U.S. DEPARTMENT OF ENERGY | Office of Science

RAPIDS

RAPIDS: The SciDAC Institute for Computer Science and Data

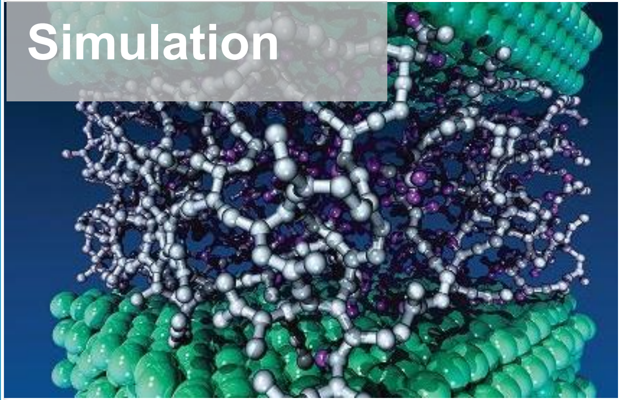
ROBERT ROSS
Institute Director
Argonne National Laboratory
rross@mcs.anl.gov

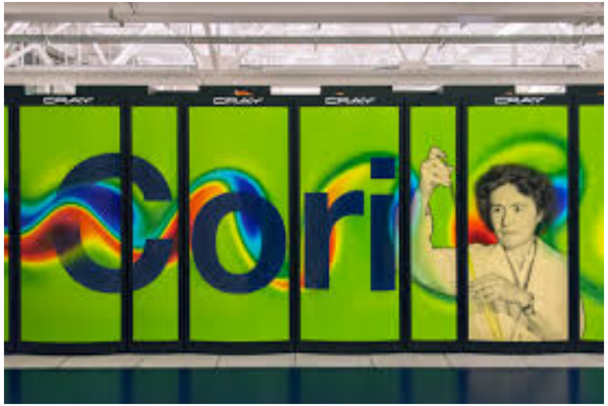
LENNY OLIKER
Deputy Director
Lawrence Berkeley National Laboratory
loliker@lbl.gov

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

Diverse Science and Systems

RAPIDS





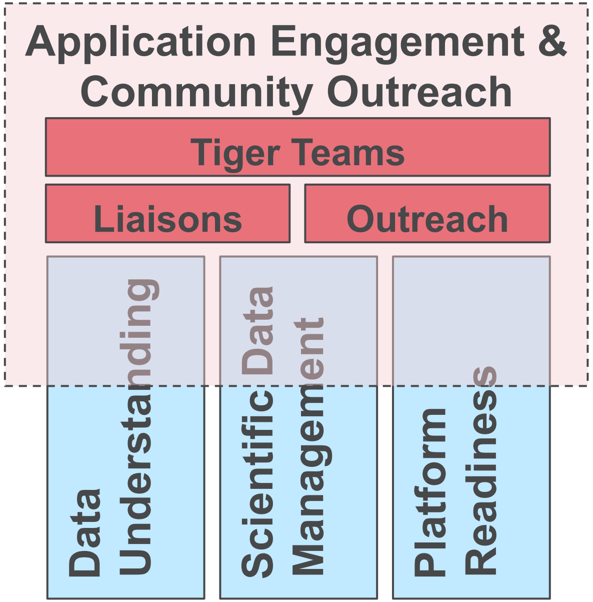
Top image credit B. Helland (ASCR). Bottom left, center, and right images credit ALCF, NERSC, and OLCF respectively.

The RAPIDS Institute

RAPIDS

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

- **Technology Focus Areas**
  - **Data Understanding** – scalable methods, robust infrastructure, machine learning
  - **Scientific Data Management** – I/O libraries, coupling, knowledge management
  - **Platform Readiness** – hybrid programming, deep memory hierarchy, autotuning, correctness
- **Application Engagement**
  - *Tiger Teams* engage experts in multiple areas
  - Software productivity: verification and validation, etc.
  - Outreach activities connect with broader community



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

Application Engagement

ANSHU DUBEY
ANL

SAM WILLIAMS
LBNL

Many Ongoing Partnership Collaborations!

Title	PI	Prog.	RAPIDS Member(s)
Coupling Approaches for Next-Gen Architectures (CANGA)	P. Jones	BER	Peterka
Prob. Sea-Level Proj from Ice Sheet & Earth System (ProSPect)	S. Price	BER	Patchett
An integrated system for optimization of sensor networks	D. Ricciuto	BER	Steed, Klasky, Podhorszki
Advancing Catalysis Modeling	M. H. Gordon	BES	Williams, Ibrahim
Comp. Framework for Unbiased Studies of Correlated Electron	T. Maier	BES	Huck
AToM: Advanced Tokamak Modeling Environment	J. Candy	FES	Bernholdt
Plasma Surface Interactions (PSI-2)	B. Wirth	FES	Bernholdt, Roth, Pugmire,
Center for Tokamak Transients Simulations (CTTS)	S. Jardin	FES	Williams
Integrated Simulation of Energetic Particles in Plasmas (ISEP)	Z. Lin	FES	Williams, Klasky, Pugmire
Multiscale Gyrokinetic Turbulence (MGK)	D. Hatch	FES	Shan
High-fidelity Boundary Plasma Simulation (HBPS)	C. S. Chang	FES	Klasky, Podhorszki
Tokamak Disruption Simulation	X. Tang	FES	Brugger, Dubey
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	G. Cerati	HEP	Norris, Lee, Vetter
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, Huck
Nuclear Low Energy Initiative (NUCLEI)	J. Carlson	NP	Norris

Visualization of Antarctica Land Ice

Scientific Achievement

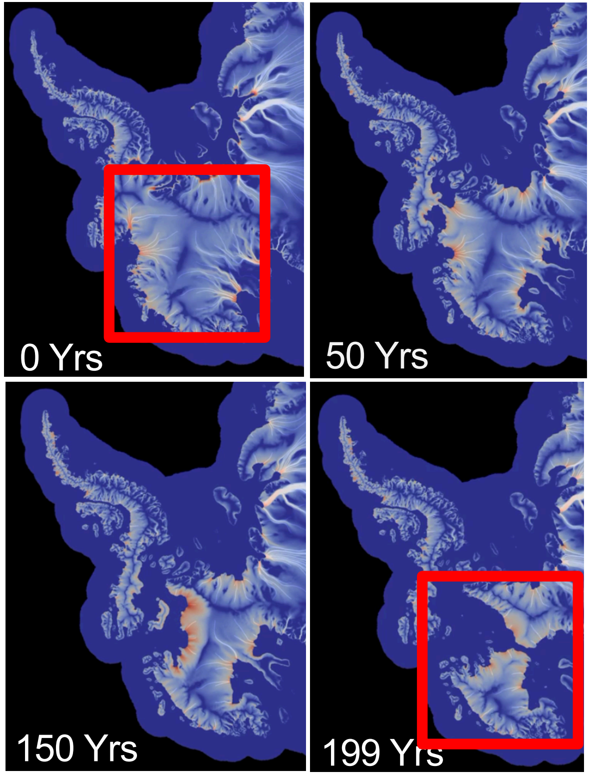
Used SciDAC ProSPect ice sheet model output and streamlines to show Antarctic ice sheet mass loss in response to extreme climate forcing scenario, using Paraview.

Significance and Impact

Visualization of key ideas in the science of land ice modeling is critical for scientific understanding within the climate research community and for the communication of climate science concepts to the general public.

Research Details

- Collaboration with SciDAC ProSPect to visualize simulations of Antarctic ice sheet evolution.
- Leveraged and improved ParaView, a SciDAC supported, open-source visualization tool.
- Ongoing support of SciDAC ProSPect for visualizing and communicating the consequences of future ice sheet evolution in response to climate change.



Evolution of the West Antarctic Ice Sheet over 200 years following the loss of all floating ice shelves.

Parallel Event Generation and Analysis on HPC Systems

Scientific Achievement

Fermilab researchers developed two scalable HPC codes using the DIY programming model that assist in rapidly generating Monte Carlo events and comparing generated events with experimental data from NOvA and LHC.

Significance and Impact

HEP workflows require generating and analyzing vast numbers of MC events. The RAPIDS DIY technology aids in development of codes that efficiently utilize HPC resources.

