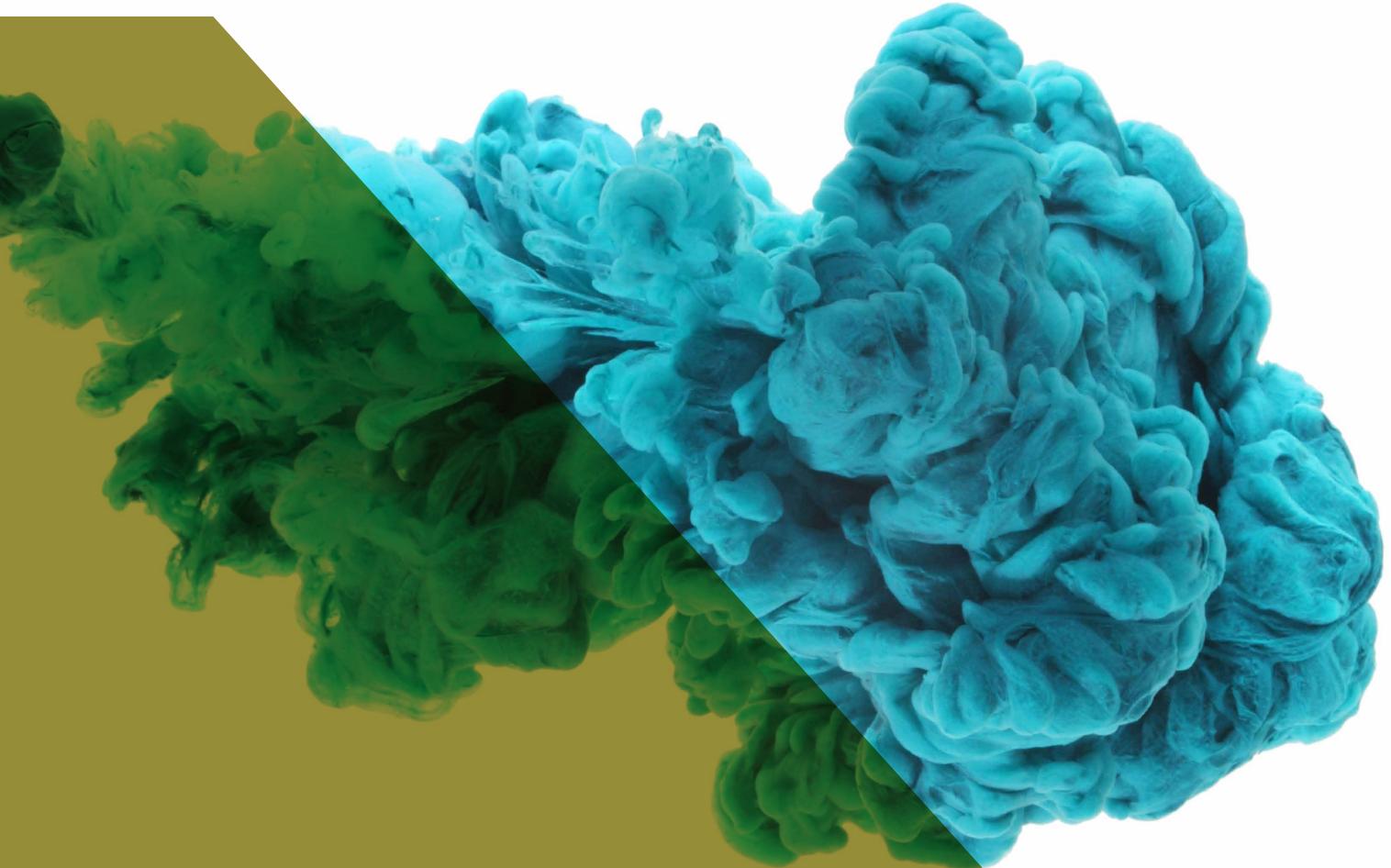


Scan this image with your SNLSimMagic App to see an associated movie clip.

USING MACHINE LEARNING FOR TURBULENCE MODEL UNCERTAINTY QUANTIFICATION

JULIA LING (PI)

JEREMY TEMPLETON



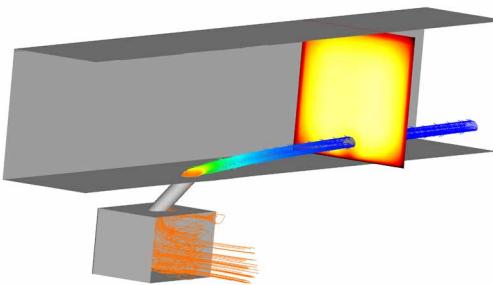
HPC resources allowed us to efficiently build up our database of flow cases on which to train and test our machine learning models. Furthermore, they enabled parallelized parametric investigations into machine learning model performance with varying input features and training hyper-parameters.

Turbulent flows occur in many applications of interest at Sandia. These include flows around wind turbines, engine flows, and aerodynamic flows relevant to national security applications, making it important to accurately simulate these flows with quantified uncertainty. Because of their computational expense, direct numerical simulations (DNS) that exactly solve the governing equations are not feasible for many flows of interest.

Instead, Reynolds Averaged Navier Stokes (RANS) models are used. These models, while more computationally efficient, suffer from model form uncertainty, which is uncertainty due to missing physics in the model. This type of uncertainty is particularly difficult to quantify, because it cannot be captured simply by varying parameter values. The goal of this research is to use machine learning methods to enable an assessment of the model form uncertainty in RANS models.

Machine learning is a set of data-driven algorithms for regression, classification, and clustering. These algorithms discern patterns and structure in large, high-dimensional data sets. The field of machine learning has undergone tremendous growth over the past two decades, and machine learning algorithms are used in a wide variety of applications including search engines, shopping recommendations, and image recognition. More recently, there has been a push to apply machine learning methods to computational simulation data sets.

We used a database of flows containing both RANS and DNS results to train our machine learning algorithms to recognize flow regimes which violated fundamental RANS model assumptions. It is in these flow regimes that the RANS model-form uncertainty is the highest and thus less accurate. With this added knowledge, we achieved error detection accuracy three times higher than previously demonstrated. These results validate the potential for machine learning methods to provide new understanding of scientific computing data sets.



Simulation of a jet in crossflow



ELECTROCHEMICAL DETECTION OF SINGLE MOLECULES IN NANOGAP ELECTRODE FLUIDIC DEVICES

RONEN POLSKY (PROJECT PI)

SUSAN REMPE (MODELING PI)

MANGESH CHAUDHARI

HPC allows the team to connect quantum, molecular, and diffusion level phenomena to assist experimental development of novel electrochemical single-molecule sensors for biological and chemical detection.

Electrochemical detection of single molecules tests the ultimate limit of analytical sensitivity. The small currents involved in electron transfer make it difficult to detect single molecules with reliable accuracy. Using the redox cycling technique, researchers have direct, real-time visualization of single biological macromolecules and their assemblies, improved understanding of interactions and mechanisms in living cells, and the ability to track the activity of drug molecules.

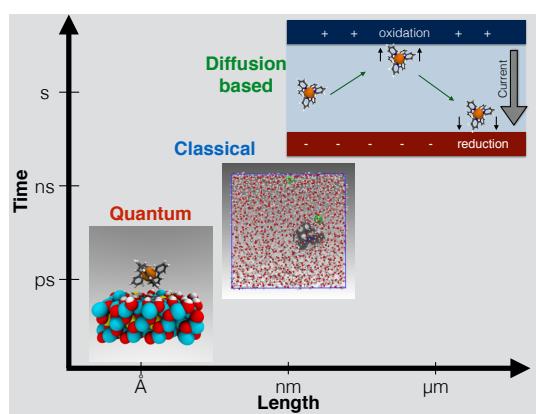
A group of researchers at Sandia is working on integrating an orthogonal validation technique, total internal reflection fluorescence microscopy and molecular simulations. The combination can clarify mechanisms leading to current build-up due to redox cycling and diffusion, and adsorption of single molecules undergoing redox reactions. The project is aimed at gaining a fundamental understanding of single-molecule diffusion in electrochemical sensors to develop a new class of highly sensitive and selective chemical and biological sensors.

High performance computing resources allow researchers to undertake computationally challenging and highly accurate *ab initio* molecular dynamics simulations. These simulations permit investigation of electron transfer between redox molecules and sensor materials.

By tracking electrons, these high-resolution simulations can provide new insights about quantum-level interactions between the sensor and molecule of interest.

Along with the *ab initio* simulations, several additional classical molecular dynamics computational studies, e.g., force fields molecular dynamics (FFMD), are underway to investigate hydration and transport of fluorescent ruthenium bipyridine (RuBpy) in water. FFMD can predict the solvation structure of RuBpy in the presence of dissolved salts. The salts can affect the diffusion of RuBpy, which can be reliably calculated and compared with direct experimental measurements to assess the quality of the classical force fields. The bulk transport properties calculated in these simulations can then be used in diffusion-based modeling.

Future work consists of diffusion-based modeling to test different sensor geometries. The input parameters for the simulations include diffusion constants, adsorption/desorption probabilities, and sensor dimensions. These simulations are designed to mimic the experimental spectral power density to calculate the adsorption/desorption rate for a particular geometry and electrode. This investigation will guide the design and fine-tuning of the sensor.

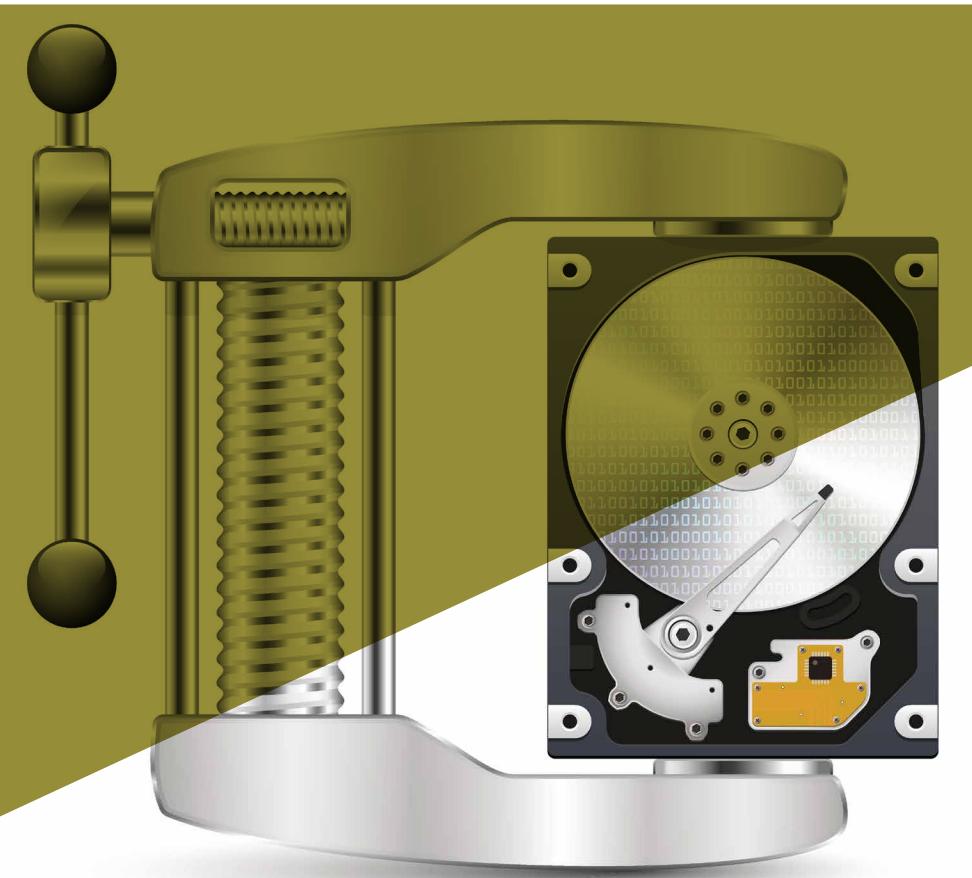


Shown are the different types of simulations performed for this project. The number of atoms and length of simulations increase from quantum to classical to diffusion based simulations. A transparent indium tin oxide electrode and a fluorescent ruthenium bipyridine molecule are also shown.



Scan this image with your SNLSimMagic App to see an associated movie clip.

IN-SITU COMPRESSED SAMPLING AND RECONSTRUCTION OF UNSTRUCTURED MESH DATASETS IN HPC APPLICATIONS



MAHER SALLOUM (PI)

NATHAN D. FABIAN

DAVID M. HENSINGER

JINA LEE

STEFAN P. DOMINO

High performance computing platforms are challenged by datasets that are becoming too large for reliable storage and transfer across networks—significantly impeding data exploration and knowledge discovery.

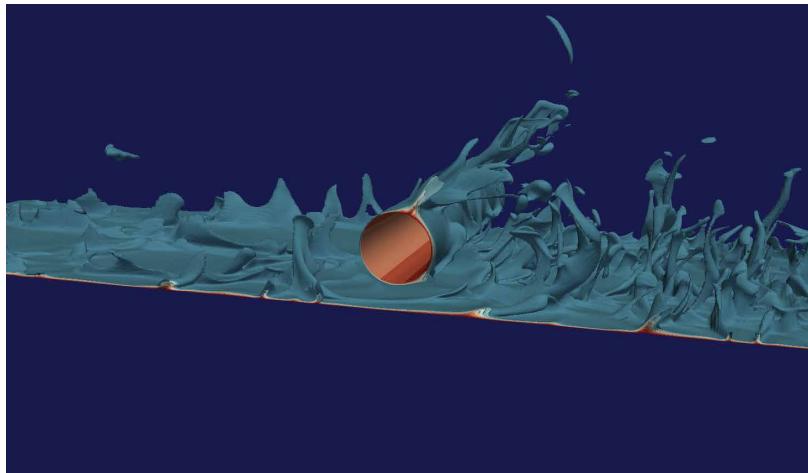
To address these data challenges, our team has been actively investigating wavelets-based compressive sensing (CS) as a method to compress scientific data. Current CS research efforts have targeted structured datasets (e.g., images). However, these efforts are unsuitable for data defined on irregularly spaced points (“point-clouds”) such as unstructured mesh datasets often encountered in finite element simulations.

