RAPIDS: The SciDAC Institute for Computer Science and Data

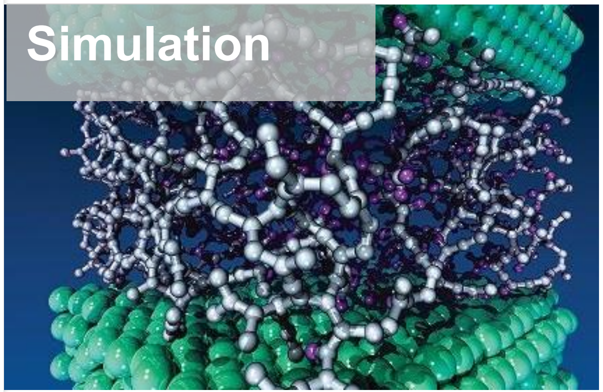
ROBERT ROSS
Institute Director
Mathematics and Computer Science Division
Argonne National Laboratory
rross@mcs.anl.gov

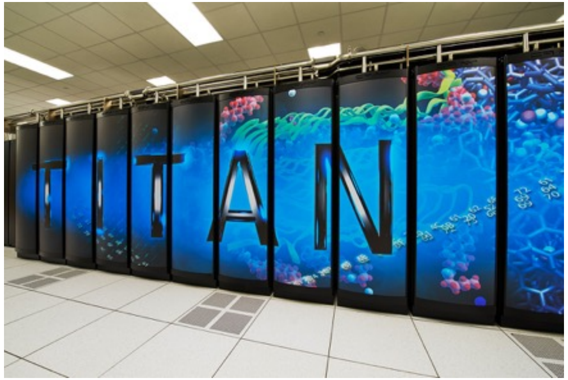
LENNY OLIKER
Deputy Director
Computational Research Division
Lawrence Berkeley National Laboratory
loliker@lbl.gov

U.S. DEPARTMENT OF ENERGY | Office of Science

Diverse Science and Systems

RAPIDS





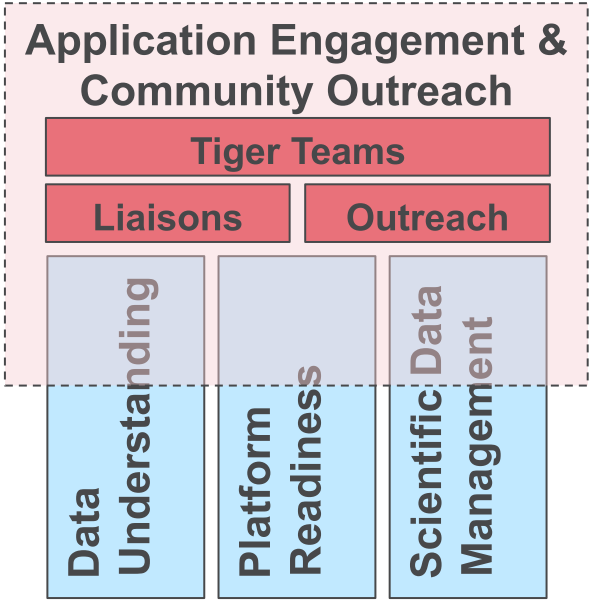
Top image credit B. Helland (ASCR). Bottom left and right images credit ALCF. Bottom center image credit OLCF.

The RAPIDS Institute

Resource and Application Productivity through computation, Information, and Data Science

Objective: Solve computer science and data challenges for Office of Science application teams to achieve science breakthroughs on DOE platforms.

