Accelerating scientific research through multidisciplinary collaboration in computation and machine learning

Esmond G. Ng (EGNg@lbl.gov)
Lawrence Berkeley National Laboratory

Putting things in context …

Putting things in context …

- I am a mathematician at Lawrence Berkeley National Laboratory (or simply Berkeley Lab).
  - My work is focused on the development of high-performance computational algorithms.

Putting things in context …

- I am a mathematician at Lawrence Berkeley National Laboratory (or simply Berkeley Lab).
  - My work is focused on the development of high-performance computational algorithms.

- Berkeley Lab is one of many R&D labs under the U.S. Department of Energy (DOE).
  - Others include, for example, Oak Ridge National Laboratory and Lawrence Livermore National Laboratory, …
- DOE is one of the many federal agencies that provide funding for supporting activities at R&D laboratories and academic institutions in the U.S.
- DOE has many "Offices", one of which is the Office of Science.
- The DOE Office of Science supports R&D in a diverse portfolio of science areas.

DOE Office of Science …

- Six major program offices.
  - Advanced Scientific Computing Research (ASCR)
  - Basic Energy Sciences (BES)
  - Biological and Environmental Research (BER)
  - Fusion Energy Sciences (FES)
  - High Energy Physics (HEP)
  - Nuclear Physics (NP)

DOE Office of Science …

- Six major program offices.
  - **Advanced Scientific Computing Research (ASCR)**
  - Basic Energy Sciences (BES)
  - Biological and Environmental Research (BER)
  - Fusion Energy Sciences (FES)
  - High Energy Physics (HEP)
  - Nuclear Physics (NP)

- Funds R&D in mathematics and computer science.

DOE Office of Science …

- Six major program offices.
  - **Advanced Scientific Computing Research (ASCR)**
  - Basic Energy Sciences (BES)
  - Biological and Environmental Research (BER)
  - Fusion Energy Sciences (FES)
  - High Energy Physics (HEP)
  - Nuclear Physics (NP)

- Funds R&D in mathematics and computer science.
- Supports almost all my work at Berkeley Lab.

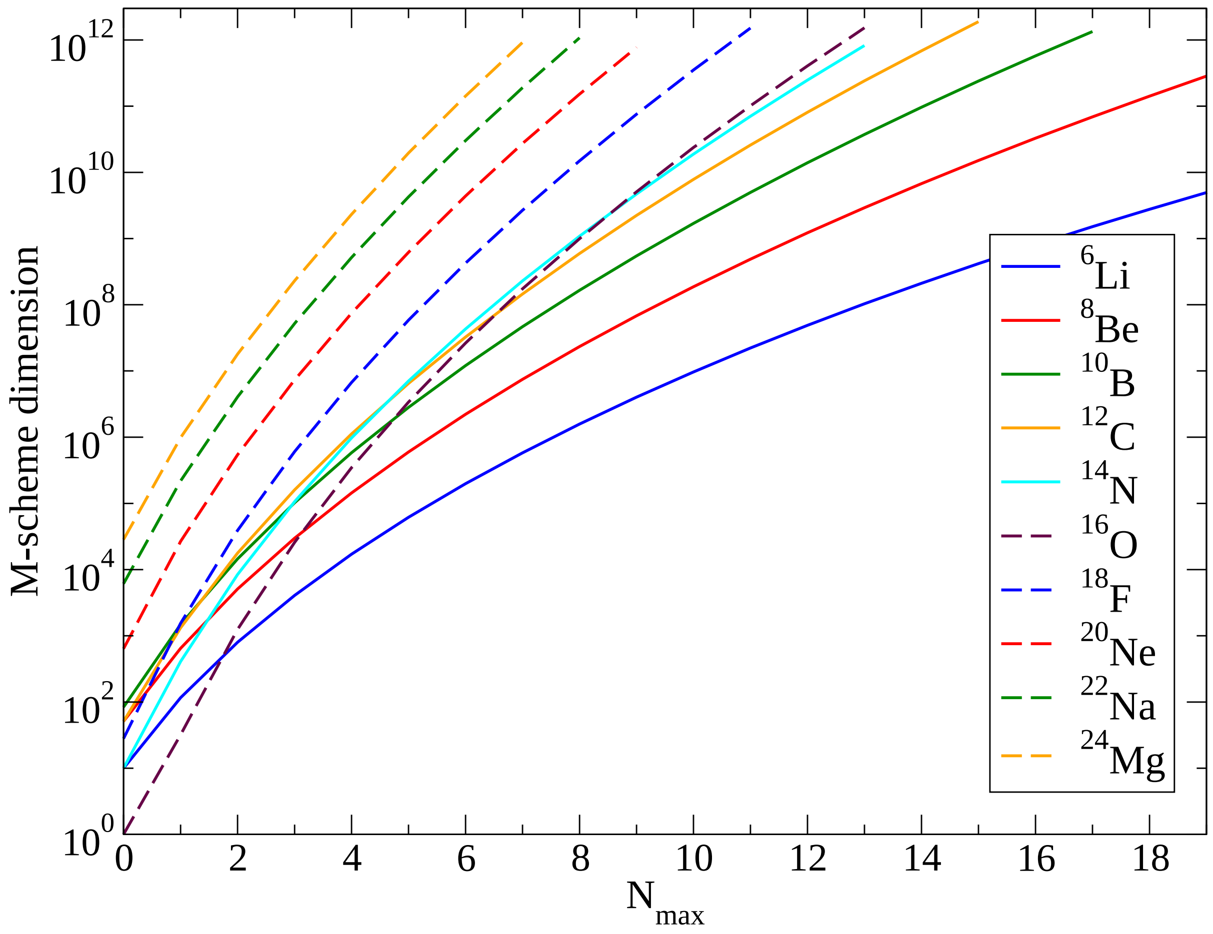
Role of computation in scientific research …