CS works in-situ on the computational cluster by drawing samples of the data according to rigorous sampling strategies. One of the main benefits of CS is the high sampling speed, thus compression overhead is minimized during simulation. Data reconstruction relies on wavelet bases that encode the information collection during the compression step. Sandia researchers explored less

common wavelet types, such as second generation wavelets and tree-wavelets, as traditional wavelets are not applicable in this case.

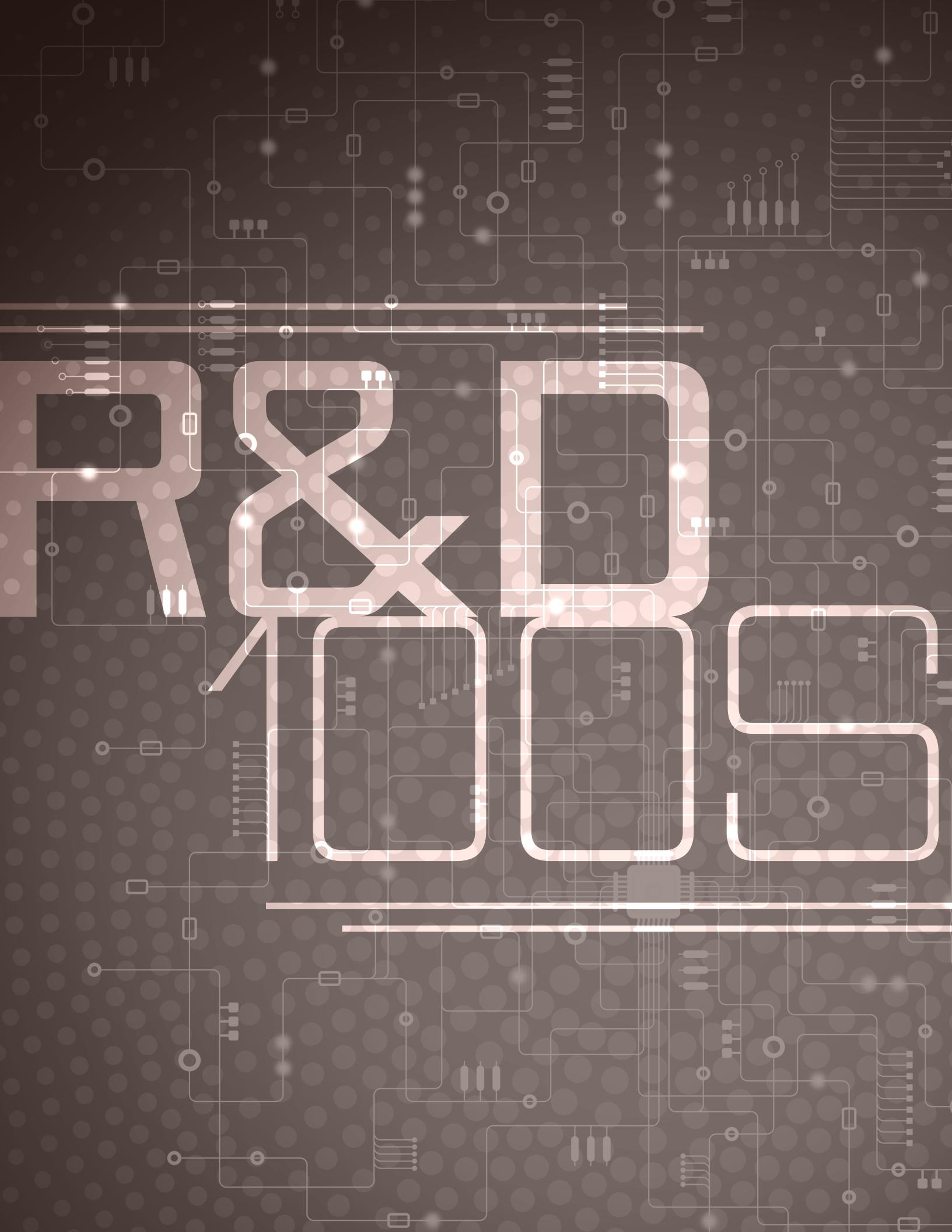
CS is attractive in cases where the compression overhead has to be minimal and where data reconstruction costs are not a major concern. For instance, reconstruction can take place in-situ, in-transit, or offline on dedicated analysis and visualization platforms. Furthermore, CS does not require any previous knowledge of the required wavelet bases during in-situ compression. The choice of wavelets is made during reconstruction, allowing interactive exploration and discovery at multiple levels of detail in the data.

Our algorithm has been tested on combustion and computational fluid dynamics simulation data and was able to compress it up to 100 times with minimal visual deterioration in the reconstructed data. Such data compression capability will be of relevance to exascale engineering simulations conducted for nuclear weapons and energy efficiency applications.



←
Scan this image
with your
SNLSimMagic
App to see an
associated
movie clip.

Regardless of how promising the new HPC data storage and transfer technologies seem to be, we will not escape passing through data compression in order to handle the huge datasets promised at the extreme scale.





INTERNATIONALLY RECOGNIZED CUTTING EDGE TECHNOLOGY

Sandia's HPC expertise, which addresses many of the most difficult problems facing our world, continues to be recognized in the prestigious R&D 100 award program. As just one example, HPC simulations were instrumental in the development of 2015 R&D 100 award winner CO₂ Memzyme, a carbon dioxide (CO₂) capture technology developed by Sandia and the University of New Mexico that could prove transformational in efforts to control global climate change.

Equally impressive are Sandia's activities—often undertaken in partnership with industry and other R&D institutions—that develop resources to unleash the full power of HPC. For example, a 2016 submission, the Rapid Optimization Library, is helping teams solve engineering optimization problems of unprecedented size by providing a programming interface and advanced algorithms that enable extremely efficient use of HPC architectures.

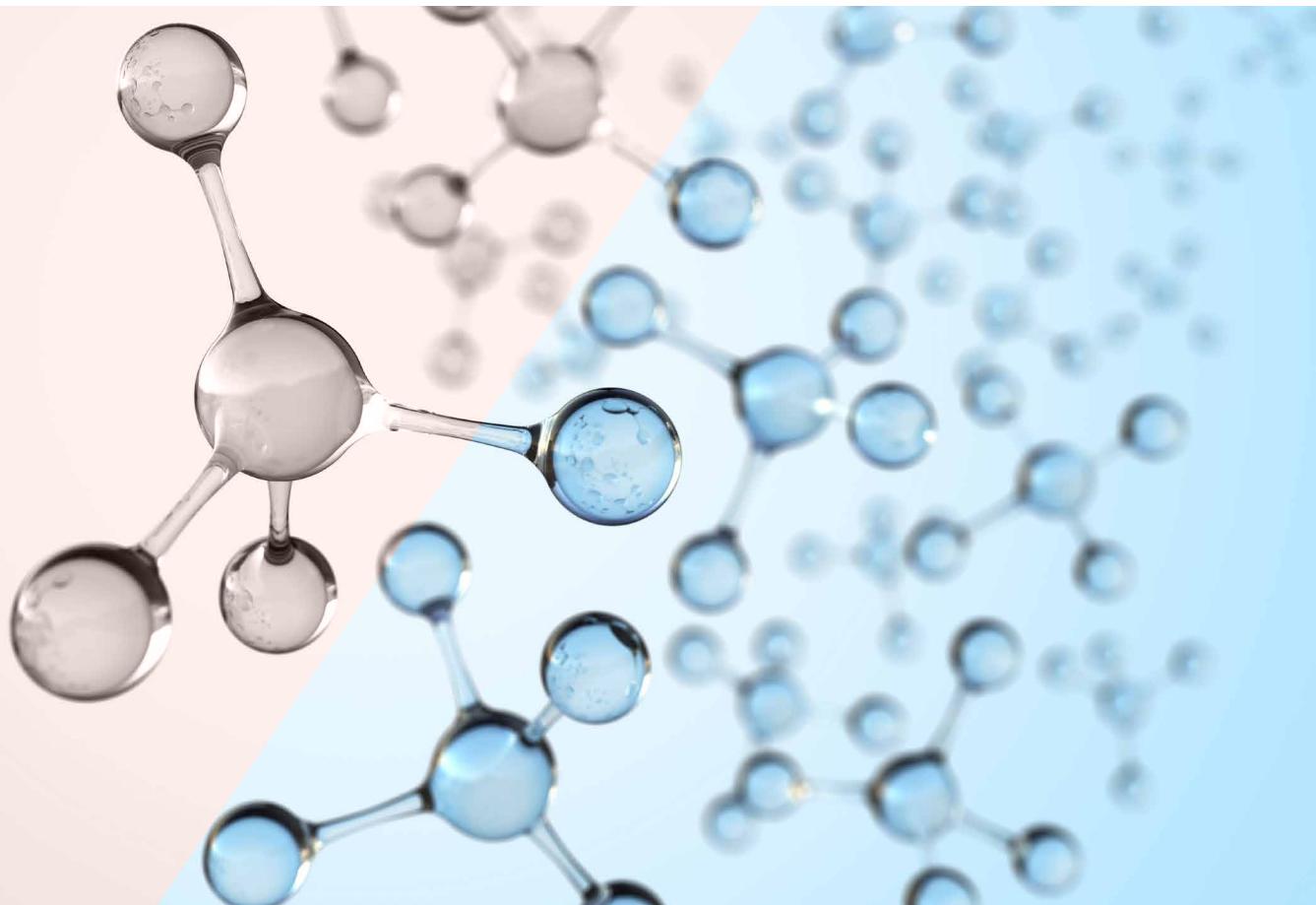
Another Sandia-led team was honored for the Lightweight Distributed Metric Service, a unique HPC monitoring tool that provides continuous right-fidelity system-wide platform awareness—information that helps administrators and users understand and unravel compute inefficiencies rooted in resource contention, network congestion, I/O bottlenecks, and other causes.

Looking ahead, a 2016 submission highlights Pyomo, a software package that helps resolve complex real-world problems by dramatically decreasing the time and complexity of HPC computation of optimization problems while providing new levels of accuracy. Through these and future submissions, and regardless of the winners and losers, I fully expect that Sandia's pioneering contributions based on HPC will continue to be recognized by the R&D 100 process.

— **Robert Q. Hwang**
Director, Research & Development

CO₂ MEMZYME

Susan Rempe (PI) Juan Vanegas



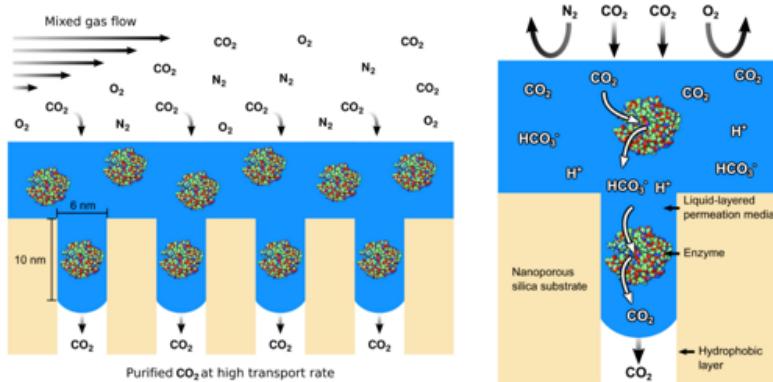
HPC is a critical resource due to the large computational costs associated with detailed atomistic models. Although the CO₂ membrane pore model spans only tens of nanometers in size, the hundreds of thousands of atoms in a single simulation can require hundreds of computing cores for days and weeks. Access to large-capacity HPC facilities allows simulations of multiple systems simultaneously for rigorous statistical analyses, and for fast progress on time-sensitive projects.

Carbon dioxide (CO_2) makes up 80% of all human-caused greenhouse gases and is a long-lived molecule that traps heat and warms the Earth for centuries. Despite investments in renewable and low-carbon fuels for electricity, domestic and global CO_2 emissions are expected to increase over the next quarter century. Conventional liquid absorption technology can achieve the 90% CO_2 reduction target set by the Department of Energy (DOE), but the cost is staggering. To address the grand challenge of efficient CO_2 capture, a team from Sandia and the University of New Mexico developed a potentially transformational technology called the CO_2 Memzyme.

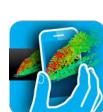
The Memzyme is a thin, enzyme-loaded, liquid membrane with the highest combined CO_2 flux and selectivity yet reported. It exceeds the DOE's standards for CO_2 capture for the first time. It was designed to achieve the high efficiency of natural biological separations that typically

take place in liquids for faster molecular diffusion than solids, through ultra-thin membranes for short diffusion pathways, and assisted by enzymes (carbonic anhydrase) that speed uptake and release of CO_2 specifically.

Sandia's HPC facilities were used to study the inner workings of the Memzyme with molecular dynamics simulations. The carbonic anhydrase enzyme was simulated within pore-like structures that resemble the real membrane. The computational model allowed exploration of interactions between the membrane pores and the enzyme to learn how membrane surface chemistry affects the overall CO_2 transport and selectivity. Favorable interactions and a high density of pores permit a remarkable 50 times higher concentration of enzymes compared with plain water. Experimental collaborators are using insights from the simulations to optimize the membrane's performance.