- **Application Engagement**
  - *Tiger Teams* engage experts in multiple technology areas to work with science teams and codes
  - Outreach activities connect with broader community
- **Technology Focus Areas**
  - **Data Understanding** – scalable methods, robust infrastructure, machine learning
  - **Platform Readiness** – hybrid programming, deep memory hierarchy, autotuning, correctness
  - **Scientific Data Management** – I/O libraries, coupling, knowledge management



The RAPIDS Team

RAPIDS

Director	Deputy
Rob Ross	Lenny Oliker

Application Engagement	Data Understanding	Platform Readiness	Data Management	Outreach
Anshu Dubey Sam Williams	Wes Bethel P. Balaprakash	Jeff Vetter Paul Hovland	Scott Klasky John Wu	David Bernholdt

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

RAPIDS Focus Areas

Application Engagement & Community Outreach

Tiger Teams, Liaisons, and Outreach

Data Understanding

- Scalable methods
- Robust infrastructure
- Machine learning

Platform Readiness

- Roofline modeling
- Hybrid programming
- Deep mem. hierarchy
- Autotuning
- Correctness

Scientific Data Management

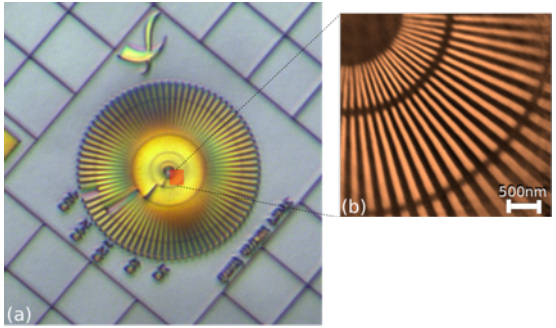
- I/O libraries
- Coupling
- Knowledge management

TECHNOLOGIES

Data Understanding

Facilitate understanding of large and complex science data through robust and scalable analysis methods, including learning approaches.

- **Tools:** AutoMOMML, DIY, GraphBLAS, Paraview and Catalyst, SENSEI, VisIt and Libsim, VTK
- **Advanced methods**: Stochastic flow maps, Lagrangian coherent structures, topological methods
- **Scalable infrastructure**: service-oriented data analysis and reduction, co-analysis with performance data
- **Learning approaches**: domain-specific applications of deep learning, predictive performance models, data- and model-parallel training



(a) Optical image of target; (b) ptychographic image generated from APS data using GPU at rate exceeding data acquisition.

Contacts: Wes Bethel (LBNL) lead, Prasanna Balaprakash (ANL) co-lead

Increasing Temporal Fidelity of Data Analysis for XGC1

- XGC1 fusion simulations use a particle representation, with many billions of particles in the largest simulations.
- Transforming from particles into a vector field reduces data size while allowing for bounded error.
- More timesteps can be captured, increasing temporal fidelity of analysis.
- Poincarè plot generated from the binned vector field data exhibits approx. 1% error using 89 MBytes of vector field data, as opposed to 500 GBytes of particle data.

  Reduced data is shown in black, original resolution data is shown in blue.

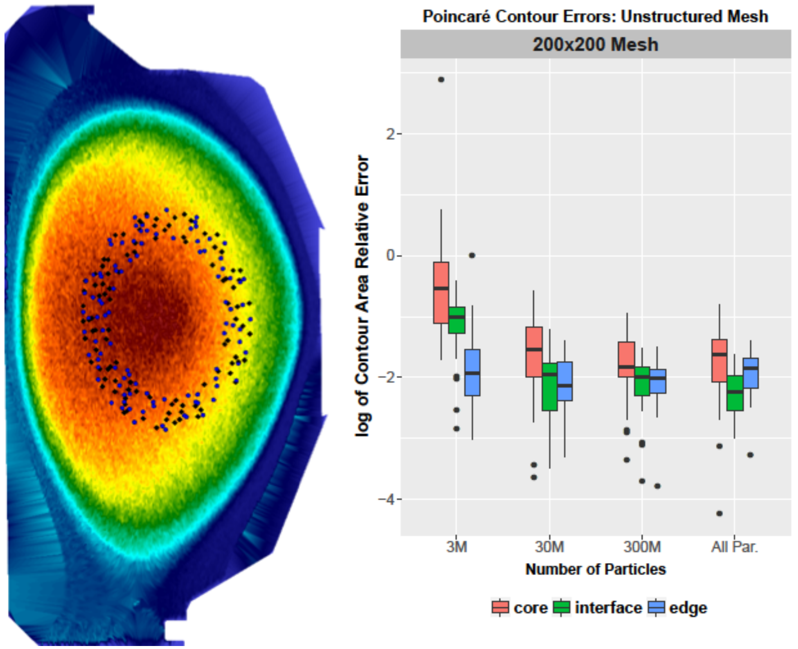
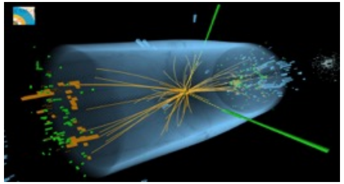