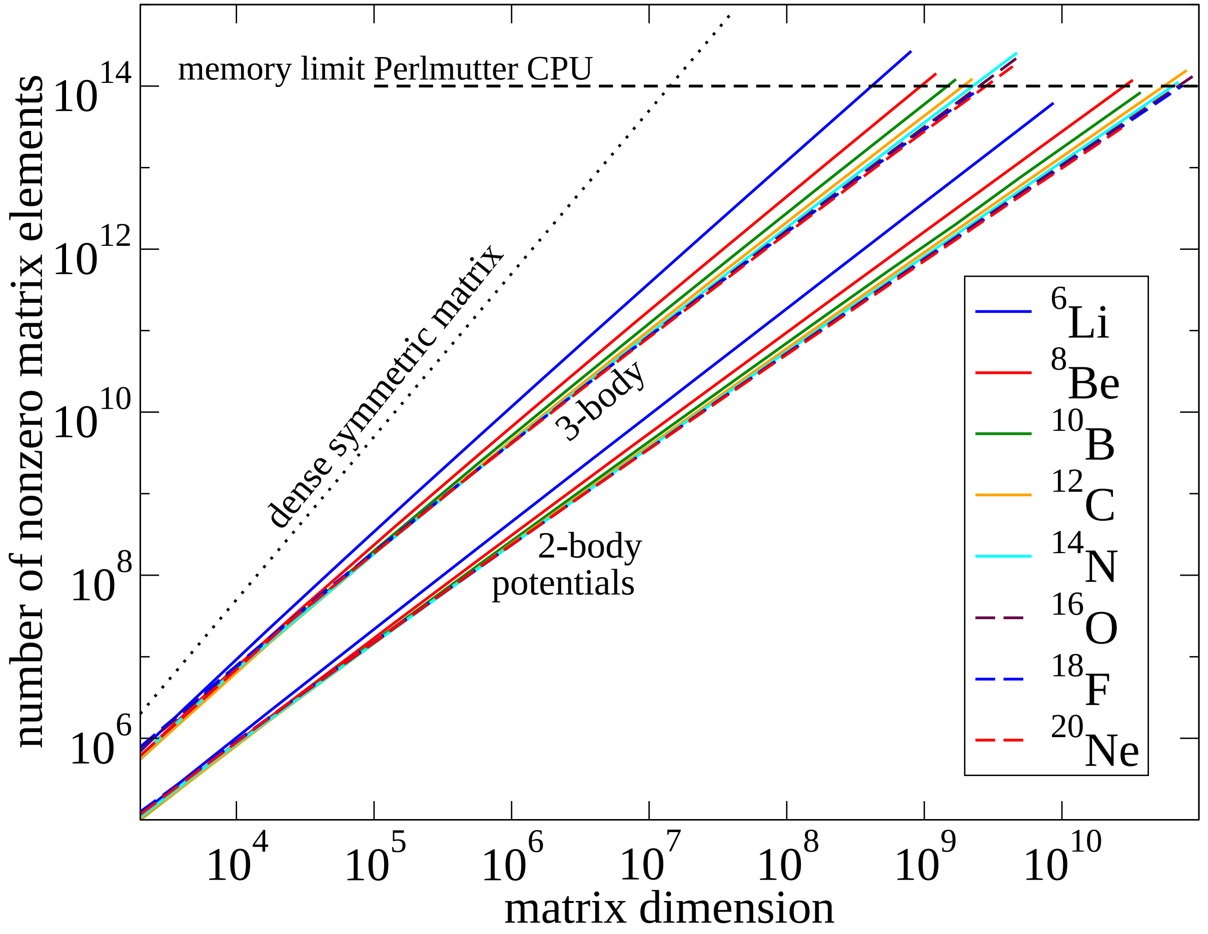
- It is undeniable by now that computation is an important approach for conducting scientific research.

- There are plenty of examples.
  - Number of conferences on computational sciences.
  - Also, the number of publications related to computational sciences.

- This conference, "Nuclear Theory in the Supercomputing Era", is a good example.
  - It started in 2012 …
  - This is the eighth (or seventh, depending on how you count) meeting.

Large-scale computation in nuclear theory can be resource demanding …





c/o Pieter Maris

Doing “big” science is not easy …

- Doing “big” science needs “big” investments.
  - People (i.e., efforts)
  - Software development
  - Computational resources (memory, time) – i.e., large-scale machines

Doing “big” science is not easy …

- Doing “big” science needs “big” investments.
  - People (i.e., efforts)
  - Software development
  - Computational resources (memory, time) – i.e., large-scale machines

- Focus on how DOE Office of Science accelerates large-scale computational science (including nuclear physics, of course).
- Caveats:
  - DOE Office of Science is **not** the only office that supports computational science.
  - And other U.S. funding agencies support computational science too.

DOE's efforts on advancing large-scale computational science …

- The DOE Office of Science has a long history of supporting computational science in basic research.

