Advanced Scientific Computing Research Update

Ceren Susut, PhD Associate Director of Science for Advanced Scientific Computing Research

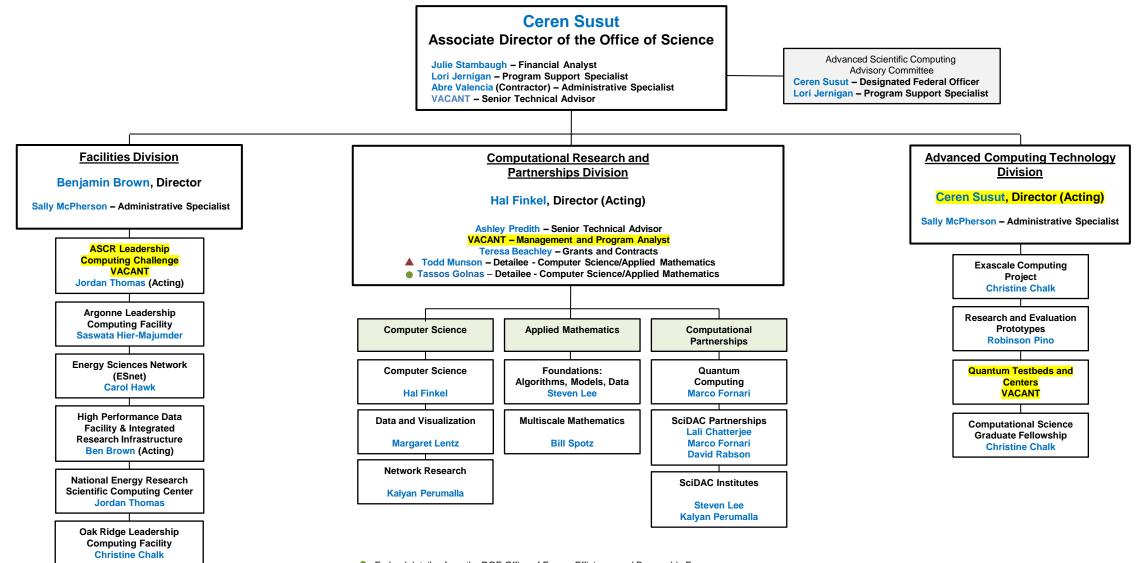


U.S. Department of Energy Office of Science

Office of

ADVANCED SCIENTIFIC COMPUTING RESEARCH

Functional Organization Chart



Federal detailee from the DOE Office of Energy Efficiency and Renewable Energy

Detailee from Argonne National Laboratory

Staffing Updates

 ASCR welcomes Tassos Golnas, on detail from DOE's Office of Energy Efficiency & Renewable Energy (EERE), Solar Energy Technologies Office (SETO) through the first week of July. Tassos is working with ASCR on Al and on ASCR engagement with DOE's applied-energy programs.

- ASCR thanks Tom Wong, who departed ASCR in December 2023 to return to Creighton University. During Tom's time at ASCR, he worked on quantum topics and also served on detail to the Office of Science and Technology Policy.
- ASCR thanks Claire Cramer, who departed ASCR in March 2024 for a position with the US Naval Observatory. Claire initiated ASCR's quantum computing testbeds program and played an instrumental role in SC's QIS initiative.





Advanced Computing Technology (ACT) Division Director Position Now Open

Senior Executive Service (SES) Position

- ACT Division is responsible for the scientific research and development activities required to implement programs in critical and emerging computing architectures and technologies, including quantum and neuromorphic computing and novel accelerators that promote advances in artificial intelligence (AI). The ACT Division Director will oversee the evaluation of research and development needs and opportunities in computational science, including hardware research and testbeds development in critical and emerging computing technologies; the development of forward-looking programs in these areas; and the formation of partnerships with other organizations including industry leading to science and technology innovations.
- Open between May 16 and June 17
- > Application through USAJOBS.gov site at <u>https://www.usajobs.gov/GetJob/ViewDetails/791395100</u>