The Memzyme was designed to achieve the high efficiency of natural biological separation processes that typically take place in liquid phase, through ultra-thin membranes, and assisted by enzymes. Computational modeling was used to understand how membrane surface chemistry affects the enzyme, which catalyzes selective uptake and release of CO_2 .



Scan this image with your SNLSimMagic App to see an associated movie clip.

SCALABLE STOCHASTIC PROGRAMMING WITH PYOMO

JEAN-PAUL WATSON (PI)

WILLIAM HART

CARL LAIRD

JOHN SIROLA

CESAR SILVA-MONROY

DAVID WOODRUFF, UNIVERSITY OF CALIFORNIA DAVIS





Scan this image with your SNLSimMagic App to see an associated movie clip.

Optimization is widely used in business, science, and engineering to minimize costs, identify worst-case scenarios, and analyze trade-offs. Sandia has developed Pyomo, an open-source software platform for developing optimization-based analytics to support complex decision making in real-world applications (see www.pyomo.org). Pyomo can leverage high performance computing to analyze large optimization problems, which was recently demonstrated on power system planning and management applications.

Pyomo can express stochastic programs, which incorporate uncertain scenario data to model risk. Each scenario represents the possible realization of system uncertainties, and a collection of scenarios can be used to model variations in system costs and related performance

measures. For example, a key power planning application is stochastic unit commitment (SUC), which considers the commitment of power generation units given uncertain demands. In this application, each scenario represents a possible set of demands that need to be met.

Pyomo features a progressive hedging (PH) solver that solves a stochastic program by decomposing it into a series of smaller subproblems that can be solved independently. Each subproblem optimizes with respect to a small number of scenarios, and subproblem results are iteratively aggregated to form a consensus solution. Since subproblems are independent, PH is easily parallelized on high performance computers. Pyomo employs a simple client-server paradigm, which leverages third-party open-source Python libraries to manage the communication between the solver process and processes that solve subproblems.

Sandia researchers have demonstrated the application of Pyomo on power grid models that provide a simplified description of the western US interconnection grid using 85 thermal generators as well as two models of ISO New England with up to 340 thermal generators. Sandia's high performance computing cluster (Red Sky) was used to analyze these models with 100 scenarios on 100 nodes. In the largest problem formulations, PH was able to find solutions within 30 minutes whose accuracy was provably within 2% of optimal. This was the first demonstration of SUC to industrial-scale problems with a realistic number of demand scenarios.

Standard methods for solving stochastic mixed-integer optimization problems are widely known to be intractable, such that even trivially small problems either take hours to days to solve on a high-end workstation—if they can be solved at all. In contrast, the ability to leverage HPC resources—enabled by the development of advanced decomposition algorithms—allows us to solve real-world problems in minutes to hours of run-time. In other words, HPC allows us to transition from “impossible” to “solvable,” ultimately enabling the application of stochastic mixed-integer optimization in industrial contexts.



RAPID OPTIMIZATION LIBRARY

DENIS RIDZAL (PI)

DREW KOURI

GREG VON WINCKEL

BART VAN BLOEMEN WAANDERS

WILKINS AQUINO

TIM WALSH

By developing the Rapid Optimization Library (ROL) and deploying it in Sandia's HPC application codes, the Center for Computing Research is helping application teams solve engineering optimization problems of unprecedented size. Most notably, ROL's unique application programming interface and advanced algorithms enable extremely efficient use of current and emerging HPC architectures when solving optimization problems with simulation constraints. As a result, ROL is used today to solve inverse problems and optimal design problems with millions of unknown parameters in applications including geophysical exploration, climate science, electrical circuit optimization, and super-resolution imaging.



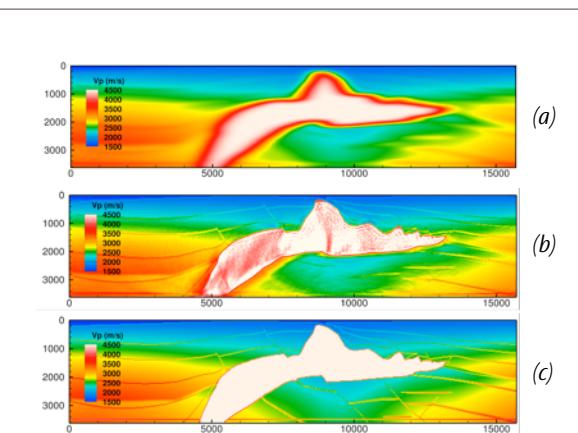
Sandia-developed Rapid Optimization Library (ROL) has proven invaluable to simulation codes throughout the Labs. This software package for large-scale numerical optimization is designed to seamlessly plug into a variety of HPC codes such as those simulating the physics of fluids and structures, electromagnetics, quantum mechanics, integrated circuits, or power grids. The simulation codes predict system state that depends on model parameters or control variables. One important use of ROL is to numerically compute the parameters that maximize the simulation's agreement with measured data, which is called an inverse problem or parameter estimation.



Another application is to compute control variables that maximize a measure of system performance, such as aerodynamic lift, light absorption, or circuit gain, which is referred to as an optimal control or optimal design problem. Other uses of ROL include image processing, image enhancement, and computational mesh generation.

Sandia researchers are using ROL in conjunction with other Sandia codes to address important national security and economic competitiveness concerns. Coupled with Sandia's Sierra Structural Dynamics code, ROL is used for large-scale, direct-field acoustic testing in Sandia's nuclear weapons mission. ROL is also used with the Albany code to estimate the basal friction of the Greenland ice sheet from data, enabling predictive estimates of ice sheet evolution and its effects on climate change. With Sandia's DGM code, ROL is used to perform advanced full-waveform inversion for estimating subsurface properties based on seismic data. It is also being used with a new super-resolution image processing toolkit developed at Sandia to combine a number of spatially displaced low-resolution images into a single high-resolution image. In the first three applications, ROL works in concert with the native CPU-based distributed memory data structures. In the latter, ROL works directly with ArrayFire data structures, enabling extremely fast super-resolution image enhancement on GPUs.

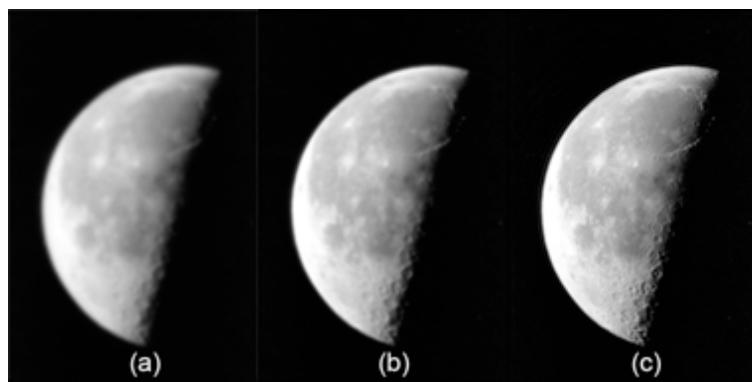
ROL makes numerical optimization practical and efficient for any application, any computational hardware, and any problem size.



Given the seismic response from the synthetic SEG 2D Salt Model, ROL is used to recover the P-wave velocity field describing the subsurface.

- (a) the initial guess of the velocity field
- (b) the recovered velocity field, i.e., the subsurface model computed by ROL
- (c) the “true” subsurface model.

The (b) and (c) images are in good agreement except for below the salt, which has not been sufficiently “illuminated” with seismic waves.



Given a set of low-resolution images of the Moon's surface, as in the sample low-resolution image (a), ROL is used to solve super-resolution optimization problems, generating image (c). This result is superior to image (b), generated by the Drizzle algorithm that was originally developed for the Hubble Space Telescope.



Scan this image with your SNLSimMagic App to see an associated movie clip.



LIGHTWEIGHT DISTRIBUTED METRIC SERVICE

Sandia National Laboratories:

Jim Brandt (PI)

Ann Gentile

Ben Allan

Open Grid Computing:

Tom Tucker

Narate Taerat

Nichamon Naksinehaboon

The evolution of fast multi-core processors, large memory, efficient operating systems, and RDMA support utilized in today's HPC environments have all played a significant role in enabling our lightweight data collection and transport infrastructure, which in turn has enabled needed insights about these technologies' interactions and their interactions with applications in the HPC environment.



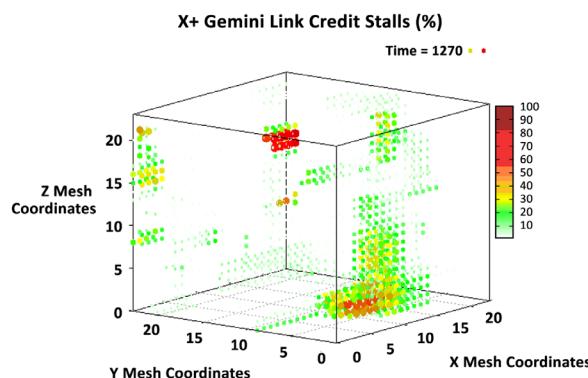
On today's distributed, massively parallel, high performance computing platforms, competition for shared network and file system resources among concurrently running applications is responsible for significant performance degradation. Currently, neither coarse-grained system monitoring tools nor fine grained application profiling tools provide sufficient data for use in anomaly detection, correlation analysis, and machine learning techniques, which could provide valuable clues to designing higher performing platform architectures. A team at Sandia collaborated with Open Grid Computing to develop Lightweight Distributed Metric Service (LDMS), the only HPC monitoring tool that provides continuous right-fidelity system-wide platform awareness allowing system administrators, application developers, and users the ability to understand and troubleshoot application resource contention, network congestion, I/O bottlenecks, and associated causes of compute inefficiencies.

This software provides continuous, high-fidelity snapshots of system status across an entire HPC platform, offering insight into how platform resources are being utilized, stressed, or depleted due to the aggregate workload. For system administrators, LDMS provides near real-time data for decision support regarding resource allocation, job placement, and scheduling. For developers, it provides application execution data that can be used to improve run-time performance. For the HPC architecture design community, it provides valuable new information about performance limiting architectural design features.

LDMS scales efficiently and affordably to large-scale HPC systems and allows synchronized, system-wide data sampling down to sub-second intervals. Its computational overhead demands are very low, so LDMS can be deployed on a continuous basis across an entire Linux-based HPC platform without adversely impacting application performance.

LDMS is included in the NNSA Tri-Lab Operating System Stack, which is the base software stack for all HPC systems at Lawrence Livermore National Laboratory, Los Alamos National Laboratory (LANL), and Sandia National Laboratories. Large-scale deployments of LDMS include the ACES Trinity supercomputer (>19,000 nodes in 2016) sited at LANL and the National Center for Supercomputing Applications' Blue Waters (27,648 nodes) system at the University of Illinois. LDMS currently collects over 15 billion data points a day on Blue Waters and will collect over a trillion data points per day on the full Trinity system.

Data provided by LDMS is unique in scope and fidelity, opening doors for researchers to investigate new and potentially impactful areas such as run-time resource-aware adaptive computing and failure prediction that will ultimately enable more energy efficient, more reliable, and more productive use of existing HPC systems—resulting in a better and more productive world science enterprise.



Scan this image
with your
SNLSimMagic
App to see an
associated
movie clip.



AROUND SANDIA



PROVIDING HPC SOLUTIONS TO THE NATION'S COMPLEX SCIENTIFIC CHALLENGES

Sandia's high performance computing capabilities have become a critical resource to our researchers as they solve the nation's most complex scientific and engineering challenges.

When coupled with traditional experimentation, our large-scale computing resources and platforms allow for new discoveries and innovation by providing unique insight. Sophisticated computational modeling allows our researchers to solve problems that have proven insoluble by traditional means, whether they are too hazardous to study in the laboratory, time-consuming, or expensive.

The following section highlights HPC's impact on projects throughout Sandia's New Mexico Laboratory. HPC's advanced model and simulation capabilities have allowed our researchers to evaluate the impact of additive manufacturing on component reliability, provide risk assessments for high consequence flight tests, and ensure the safety of the US warfighter through blast simulations.