DOE's efforts on advancing large-scale computational science …

- The DOE Office of Science has a long history of supporting computational science in basic research.
- Individual science programs have funded computational projects, both large and small.
  - Basic Energy Sciences (BES)
    - Computational Chemical Sciences (CCS) Centers
    - Computational Materials Sciences (CMS) Centers
  - Biological and Environmental Research (BER)
    - Energy Exascale Earth System Model (E^3SM)

DOE's efforts on advancing large-scale computational science …

- The DOE Office of Science has a long history of supporting computational science in basic research.
- Individual science programs have funded computational projects, both large and small.
  - Basic Energy Sciences (BES)
    - Computational Chemical Sciences (CCS) Centers
    - Computational Materials Sciences (CMS) Centers
  - Biological and Environmental Research (BER)
    - Energy Exascale Earth System Model (E^3SM) Project
- The **Scientific Discovery Through Advanced Computing (SciDAC) Program** is probably the most important computational science program in the Office of Science at DOE.

The DOE Scientific Discovery Through Advanced Computing (SciDAC) Program …

- **The SciDAC Program mandated collaborations of applied mathematicians, computer scientists, and domain scientists to advance scientific frontiers through modeling, simulation and analysis.**
  - It is a 5-year program (most of the time and most of the projects).
  - The first program started in 2001.
  - It is now in its **fifth** cycle!
  - Focus on large-scale computational science using high-performance computing resources available today.
    - Every science project is required to take a multidisciplinary approach.

The DOE SciDAC Program ...

Program	Time Frame	Description	Result
SciDAC-1	2001 – 2006	Created scientific software infrastructure for parallel computing; Funded collaborations in DOE science domains	Science at the Terascale
SciDAC-2	2006 – 2011	Added DOE science domains; Enhanced university involvement; Outreach to broader scientific community; Added Data & Visualization	Science to the Petascale
SciDAC-3	2011 – 2016	Improved collaborations among the Institutes and between ASCR-SC programs; Enhanced architecture- and applications-awareness within each Institute; Added Uncertainty Quantification	Science on multi-core and emerging hybrid architectures
SciDAC-4	2017 – 2021	Outreach to broader scientific community; First connection to Applied Energy; Built-in flexibility; Added Machine Learning	Science on pre-exascale architectures
SciDAC-5	2020 – >2025	Continued outreach to broader scientific community including Applied Energy; Rise of Artificial Intelligence and Machine Learning; End of Exascale Computing Project	Science at the Exascale

c/o Ceren Susut-Bennett, ASCR Associate Director

The DOE SciDAC Program is consisted of two components …

The DOE SciDAC Program is consisted of two components …

- **Institutes**
  - Focus on R&D and software development in applied mathematics and computer science.
  - Two SciDAC Institutes at the moment:
    - **FASTMath**: computational mathematics and numerical libraries.
    - **RAPIDS2**: computer science, including performance analysis and data management.
  - Engagement with domain scientists in **Partnerships**.
  - *Funded **solely** by ASCR.*

The DOE SciDAC Program is consisted of two components …

- **Institutes**
  - Focus on R&D and software development in applied mathematics and computer science.
  - Two SciDAC Institutes at the moment:
    - **FASTMath**: computational mathematics and numerical libraries.
    - **RAPIDS2**: computer science, including performance analysis and data management.
  - Engagement with domain scientists in **Partnerships**.
  - *Funded **solely** by ASCR.*
- **Partnerships**
  - Focus on the solution of science problems on today's high-performance computers.
  - Multidisciplinary collaborations in applied mathematics, computer science, and domain science are mandatory.
  - *Funded **jointly** by ASCR and a science program office* (e.g., the Nuclear Physics Program Office under the Office of Science).

The DOE SciDAC Program is consisted of two components …

Institutes

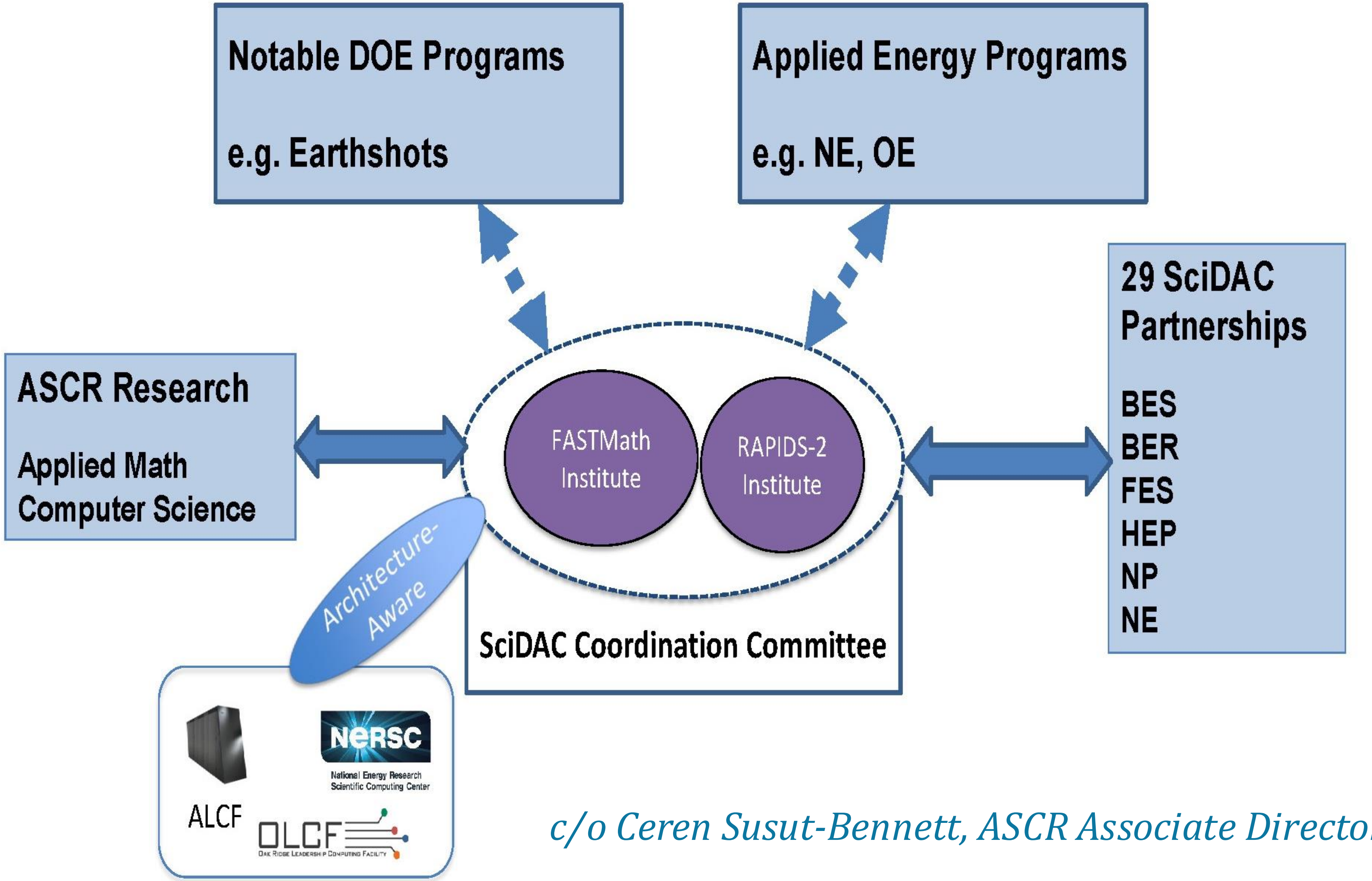
- Focus on R&D and software development in applied mathematics and computer science.
- Two SciDAC Institutes at the moment:
  - **FASTMath**: computational mathematics and numerical libraries.
  - **RAPIDS2**: computer science, including performance analysis and data management.
- Engagement with domain scientists in **Partnerships**.
- *Funded **solely** by ASCR.*