ASCR FY 2025 Request

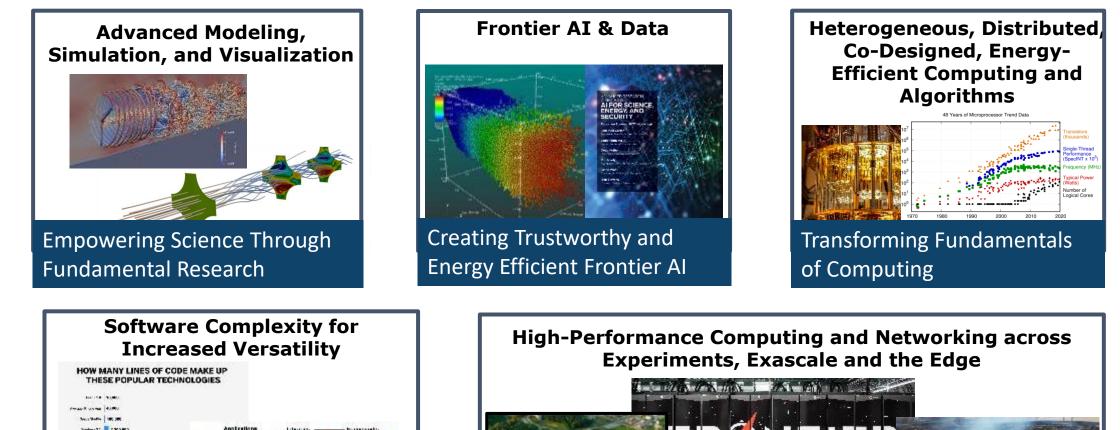
	(dollars in thousands)						
	FY 2023 Enacted	FY 2024 Enacted	FY 2025 Request	FY 2025 Request vs FY 2024 Enacted		FY 2025 Request vs FY 2023 Enacted	
Advanced Scientific Computing Research							
Applied Mathematics Research	61,035	52,182	77,565	+25,383	+48.64%	+16,530	+27.08%
Computer Sciences Research	60,667	66,718	86,736	+20,018	+30.00%	+26,069	+42.97%
Computational Partnerships	95,875	75,182	93,449	+18,267	+24.30%	-2,426	-2.53%
Advanced Computing Research	108,920	108,918	148,197	+39,279	+36.06%	+39,277	+36.06%
Energy Earthshot Research Centers	12,500	5,000	12,500	+7,500	+150.00%	-	-
Total, Mathematical, Computational, and Computer Sciences Research	338,997	308,000	418,447	+110,447	+35.86%	+79,450	+23.44%
High Performance Production Computing	132,003	142,000	146,500	+4,500	+3.17%	+14,497	+10.98%
Leadership Computing Facilities	430,000	474,000	475,195	+1,195	+0.25%	+45,195	+10.51%
High Performance Network Facilities and Testbeds	90,000	91,000	93,540	+2,540	+2.79%	+3,540	+3.93%
Integrated Research Infrastructure	-	-	3,000	+3,000	-	+3,000	-
Total, High Performance Computing and Network Facilities	652,003	707,000	718,235	+11,235	+1.59%	+66,232	+10.16%
17-SC-20, SC Exascale Computing Project	77,000	-	-	-	-	-77,000	-100.00%
Subtotal, Advanced Scientific Computing Research	1,068,000	1,015,000	1,136,682	+121,682	+11.99%	+68,682	+6.43%
24-SC-20 High Performance Data Facility	_	1,000	16,000	+15,000	+1,500.00 %	+16,000	-
Subtotal, Construction	-	1,000	16,000	+15,000	+1,500.00 %	+16,000	-
Total, Advanced Scientific Computing Research	1,068,000	1,016,000	1,152,682	+136,682	+13.45%	+84,682	+7.93%



FY 2024 Enacted Highlights - \$1,016M (-\$52.0M or 4.7% below FY 2023 Enacted)