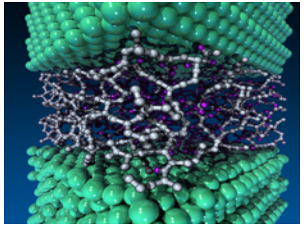
Image credit James Kress (ORNL).

Learning Approaches for Scientific Discovery

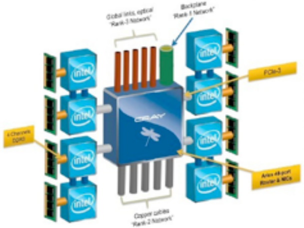
RAPIDS



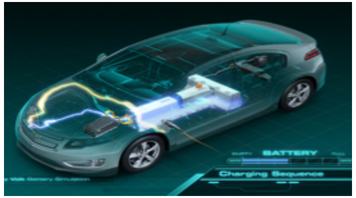
Deep learning for object identification and image classification for **LHC** and **APS**



Iterative learning and model-based sampling for **molecular dynamics simulations**



Performance, power, energy modeling of novel **HPC architectures**

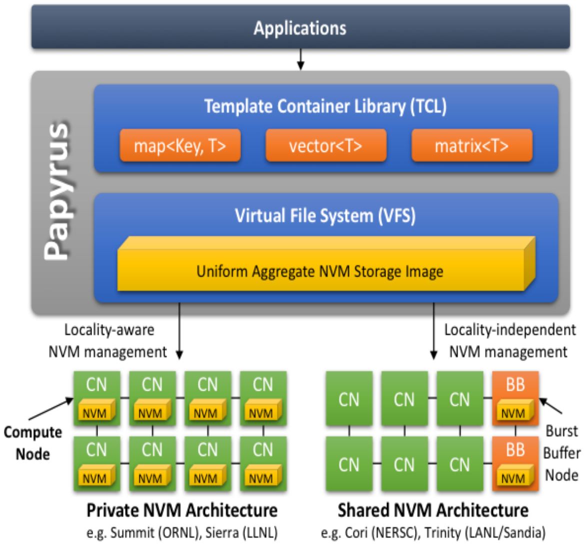


Machine learning for **urban mobility planning and vehicle technology assessment**

Platform Readiness

Preparing scientific codes for current and upcoming system through application of best-in-class expertise and tools.

- **Tools:** CHiLL, CIVL, IMPACC compiler, Intel Advisor (Roofline), OpenARC, Orio, Papyrus, ROSE, SCR, TAU
- **Performance modeling and analysis** for identifying optimization opportunities
- **Code generation and autotuning** for computation and communication
- **Portable programming** for heterogeneous and many-core systems, deep memory hierarchies
- **Correctness** of programs (e.g., when moving to new platforms)



Papyrus provides abstractions for large shared data structures using map, vector, and matrix modalities

Contacts: Jeff Vetter (ORNL) lead, Paul Hovland (ANL) co-lead

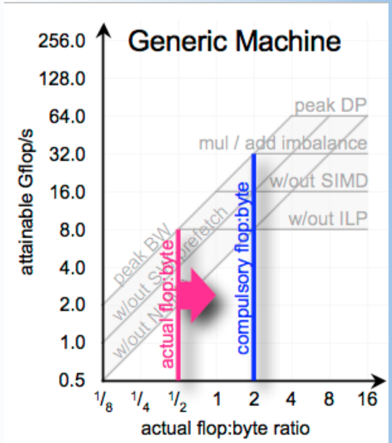
Roofline Performance Modeling

RAPIDS

ASCR Base & LDRD

Developed Roofline concept
2006-2011:

- Easy-to-understand, visual performance model
- Offers insights to programmers and architects on improving parallel software and hardware.