NUCLEI
Nuclear Computational Low-Energy Initiative

Partnerships

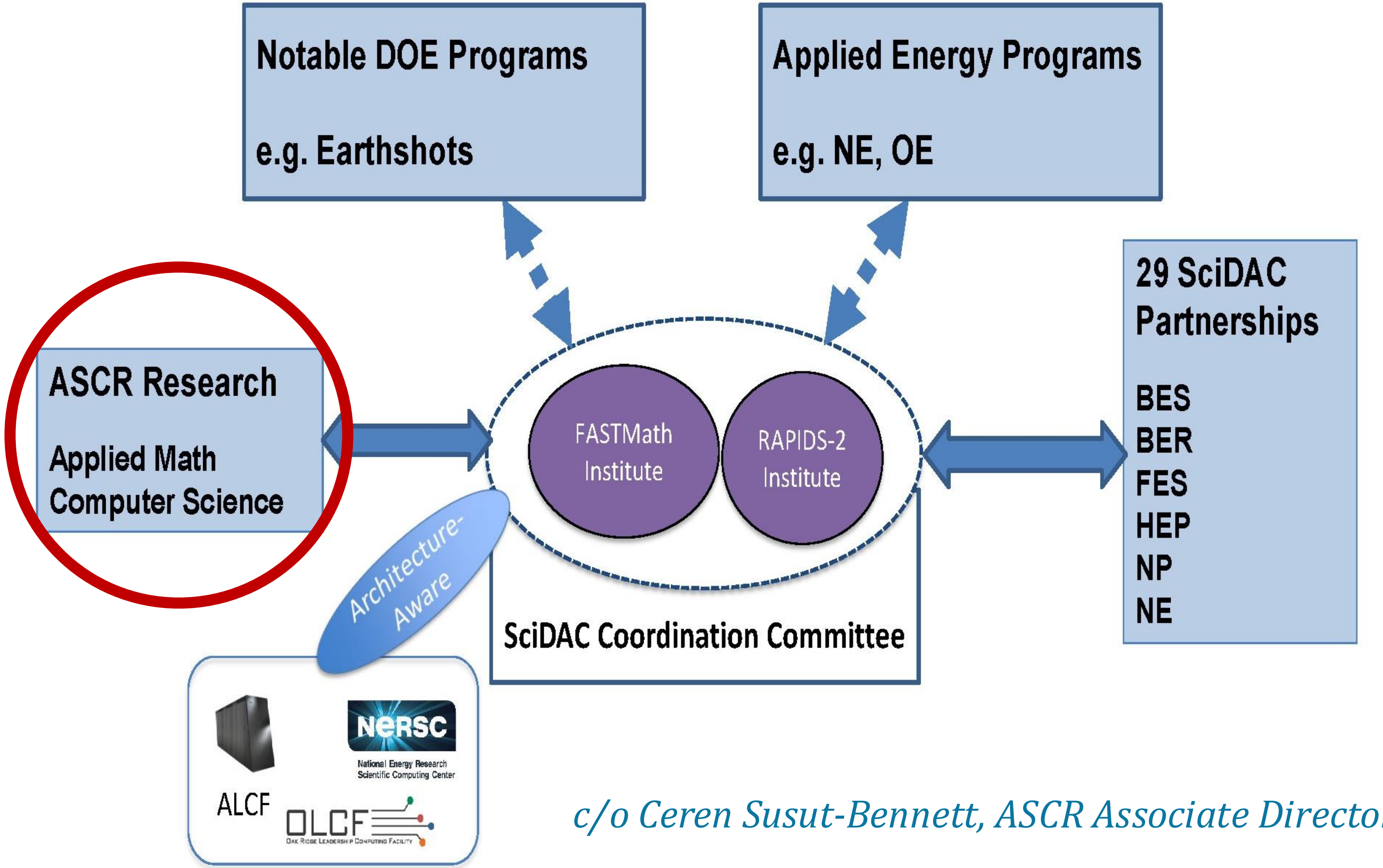
- Focus on the solution of science problems on today's high-performance computers.
- Multidisciplinary collaborations in applied mathematics, computer science, and domain science are mandatory.
- *Funded **jointly** by ASCR and a science program office* (e.g., the Nuclear Physics Program Office under the Office of Science).