- Ensure successful completion of the Exascale Computing Project (ECP) on time and within budget, capturing lessons learned and highlighting scientific accomplishments. Prioritize transitioning ECP researchers, software, and technologies into core research efforts and Department of Energy (DOE) priority research areas.
- Maintain facility operations and existing upgrade projects and continue planning activities for National Energy Research Scientific Computing (NERSC)-10 and Leadership Computing Facility (LCF) upgrades.
- Continue to emphasize foundational applied mathematics and computer science research addressing the combined algorithm and software challenges of increasingly heterogeneous architectures that can effectively leverage AI and ML techniques, and emerging technologies such as quantum computing.
- Initiate Microelectronics Science Research Centers that will focus on broad basic research in promising directions in computing systems and architectures for both traditional and novel computing paradigms to lead to microelectronics energy efficiency meeting the future needs of the scientific enterprise, and the nation.
- Continue to broaden participation and retention in ASCR's programs and workforce by under-represented groups through increases in Reaching a New Energy Sciences Workforce (RENEW), and in Computational Science Graduate Fellowships (CSGF), and continued investments in Funding for Accelerated, Inclusive Research (FAIR) and Established Program to Stimulate Competitive Research (EPSCoR).
- **Continue planning the High-Performance Data Facility (HPDF)** in support of the Integrated Research Infrastructure activity.
- **Maintain strategic Computational Partnerships** including support for SciDAC, Biopreparedness Research Virtual Environment (BRaVE) as well as the on-going partnership with the National Cancer Institute.
- Continued investments in SC Energy Earthshots initiative.

Reinvigorating ASCR Research To Respond to Critical Technology Trends



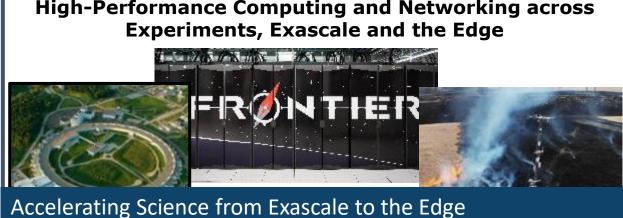
Enhancing Scientific Programming

Campoortes.

Office of

Science

DEPARTMENT OF



Building a Core Research Base: Competitive Portfolios for Fundamental Research

Solicitation Scope:

- Basic applied-mathematics research building on research expertise in a core area such as linear algebra, nonlinear solvers, discretization and meshing, multiscale mathematics, discrete mathematics, optimization, differential equations, scientific machine learning, interactive data analysis, randomized algorithms, differentiable and probabilistic programming, complex systems, computational toolkits, or related topics.
- Basic computer-science research building on research expertise in a core area such as programming languages, high-performance computing tools, scientific data management and scientific visualization, distributed computing infrastructure, programming models for novel computer architectures, automatic tuning for improving code performance, or related topics.

Program Planning/Context:

 Includes recompete of ASCR's "base math" portfolio, broadened to computer science, but also broadened to fund academic collaborators.

Proposals under review.



SC Energy Earthshots

- The SC Energy Earthshots Initiative addresses key scientific challenges that underpin stretch goals for the first 6 DOE Energy Earthshots.
 - Closely coordinated with the Energy Technology Offices.
- SC announced 29 awards in FY 2023: 18 scientific foundations grants (~\$2-5M/award over 3 yrs) and 11 EERCs (recommended ~\$4.8M/yr/EERC).
- ASCR supports 8 Energy Earthshot Research Centers (EERCs).
 - An example: NREL leads the Floating Offshore Wind Modeling and Simulation (FLOWMAS)
 - Focuses on addressing DOE's Floating Offshore Wind Shot—reduce the levelized cost of energy of floating offshore wind by 70% by 2035
 - Overcome the fundamental mathematical and high-performance-computing challenges to predicting and understanding offshore wind energy systems.
 - Leverages ECP developed ExaWind software

Office of Science



Many projects will build on ECP investments!

https://www.energy.gov/policy /energy-earthshots-initiative



Photo by Brent Rice, NREL

https://science.osti.gov/



FY 2024 Lab Call: Microelectronics Science Research Center Projects For Energy Efficiency And Extreme Environments

- \$160 million lab funding opportunity announcement (5/8/2024) for Microelectronics Science Research Centers (MSRCs), anticipating 4-year awards.
- This funding partially implements Sec. 10731 Microelectronics Research for Energy Innovation (or Micro Act) from the CHIPS and Science Act.
- This DOE-led effort will complement the work of the broader interagency and especially Commerce, DOD, and NSF to advance microelectronics innovation.
- The centers will bring together teams led by our national labs to tackle fundamental science research in microelectronics with a focus on **energy efficiency and extreme environments**.
- These centers will be composed of networks of smaller awards, which will total no more than \$25 million per "center." Each project ceiling is \$3M/year, floor \$750k/year.
- Each awarded project will be assigned either to an MSRC focused on energy efficiency or an MSRC focused on extreme environments.
- Pre-proposals due May 30, pre-proposal response June 20, proposal submission due July 25.