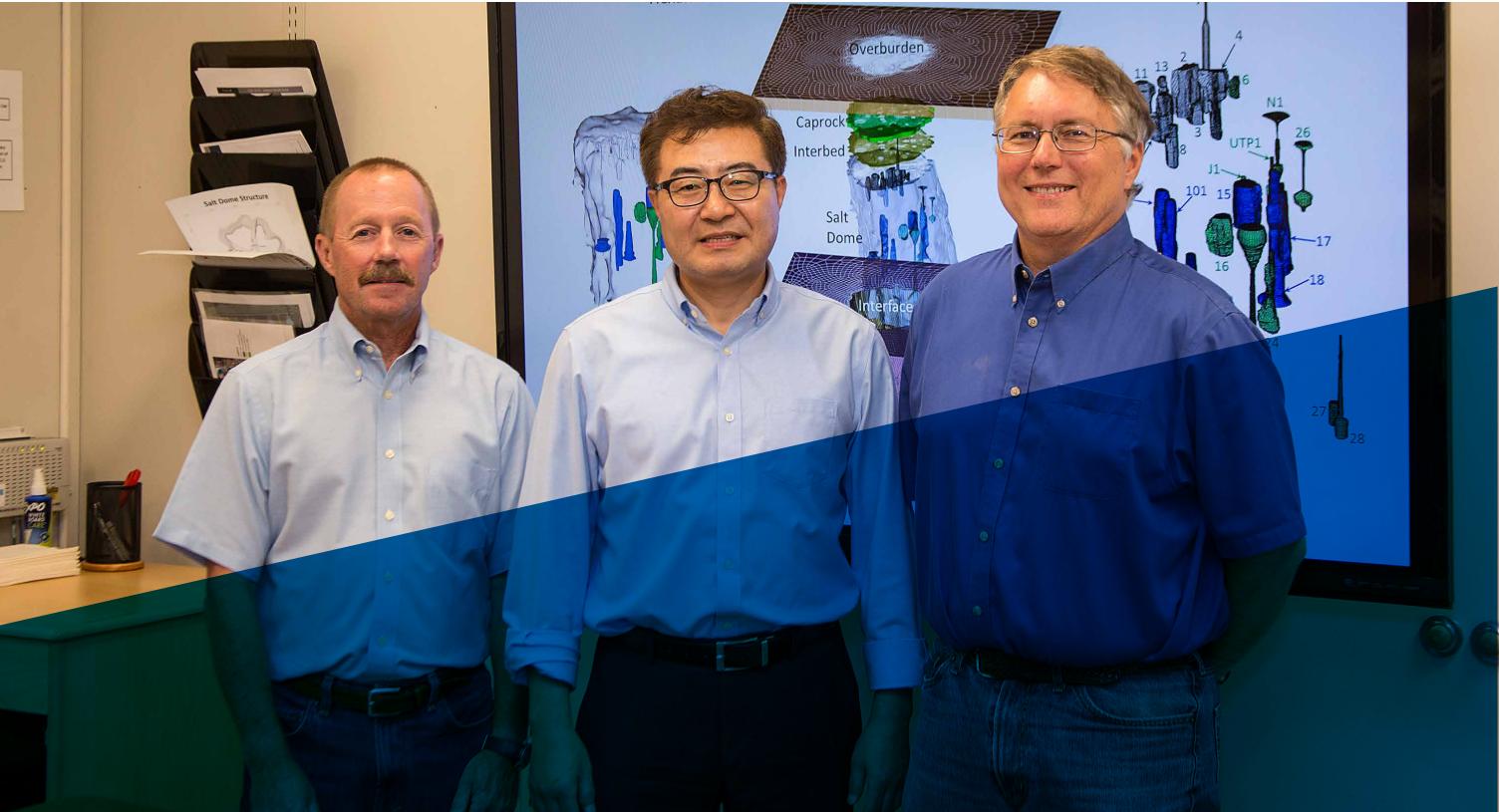
Other researchers in New Mexico are seeking to improve HPC performance. From analyzing the effect switch and node arrangement has on system performance to using our Test Bed facility to analyze and evaluate architectures, we are all working toward exascale computing. This continued research and investment into HPC will continue to allow our researchers to push the boundaries of science and technology.

— **John D. Zepper**

*Director, Research & Development Science & Engineering
Former Director, Computing & Network Services*

HEXAHEDRAL FINITE ELEMENT MESH TO FIT COMPLEX GEOLOGICAL GEOMETRIES

BYOUNG YOON PARK (PI) BARRY L. ROBERTS STEVEN R. SOBOLIK



Sandia's HPC allows us to leverage detailed information regarding the Strategic Petroleum Reserve (SPR) caverns into more accurate simulations of salt dynamics. This provides information critical in fulfilling Sandia's role as geotechnical advisors to the SPR, and assuring the integrity of this national asset.

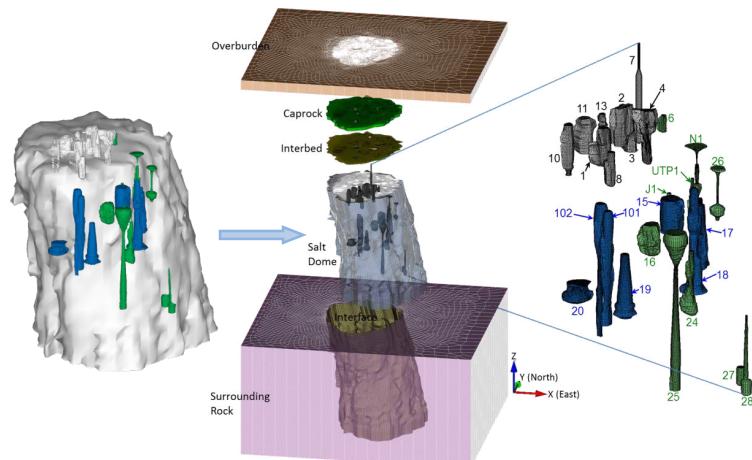


The US Strategic Petroleum Reserve (SPR) stores approximately 695 million barrels of crude oil in 60 caverns located at four sites along the Gulf Coast. As a matter of normal operation of caverns in a salt dome, the continuous mechanical creep of salt, along with the change in internal cavern and casing pressure due to cavern closure and fluid exchanges, impose several mechanical conditions on the skin, well, and casing of a cavern that could potentially create damage. Sandia, on behalf of DOE, is evaluating the structural integrity of the salt surrounding existing caverns in the Bayou Choctaw (BC) salt dome in Louisiana.

We constructed a three-dimensional finite element mesh capturing realistic geometries of the BC site using seismic and sonar survey data obtained from the field (see image below). It is difficult to realize the naturally and artificially formed caverns and salt dome for numerical analysis. It is harder to convert the geometries into the meshed mass consisting of only hexahedral finite elements.

The mesh consists of hexahedral elements because the salt constitutive model is coded using hexahedral elements. We also developed techniques to reduce the number of elements as much as possible to save on computer run time while maintaining computational accuracy. The previously existing computational mesh for BC used simplified elliptical and cylindrical shapes for dome and caverns, respectively, yielding rough predictions. The newly developed mesh is expected to provide more accurate solutions. The steps and methodologies are being applied to construct the meshes of Big Hill, Bryan Mound, and West Hackberry SPR sites. The methodology could be applied to the complicated shape masses for not only various civil and geological structures, but also biological applications such as artificial limbs.

The intent of this work is to predict the salt and surrounding lithologies behavior from the year 1990 through 2045 (29,980 load steps) to examine geological concerns. A numerical simulation using Sierra Adagio requires 1024 processors, running the calculation for 89 CPU hours on Red Sky, and generating over 1,146 GB of raw data that is post-processed at a later time.

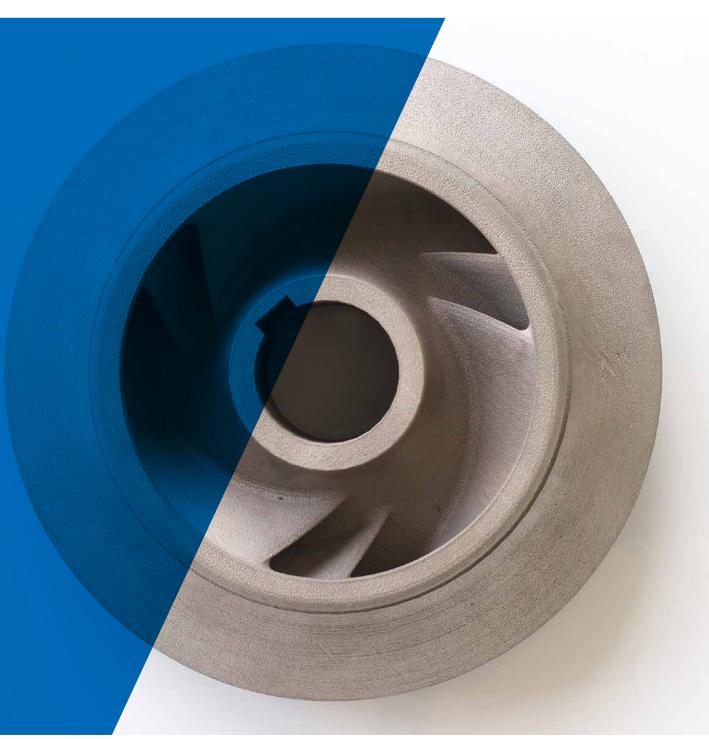


Images of salt dome and caverns obtained from the seismic and sonar surveys, respectively (left) and overview of the hexahedral finite element meshes of the stratigraphy and cavern field at Bayou Choctaw. The US Strategic Petroleum Reserve stores crude oil in the seven blue caverns. Green shows privately owned caverns, and grey depicts abandoned caverns.

The cavern ID numbers are also shown.



Scan this image with your SNLSimMagic App to see an associated movie clip.



SIMULATING SHOCK WAVES IN ADDITIVELY MANUFACTURED COMPONENTS

STEWART SILLING (PI)

HPC can help us explore regimes of high-rate, high-stress, high-consequence physical phenomena that are inaccessible directly through existing experimental techniques. Although it cannot replace experimental science, it can help to explain observations and to guide the development of experimental techniques and programs.

Additive manufacturing offers engineers the ability to produce parts with complex shapes, however, before utilizing additive manufactured parts in high-consequence systems, it is important to understand how shock waves affect the material within the components. Many components in the National Nuclear Security Administration (NNSA) and Department of Defense (DoD) systems are subjected to high-pressure, high-rate loading resulting from detonation waves or high velocity impact, creating shock waves within a material. Shock waves are very large stress pulses that can move at speeds far in excess of the speed of sound in the material. A particularly important scenario is the interaction between shock waves and the surface of an additively manufactured component.



Under the sponsorship of the Air Force Research Laboratory (AFRL) and the Joint Munitions Program, Sandia researchers are investigating how shock and impact loading interacts with complex microstructures and mesoscale structures. Sandia's HPC capabilities provide researchers with a resource to simulate damage within non-uniform material composition and morphology. The computational model is based on the peridynamic theory of solid mechanics, an extension of the standard theory of solid mechanics, that allows for increased generality in the prediction of shock wave propagation, fracture, and fragmentation.

Sandia researchers can simulate the response of an additively manufactured specimen to a strong shock wave. This simulation allows researchers the ability to observe how shock waves move through the object and interact with any irregularities in the surface created by the manufacturing process. These shock waves have the ability to form jets, which can create extensive damage and fragmentation in the interior of the specimen. The jetting phenomenon can be important in many applications, particularly if the jets could potentially strike some other component.

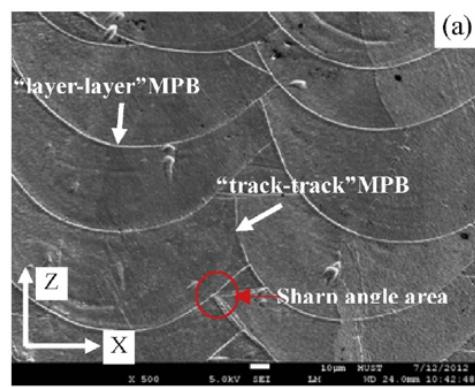


Figure 1. A microstructure of this nature can result from an additive manufacturing technique called selective laser melting. The “fish scale” pattern shown in the figure results from the melting, pooling, and resolidification of the material that is heated by a moving laser beam. Each “fish scale” has a diameter of about 0.1 millimeter.

Image credit: W. Shifeng, L. Shuai, W. Qingsong, C. Yan, Z. Sheng, and S. Yusheng, “Journal of Materials Processing Technology Effect of molten pool boundaries on the mechanical properties of selective laser melting parts,” *J. Mater. Process. Tech.*, vol. 214, no. 11, pp. 2660–2667, 2014.

Sandia's HPC capabilities and this research can provide insight—with minimal cost—into the design implications of using additive manufacturing in a variety of systems.

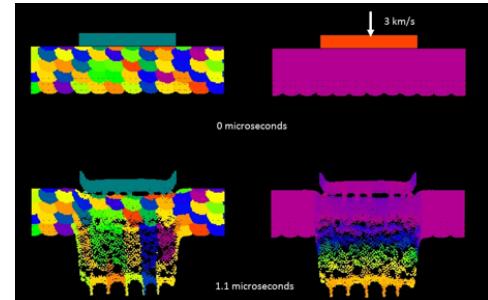


Figure 2. The simulated response of an additively manufactured specimen to a strong shock wave generated by the impact of a plate. The shock wave is created at the upper surface of the specimen where it is impacted. The wave moves downward and reflects off the lower surface. It interacts with the irregularities in the surface created by the manufacturing process. These irregularities form jets at the points of intersection between the “fish scales.” Extensive damage and fragmentation occurs in the interior of the specimen.

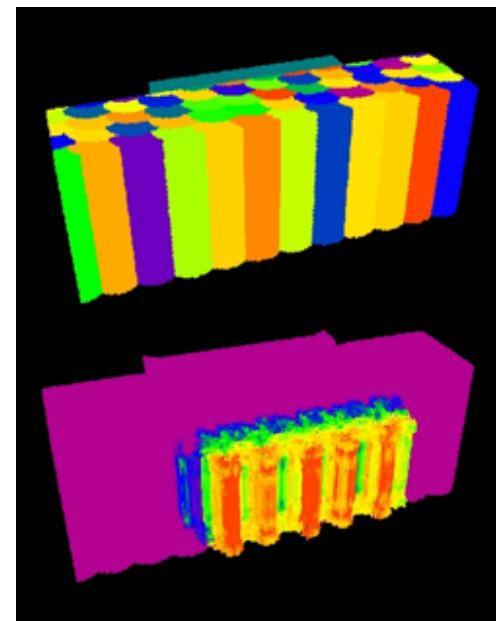


Figure 3. A cross-sectional view of the specimen from outside.



REVOLUTIONIZING DESIGN THROUGH TOPOLOGY OPTIMIZATION AND ADDITIVE MANUFACTURING

BRETT CLARK (PI)

JOSH ROBBINS

BRETT SNEED

TED BLACKER

TOM VOTH

EMILY DANIELS

MIGUEL AGUILLO

LOREN BEGHINI

With the new topology optimization capabilities, we now have the potential for vast exploration of design spaces never conceived before. However, each design requires its own set of optimization calculations. Getting uncertainty bounds on each design may require an exploration space around the design optimization in the hundreds. Therefore, with all the multipliers (1000 possible designs with the space being explored with hundreds of optimizations required to bound uncertainties per design) the criticality and the powerful impact of HPC computing is obvious.