The SciDAC model (as envisioned by DOE) …



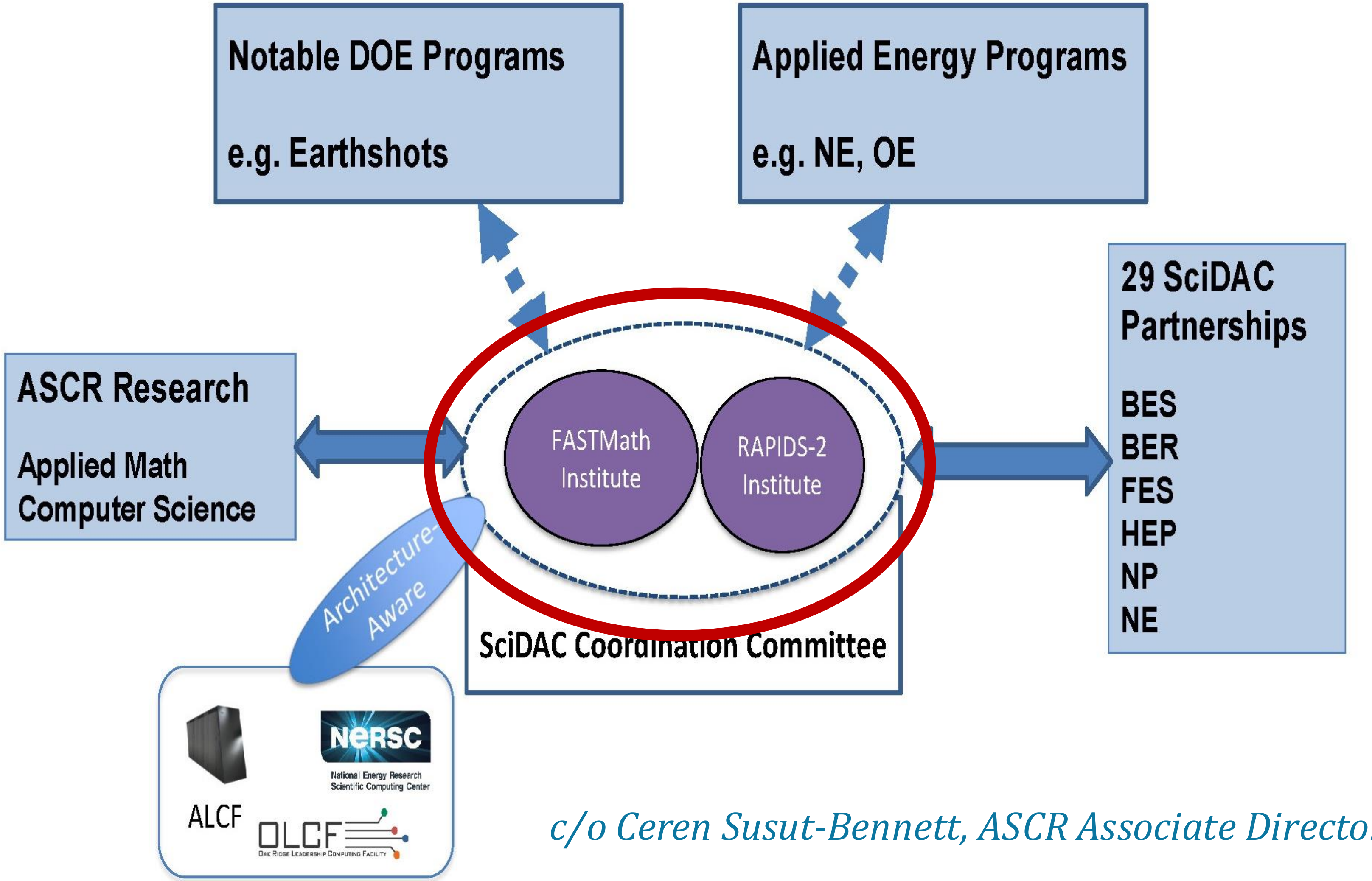
c/o Ceren Susut-Bennett, ASCR Associate Director