Other FY 2024 ASCR Funding Opportunities – Released

- EPSCoR-State/National Laboratory Partnerships. Supports collaborations between EPSCoR-eligible institutions and DOE national laboratories on fundamental, early-stage energy science research. Proposals under review.
- Annual Early Career FOA: Supports outstanding scientists early in their careers in research areas supported by the Office of Science. Proposals under review.
- EXPRESS: 2024 Exploratory Research for Extreme-Scale Science: Supports single investigator or small team high risk and high reward projects in the following research topics: Harnessing Technology Innovations to Accelerate Science through Visualization, Scalable Space-Time Memories for Large Discrete/Agent-Based Models, Neuromorphic Computing, Advanced Wireless, Quantum Hardware Emulation. Proposals under review.
- Advancements in Artificial Intelligence for Science: Basic computer science and applied mathematics research in the fundamentals of AI for science. Proposals under review.
- Accelerated Research in Quantum Computing: Supports large cross-disciplinary teams that will advance computer science toward a
 software stack that is ready to leverage multiple quantum technologies, or will develop mathematical foundations, algorithms, and
 software tools toward quantum utility demonstration for applications within the DOE mission. Proposals under review.
- Data Reduction for Science: Research to explore potentially high-impact approaches in the development and use of data reduction techniques and algorithms to facilitate more efficient analysis and use of massive data sets produced by observations, experiments and simulation. Proposals under review.
- **RENEW, FAIR**: Supports efforts to broaden participation within the SC research community through training opportunities leveraging the DOE complex and direct research support at non-R1 MSIs and ERIs (incl. DOE lab and R1 MSI partners). **Pre-applications under review.**



Other FY 2024 ASCR Funding Opportunities – Released

- Scientific Discovery through Advanced Computing (SciDAC): Partnership in Electricity. Supports SciDAC-5 partnerships between electric-grid scientists/engineers and applied mathematicians and computer scientists and that fully exploit the capabilities of DOE High-Performance Computing (HPC) in order to accelerate and enhance Office of Electricity objectives. Pre-applications due on May 30, 2024.
- Microelectronics Science Research Center Projects for Energy Efficiency and Extreme Environments: See previous slide...

ASCR Participation in Solicitations Led by Other Offices and Agencies

- With DOE's Office of Energy Efficiency & Renewable Energy (EERE), Advanced Materials & Manufacturing Technologies Office (AMMTO): Data, Analysis, and Modeling Tools. Building on a history of collaboration between ASCR and EERE on their HPC4Mfg effort, ASCR is planning to co-fund this year with AMMTO projects selected in response to the "Codes for High Performance Computing for Manufacturing" topic. Proposals due June 11, 2024.
- With the National Science Foundation (NSF, NIH, and international partners): Collaborative Research in Computational Neuroscience (CRCNS). Basic collaborative research in computational neuroscience. Proposals under review.
- With the National Science Foundation (NSF): Correctness for Scientific Computing Systems (CS²). Basic collaborative research in correctness for scientific computing. Proposals due August 13, 2024.



ASCR-Funded Software Stewardship Organizations (SSOs)



COLABS	CORSA	FASTMATH	PESO
Training, workforce development, and building the RSE community	Partnering with foundations to provide sustainable pathways for scientific software	Stewardship, advancement, and integration for math and ML/AI packages	Stewarding, evolving and integrating a cohesive ecosystem for DOE software
RAPIDS	S4PST	STEP	SWAS

- ASCR provided phase I funding to six pilot SSOs in FY23 after an open, competitive review process.
- ASCR provided phase II funding to six SSOs for January 2024 after an additional round of proposal review.