Research Details

– Pythia8 generates Monte Carlo events [1]

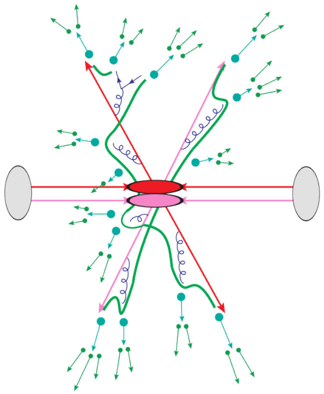
– Feldman-Cousins correction used for comparison [2]

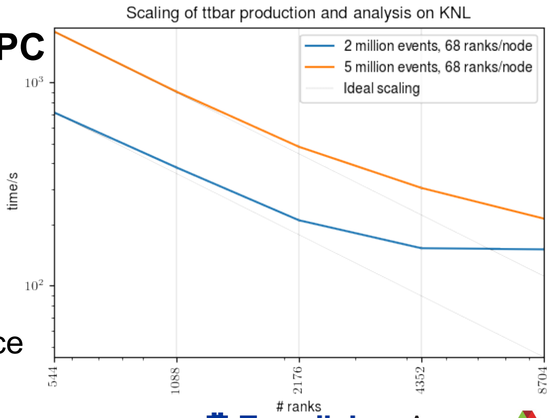
– DIY encapsulates communication in a productive model

– Allows for extremely short turn-around of large parameter space explorations (e.g. generator tuning)

[1] Hoche et al., *arXiv 2019*. [2] Sousa et al., *CHEP 2018*.

Event generator model for proton-proton collision: Robust predictions of collider events are needed to search for new physics effects. Much of the dynamics is described by tunable parameters. The calculation of event generator predictions is expensive, and must be done for each choice of parameters. A full detector simulation of these calculations is even more expensive, requiring parallel HPC codes.





Scalability: strong scaling of the HEP's Pythia8 event simulation with ASCR's DIY up to 8704 KNL cores on Cori. Up to 5 million proton-proton collisions. The deviation from ideal scaling is due to diminishing work per core at high core counts. (Images: Holger Schulz, U Cicinicnatti, Fermilab)

Fermilab · Argonne National Laboratory · University of Cincinnati · Colorado State University

Performance Optimization for Multiscale Gyrokinetic Turbulence

Scientific Achievement

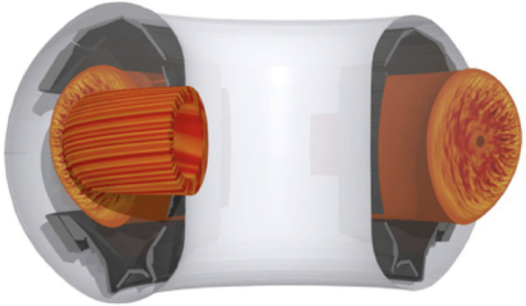
Improved the Gyrokinetic Electromagnetic Numerical Experiment (GENE) of SciDAC MGK throughput up to 30%, which has been used to achieve scientific breakthroughs on frontier multiscale turbulent transport problems

Significance and Impact

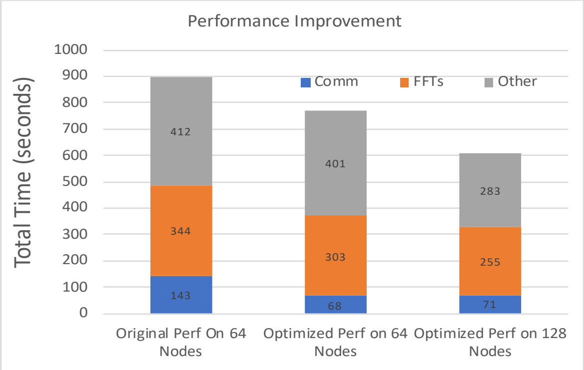
The simulation experiments often last days current HPC platforms, and reducing simulating running time accelerates scientific discovery.

Research Details

- Collaboration between LBNL and U. Texas-Austin
- Developed with Fortran 2003+MPI+OpenMP, using BLAS, LAPACK, FFT, PETSc, SLEPc, ScaLAPACK, HDF5 libraries,
- Communication performance significantly improved up to 50%
- Recent effort focusing on GPU optimization



Snapshot from a numerical simulation of plasma turbulence by GENE



Performance improvement due to communication and FFT optimizations

Accelerating Fusion/Fission Simulations

RAPIDS

Scientific Achievement

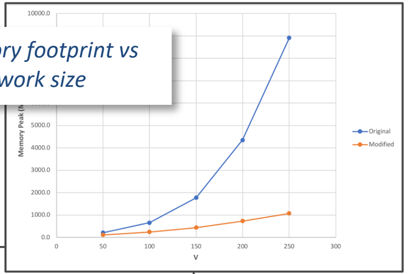
Improved science throughput for several applications used in SciDAC fusion and fission projects

Significance and Impact

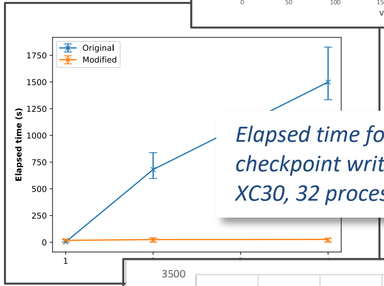
Can simulate much larger Xolotl reaction networks, with greater simulation throughput than before; established baseline for KORC scientific throughput, and then improved it.

Research Details

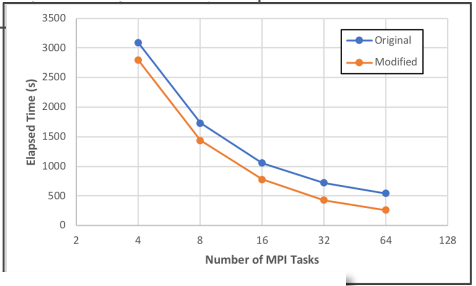
- Xolotl: cluster dynamics simulator for predicting gas bubble evolution in solids
- Reduced memory footprint of Xolotl version used for FY2018 FES theory milestone experiments by up to 88% on tested configurations (with contribution to PETSc code)
- Identified and fixed multiple Xolotl I/O performance problems: writing checkpoints 57x faster at scales, in FES theory milestone runs.
- Identified optimization enabling Kinetic Orbit Runaway Code (KORC) to run up to ~2× faster (from SciDAC SCREAM project)



Xolotl memory footprint vs reaction network size



Elapsed time for Xolotl checkpoint writing on Eos Cray XC30, 32 processes per node



KORC elapsed time vs number of MPI tasks (64 total threads for all runs)

OAK RIDGE National Laboratory

U.S. DEPARTMENT OF ENERGY | Office of Science

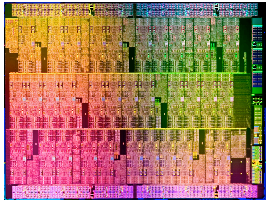
RAPIDS

Platform Readiness

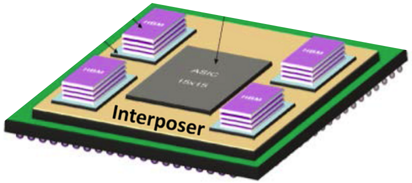
JEFFREY VETTER
ORNL

PAUL HOVLAND
ANL

Exascale: End of the CMOS Era



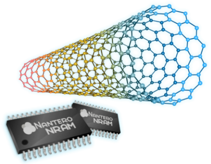
O(100 TF) per node
Wide vectors / GPUs
More for specialized procs



Memory system
on package

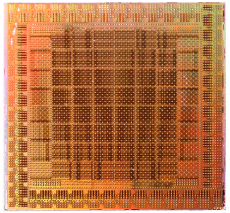
Disk for
archive

Flash for
main storage



NVM for
storage cache

O(100k) nodes, ~30-50 MW
O(10 Exaflops)



Low diameter networks with optics

Over the next decade, computers won't change that much from the current model.

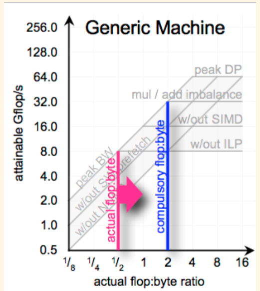
S. Scott, "Beyond Exascale: Playing the CMOS Endgame," SOS23, March 2019.