The SciDAC model (as envisioned by DOE) …



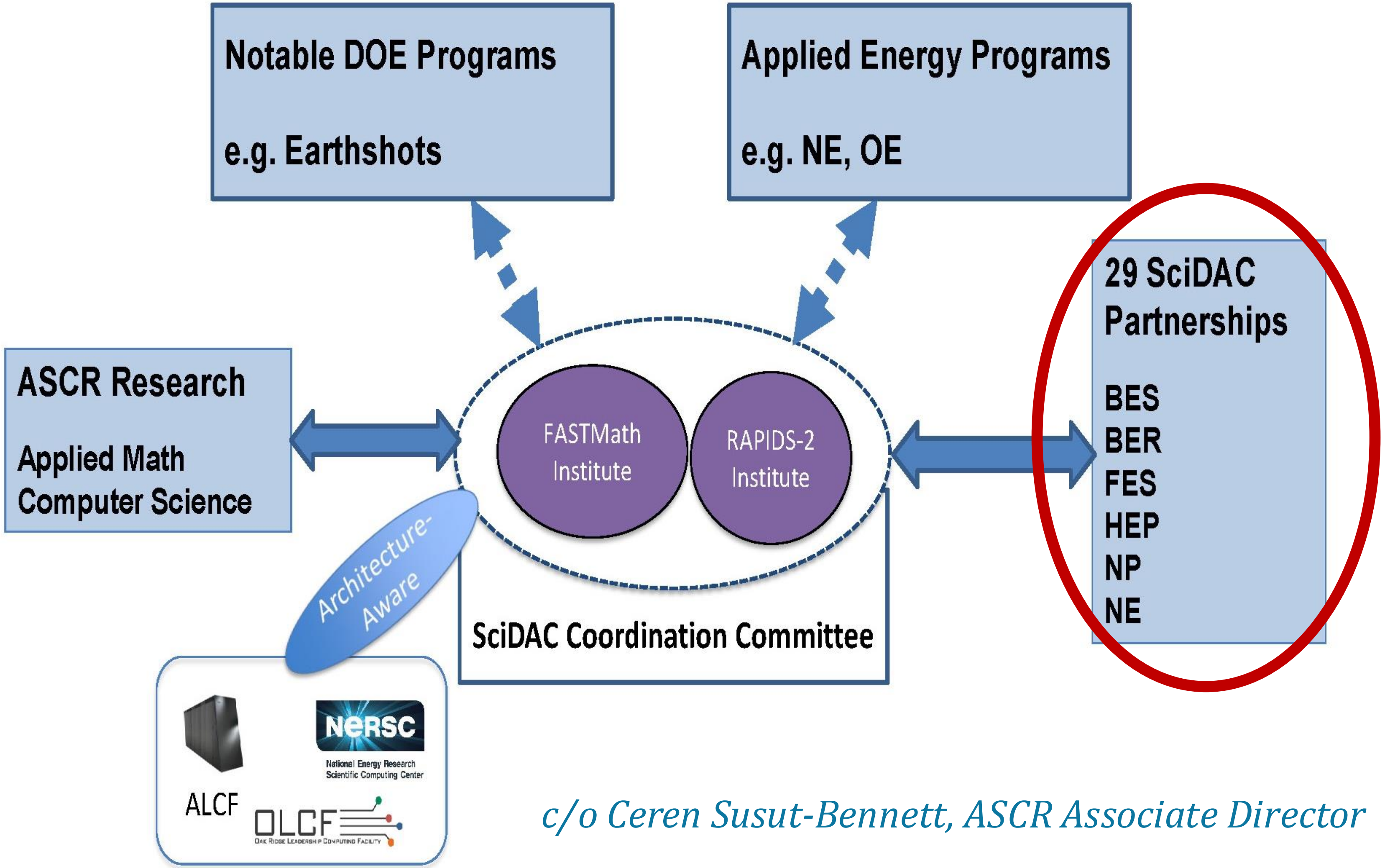
c/o Ceren Susut-Bennett, ASCR Associate Director