The Consortium for the Advancement of Scientific Software



CASS Basics

- A newly-formed organization
- Sponsored by DOE Office of Advanced Scientific Computing Research (ASCR)
- Established by DOE Software Stewardship Organizations (SSOs)

CASS Goals

- Forum for SSO collaboration and coordination
- Bigger than the sum of its parts
- Vehicle for advancing the scientific software ecosystem

CASS Status

- Defining governance structure
- Establishing community awareness
- Building a team of teams
- Collaborating on outreach

Software Stewardship Organization (SSO) Basics

- Each SSO represents a specific software ecosystem concern
- Product SSOs: Programming systems, performance tools, math packages, data/viz packages
- Portfolio SSO: Curating & delivering software stack to the community
- Community SSOs: Workforce, partnerships

Engage with CASS

- Participate in June 11-13 CASS Community BOF Days: <u>https://cass.community/bofs</u>
- Visit <u>https://cass.community</u>

DOE stewards the leading edge high performance computing capability for the nation

- *Aurora*, at ANL debuted as the second fastest supercomputer in the world
- Frontier, at ORNL remains the world's fastest, but only until Aurora completes commissioning and "shake out"
- *El Capitan,* at LLNL will be the third exascale computer, supported by NNSA

Frontier Holds Top Supercomputer Spot, Aurora Closes in on Al Power

See who else made the biannual Top500 world's supercomputer list



Ben Wodecki, Jr. Editor May 15, 2024

④ 4 Min Read





The Aurora Supercomputer at Argonne National Laboratory broke the exascale barrier and took the top spot in Al performance benchmark

TOP500

- #2 Ranking on Top500
- Using only 87% of the system
- DOE's Frontier retains #1
- Crossed Exascale barrier with 1.012 EF/s

HPL-MxP (Al Benchmark)

- #1 Ranking fastest
 Al supercomputer
- Only 89% of the system used
 - 10.6EF/s

Aurora contains 63,744 graphic processing units making it the world's largest GPU-powered system.

International High Performance Computing conference

HPLMXP

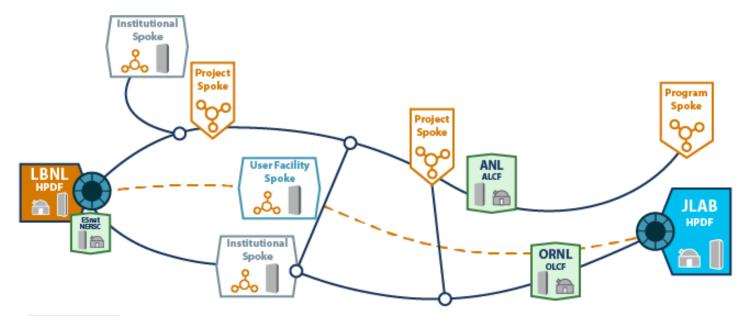
May 13, 2024

Hamburg, Germany

The High Performance Data Facility (HPDF) project

FY 2024 appropriations provided \$8.0M

- A DOE Order 413 project to create a new Office of Science User Facility
- JLab is the lead, with LBNL as partner; the project team is integrated with JLab and LBNL staff, with support from both DOE Site Offices
- HPDF will have a Hub-and-Spoke distributed operations model. Primary infrastructure will be at JLab with resilience site at LBNL



Conceptual representation of the HPDF Hub-and-Spoke operations model



The HPDF mission:

To enable and accelerate scientific discovery by delivering state-of-the-art data management infrastructure, capabilities, and tools

> HPDF will host a community webinar on June 11, 2024 at 2-3pm ET

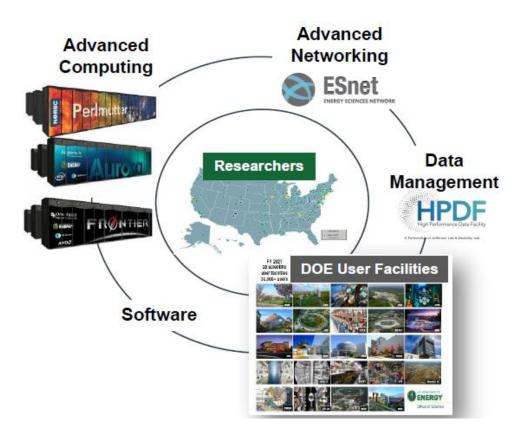




DOE's Integrated Research Infrastructure (IRI)