Platform Readiness

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

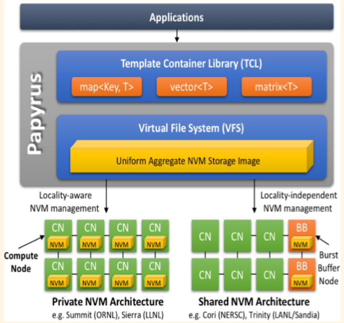
Performance Modeling/Analysis

- TAU: Performance Analytics & Tuning for Heterogeneous HPC
- Roofline: Easy-to-understand, visual performance model



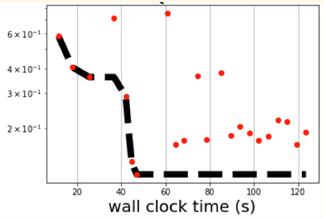
Portable Programming

- For heterogeneous systems, deep memory hierarchies: OpenMP, OpenACC
- AML, Papyrus: abstractions for shared data across deep memories



Autotuning

- Rigel: Vary OpenMP pragmas to investigate search space of
- SuRF: Uses ML to search optimal autotuning parameters
- Enables effective use of accelerators without multiple code versions



Program Correctness

- CIVL: Static verification of HPC programs
- Uses static analysis techniques over well-defined input ranges to do symbolic execution
- Enables verification equivalence of two implementations

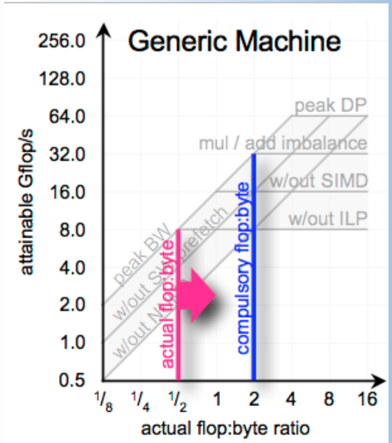
```
> civl compare
  -spec diffusion1d.c
  -impl diffusion1d_mpi.c

=== Result ===
The standard properties hold
for all executions.
```

Roofline Performance Modeling

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.

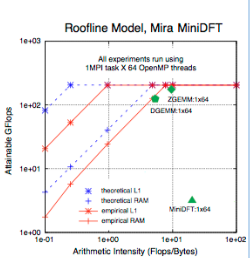


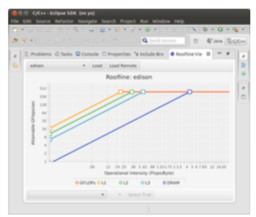
Proof of concept successfully applied to numerous computational kernels and emerging computing systems.

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

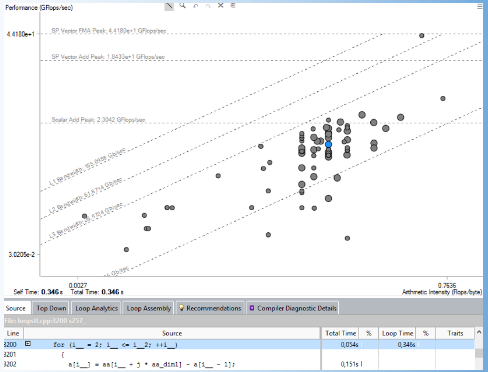




Automated Roofline code used to diagnose performance problems for DOE and SciDAC codes.

Outcome & Impact

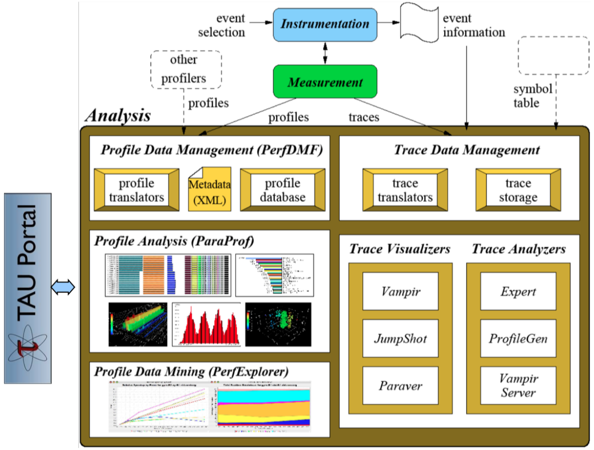
- Roofline has become a broadly used performance modeling methodology across DOE
- Intel has embraced the approach and integrated it into its production Intel® Advisor
- **Collaboration with NERSC to instrument and analyze execution of real applications on machines such as Edison and Cori**

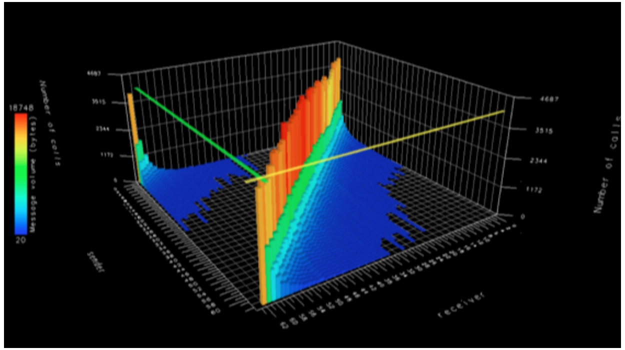


Snapshot of existing Intel Roofline tool in practice.

TAU: Performance Observation, Analytics, and Tuning for Heterogeneous Platforms

- Heterogeneous software stacks
  - Languages: OpenMP, OpenACC, CUDA, ROCm
  - Libraries/Metaprogramming: Kokkos, RAJA
  - Hybrid: MPI+X
- Runtimes
  - OpenMP, MPI, I/O, asynchronous multitasking
- Compilers and autotuners
  - LLVM, Chill, Oreo, Active Harmony, OpenARC
- Heterogeneous hardware measurement
  - Memory, Power, Network
- Integration with ADIOS2 for both I/O library measurement and ADIOS2 output of application performance data





Reproducible Performance Analysis
with HEP and NUCLEI

RAPIDS

Scientific Motivation

Develop a data analytics platform for creating and reusing the performance analysis workflows for improving the performance of DOE science codes.

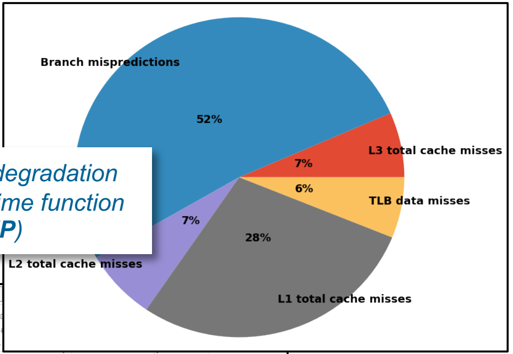
- **HEP Event Tracking:** Effective utilization of many-core SIMD and SIMT
- **NUCLEI:** Exploiting hybrid distributed- and shared-memory parallelism in integrated legacy and newly developed codes

Significance and Impact

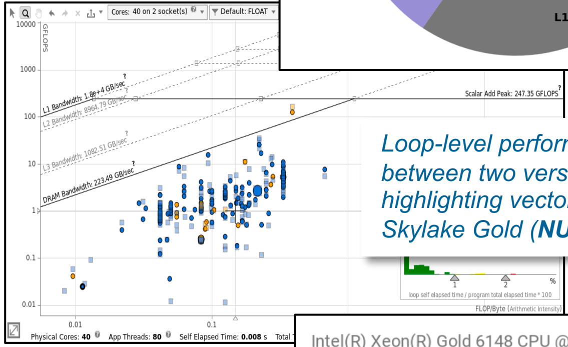
By enabling customizable, reusable performance analysis that can be maintained and extended by application teams, we can reduce the reliance on expert help and speed up performance optimization.

Research Details

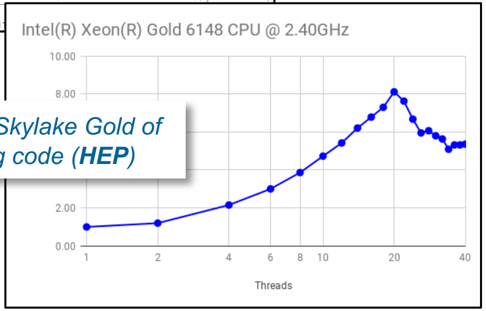
- Parallelized Kalman filter tracking, mkFit (**HEP Event Tracking**): Used TAU Commander and Python Pandas to create performance analysis "recipes". Achieved 2.7x speedup from explicit vectorization and > 10x from shared-memory parallelization on KNL; 4.4x speedup when integrated into main CMSSW framework (without optimizing data conversion).
- HFODD (**NUCLEI**): Used TAU and Intel's VTune and Advisor tools to create automated analysis workflows for shared-memory scaling and vectorization.