The SciDAC model (as envisioned by DOE) …



c/o Ceren Susut-Bennett, ASCR Associate Director

The SciDAC model (as envisioned by DOE) …

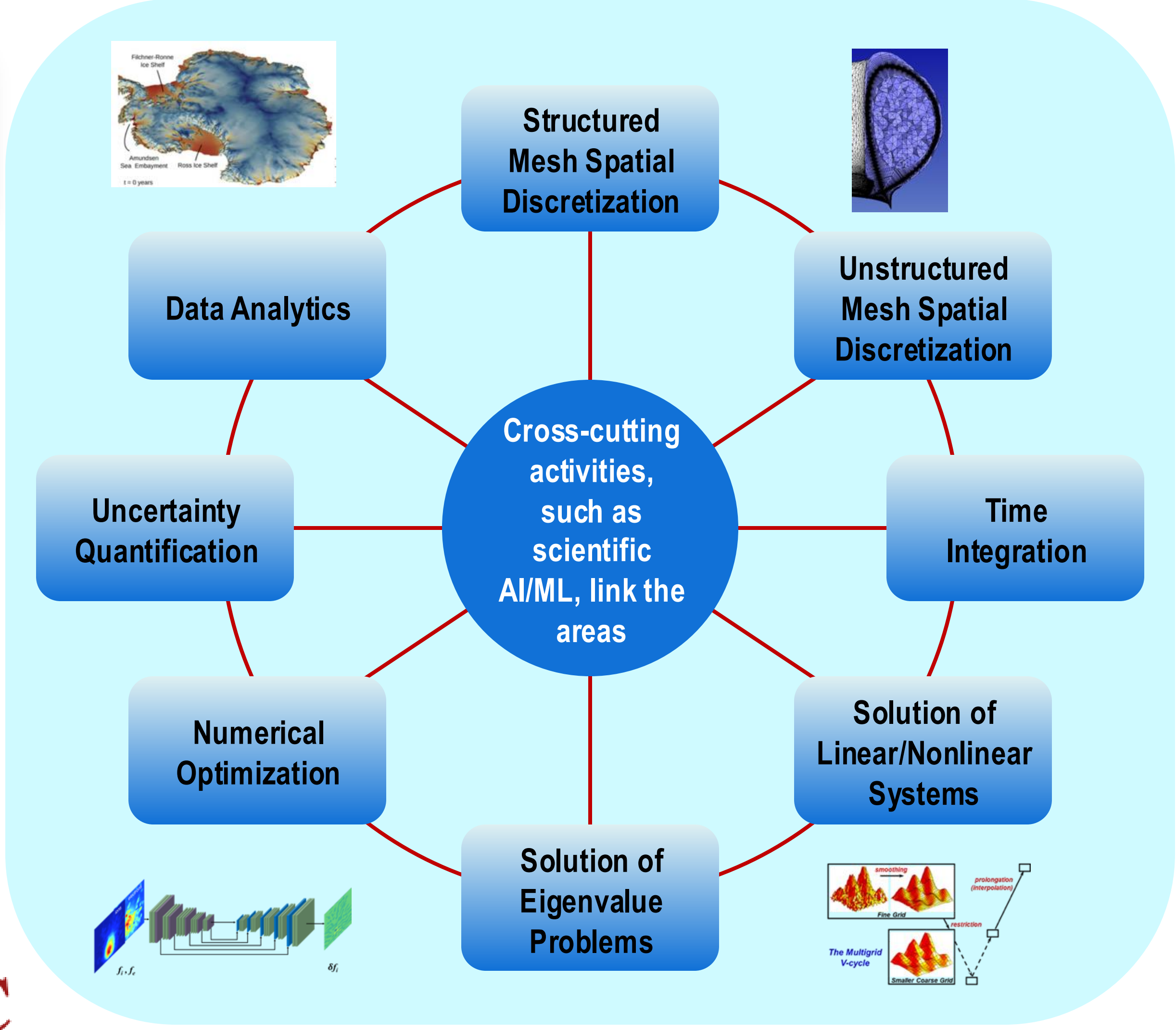


c/o Ceren Susut-Bennett, ASCR Associate Director

FASTMath – SciDAC Institute on Frameworks, Algorithms and Scalable Technologies for Mathematics

Provide leading edge applied and computational mathematics to remove the barriers facing computational scientists

- Develop robust math techniques and numerical algorithms for DOE science problems
- Deliver highly performant software with strong software engineering to run efficiently and scalably on DOE supercomputers
- Work closely with domain scientists to leverage our math and ML expertise and deploy our software in large-scale scientific codes
- Build and support the broader computational math and computational science communities across DOE



scidac5-fastmath.lbl.gov

Director: Esmond G. Ng (LBNL)
Deputy: Todd Munson (ANL)

The RAPIDS(2) Institute

RAPIDS

Enabling scientific breakthroughs using DOE supercomputers by assisting SciDAC and DOE scientists and engineers to solve computer science, data, and artificial intelligence (AI) challenges.

- **Five primary thrust areas, including engagement (top), covering key needs**
- **New emphasis on AI: enabling use by science teams and leveraging within RAPIDS tools**
- **Coordination with facilities and ECP**

Application, Institute, Facility, and Community Engagement

Application enhancement, RAPIDS/FASTMath co-design, tutorials

Data Understanding	Platform Readiness	Scientific Data Management	Artificial Intelligence
• Ensemble analysis • Feature detection • Production vis. • In situ analysis	• Heterogeneous programming • Performance modeling and analysis • Autotuning • Correctness	• Storage systems and I/O • Knowledge management • Workflow automation	• Representation learning • Surrogate modeling • Automation

Integration of AI within RAPIDS technologies

Argonne NATIONAL LABORATORY · BERKELEY LAB · UNIVERSITY OF DELAWARE · Lawrence Livermore National Laboratory · OAK RIDGE National Laboratory · RUTGERS · BROOKHAVEN NATIONAL LABORATORY · Kitware · Los Alamos NATIONAL LABORATORY EST.1943 · Northwestern University · THE OHIO STATE UNIVERSITY · UF · UNIVERSITY OF OREGON

Large-scale computational science needs large computing resources …

- So, DOE Office of Science has programs that provide research funds for R&D in computational science.
  - These funds are solely for supporting people and software development.
  - It **does not** provide funds for procuring high-performance computers. It also **does not** provide machine cycles for the computations.

Large-scale computational science needs large computing resources …

- So, DOE Office of Science has programs that provide research funds for R&D in computational science.
  - These funds are solely for supporting people and software development.
  - It **does not** provide funds for procuring high-performance computers. It also **does not** provide machine cycles for the computations.

- This is because DOE has several high-performance computing facilities, which have their own time allocation processes.

- For the DOE Office of Science:
  - National Energy Research Scientific Computing Center (NERSC) … a national facility
  - Argonne Leadership Computing Facility (ALCF)
  - Oak Ridge Leadership Computing Facility (OLCF)

The National Energy Research Scientific Computing Center (NERSC) …

- One of the most scientifically productive high-performance computing (HPC) centers in the world:
  - Provide open computing resources for general, unclassified computation.
  - ~10,000 users in ~1,000 projects.
  - Resulting in a large number of publications in refereed journals.
  - Several Nobel laureates have performed calculations on the NERSC machines.
  - Current system:
    - Perlmutter – HPE Cray EX 235n, AMD, NVIDIA (888,832 cores, 113 PF ; 1,536 TB on CPUs).
      - #19 on TOP500 (November 2024).

Leadership Computing Facilities (LCF's) …

- The LCF's are providing resources for high-end computational projects that require extreme scale computational resources.

  - LCF at Argonne National Laboratory (ALCF).
    - Aurora: HPE Cray EX – Intel Xeon CPU, Intel GPUs (9,264,128 cores; 1,012 PF ; >7 PB).
      - #3 on TOP500 (November 2024).