<complex-block><section-header><section-header><section-header>

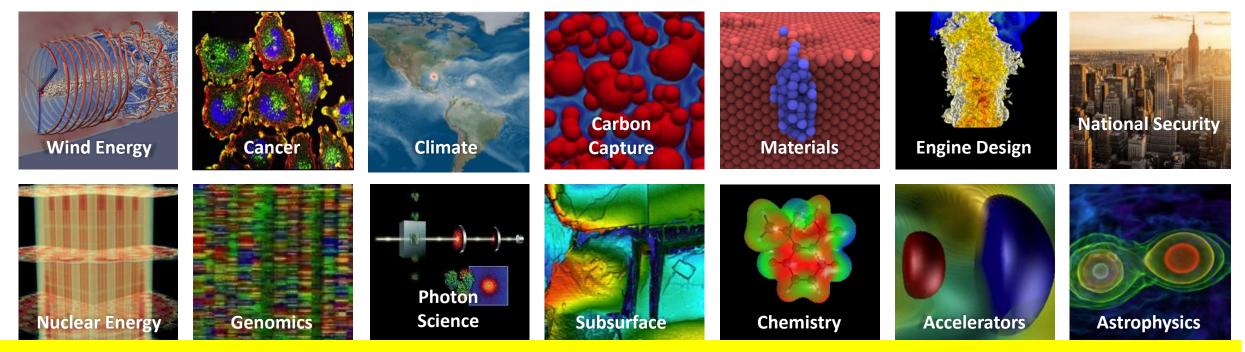
The DOE envisions a revolutionary ecosystem – the Integrated Research Infrastructure – to deliver seamless, secure interoperability across the DOE scientific user facilities, the National Laboratories, and other research infrastructure.



The ASCR Facilities and HPDF project team are collaborating to jump-start the IRI program governance, and will soon broaden to include other labs, user facilities, and stakeholders.

>> An update on progress is on the agenda today from 2:30-4:00pm ET

The breadth of exascale-ready applications is indicative of a sea change in computing abilities for DOE and the Nation



ECP CD-4 IPR was completed in April. The project will officially close out in July.





Energy.gov/science

A New Science Era: Exascale Today Enables the AI of Tomorrow

Long-term investments in applied mathematics and computer science enabled exascale.





Frontier and Aurora #1 and #2 on the Top500, **lead the** world in computational capability. Aurora is #1 in the world for Al capability, and both systems exceed 10 exaops of mixed precision performance.

The exascale and AI-enabled science era will lead to dramatic capabilities to predict extreme events and their impacts on the electric grid across weather and climate time scales...



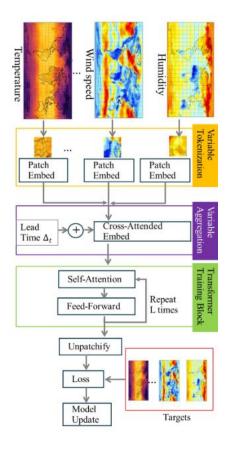
and will accelerate the design and deployment of clean-energy technologies to create a better future.



Energy.gov/science



Highlight: Exascale AI on Frontier Advancing Science



ORBIT: Oak Ridge Base Foundation Model for Earth System Predictability

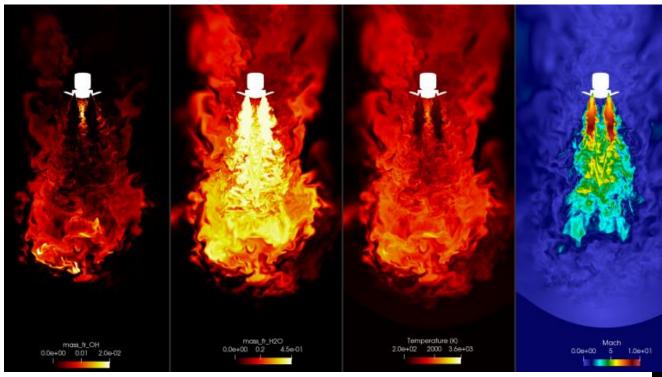
- An advanced vision-transformer model that scales up to 113 billion parameters.
- ORBIT achieves 230 to 707 PFLOPS, with scaling efficiency maintained at 78% to 96% across 24,576 AMD GPUs.
- Preprint posted April 24, 2024 arXiv:2404.14712

Optimizing Distributed Training on Frontier for Large Language Models

- Identified efficient strategies for training large LLMs of varying sizes through empirical analysis and hyperparameter tuning.
- In collaboration with AMD, developed software for AMD GPUs to enable highperformance at-scale AI training.
- For <u>1 trillion parameters</u>, achieved GPU throughputs of 31.96%, 100% weak scaling efficiency on 3072 MI250X GPUs, and strong scaling efficiency of 87%.
- Dash, et al. ISC 2024. DOI: 10.23919/ISC.2024.10528939