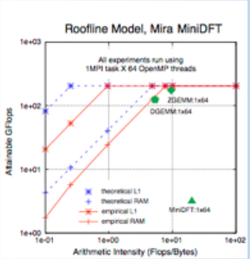


Proof of concept successfully applied to numerous computational kernels and emerging computing systems.

SciDAC3 Development

Roofline augmentation under SciDAC3
2013-2017:

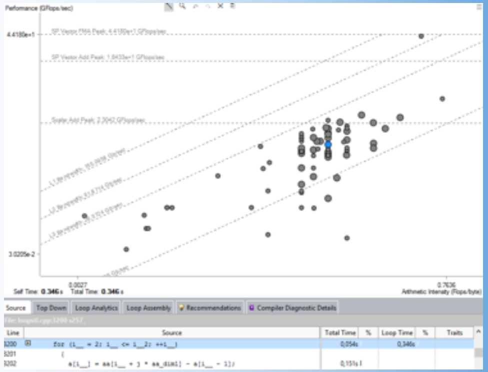
- Collaboration with FASTMath SciDAC Institute
- Developed Empirical Roofline Toolkit (ERT) with public release 03/2015, with Roofline Visualizer
- Created community tool for automatic hardware introspection and analysis



Automated Roofline code used to diagnose performance problems for DOE and SciDAC codes.

Outcome & Impact

- Roofline has become a broadly used performance modeling methodology across DOE
- Intel has embraced the approach and integrated it into its production Intel® Advisor
- **Collaboration with NERSC to instrument and analyze execution of real applications on machines such as Edison and Cori**

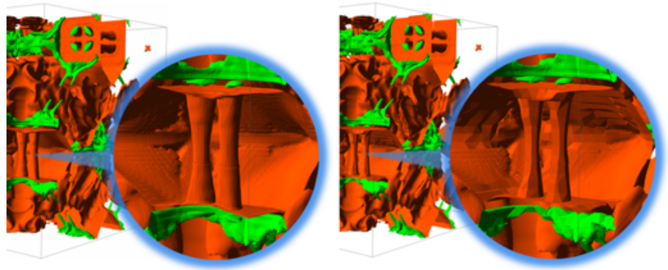


Snapshot of existing Intel Roofline tool in practice.

Scientific Data Management

Assist teams in storing, accessing, managing, and sharing large volumes of scientific data at DOE compute facilities.

- **Tools:** ADIOS, Darshan, DataSpaces, FastBit, HDF5, Parallel netCDF, TAU
- **I/O libraries:** Efficiently interfacing to complex and varied storage resources in HPC platforms
- **Coupling codes:** Effective communication and data flow for (near-)real-time science workflows
- **Knowledge management:** Capturing provenance of data, indexing and feature identification, facilitating sharing within and among teams



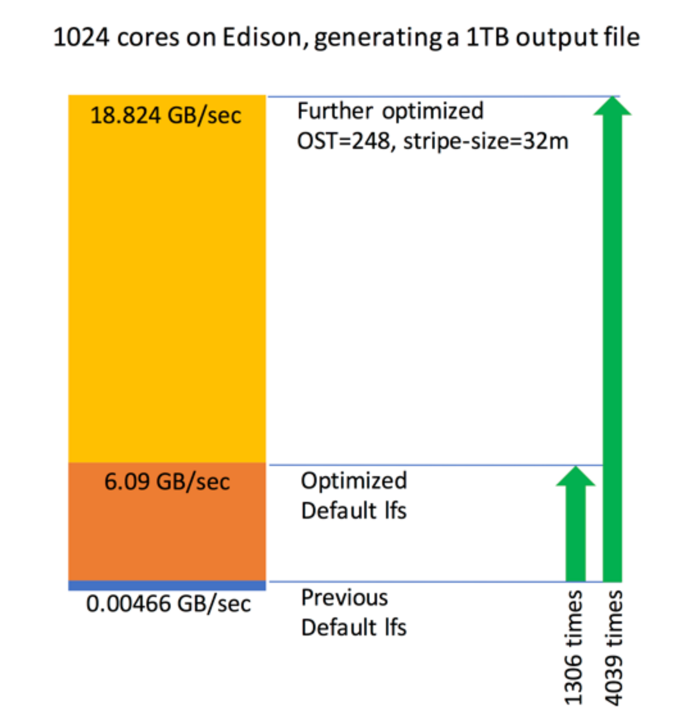
Polytropic gas simulation for modeling tokomak edge plasma. Data automatically translated from full resolution (left) to reduced resolution (right) to meet memory constraint.