  - LCF Oak Ridge National Laboratory (OLCF).
    - Frontier: HPE Cray EX, AMD CPUs, AMD GPUs (8,699,904 cores; 1,206 PF ; 4.5 PB).
      - #2 on TOP500 (November 2024).

Getting time on the DOE Office of Science high-performance computers …

- **NERSC**
  - Managed by the Office of Science.
  - Open allocation applications every year.
    - Modest amount of time for code development, testing, and production runs.
- **DOE Innovative and Novel Computational Impact on Theory and Experiment (INCITE)**
  - Managed entirely by ALCF and OLCF.
  - Open call each year, where allocation decisions are made based on scientific and computational merit. The projects are also required to use a large percentage of the machine (no small runs). They can be for up to three years.
- **ASCR Leadership Computing Challenge (ALCC)**
  - Managed by ASCR.
  - Geared towards DOE mission science. Allocation decisions are made by DOE program managers based off merit and what supports their scientific portfolios. Only one-year duration.

DOE Innovative and Novel Computational Impact on Theory and Experiment Program (INCITE) ...

- Aims to accelerate scientific discoveries and technological innovations by awarding time on supercomputers to researchers with large-scale, computationally intensive projects that address "grand challenges" in science and engineering.
- Only available on ALCF and OLCF.
- Decisions primarily made by ALCF and OLCF.

Resource	Aurora (ALCF)	Frontier (OLCF)
Number of projects	24	59
Smallest Project (node-hrs)	50,000	175,000
Largest Project (node-hrs)	1,200,000	1,000,000
Average Project (node-hrs)	382,500	656,864
Median Project (node-hrs)	375,000	700,000
Total Awards (node-hrs)	9,180,000	38,755,000

Physics
Aurora: 43.85%
Frontier: 33.00%

DOE Leadership Computing Challenge Program (ALCC) ...

- Emphasis on high-risk, high-payoff scientific campaigns enabled via high-performance computing (HPC) in areas directly related to the DOE mission, that respond to national emergencies, or that broaden the community of researchers capable of using leadership computing resources.
- Decisions made by DOE Office of Science program managers.

Resource	Perlmutter-CPU (NERSC)	Perlmutter-GPU (NERSC)	Aurora (ALCF)	Frontier (OLCF)
Number of projects	11	10	13	29
Smallest Project (node-hrs)	30,000	50,000	8,960	15,300
Largest Project (node-hrs)	500,000	280,000	2,000,000	3,846,000
Average Project (node-hrs)	212,201	125,664	381,122	733,868
Median Project (node-hrs)	200,000	110,000	145,628	425,000
Total Awards (node-hrs)	2,334,208	1,256,640	4,954,588	21,282,186

Physics
NERSC-CPU: 31.72%
NERSC-GPU: 5.17%
Aurora: 32.29%
Frontier: 24.84%

Challenges of using these large-scale high-performance computer systems …

- Complexity in architectures.
  - Heterogeneous: CPUs, multicores, GPUs.
    - Different computational rates – work scheduling.
    - Mixture of parallel program constructs – MPI, OpenMP, CUDA, + many others.
    - Load balancing.
  - Memory hierarchy.
    - Memory on CPUs and GPUs.
- Communication overhead.
  - High communication-to-computation ratio.
    - Runtime could be dominated by communication.
- Primary reason why SciDAC, which enforces multidisciplinary collaborations of mathematicians, computer science, and domain scientists, is considered by DOE to be the right way to accelerate science using advanced computing.

Advancing computational science using machine learning …

- Based on Wikipedia …

  Machine learning (ML) is about the development and study of statistical and numerical optimization algorithms (models) that can learn from data and generalize to unseen data, and thus perform tasks without explicit instructions.

- ML approach provides a **predictive** tool for finding **possible** solutions to very hard problems that don't have closed form solutions or require way too much computation to solve.

Advancing computational science using machine learning …

- Based on Wikipedia …

  Machine learning (ML) is about the development and study of statistical and numerical optimization algorithms (models) that can learn from data and generalize to unseen data, and thus perform tasks without explicit instructions.

- ML approach provides a **predictive** tool for finding **possible** solutions to very hard problems that don't have closed form solutions or require way too much computation to solve.

- There have been some successes …
  - This conference also features work using ML.
  - This year's Nobel Prizes in Chemistry and Physics.

- **Lot of interests in using ML in many SciDAC Partnership projects.**

Challenges in using machine learning in scientific research …

- Someone has to design ML model.
  - Physics-informed ML seems to be the right approach and has demonstrated successes.
- Need to provide useful/reliable data that can be used in training.
  - Modeling and simulation will never go away.
- Understand what the model does is important.
  - Being a statistical approach, statistical errors are expected, so care is needed in the predicted outcome.
- Being able to explain the predicted outcome is important.
- May not tell you why or give you insights.

Take-home messages …

- Conducting large-scale computational science requires a significant infrastructure.
  - Funding for supporting people.
  - Skill for software development.
  - Access to large-scale computing platforms.
- Collaboration with mathematicians and computer scientists is considered important in getting the most out of today's computers, which have growing complexity in architectures.
  - More scalable, more robust, more accurate numerical algorithms.
  - More optimized codes.
  - Perhaps even alternative approaches.
- Use of machine learning has become fashionable.
  - More understanding of how ML works and how to interpret results is needed.
  - Not perfect at the moment but may have potential.
- **Questions/Comments: Send email to EGNg@lbl.gov**