*Sources of performance degradation in the highest execution-time function in mkFit on Intel KNL (**HEP**)*



*Loop-level performance comparison between two versions of HFODD, highlighting vectorization on Intel Skylake Gold (**NUCLEI**)*



*Strong scaling on Skylake Gold of OpenMP hit finding code (**HEP**)*

Lawrence Livermore National Laboratory
Fermilab
University of Oregon

Leveraging One-Sided Communication for Sparse Triangular Solvers

Scientific Achievement

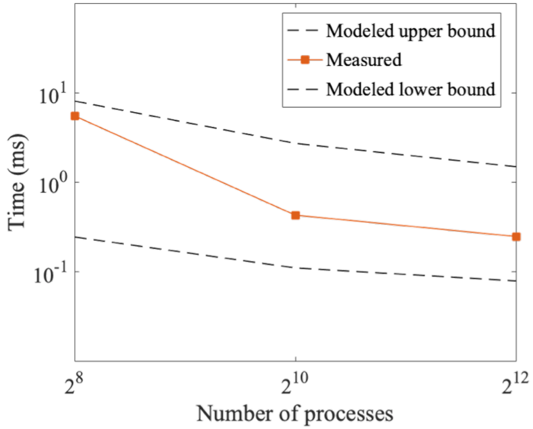
- Our one-sided MPI version of SpTRSV attains a 2.2x speedup at 4,096 processes on Cori (NERSC) over the existing two-sided in SuperLU_DIST.
- We constructed a critical path model to assess the observed performance relative to machine capabilities.

Significance and Impact

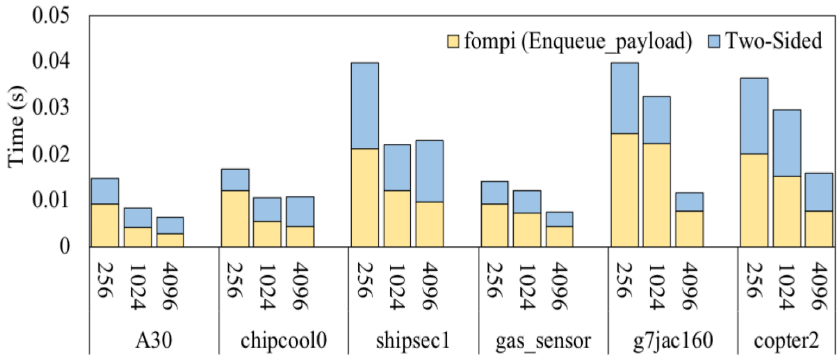
SuperLU preconditioners are essential for M3D-C1 and NIMROD solver, where overhead is dominated by communication. Our work directly improves performance and scalability of SpTRSV. Moreover, our one-sided implementation provides a pathway to accelerator-based exascale solvers.

Research Details

- Collaboration between CTTS, RAPIDS, and FASTMath
- Analyzed M3D-C1 and NIMROD scaling performance
- Evaluated Cray's vs ETH's foMPI one-sided MPI (8x speedup)
- Integrated ETH's foMPI into SpTRSV (2.2x speedup)
- Improved runtime/scalability on NERSC's KNL & Haswell



Model of M3D-C1 matrix. Gap between measured (orange) and lower bound indicates the potential benefits gained from better task placement.



Blue is SpTRSV total solve time of two-sided MPI, Yellow is one-sided. Performance improved 2.2x, and SpTRSV shows good scalability (making it friendly for exascale solvers)

Improving Collective Reduction Performance On Manycore Architectures

Scientific Achievement

Improve collective reduction performance on manycore architectures with threading and data compression. Improves reductions by up to 4X on Cori KNL for vectors, and 2.6x for overall BIGSTICK application (part of the CalLat NP SciDAC3 partnership).

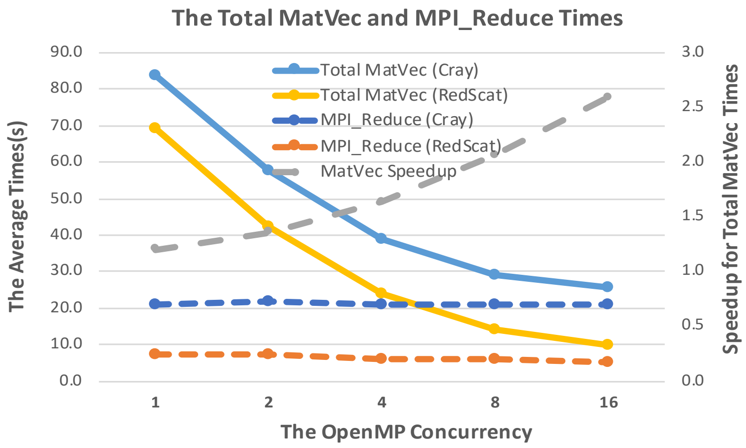
Significance and Impact

Global reduction is the top collective functions widely used on HPC platforms. Improving its performance has fundamental implications for many scientific codes.

Research Details

- Most optimizations focus on latency and bandwidth. This is not enough on manycore architectures.
- Use idle threads to accelerate local reductions, which often become the performance bottleneck for large vectors
- Data compression for reduced communication of sparse data
- Developed hierarchical algorithms which combines the advantages of both algorithms

MPI_Reduce performance as a function of algorithms, MPI concurrencies, and OpenMPs on NERSC Cori-KNL



Performance Improvement for BIGSTICK's matvec showing 2.6x speedup on 2048 processors

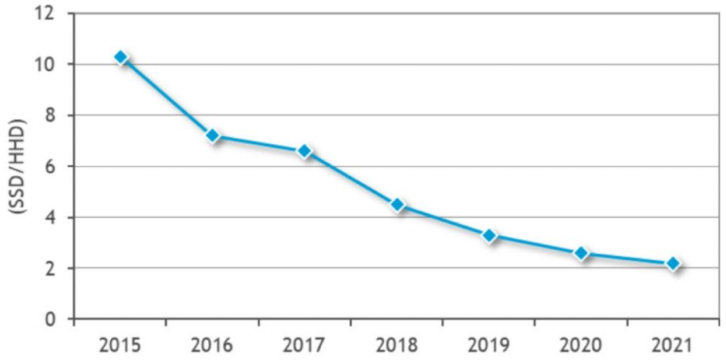
Data Management

SCOTT KLASKY
ORNL

JOHN WU
LBNL

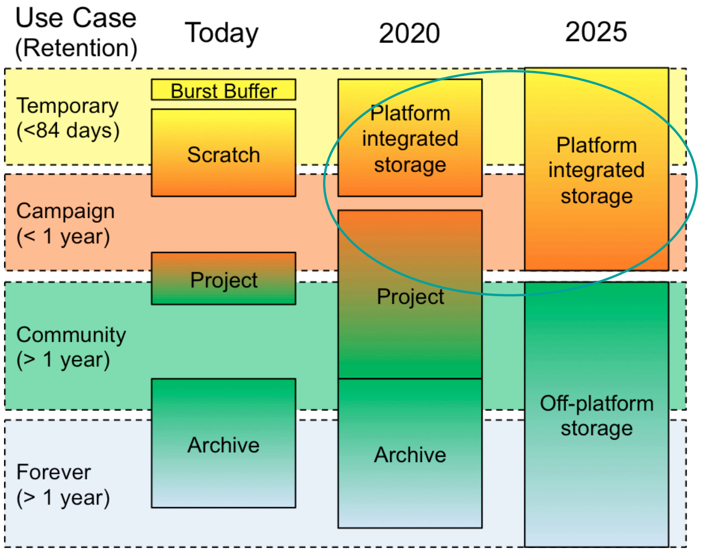
Technology Change in Storage Systems

Solid-state disk vs. hard disk drive pricing (per GB ratio)



Source: Hyperion research
https://www.storagenewsletter.com/2018/08/07/flash-storage-trends-and-impacts

Evolution of the NERSC storage hierarchy between today and 2025



Continued decline in cost of SSD capacity relative to HDD has led to plans to employ SSD-backed platform storage, integrated into the platform. *G. Lockwood et al. "Storage 2020: A Vision for the Future of HPC Storage," October 2017,*

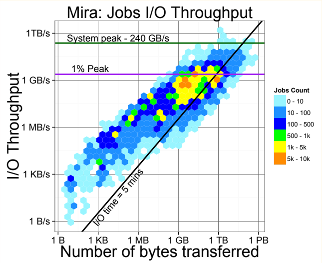
RAPIDS

Data Management

Deploying and supporting efficient methods to move and manage data in a scientific campaign.