Contacts: Scott Klasky (ORNL) lead, John Wu (LBNL) co-lead

Accelerating I/O for WARP

RAPIDS

- **Warp** is an open-source particle-in-cell code designed to simulate charged particle beams with high space-charge intensity
- Performs I/O using **Python interface to HDF5**
  - Initial performance in the MB/sec!
- Analysis identified two points for improvement, type conversion and Lustre striping
- 18 GB/sec writes attained after optimization



I/O performance of WARP code after two rounds of improvements.

Building on Success

- **Institute for Sustained Performance, Energy, and Resilience (SUPER)**
  - Performance Engineering
  - Autotuning
  - Performance Tools
- **Scalable Data Management, Analysis, and Visualization (SDAV) Institute**
  - Data Management
  - Data Analysis and Visualization
- **FastMath Collaboration**
  - Performance improvements in solvers and applications
- **Machine Leaning and Deep Learning**
  - Applied to science domains

MPAS-O 2x *efficient use of OpenMP* **BGQ** GPU **KNL**	AMR Dycore 12x *vectorized & autotuned column solves* BGQ GPU **KNL**	PISCEES 2x *eliminated scalability bottlenecks* BGQ GPU KNL	LibTensor 150x *was single node; added MPI support* **BGQ** **GPU** **KNL**
CGYRO 3x *LAPACK optimizations, use of cuFFT* BGQ **GPU** KNL	XGC1 2x *load balancing & efficient OpenMP* BGQ **GPU** **KNL**	XOLOTL init. $>10^4$x *efficient hashing work pruning* BGQ GPU KNL	NWChem 65x *OpenMP in lieu of memory-limited flat MPI* BGQ GPU **KNC**
MFDn 1.8x *efficient SpMM, block eigensolver* BGQ GPU **KNL**	BIGSTICK 8x *load balancing & efficient OpenMP* **BGQ** GPU **KNL**	Geant4 12.6% *code optimization algorithmic innovation* BGQ GPU KNL	ParaDis 2x *efficient use of OpenMP* **BGQ** GPU KNL
GTC 35x *ADIOS I/O optimizations* BGQ **GPU** KNL	XGC1 27x *ADIOS I/O optimizations* BGQ GPU KNL	PiconGPU 25x *ADIOS I/O optimizations* BGQ GPU KNL	FireFOAM 12x *ADIOS I/O optimizations* BGQ GPU KNL
BISICLES Calving 6x *better graph algorithms* BGQ GPU KNL	AMR Index Query 500x *new indexing techniques* BGQ GPU KNL	HDF5 writes 25x *HDF5 autotuning* BGQ GPU KNL	Power Grid 200x *less data thru better compression algorithm* BGQ GPU KNL

Building on Success

- **Institute for Sustained Performance, Energy, and Resilience (SUPER)**
  - Performance Engineering
  - Autotuning
  - Performance Tools
- **Scalable Data Management, Analysis, and Visualization (SDAV) Institute**
  - Data Management
  - Data Analysis and Visualization
- **FastMath Collaboration**
  - Performance improvements in solvers and applications
- **Machine Leaning and Deep Learning**
  - Applied to science domains