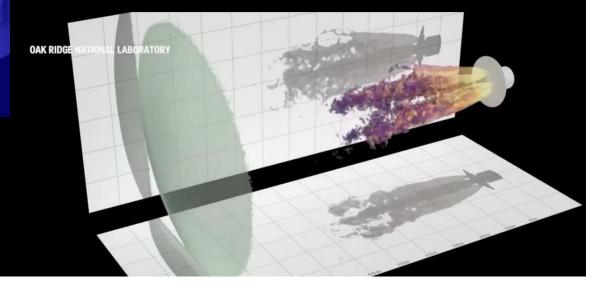
Highlight: NASA uses Frontier to simulate Mars landing

NASA's FUN3D computational fluid dynamics code utilized 90% of Frontier's GPUs to model human-scale landing



Credit: Gabriel Nastac/NASA

The simulation utilized a static Mach 1.4 solution on a mesh consisting of 33 billion elements using 33,880 GPUs. From left to right, contours show the mass fractions of the hydroxyl radical and H_2O , the temperature in Kelvin, and the local Mach number.





Overview of AI Executive Order : Safe, Secure, and Trustworthy Development and Use of AI

Issued October 30th, 2023 with directives to over 20 federal agencies, and 30-365 day deadlines.

DOE has a critical role to play in the highlighted sections:

- Guidelines, standards, best practices for AI safety and security
- Promoting innovation and competition
- Supporting workers
- Advancing Equity and Civil Rights
- Protecting consumers, patients, passengers, students
- Protecting privacy
- Federal Government use of AI
- Strengthening Leadership Abroad

ASCR has been leading SC's implementation of the Executive Order.



DOE contributions to the NAIRR Pilot

- OLCF Summit: To provide computational resources to the research community in support of President Biden's Executive Order, DOE extended operations of ORNL's Summit supercomputer through October 2024 to support the NAIRR Pilot and other high priority DOE research projects.
- ALCF AI Testbed: NAIRR Pilot users may apply to access the ALCF AI Testbed, a resource platform supporting AI and data-centric workloads on next-generation AI hardware.
- DOE national laboratory experts are sharing knowledge and participating in communitydriven AI software and AI user outreach activities.
- DOE is participating in the federal NAIRR Pilot Program Management Office.

DOE, with NIH, is co-leading the NAIRR Secure Pilot to inform design of NAIRR Secure:

- DOE is expanding its efforts in trustworthy and responsible AI research and will showcase ORNL's CITADEL framework as an example of research infrastructure to enable researchers to safely and securely process protected data at scale.
- DOE is expanding its research efforts under its Biopreparedness Research Virtual Environment (BRaVE) program with a new project that will leverage generative AI to create a high-quality set of synthetic pathology reports. This effort builds on a longstanding partnership between DOE, the National Institutes of Health, the U.S. Department of Veterans Affairs, and the Centers for Disease Control and Prevention.

More information is available at https://nairratdoe.ornl.gov/







FY 2025 Request Highlights - \$1,153M (+\$136.7M or 13.5% above FY 2024 Enacted)

- Advancing the Exascale and AI-Enabled Science Era: Critical basic research investments in applied mathematics and computer science to combine the power of exascale computing and artificial intelligence for a new science era.
- **Extending Frontiers in AI for Science, Security, and Technology:** Increased support for AI research that leverages the unique capabilities of the DOE ecosystem to expand responsible, trustworthy and secure application of AI technologies across DOE's broad mission needs.
- Next-Generation Facilities: Maintaining facility operations and existing upgrade projects and continued planning activities for NERSC-10 and the LCF upgrades. Increased support for the planned advancement of the High Performance Data Facility (HPDF).
- Leveraging Energy-Efficient Heterogeneous Architectures: Continued support in applied mathematics, computer science, networking, hardware and microelectronics research to advance and leverage energy-efficient emerging heterogeneous architectures including quantum.
- Integrated Research Infrastructure (IRI): Continued implementation efforts for the IRI initiative that will provide researchers with the ability to seamlessly and securely meld DOE's unique data, user facilities, and computing resources including the planned HPDF.
- **Strategic Partnerships:** Maintaining our partnerships within DOE and at the interagency level to accelerate scientific discovery and to support energy earthshots and national emergency and bio-preparedness efforts.
- **Broadening ASCR's impact** through increased investments in RENEW and FAIR, and continued support for Computational Science Graduate Fellowship and EPSCoR.