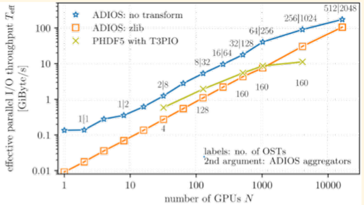
Performance Monitoring

- Understanding of I/O performance at scale
- Darshan: "Always on" statistics gathering
- TAU: Fine-grained I/O tracing of operations at multiple layers



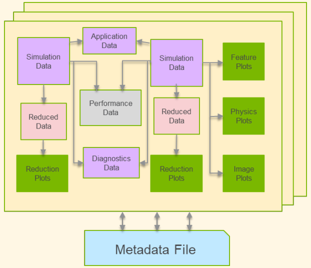
Storage and I/O

- HDF5: A data model, parallel I/O library, and file format for storing and managing data
- Parallel netCDF: Provides parallel access to traditional netCDF datasets
- ADIOS: community I/O framework to enable scientific discovery



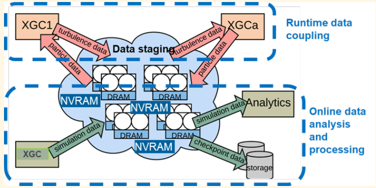
Knowledge Management

- FastBit: Organize and quickly find records across files generated and used during a scientific campaign
- Manage and query the data across a scientific campaign



Code Coupling

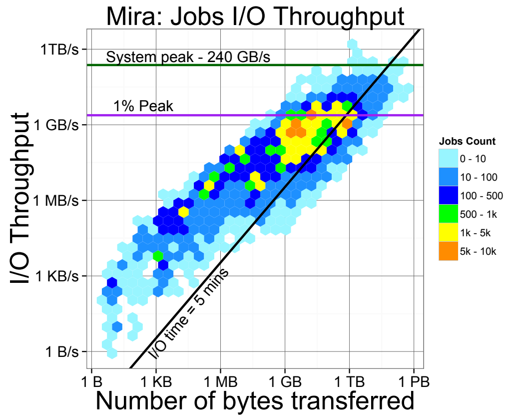
- Dataspaces & SST2: In-memory storage distributed across set of cores/nodes, using RAM and/or NVRAM
- Fast I/O to couple codes together asynchronously
- In-staging data processing, querying, sharing, and exchange

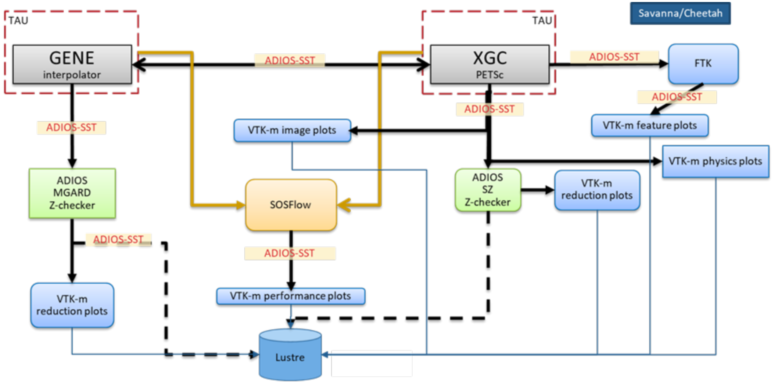


Performance Monitoring

Enabling understanding of I/O performance at scale

- Darshan
  - **“Always on” statistics gathering**
  - Observes I/O patterns of applications running on production HPC platforms, without perturbing execution, with enough detail to gain insight and aid in performance debugging
- TAU
  - **Fine-grained tracing of I/O operations at multiple layers**
  - ADIOS2 integration: integrated profile instrumentation of ADIOS2 and ability to stream TAU application performance data directly out to ADIOS2 at runtime

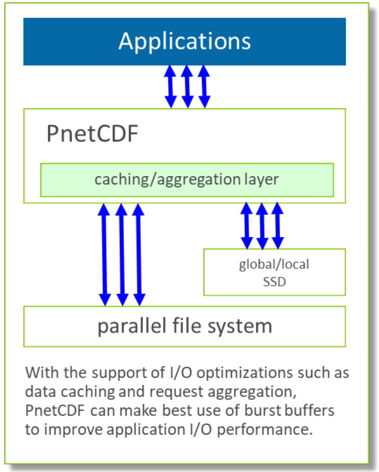




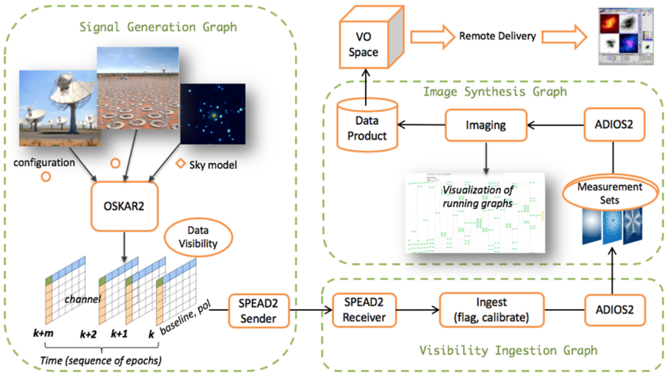
Storage and I/O

Libraries/frameworks to assist in fast and portable I/O

- **HDF5**
  - A data model, parallel I/O library, and file format for storing and managing data
  - Flexible, self-describing, portable, high performance
- **Parallel netCDF**
  - Provides parallel access to traditional netCDF datasets
  - Includes algorithms for accelerating common patterns such as multi-variable writes
- **ADIOS**
  - A community I/O framework to enable scientific discovery
  - In-memory code coupling for applications to other applications and/or analysis/visualization
  - Incorporates the state of the art I/O techniques for checkpoint, self describing data, and in situ data movement between codes



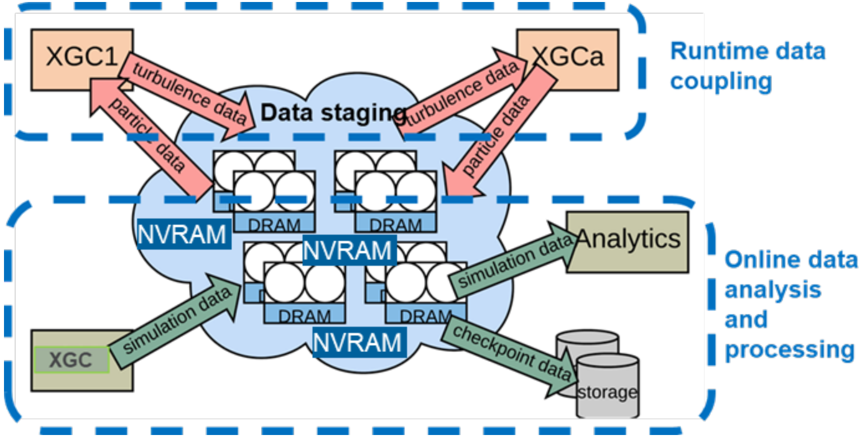
ADIOS is used for the backend for SKA data movement/storage



Code Coupling: DataSpaces

In-memory storage distributed across set of cores/nodes, using RAM and/or NVRAM

- Fast I/O to asynchronously couple codes together
- Couple simulation, visualization, analysis, and performance monitoring
- In-staging data processing, querying, sharing, and exchange
  - Virtual shared-space programming abstraction
  - Provides an efficient, high-throughput/ low-latency asynchronous data transport
  - Predictive data movement & layout



Staging-based in-situ workflows using DataSpaces: DataSpaces provides a semantically specialized shared-space abstraction using staging resources to support dynamic and asynchronous coordination, interactions and data exchanges between components of an in-situ workflow. The figure shows an **in-situ fusion simulation workflow and illustrates code coupling and in-situ data processing**.