MPAS-O 2x efficient use of OpenMP BGQ GPU KNL	AMR Dycore 12x vectorized & autotuned column solves BGQ GPU KNL	PISCEES 2x eliminated scalability bottlenecks BGQ GPU KNL	LibTensor 150x was single node; added MPI support BGQ GPU KNL
CGYRO 3x LAPACK optimizations, use of cuFFT BGQ GPU KNL	XGC1 2x load balancing & efficient OpenMP BGQ GPU KNL	XOLOTL init. $>10^4$x efficient hashing work pruning BGQ GPU KNL	NWChem 65x OpenMP in lieu of memory-limited flat MPI BGQ GPU KNC
MFDn 1.8x efficient SpMM, block eigensolver BGQ GPU KNL	BIGSTICK 8x load balancing & efficient OpenMP BGQ GPU KNL	Geant4 12.6% code optimization algorithmic innovation BGQ GPU KNL	ParaDis 2x efficient use of OpenMP BGQ GPU KNL
GTC 35x ADIOS I/O optimizations BGQ GPU KNL	XGC1 27x ADIOS I/O optimizations BGQ GPU KNL	PiconGPU 25x ADIOS I/O optimizations BGQ GPU KNL	FireFOAM 12x ADIOS I/O optimizations BGQ GPU KNL
BISICLES Calving 6x better graph algorithms BGQ GPU KNL	AMR Index Query 500x new indexing techniques BGQ GPU KNL	HDF5 writes 25x HDF5 autotuning BGQ GPU KNL	Power Grid 200x less data thru better compression algorithm BGQ GPU KNL

ENGAGEMENT

Community Engagement

Goal: Help application teams realize highest possible scientific impact.

- **Application Engagement** (Anshu Dubey (ANL) lead, Sam Williams (LBNL) co-lead)
  - Direct assistance to application teams running on DOE compute platforms
  - *Tiger teams* with RAPIDS technology experts tackle challenges
  - Provide consultation for productivity issues
- **Community Outreach** (David Bernholdt (ORNL) coordinating)
  - Educating and coordinating with broader DOE community, including SciDAC-4 Math Institute
  - Workshops, tutorials, and webinars
- **Facility Coordination** (Kevin Harms (ANL), Oliver Rubel (LBNL), David Bernholdt (ORNL))
  - Ensure two-way communication with facilities
  - Coordinate software deployment, assist with technical challenges

Tiger Teams

Goal: Focused engagements to resolve specific application challenges

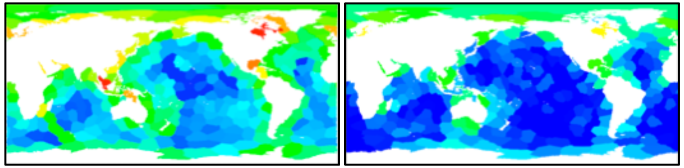
- **Engagement with one team for 3-6 months**
  - Needed expertise drawn from all the focus areas as needed
  - Plan ahead for scope and resources
  - Clearly articulate expected outcomes
- **Three to four Tiger Team activities per year**
  - Prioritization based on available expertise and potential impact on the application
  - Resources budgeted in RAPIDS appropriately.

Contacts: Anshu Dubey (ANL) lead, Sam Williams (LBNL) co-lead

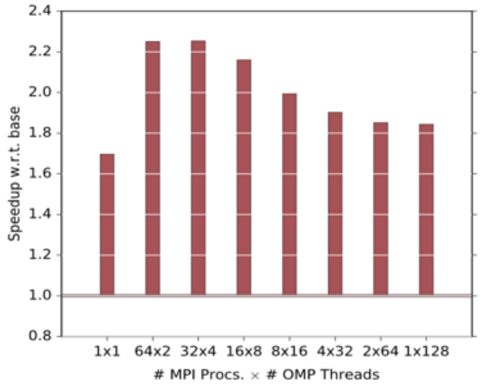
MPAS-Ocean Optimization for Emerging DOE Systems

RAPIDS

- Improved performance and energy of MPAS-Ocean, an unstructured mesh-based ocean modeling code, on leading DOE platforms using a variety of innovative approaches.
- Multi-disciplinary optimization: SUPER, SDAV, and FASTMath personnel, VisIT and TAU tools
- Implemented hybrid OpenMP/MPI version of code to target new node technologies
  - e.g., KNL, up to 2.2x speedup
- Code also used as driver for other SUPER CS activities



Visualization of load imbalance using VisIT (left) leads to new partitioning scheme for improved load distribution (right)



Speedup on Knights Landing node using optimized, threaded version versus original flat MPI, showing up to 2.2x speedup