Additive manufacturing offers profoundly new design flexibility, ushering in a new approach to shape design utilizing topological optimization. A team of Sandia engineers developed a new set of tools within the PLATO (PLatform for Advanced Topology Optimization) framework which leverage the power of HPCs to optimally define shape based on functional requirements. It requires multiple simulations (finite element calculations) which iterate based on an objective function and allowable design space. An objective function can be as simple as a part less than a certain weight and is maximally stiff, or as complex as prioritizing heat transfer, compliance, and stress at varying ratios, to produce a family of designs. Each iteration of the optimization requires both a forward (normal solution) and reverse (derivatives) calculation to allow a density function to be adjusted which ultimately defines the shape. The program then uses approximately 50 iterations to converge the shape, and significantly more to provide some bounds on uncertainty of that shape to variances in the materials, the loading directions and magnitudes, and even randomly generated residual stress states.



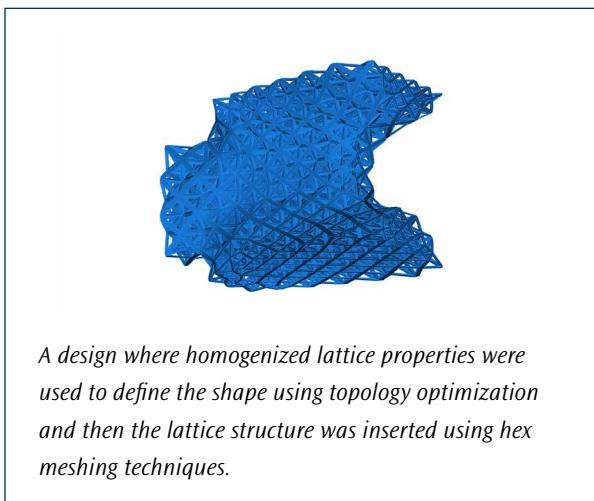
A topology optimized shape for a bracket loaded vertically and bolted to the wall horizontally.



A topology optimized round table to support a uniformly distributed load. The table legs take on the shape of branching tree trunks simply by specifying material to be used optimally within the allowable space to maximize stiffness.

This tool also allows for design optimization using lightweight metamaterials. The tool calculates a homogenized set of properties for the unit cell of the lattice(s) chosen for the design. This set of homogenized properties is then used to perform a multi-material optimization where PLATO calculates optimal shape as well as where to put specific metamaterials within this shape. This allows for a mix of solids to connect with mating parts of an assembly, and lattices where material properties themselves can be optimized for such things as bending, shear, or energy absorption.

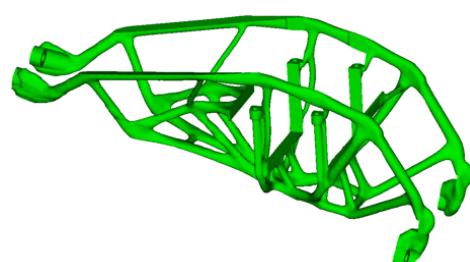
This new “generative” design will transform the role of the designer from defining shapes using CAD packages, to evaluating immense families of generated designs where performance priorities and confidence factors become the trade space being explored.



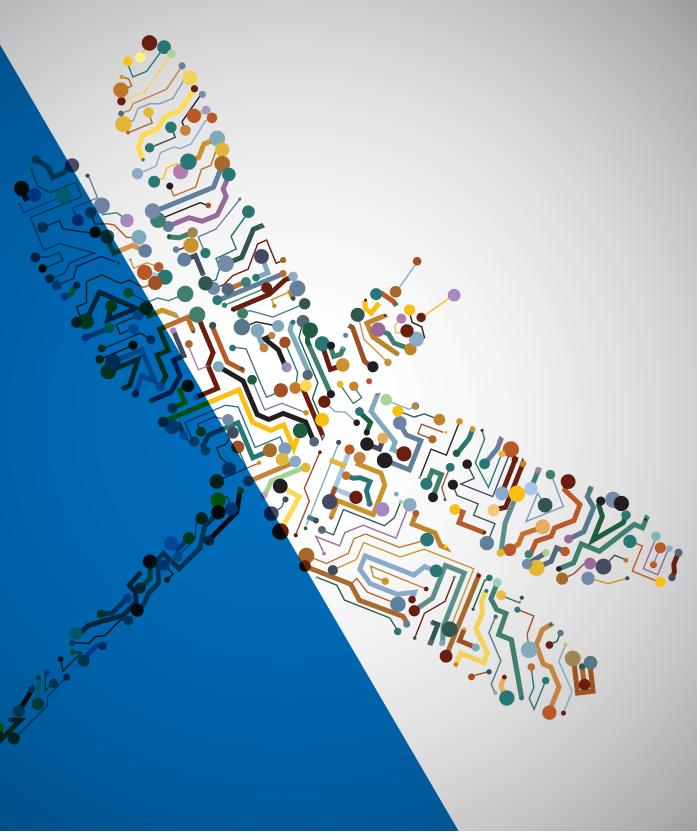
A design where homogenized lattice properties were used to define the shape using topology optimization and then the lattice structure was inserted using hex meshing techniques.



Scan this image with your SNLSimMagic App to see an associated movie clip.



A complex mounting assembly designed to hold a mass as rigid as possible at a given location with a minimal weight.



EVALUATING DRAGONFLY GLOBAL LINK ARCHITECTURES

VITUS LEUNG (PI)

KNOX COLLEGE:

DAVID BUNDE

EMILY HASTINGS

DAVID RINCON-CRUZ

MARC SPEHLMANN

SOFIA MEYERS

ANDA XU

The challenges of HPC resource management have provided many interesting research problems including network architecture, routing, scheduling, allocation, and task mapping. The techniques used to evaluate solutions to these problems have included analysis, simulation, and real-machine experiments. Thus, HPC has provided a very satisfying range of computer science work.

There are several key challenges that must be addressed to continue improving the capabilities of high performance computing systems—one of these challenges concerns the node interconnects. Scaling up traditional interconnects creates a network that requires too much power and faces severe bandwidth limitations as an increasing number of nodes attempt to split link bandwidth between more communicating pairs. New interconnect topologies are being developed to address these challenges by exploiting two technology changes, the development of high-radix routers and economical optical links.



Focusing on Dragonfly topology, a team of researchers from Sandia and Knox College are analyzing how node connection impacts system performance. A Dragonfly system is composed of switches organized into groups. Each switch has electrical links attaching it to nodes and to every other switch in its group. Each switch also has optical links to switches in other groups, organized so that every pair of groups has a link between them. The switches of a group effectively form a virtual switch of very high radix, connecting all nodes associated with its group to the virtual switches associated with each other group. Because of their different roles, a Dragonfly's electrical links are called local links while its optical ones are called global links. The shortest path between any pair of switches is at most three hops: a local link in the source group, a global link between groups, and a local link in the destination group. These shortest path routes are called direct routes. When direct routes are overly congested, a message can take an indirect route, traveling by way of a randomly chosen intermediate group in a variation of Valiant's randomized routing algorithm. Even with indirect routing, all messages traverse at most five hops.

The team analyzed three specific ways to make these connections: absolute, relative, and circulant-based arrangements. Each arrangement was evaluated for their bisection bandwidth, the minimum bandwidth between two equal-sized parts of the system. This common network metric captures the worst-case communication bottleneck in a large computation. It is also particularly suitable for Dragonfly networks since randomized indirect routing will tend to equalize link utilizations. Their analysis concluded that the choice of global link arrangement changes the achievable network performance. Ongoing research and simulation is currently underway to analyze enough structures to determine which arrangement can markedly improve current HPC systems.

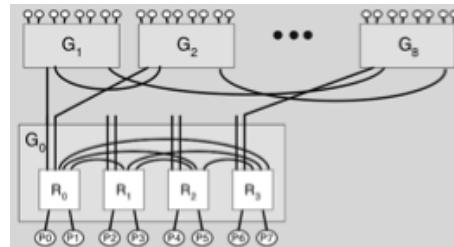


Figure 1. Dragonfly architecture

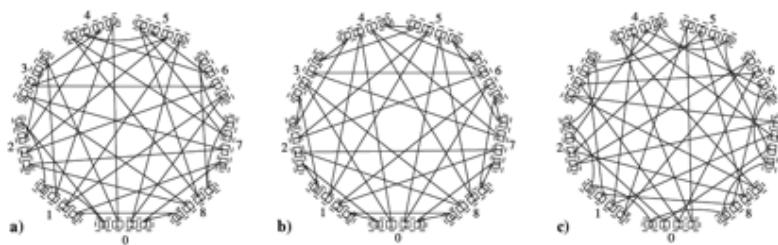


Figure 2. Three global link arrangements for
 (a) Absolute
 (b) Relative
 (c) Circulant-based

Range of α	Bandwidth by link arrangement		
	Absolute	Relative	Circulant
0 – 1.25	4 + 16 α	4 + 16 α	4 + 16 α
1.25 – 1.5		14 + 8 α	
1.5 – 2			16 + 8 α
2 – 2 $\frac{2}{3}$	24	20 + 4 α	20 + 6 α
2 $\frac{2}{3}$ – 4			36
> 4		36	

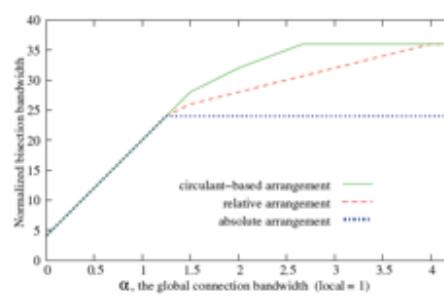
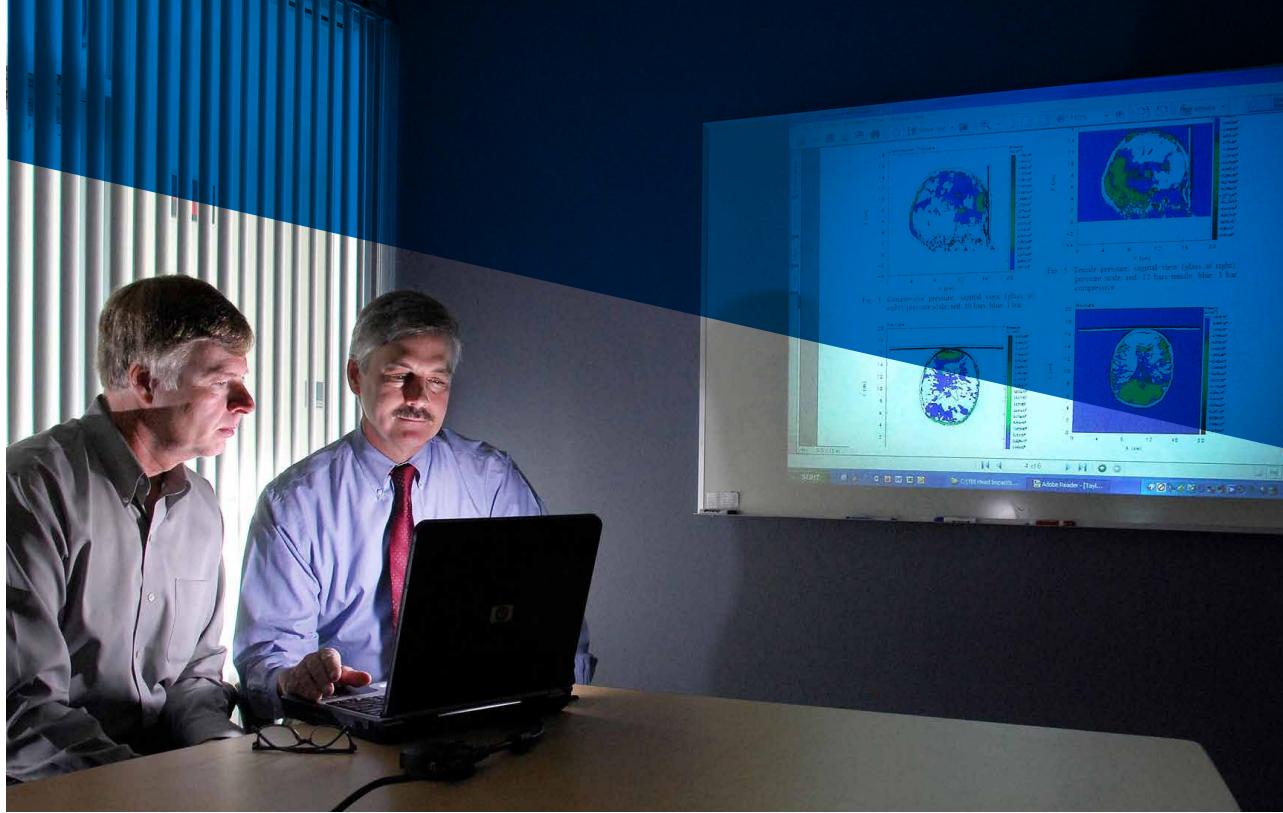


Figure 3. Normalized bisection bandwidth for the global link arrangements. Each cut has bandwidth (# local links) + $a(\# \text{global links})$. The table (left) and graph (right) show the bandwidth of the min-bandwidth cut.