Global Particle-in-Cell Simulation of Fusion Plasmas

Scientific Achievement

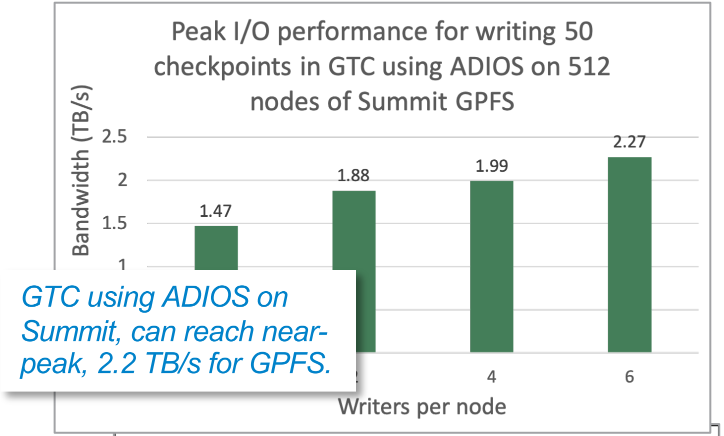
Energetic particle (EP) confinement is a key physics issue for the burning plasma experiment ITER. By enabling GTC with the ADIOS framework, we can finally write the majority of the physics data with minimal impact on the code performance on the Summit HPC resource at the OLCF

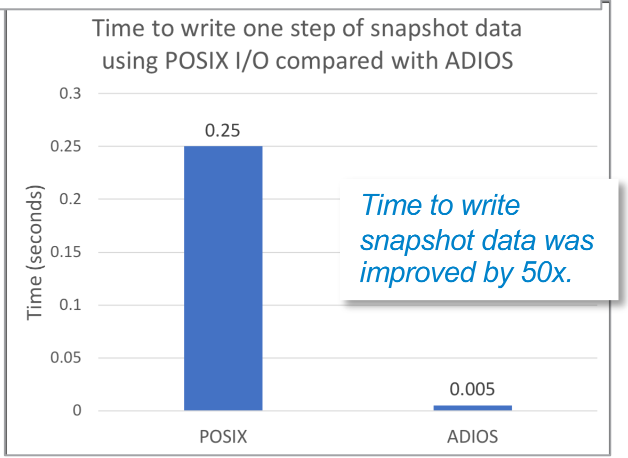
Significance and Impact

- GTC can generate over 100 TB of physics data every hour
- GTC has been equipped with ADIOS to allow all of the relevant physics information to be written to the Summit GPFS file system in less than 3% of the total runtime
- New data analytics is being written for GTC to work in both post-processing and in situ workflows

Research Details

- A new “engine” inside of ADIOS was developed to allow for extreme performance for Particle In Cell code I/O





L. Wan, K. Mehta, et al, “Data Management Challenges of Exascale Scientific Simulations: A Case Study with GTC and ADIOS”, ICCM 2019 (accepted)

Fusion Coupling Workflow

Scientific Achievement

The Fusion HPBS project is focusing on researching multi-way coupling science to study multi-scale/multi-physics. ADIOS enables flexible data movement and management between coupled codes.

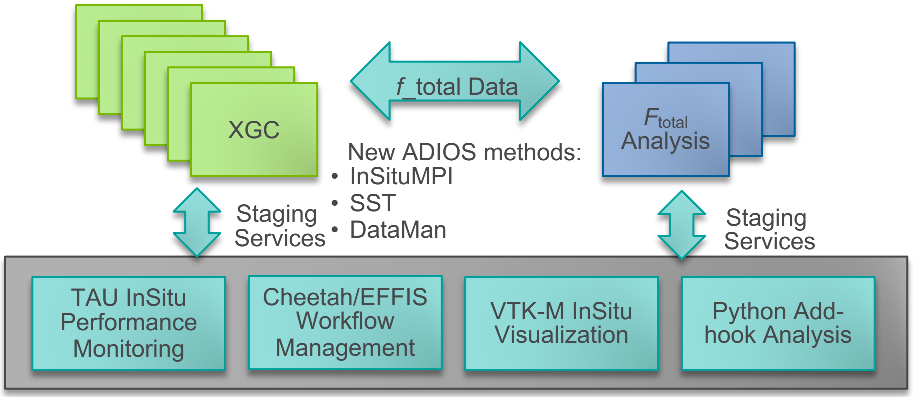
Significance and Impact

In XGC and *f* analysis coupling, we move *f-analysis* computation, with full distribution of function *f* and electrostatic field, to a dedicated analysis code. XGC offloads these with ADIOS and improves core computational capability.

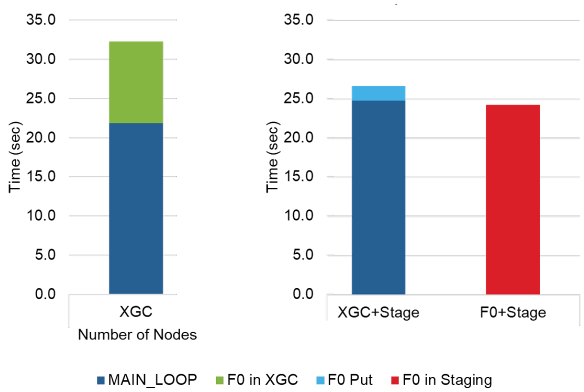
Research Details

- To improve data movement and flexibility, HBPS integrated with ADIOS for I/O management.
- Developing multi-way coupling science cases to study multi-scale/multi-physics study.
- Further investigating complex multiple coupling scenarios: XGC-hPIC (plasma-material interaction) and XGC and M3D-C1 for a higher fidelity MHD calc

HBPS Coupling though ADIOS



XGC Cyclone Case on Summit (1024 nodes) with 9X Analysis computations



HBPS's XGC F-analysis coupling diagram (up) and performance improvement (left). Performance of F-analysis routine in XGC improved with ADIOS in-situ coupling framework.

RAPIDS

Oak Ridge National Laboratory

Accelerating Earthquake Detection

Scientific Achievement

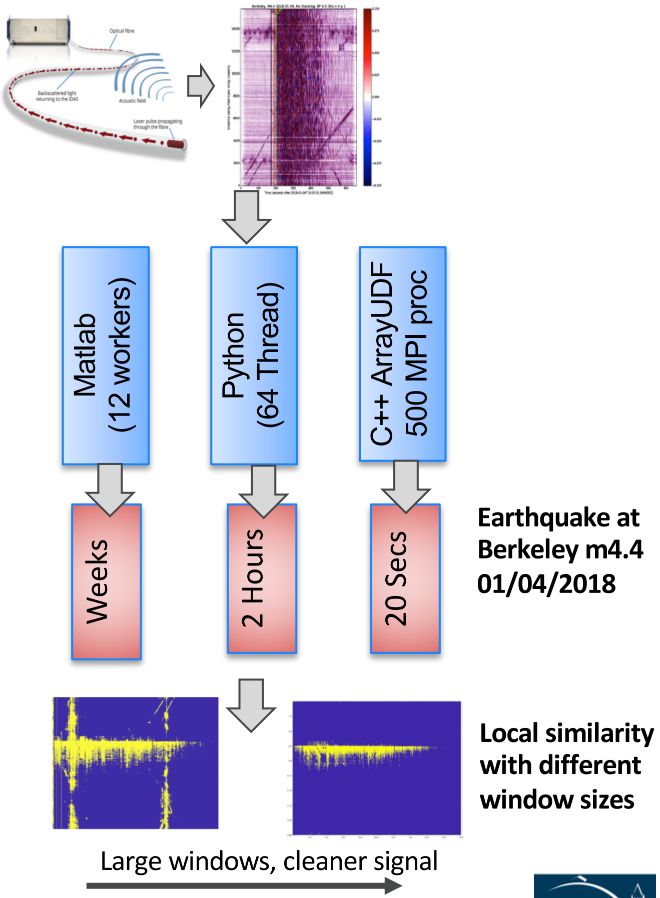
Developed a data parallel mechanism for analyzing the large data sets from distributed acoustic sensing

Significance and Impact

Demonstrate automatic parallelization applied to complex analysis tasks in earthquake detection, achieving significant speedup.

Research Details

- DAS: Distributed Acoustic Sensing is using Dark Fiber to collect ground motion data, about 500TB in past couple of years
- Due to the noisy nature of the data, significant computation is needed to extract signals such as minor earthquakes
- Earth scientists currently use matlab to do their analyses: Converting analyses to python reduces computation time from weeks to hours
- Using our automatic parallelization framework, ArrayUDF, further reduces the execution time to 20 seconds



Xing, et al. Automated Parallel Data Processing Engine with Application to Large-Scale Feature Extraction. In: MLHPC. 2018.