SciDAC-4 Partnership Connections

Title	PI	Prog.	RAPIDS Member(s)
Coupling Approaches for Next-Generation Architectures (CANGA)	P. Jones	BER	Peterka, Ross
Prob. Sea-Level Projections from Ice Sheet and Earth System Models	S. Price	BER	Patchett
Advancing Catalysis Modeling	M. H. Gordon	BES	Williams
AToM: Advanced Tokamak Modeling Environment	J. Candy	FES	Bernholdt
Plasma Surface Interactions: Predicting the Performance and Impact of Dynamic PFC Surfaces (PSI-2)	B. Wirth	FES	Bernholdt, Roth, Pugmire
Center for Tokamak Transients Simulations (CTTS)	S. Jardin	FES	Williams
Integrated Simulation of Energetic Particles in Burning Plasmas (ISEP)	Z. Lin	FES	Williams, Klasky, Pugmire
High-fidelity Boundary Plasma Simulation (HBPS)	C. S. Chang	FES	Klasky, Podhorszki
Tokamak Disruption Simulation	X. Tang	FES	Patchett
Inference at Extreme Scale	S. Habib	HEP	Yoo, Morozov, Balaprakash
HEP Data Analytics on HPC	J. Kowalkowski	HEP	Peterka, Ross
HPC Framework for Event Generation at Colliders	S. Hoeche	HEP	Hovland
HEP Event Reconstruction with Cutting Edge Computing Architectures	G. Cerati	HEP	Norris
Simulation of Fission Gas in Uranium Oxide Nuclear Fuel	D. Andersson	NE	Bernholdt, Roth
Towards Exascale Astrophysics of Mergers and SuperNova	W. R. Hix	NP	Dubey
Nuclear Low Energy Initiative (NUCLEI)	J. Carlson	NP	Norris

SciDAC-4 Coordination Committee

First meeting December 7th, 2017

- **Goals:**
  - **Facilitate interactions between SciDAC Institutes and Partnerships, and with broader science community**
  - **Assess emerging needs across SciDAC science teams**

- Committee:

  Rob Ross (ANL, RAPIDS lead, chair)
  Lori Diachin (LLNL, FASTMath lead)
  Katherine Evans (ORNL, BER POC)
  Teresa Head-Gordon (LBNL, BES POC)
  David Hatch (U. Texas-Austin, FES POC)
  Jim Amundson (FNAL, HEP POC)
  Brian Wirth (UTK-ORNL, NE POC)
  Martin Savage (INT & U. Washington, NP POC)

- DOE:

  Randall Laviolette (ASCR)
  Ceren Susut-Bennett (ASCR)
  Dorothy Koch (BER)
  Jim Davenport (BES)
  John Mandrekas (FES)
  Lali Chatterjee (HEP)
  Dan Funk (NE)
  Ted Barnes (NP)

Thanks to the RAPIDS Team

RAPIDS

Argonne National Laboratory	Rob Ross (Director)	Los Alamos National Laboratory Est. 1943	Jim Ahrens (PI)
Berkeley Lab Lawrence Berkeley National Laboratory	Lenny Oliker (Deputy)	Northwestern University	Wei-keng Liao (PI)
Brookhaven National Laboratory	Shinjae Yoo (PI)	Oak Ridge National Laboratory	Scott Klasky (PI)
University of Delaware	Stephen Siegel (PI)	The Ohio State University	Han-Wei Shen (PI)
The HDF Group	Neil Fortner (PI)	University of Oregon	Kevin Huck (PI)
Kitware	Berk Geveci (PI)	Rutgers	Manish Parashar (PI)
Lawrence Livermore National Laboratory	Bronis de Supinski (PI)	The University of Utah	Mary Hall (PI)

RAPIDS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) program.

Argonne National Laboratory · Berkeley Lab · University of Delaware · Lawrence Livermore National Laboratory · Oak Ridge National Laboratory · University of Oregon · Rutgers · Brookhaven National Laboratory · The HDF Group · Kitware · Los Alamos National Laboratory · Northwestern University · The Ohio State University · The University of Utah