SIMULATION OF BLAST AND BEHIND-ARMOR BLUNT TRAUMA TO LIFE-CRITICAL ORGANS IN THE HUMAN BODY

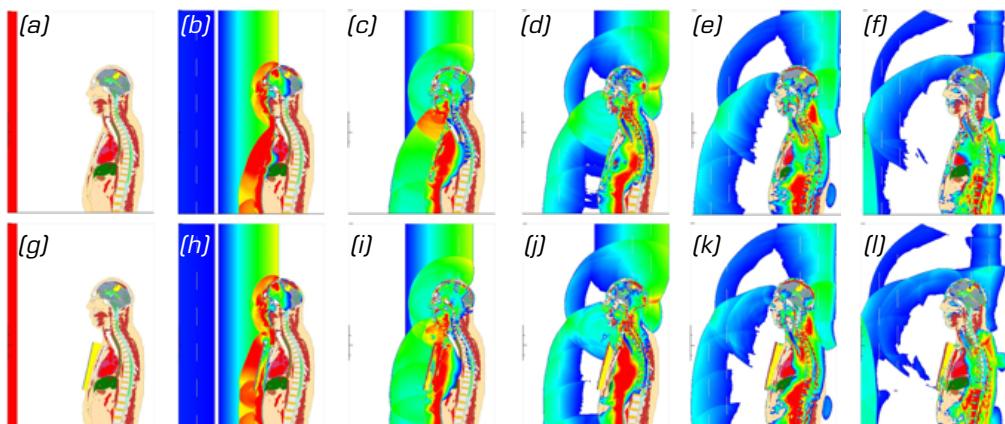
PAUL A. TAYLOR (PI)

CANDICE F. COOPER

Our virtual simulation of warfighter injury scenarios and protective armor assessment relies heavily on our use of the Sandia HPC facilities. These facilities play an essential role in our efforts to protect the warfighter against debilitating and life threatening battlefield conditions.

Researchers at Sandia have developed a high-fidelity virtual model of the human head, neck, and torso to investigate the details of life-threatening injury to the central nervous, respiratory, and cardiovascular systems as a result of blast exposure and behind-armor blunt trauma. This model set comprises separate head-neck and torso models that can be used independently or combined to investigate comprehensive injury to life-critical organs as a result of blast, blunt impact, and/or projectile penetration. The Sandia head-neck-torso model represents a 60th percentile human male from the waist up possessing anatomically correct distributions of bone, white and gray brain matter, falx and tentorium membranes, spinal cord, intervertebral disks, cartilage, vasculature, blood, airways, lungs, heart, liver, stomach, kidneys, spleen, muscle, and fat/skin.

This year, researchers built upon previous simulation capability by adding the ability to assess virtual armor protection of the torso model. The Sandia wave physics code, CTH, was used in conjunction with this model to simulate blast loading and ballistic projectile impact to a warfighter with and without protective armor, in order to investigate the details of injury to life-critical organs such as the brain, spinal cord, lungs, airways, heart, blood vessels, and liver as a result of the intrathoracic pressure waves that are generated from a blast or impact. These simulations can be used to assess new armor design and facilitate development of advanced armor.



Frontal blast wave exposure of Sandia head-neck-torso model unprotected (a-f) and with notional chest armor (g-l).
 (a & g) Initial setup. Blast wave is generated from reservoir (denoted in red) of high pressure air on left side of image.
 (b & h) Blast wave propagating to right with wave front impacting human model.
 (c & i) Blast wave has generated transmitted pressure waves into the body with reflected air wave traveling back toward blast source.
 (d-f) and (j-l): Time-evolution of wave transmission within the body and reflected wave in air. Differences in pressure plots between (a-f) and (g-l) are due to the presence of the chest armor. White space outside human model indicates regions experiencing pressures below 1.1 atmospheres (110 KPa) in magnitude.



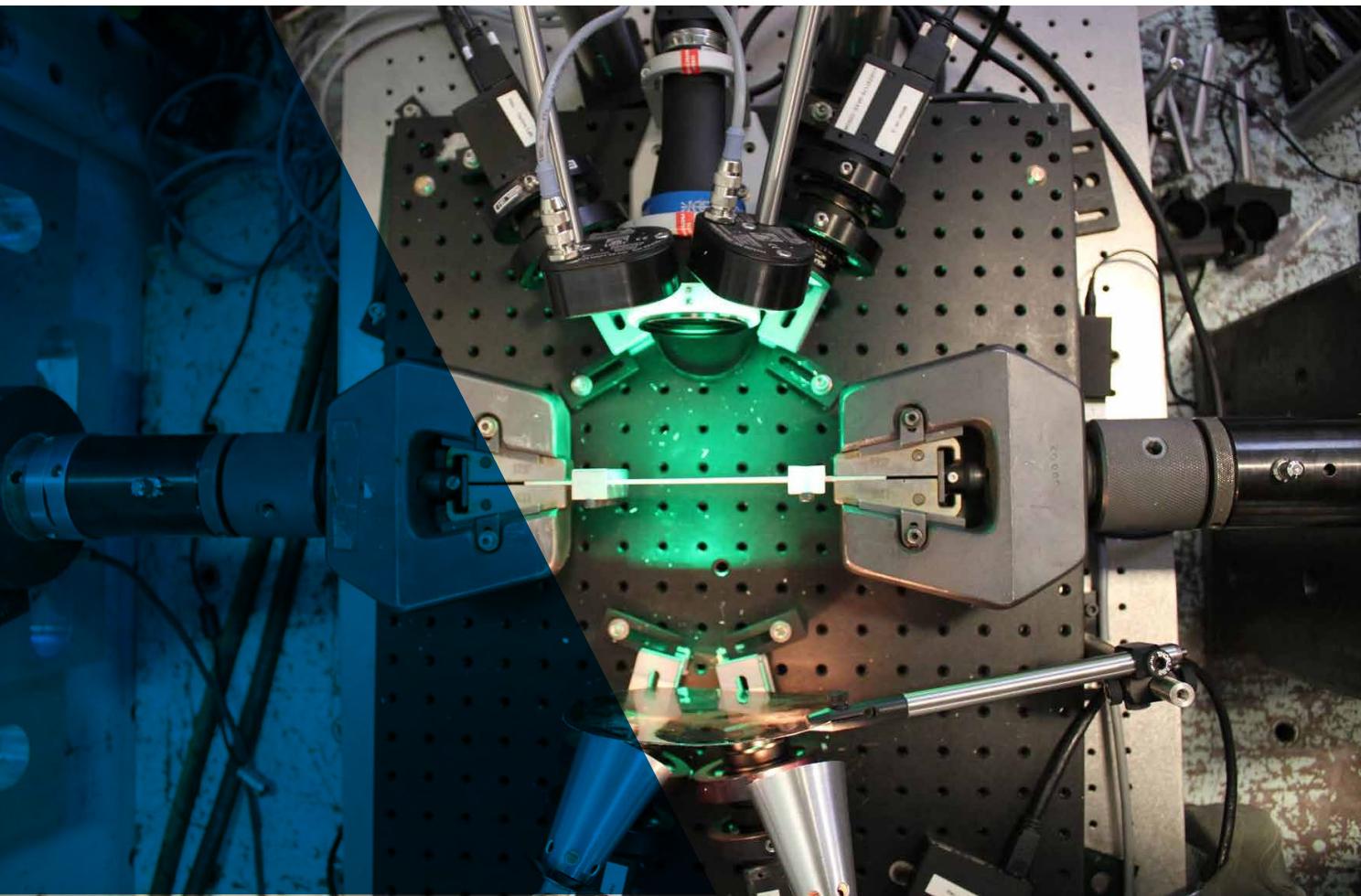
Scan this image with your SNLSimMagic App to see an associated movie clip.

HIGH FIDELITY MODEL CALIBRATION USING DIGITAL IMAGES

DAN TURNER (PI)

RICH LEHOUcq

PHIL REU

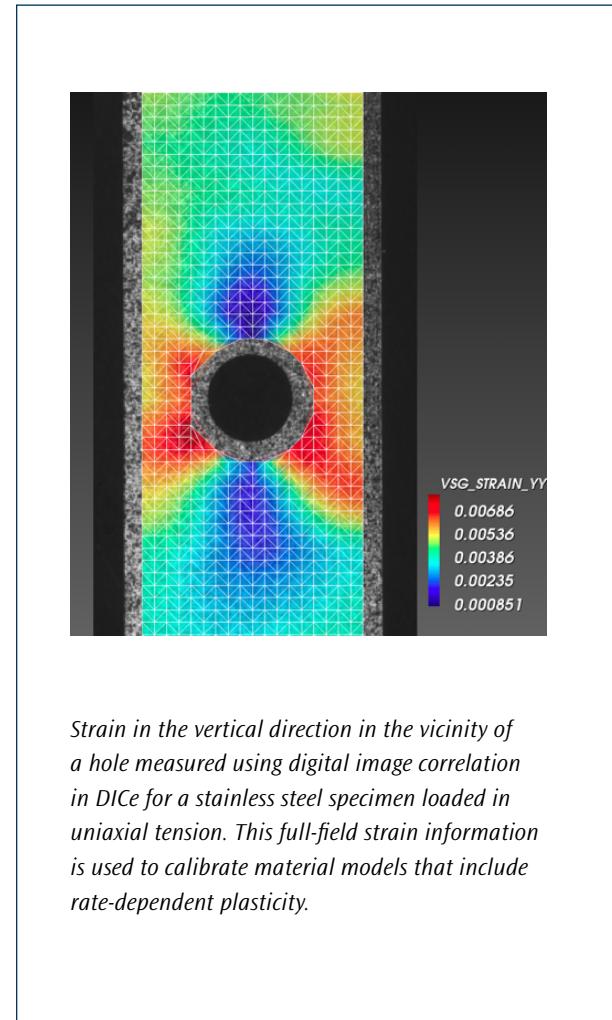


HPC has enabled the analysis process to keep pace with the rapidly growing volume and heterogeneity of data provided by modern experimental techniques in material characterization. Contemporary model calibration involves the fusion of highly resolved numerical models, data processing, and optimization routines—each of these being very computationally intensive individually. This coordinated integration, which has led to unsurpassed scientific discovery, would not be possible without substantial investment in HPC.

Digital image correlation (DIC) is rapidly transforming the way material properties are obtained and gaining traction in other areas such as manufacturing quality assurance. Contemporary DIC methods involve a tremendous amount of data, acquired at rates of over tens of gigabytes per second, and require high-resolution discretizations to capture important phenomena such as strain localizations or discontinuities. At Sandia, DIC is used to calibrate material models for weapon component materials such as stainless steel. DIC is also used to evaluate the performance of component systems like the nuclear强链接, which prevents accidental detonation of nuclear weapons.

Sandia's Digital Image Correlation Engine (DICe) project is developing advanced, next-generation DIC algorithms that can tackle both the volume of data involved and the need for high-fidelity measurements. The DICe analysis code is built on the Kokkos library from Trilinos that enables platform-portable, on-core parallelism as well as inter-processor parallelism. In addition, the DICe project has developed a novel, unconditionally stable formulation for DIC that circumvents the spurious oscillations present in existing methods. This is the first DIC formulation limited by the camera resolution rather than the robustness of the numerical algorithm.

The DICe project represents a truly unique collaboration that seamlessly fuses experimental measurement, applied mathematics, and high performance computing. The integration of these disciplines is paving the way toward unsurpassed fidelity in model calibration and opening the door to a number of previously intractable scientific endeavors.





RISK ASSESSMENTS FOR HIGH CONSEQUENCE FLIGHT TESTS

LARRY YOUNG (PI)



High consequence flight tests for Strategic Partnership Projects and nuclear weapons applications require review and approval of the risks associated with any flight program involving Sandia personnel or assets.

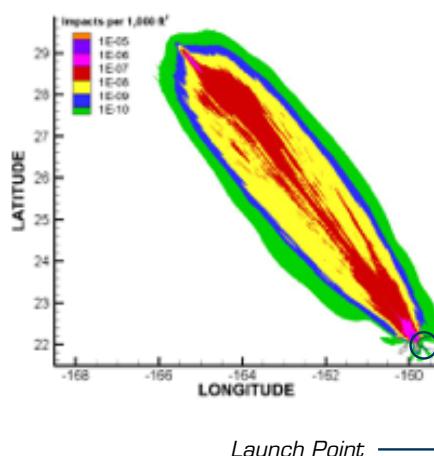
The Range Commanders Council 321-10 Standard has provided guidelines to quantify uncertainty in probabilistic risk assessments. High consequence flight tests include missile launches. If a flight has to be terminated during a launch, the impact locations of a large amount of debris have to be computed to assess the potential probability of impact at any particular location. The many models and approximations for debris, trajectories, motor and component malfunctions, and population/asset demographics used in flight risk assessment have considerable uncertainty. The DoD continues to fly increasingly complex flight tests, which will be better designed when informed about uncertainties in the failure risks.

In the past, flight safety risk assessments only provided a mean estimate of the risk to any potential population center. The uncertainty around that mean value has now been added to the risk assessment tool to show what an estimate of an upper limit to the risk assessment would be. The need for this assessment was demonstrated during a recent launch where the mission program wanted to move an observation ship as close to the nominal missile impact area as possible. The estimated risk at the proposed location mean impact probability was within the risk

limits; however, the two sigma impact probability was very close to the risk limit. This showed the decision makers that the chosen location was as close to the nominal impact area as safely possible.

The computational needs are demonstrated during the analyses of rocket launches. The impact locations of over eighteen billion debris pieces were computed during this recent flight safety risk assessment. This would not have been possible without the resources of high performance computing.

As missile tests have become more and more complex, the need to estimate the impact locations along the order of tens—and approaching hundreds of billions—of debris pieces is becoming the norm. Without the HPC resources, the computation times for a pre-flight analysis would be prohibitive.



Scan this image with your SNLSimMagic App to see an associated movie clip.