Data Understanding

DMITRIY MOROZOV
LBNL

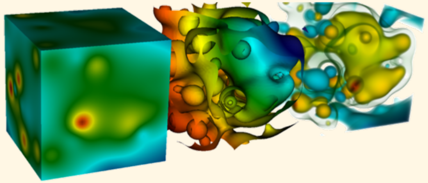
PRASANNA BALAPRAKASH
ANL

Data Understanding

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

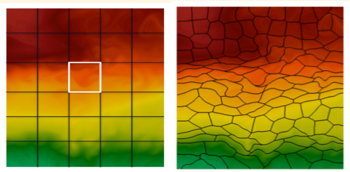
Visualization

- Visualization tools that leverage modern HPC
- In situ frameworks, to enable efficient system usage
- Scalable infrastructure: service oriented data analysis and reduction
- Leveraging deep memory hierarchy, on-node parallelism
- Analysis/visualization of high dimensional datasets



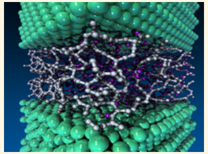
Scientific Data Analysis

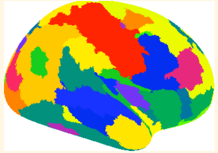
- Feature detection for visualizing and comparative analysis
- Geometric analysis: Delaunay/Voronoi tessellation
- Statistical analysis of ensemble and uncertain data
- Uncertain flows from ensemble modeling
- Topological features in scalar fields



Machine Learning

- Supervised learning methods, including deep learning for object classification
- Unsupervised learning methods, including dimension reduction
- Scalable parallel graph algorithms
- Sparse inverse covariance matrix estimation

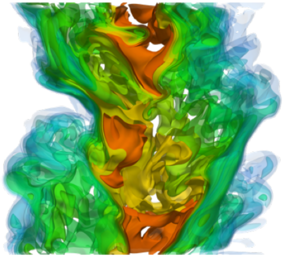


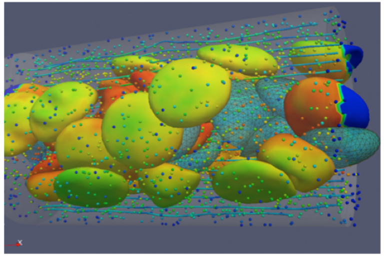


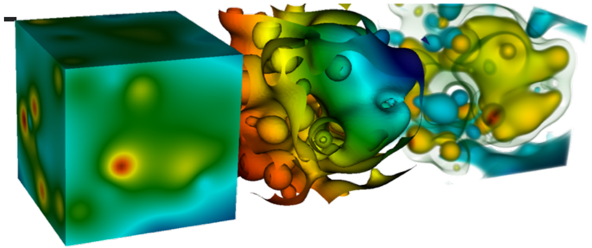
Visualization

Feature-rich visualization tools that can be run at scale, in situ

- Successful existing tools: **ParaView** and **VisIt,** both built on top of **VTK**, take advantage of massively parallel architectures of modern super-computers
- In situ frameworks, **VisIt/libsim, ParaView/Catalyst, ADIOS, Sensei, Henson**, enable using these systems efficiently with the simulations, e.g., to visualize live simulations avoiding the I/O bottleneck
- **Scalable infrastructure**: service-oriented data analysis and reduction, co-analysis with performance data
- Major focus on adapting to the deep memory hierarchies and massive on-node hybrid parallelism **(VTK-m)**
- Also useful information visualization techniques **(EDEN, CrossVis)**, techniques for analysis and visualization of high-dimensional datasets





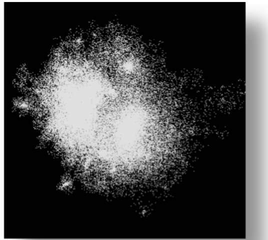
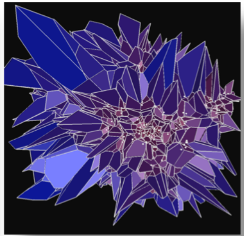
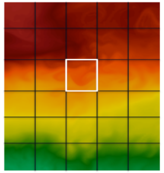
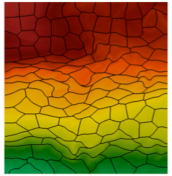
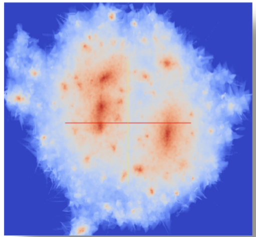
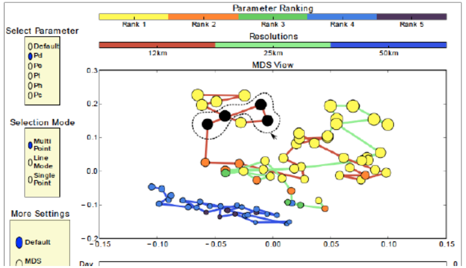


Scientific Data Analysis

RAPIDS

Scalable methods for finding and analyzing features of importance

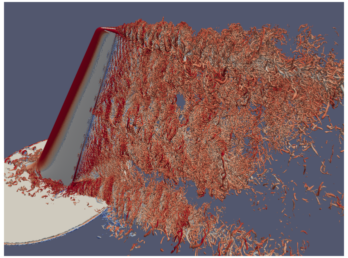
- Expertise in **feature detection (ftk)**, traditionally for visualization and comparative analysis. Moving forward as input to machine learning methods.
- **Geometric analysis (tess):** scalable computation of Delaunay and Voronoi tessellations, e.g., for density estimation in cosmological data
- Statistical analysis of ensemble data **(edda)**:
  - representation of large scale uncertain data
  - analysis of ensemble and uncertain features
  - exploration of parameter space for ensemble simulations
- **Uncertain flows** from ensemble modeling (fluid dynamics, climate, weather)
  - Generalizing flow features for uncertain data
  - Surface Density Estimates to quantify uncertainty
  - Scalable algorithms to stochastically trace particles
- **Topological features** in scalar fields
  - Scalable computation of merge trees, contour trees, persistence diagrams (used in cosmology, combustion, materials science, etc.)
  - Useful both for visualization and for comparison of simulations, to each other and to experiments

In situ Viz Unlocks Unsteady Dynamics at Extreme Scale

Scientific Achievement

Unsteady synthetic jet flow control simulations create data streams so large that dynamics are only practical to access with *in situ* visualization.

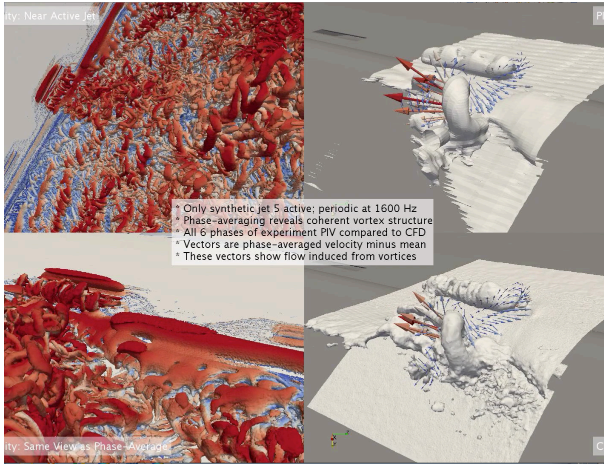


Significance and Impact

***In situ* visualization allows comparison of instantaneous vortical structures with phase averaged quantities from experiment and simulation. Animations of images help engineers understand how jets improve flow and scale that improvement to flight/turbine conditions.**

Research Details

- ParaView Catalyst compiled into PHASTA, provides isosurfaces at every time step, at just 3% overhead to simulation – far less than writing full data.
- Communication of parameter changes back to simulation (computational steering) in progress
- Demonstrated scaling at over 1M MPI ranks on BG/Q using SENSEI interface for *in situ* analysis and viz.



Comparison to experimental results.