Office of

Science





Al enabled rapid threat detection, mitigation, response, and recovery in the power grid

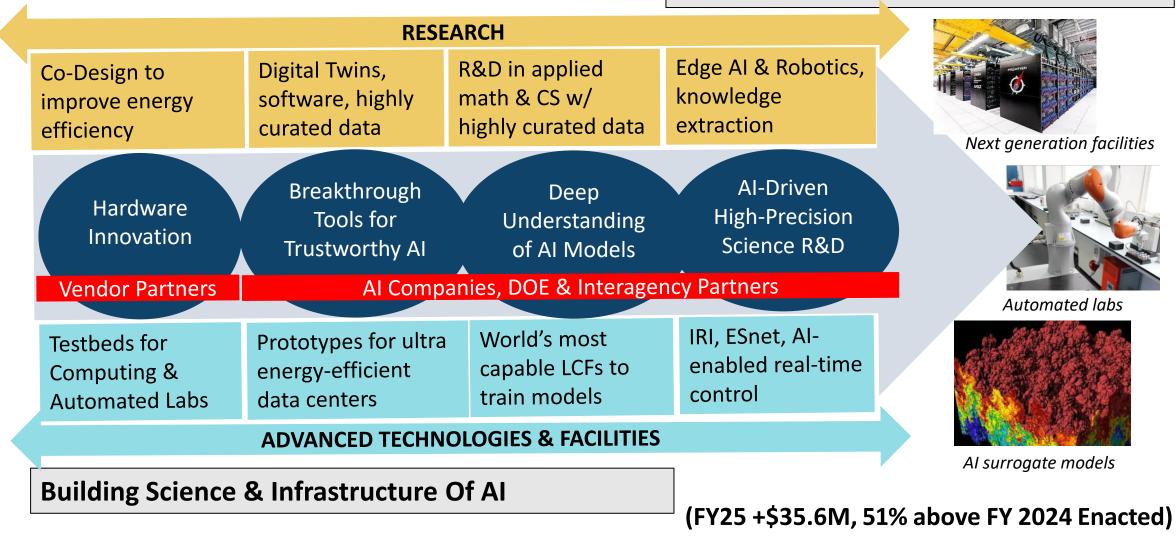


Quantum Testbeds

Energy.gov/science

ASCR's End-to-End Approach for Trustworthy and Energy-Efficient Frontier AI

Transforming Discovery Science & Infrastructure by AI





ASCR Community News

MEXT

- DOE/MEXT HPC project arrangement (MOU) signed by Deputy Secretary Turk and Japan's Minister of Education, Culture, Sports, Science and Technology, Moriyama Masahito, April 9, 2024. Enabling collaboration in HPC, AI, and related research areas. Updates the prior agreement from 2014.
- Research workshops in Summer 2024
- PI Meetings:
 - Applied Mathematics PI Meeting, January 2024, held in Albuquerque, NM.
 - Computer Science PI Meeting, February 2024, held in Atlanta, GA
 - Outreach to and participation from surrounding universities, including minority-serving and emerging research institutions, attended poster sessions and other meeting activities.



ASCR Office Hours

- Starting in March, ASCR started holding virtual office hours on the second Tuesday of the month, 2 PM ET
- Researchers, educators, and leaders within research administration from all institutional types are encouraged to join
- A primary goal of the virtual office hours is to broaden awareness of our programs; no prior history of funding from DOE is required to join
- Program managers are available to answer questions
- Past topics:
 - Tuesday, March 12, 2024, at 2pm ET Introduction to ASCR and its program mission and history
 - Tuesday, April 9, 2024, at 2pm ET Introduction to ASCR's Computer Science research program
 - Tuesday, May 14, 2024, at 2pm ET Introduction to ASCR's Applied Mathematics research program
- Upcoming topics include:
 - Tuesday, June 11, 2024, at 2pm ET Introduction to ASCR's user facilities and their allocation programs
 - Tuesday, July 9, 2024, at 2pm ET Overview of the ASCR research proposal and review process

See the ASCR website (<u>https://science.osti.gov/ascr/officehours</u>) for more information, including slides and recordings of past office hours.