PREDICTING FRAGMENTATION DISTRIBUTION USING PERIDYNAMIC THEORY



PAUL DEMMIE (PI)



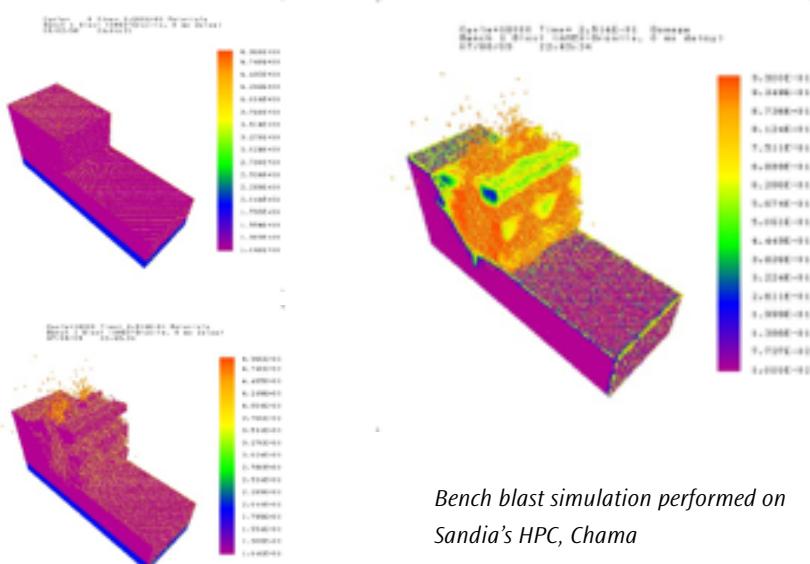
Using Sandia's HPC capabilities, we were able to perform large-scale simulations of rock blasting in a reasonable timeframe. Hence, we were able to explore the effects of varying the initiation of detonation in multiple blast holes and uncertainties in model parameters.

Bench blasting is commonly used in hard rock mining to remove a host rock with targeted minerals from in-situ. This method is carried out using explosives placed into holes drilled throughout the host rock and detonated at varying intervals. Sandia researchers developed a simulation of this process to predict the distribution of blast fragmentation.

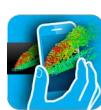
The objective of this project was to perform code enhancements and demonstration problems to assess the capability of a computer code based on peridynamic theory for the modeling of rock blasting. Peridynamic theory is a theory of continuum mechanics based on integral equations rather than the standard approach, which is based on partial differential equations.

Partial differential equations fail to apply when cracks are present, due to the mathematical singularities that the cracks represent, necessitating the use of peridynamic theory. Simulations were performed on Sandia's HPC, Chama, using 256 processors. These simulations modeled a section of a granite bench with two blast holes each containing an AMEX™ emulsion (explosive). They predicted fragment distributions and were able to demonstrate how timing of blast initiations affected the fragment distribution.

These simulations can provide valuable information for those in the mining industry who want to optimize blasting to decrease costs and to protect structures near the blasting site.



Bench blast simulation performed on Sandia's HPC, Chama



Scan this image with your SNLSimMagic App to see an associated movie clip.



SANDIA'S ADVANCED ARCHITECTURE TEST BEDS: PARTNERING WITH INDUSTRY TO SCOUT A PATH FOR FUTURE HPC SYSTEMS

JAMES H. LAROS III (PI)

ROBERT HOEKSTRA

JAMES BRANDT

JASON REPIK

BEN ALLEN

SIMON HAMMOND

SUZANNE KELLY

ANN GENTILE

NATHAN GAUNTT

JAMES A. ANG

JOHN NOE

VICTOR KUHNS

CONNOR BROWN



PREPARING FOR THE FUTURE OF HPC

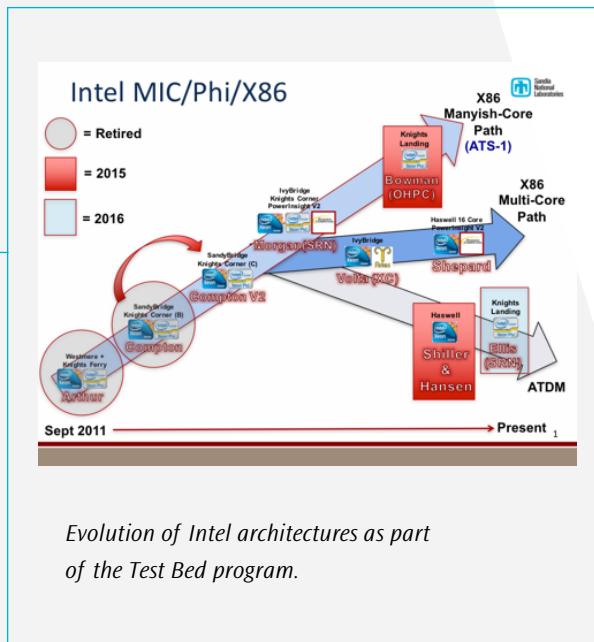
The Test Bed program has provided IBM's developers with fact-based feedback on many aspects of the design, usability, performance, and energy management aspects of the hardware and software platforms involved. Sandia's feedback is crucial to our success in developing and validating configurable, flexible exascale architecture solutions that meet the needs of the Labs and the wider HPC community. The Test Bed program provides an ideal vehicle by which IBM and Sandia can collaborate on improvements needed and set the correct direction for the next generations of high performance systems.

—Steve Fields, IBM Fellow & Chief Engineer of Power Systems

In 2011, Sandia established the Advanced Architecture Technology Research and Development project, more simply known as the Test Bed project. The Test Bed project was originated as an element of the NNSA's Advanced Simulation and Computing (ASC) program. Decades-long relative stability in node and system architectures are currently being replaced by a diverse set of disruptive technology changes, primarily in node architectures. While it seemed clear that significant changes to Tri-Lab production application codes would likely be required, which node architecture and what changes, (e.g., which new programming model to target) was unclear. Sandia's Test Bed project was chartered to facilitate analysis to become a scout for future architectures, as production applications are too complex and costly for this type of experimentation and analysis.

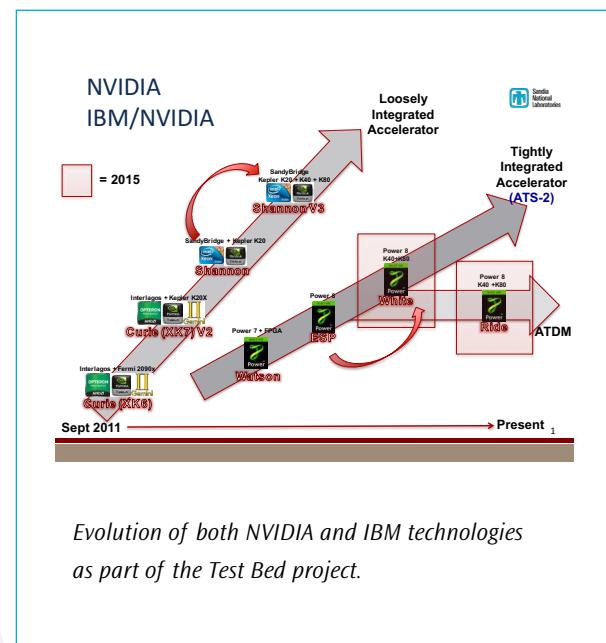
The Test Beds, in conjunction with mini-apps developed as part of Sandia's Mantevo project, are used to deploy and analyze node and rack scale preproduction and prototype advanced architectures. These path-finding explorations include, but are not limited to: application performance analysis, programming model investigations, memory sub-system development, advanced systems software, and power and energy research. Early access to emerging technologies (before general availability) and the dynamic analysis capability afforded by mini-apps—which represent the most important aspects of our large production applications—has enabled the Test Bed program to maneuver an unpredictable landscape, charting a clearer vector for larger production applications.

At the initiation of this program it was unclear which architectures would emerge as viable candidates for an eventual exascale platform. This year, potential technology paths for an exascale platform are beginning to solidify, in large part, due to early analysis performed on preproduction/prototype node architectures deployed as part of the Test Bed program. Arthur, one of the earliest Test Bed systems, was the first—and at the time largest—deployment of Intel’s Many Integrated Core (MIC) technology. MIC prototype accelerator cards were delivered to Sandia a full two years prior to general availability of what would be re branded as Xeon Phi. In 2015, the Bowman Test Bed system was the first A0 silicon Knights Landing (KNL) Phi cluster Intel delivered to a customer worldwide. Bowman was refreshed with B stepping silicon in April 2016 and a second KNL cluster (Ellis) was delivered to Sandia in July. This early exploration of MIC/Phi technology directly influenced the decision to deploy Intel KNL processors for the first Advanced Technology System (ATS-1 or Trinity) in 2016.



From inception, the Test Bed project included a broad range of architectures as part of an overall evaluation of industry directions. Investigation of NVIDIA GPUs, likely the most prominent GPU technology, began at Sandia prior to the Test Bed project. Initially, NVIDIA GPUs were investigated in combination with x86 processors from

both AMD (Curie) and Intel (Shannon). While it began as an orthogonal path, IBM Power processors were also investigated. The Test Bed project provided the conduit for Sandia and IBM to develop a close relationship, which led to Sandia's selection to participate in IBM's Early Ship Program that delivers pre-general availability hardware to a small select group of collaborators many months in advance. The Test Bed project began analysis of the combination of IBM Power processors and NVIDIA GPU accelerators in 2014, with the promise that an IBM/NVIDIA collaboration would provide a high-speed link (NVLINK) between the Power host processor and the NVIDIA accelerator. This change in technology would make discrete accelerator solutions viable for Tri-Labs weapons codes. Since then, multiple generations of IBM/NVIDIA node architectures have been delivered as part of the White and Ride Test Bed clusters. A platform combining IBM Power host processors and NVIDIA GPU accelerators was selected for Sierra (ATS-2), which will be deployed in 2018.



When these technologies (Intel MIC/Phi, IBM Power, and NVIDIA GPUs) were deployed as part of the Test Bed project, their future viability was unknown. The early MIC line of accelerators was a risky path for Intel and an IBM/NVIDIA collaboration to integrate their technologies was unlikely at the time. Following all of these technologies from the start has allowed the NNSA/ASC program to make important

The Sandia Advanced Architecture Test Bed program has been invaluable for evaluations of our products and workloads. We were able to debug software issues and validate important DOE applications. This continued collaboration is an example of the partnership between our two organizations in creating leadership technologies for the computing industry.

—Barry Davis, General Manager of Intel's Accelerated Workloads Group

strategic decisions and navigate a rapidly changing technology environment, which informed the selection of the first two Advanced Technology Systems. The Test Bed project is now providing critical information that will be used to evaluate technologies for Crossroads (ATS-3), which will be deployed in 2020.

The Test Bed project continues to deploy architectures that have the potential to impact the NNSA/ASC mission space. ARM processors have recently become potentially viable for high performance computing. The Test Bed project was the first in the world to deploy a 64-bit ARM processor cluster (Hammer) using X-Gene 1 processors from Applied Micro. While these processors were not targeted toward HPC, this deployment allowed an early look at new technology and an evaluation of potential gaps in the ARM software eco-system for HPC. In collaboration with Cavium, the Test Bed project deployed the first in the world Cavium-based 64-bit ARM cluster (Sullivan). These server class processors are allowing Sandia to continue work to close the gaps in the ARM

software ecosystem and begin performance analysis using the Manteye mini-app suite. Delivery of next generation ARM processors is expected in the latter part of 2016.

Sandia has fostered relationships for many years with a wide range of technology providers and original equipment manufacturers. The Test Bed project has leveraged and strengthened these relationships by providing a mutually beneficial environment. Companies like IBM and Intel, among others, provide early prototypes of emerging node and system architectures to Sandia Labs. Sandia benefits by getting an early look at architectures that have been, and may be, part of capability class systems that host Tri-Lab production scientific simulation applications. Sandia's partners benefit from early identification of hardware and software issues. Strategic relationships like these are critical to the Tri-Lab weapons mission. Sandia has and will continue to lead the way in investigating the most beneficial path to achieve exascale processing capability for the NNSA's mission applications, and beyond.

The Test Beds are largely named after notable Americans (with some exceptions).

Arthur Compton – American Physicist, Nobel Prize 1927

Alice Bowman – First female Missions Operations Manager, Pluto Mission

Clarence Ellis – First African American to earn a PhD in Computer Science

Marie Curie – First woman to win a Nobel Prize 1903, first person to win twice (again in 1911)

Claude Shannon – American, father of information theory

Helen Magill White – First American woman to earn a PhD in the United States

Sally Ride – First American woman in space

Hammer – Exception to the rule, comically named after Arm & Hammer

Kathryn Sullivan – First American woman to walk in space

MISSION

COMPUTING

SYSTEMS & SUPPORT

CHAMA

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	392	172,677,120
Nodes	Cores	Memory/Core	
1,232	19,712	2.0 GB	

Chama, along with Pecos, is a NW/ASC HPC system deployed in 2012 as part of the DOE/NNSA Tri-Labs TLCC2 procurement.

CIELO DEL SUR

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	86	77,928,960
Nodes	Cores	Memory/Core	
556	8,896	2.0 GB	

Cielo del Sur is a Cray XE6 system that supports users of classified computing as part of the National Security Computing Center operated by Sandia on behalf of DOE.

DARK BRIDGE

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	294	129,507,840
Nodes	Cores	Memory/Core	
924	14,784	4.0 GB	

Dark Bridge is a TLCC2 system that supports users of classified computing as part of the National Security Computing Center operated by Sandia on behalf of DOE.

DARK SAND

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	294	129,507,840
Nodes	Cores	Memory/Core	
924	14,784	4.0 GB	

Dark Sand is a TLCC2 system that supports users of classified computing as part of the National Security Computing Center operated by Sandia on behalf of DOE.

JEMEZ

Usage	Program	TFLOPS	Process Hours/yr
HPC	Institutional	95	40,366,080
Nodes	Cores	Memory/Core	
288	4,608	2.0 GB	

Jemez is an Institutional HPC system that was released for general use in FY14. At 96TFlops peak, this system is a major and cost-effective addition to the Institutional resources available to users who are doing classified computing.

MINI SEQUOIA

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	107	71,761,920
Nodes	Cores	Memory/Core	
512	8,192	1.0 GB	

Mini Sequoia is a small version of the Tri-Lab Sequoia system sited at LLNL. The purpose of Mini Sequoia is to provide local code-development and checkout system for the application teams that support Sandia users on Sequoia.

MUTRINO

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	NA	28,032,000
Nodes	Cores	Memory/Core	
100	3200	4.0 GB	

Mutrino is a small-scale Cry XC system that is one of two Application Regressions Testbed (ART) systems, supporting the Sandia/LANL ACES partnership Trinity platform located at LANL.

MUZIA

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	NA	2,803,200
Nodes	Cores	Memory/Core	
20	320	2.0 GB	

Muzia is a small-scale Cray XE6 system that was acquired as part of the Sandia/LANL partnership that manages the Cielo platform at LANL.

OPENSTACK

Usage	Program	TFLOPS	Process Hours/yr
Cloud	Institutional	NA	14,016,000
Nodes	Cores	Memory/Core	
100	1,600	8.0 GB	

OpenStack is an Institutional Cloud system that was acquired at the end of FY13 for use as a research cloud. The system comprises 100 Dell r720 nodes having local disk and bound together with a highly configurable Ethernet fabric. An OpenStack cloud environment will support multiple research groups who are experimenting with “Infrastructure as a Service.”



PECOS

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	392	172,677,120
Nodes	Cores	Memory/Core	
1,232	19,712	2.0 GB	

Pecos is a NW/ASC HPC system deployed in 2012 as part of the DOE/NNSA Tri-Labs TLCC2 procurement. Pecos is a primary resource for NW/ASC users.

PLATO

Usage	Program	TFLOPS	Process Hours/yr
Analytics	Institutional	NA	7,148,160
Nodes	Cores	Memory/Core	
51	816	6.0 GB	

Plato is a Hadoop cluster, which entered production in FY14. The HP-based system runs out-of-the box Cloudera's CDH enterprise product.

RED MESA

Usage	Program	TFLOPS	Process Hours/yr
HPC	Institutional	180	134,553,600
Nodes	Cores	Memory/Core	
1,920	15,360	1.5 GB	

Red Mesa is an Institutional HPC system on the collaborative network. Based on the Red Sky platform architecture, Red Mesa is used by EC PMU and other Partners.

RED SKY

Usage	Program	TFLOPS	Process Hours/yr
HPC	Institutional	264	197,835,840
Nodes	Cores	Memory/Core	
2,823	22,584	1.5 GB	

Deployed in 2010, Red Sky and Red Sky (C) have been the workhorses of Institutional HPC computing. Developed in collaboration with Sun Microsystems, this cluster was the first large-scale HPC system to deliver an Infiniband interconnect based on a Torus network topology.

SKY BRIDGE

Usage	Program	TFLOPS	Process Hours/yr
HPC	Institutional	588	259,015,680
Nodes	Cores	Memory/Core	
1848	29,568	4.0 GB	

Based on the same hardware as TLCC2, Sky Bridge is the latest Institutional compute cluster. At 588 TF peak, the water-cooled Cray CCS cluster is ranked 123 in the June 2015 TOP500 HPC list. By partnering with Cray and Asetek, Sandia was able to show that a water-cooled HPC cluster can be both reliable and energy efficient. This success has set the stage for future water-cooled HPC clusters at Sandia and elsewhere.

TWINFIN

Usage	Program	TFLOPS	Process Hours/yr
Analytics	Institutional	NA	0
Nodes	Cores	Memory/Core	
1848	29,568	4.0 GB	

Twinfin is an IBM/Netezza analytics appliance for structured- and semi-structured search. Twinfin came online early in FY13. The system integrates proprietary hardware and software to accelerate structured search integrated with data analytics.

UNO

Usage	Program	TFLOPS	Process Hours/yr
HPC	NW/ASC	58.0608	23,546,880
Nodes	Cores	Memory/Core	
168	2,688	4.0 GB	

Uno is the first high-throughput cluster deployed at Sandia and features three capabilities: 168 traditional compute nodes for a single node jobs, 25 compute+GPU nodes and 8 many core nodes. Based on a Dell compute node, Uno is designed to provide high-throughput and fast turnaround for single-node jobs. Uno provides a variety of heterogeneous nodes (small and large memory, processors and accelerators) with its interconnect and file systems tuned for single-node activities.

HPC OneStop

SERVICE DESK

Sandia's HPC OneStop is an entry point for all your scientific, engineering, and high performance computing needs. HPC OneStop provides lots of useful information, outstanding user support, and rapid problem solving.

Access HPC OneStop using the venue that works best for you:

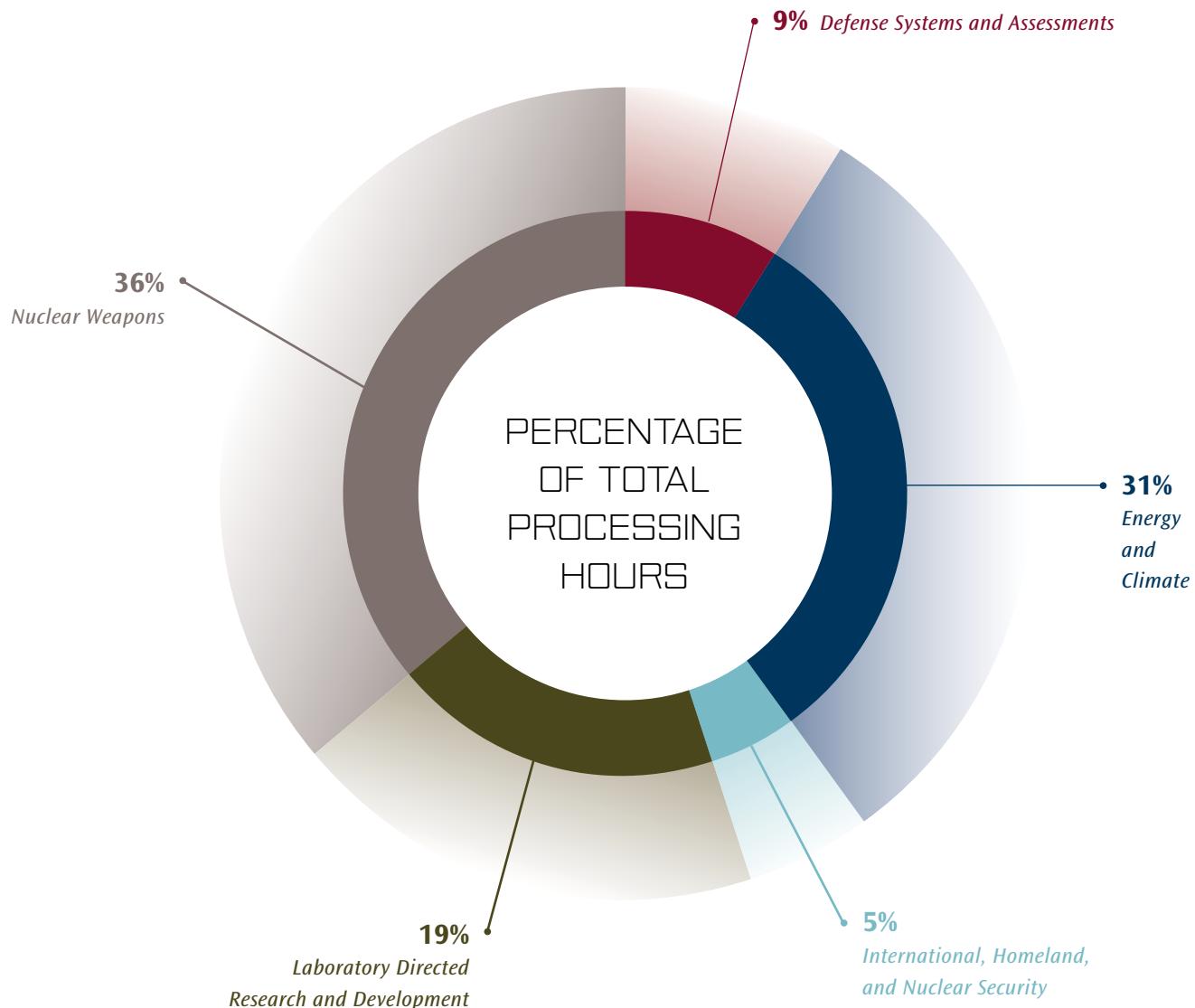
- Call HPC@CCHD: 845-2243 Select Option 8
- Call HPC Service Desk: 844-9328
- E-mail: hpc-help@sandia.gov
- <https://computing.sandia.gov>

[HTTPS://computing.sandia.gov](https://computing.sandia.gov)

SANDIA

PROGRAM MANAGEMENT UNITS

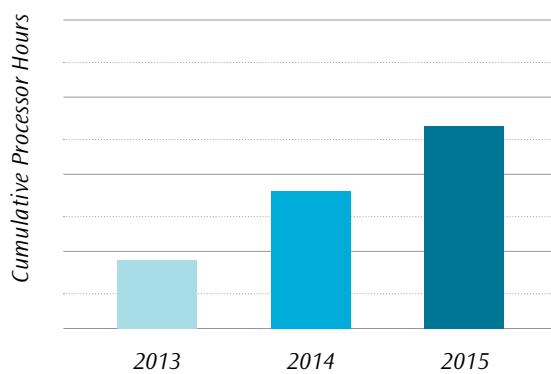
The ability to routinely access computers, data, and expertise are now essential components of scientific research, but they would not be possible without the support of Labs Leadership and Program Management Units (PMUs), including Defense Systems and Assessments, Energy and Climate, International, Homeland, and Nuclear Security, and Laboratory Directed Research and Development. Collectively, the PMUs, LDRD, and scientific computing have become partners in providing Mission Computing to Sandia and our collaborators.



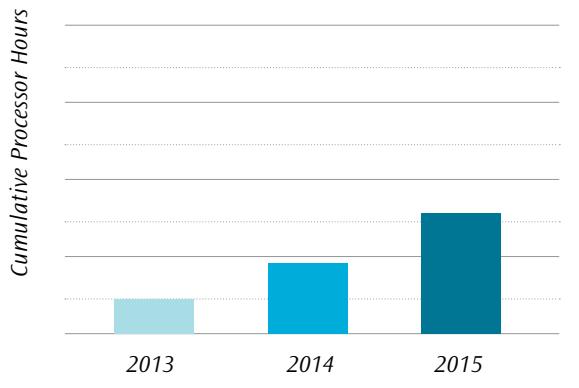
CUMULATIVE PROCESSOR HOURS PER FISCAL YEAR

**2016 data was not yet available at the time of press.*

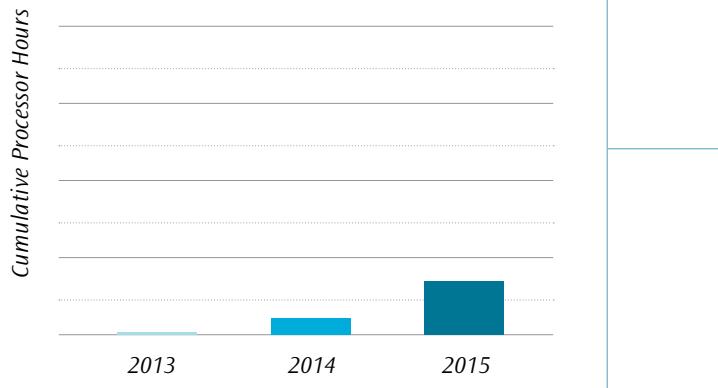
Defense Systems and Assessments



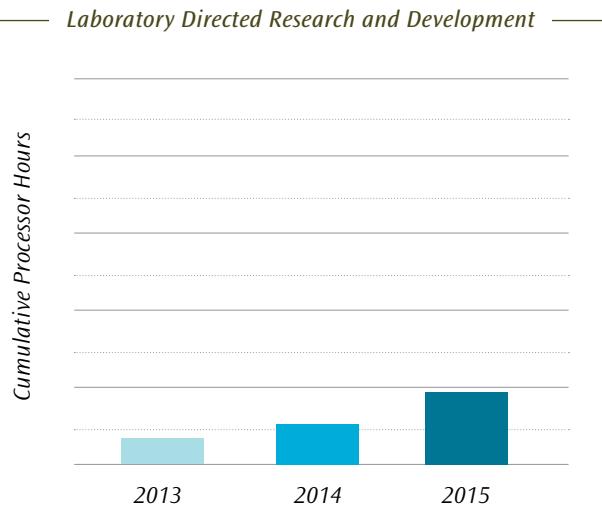
Energy and Climate



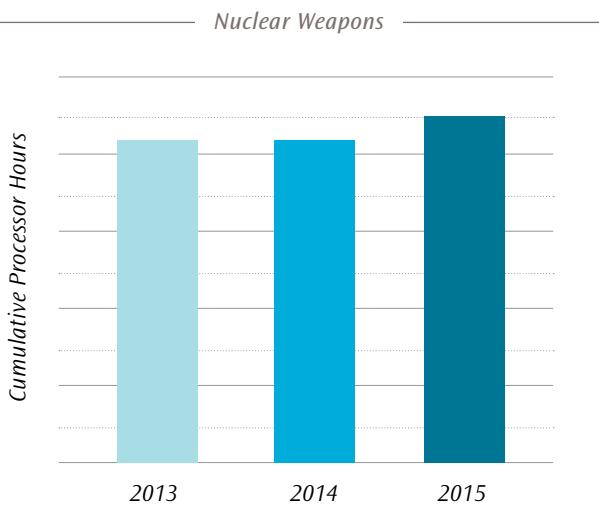
International, Homeland, and Nuclear Security



Laboratory Directed Research and Development



Nuclear Weapons





Sandia
National
Laboratories



U.S. DEPARTMENT OF
ENERGY

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND XXXXXX