RAPIDS · CU Boulder · Kitware

Integrating Human Perception with Computation for Guided Exploratory Data Analysis

Scientific Achievement

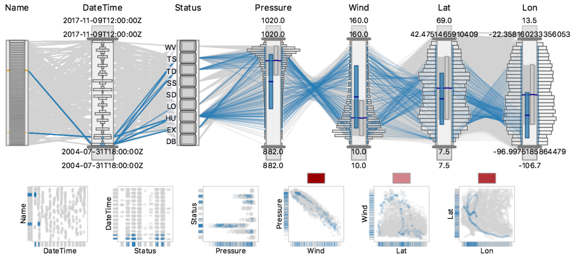
CrossVis helped materials scientists at ORNL CNMS understand and improve a neural network classification process for microscopic imagery and allowed ORNL BER climate scientists to consider more variables from large scale, land model parameter sensitivity analyses and improve model accuracy.

Significance and Impact

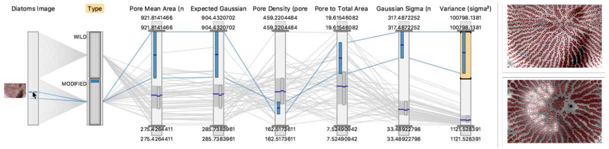
CrossVis enables flexible exploration and comprehensive understanding of large, heterogeneous, and multivariate data by integrating interactive visualizations and statistical analytics.

Research Details

- Implements theoretical information foraging concepts, where information dynamically derived from statistical analytics are used to augment interactive data visualizations to make key patterns visually salient.
- Provides an advanced multivariate visual analytics framework supporting heterogenous data types (e.g., images, temporal, categorical), progressive high-performance rendering, and a scalable data model.



Exploration of Historical Hurricane Observations



Analyze Neural Networks for Microscopic Imagery

CrossVis is a visual analytics tool that integrates statistical analytics and an extended version of parallel coordinates to allow flexible exploratory of large and heterogenous multivariate data. In addition to climate and materials science, CrossVis has been applied to cyber security, manufacturing, power grid, and system performance projects.

Statistical Super Resolutions for Large Scale Ensemble Cosmological Simulations

Scientific Achievement

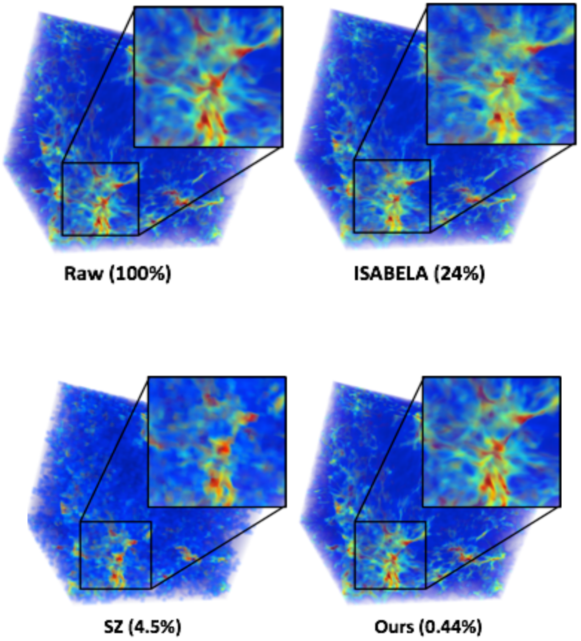
Enable scientists to reduce the storage space requirement when running large ensemble simulations, while still make it possible to perform full scale simulation parameter exploration for post-hoc analysis

Significance and Impact

With the statistical signatures, it is now possible to reconstruct simulation output of novel parameters that was not saved during simulations. The space saving can be more than 95%.

Research Details

- Store a small number of simulation results at full resolutions into a code book as prior knowledge
- Down sample the remaining data into GMMs as the statistical signatures
- Data at an arbitrary parameter configuration can be reconstructed from the prior knowledge and the statistical signatures
- The priori knowledge only takes 0.44% of the original data for a cosmology simulation using Nyx



Images produced by our super resolution representations

U.S. DEPARTMENT OF ENERGY | Office of Science

RAPIDS

Machine Learning and AI

Executive Order 13859 of February 11, 2019

Maintaining American Leadership in Artificial Intelligence

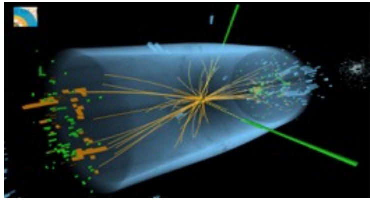
By the authority vested in me as President by the Constitution and the laws of the United States of America, it is hereby ordered as follows:

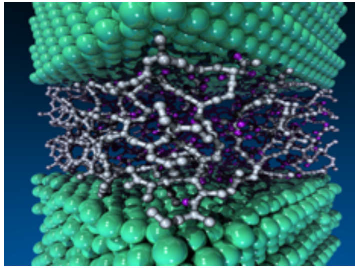
Section 1. *Policy and Principles.* Artificial Intelligence (AI) promises to drive growth of the United States economy, enhance our economic and national security, and improve our quality of life. The United States is the world leader in AI research and development (R&D) and deployment. Continued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent with our Nation's values, policies, and priorities.

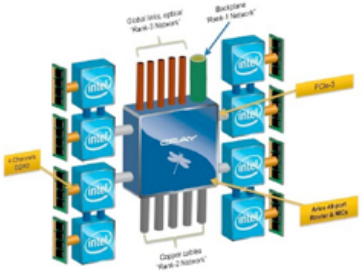
Machine Learning and AI

RAPIDS

Domain-specific applications of deep learning, predictive performance models, data- and model-parallel training







- Supervised learning methods:
  - Deep learning for object classification and identification
  - Large-scale training of convolutional NNs
  - Automatic multiobjective modeling **(AutoMOMML)** to simplify model selection
  - Asynchronous hyper-parameter and neural arch search **(DeepHyper/HPS)**
  - Autotuning parameters for code/application **(SuRF)**
  - Performance, power, and energy modeling of novel HPC architectures;
- Unsupervised learning methods:
  - Manifold learning/dimensionality reduction; approximation algorithms to cope with streaming data, streaming spectral clustering
  - Useful for adaptive sampling (e.g., for molecular dynamics trajectories)
- Reinforcement learning
- Scalable parallel **graph algorithms (LAGraph)**:
  - recast graph algorithms into linear algebra operations
  - building blocks and communication-avoiding algorithms for key functions
- Tools for understanding ML models (DeepVid, GANViz, DQNViz)

Using Roofline to Characterize Tensor Flow on GPUs

RAPIDS

Scientific Achievement

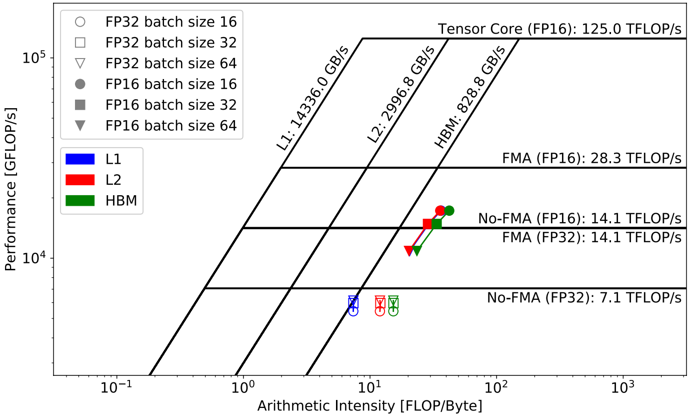
Created a methodology for analyzing the execution of GPU Tensor Core-accelerated DL/AI applications using Roofline.

Significance and Impact

This work enables Roofline-based analysis of NVIDIA Tensor Core accelerated AI/DL Applications including quantitative assessments of TensorFlow performance on NVIDIA Volta GPUs.

Research Details

- Collaboration between RAPIDS, NERSC, and NVIDIA
- Formulated methodology for using NVProf to analyze tensor-core accelerated applications using Roofline
- Used Roofline to analyze the forward and backward phases in TensorFlow as a function of FP16 and FP32.
- TensorFlow cannot sustain the theoretical 125TF/s due to a lack of locality and data permutation overheads.



TensorFlow (forward pass) on Volta V100

Results shown are relative to precision (32b and 16b tensor cores) and batch size (16,32,64). Although tensor cores deliver >2x performance, performance is far from theoretical 125TF/s

Yang et al., "Hierarchical Roofline Analysis for GPUs: Accelerating Performance Optimization for the NERSC-9 Perlmutter System", CUG, 2019.

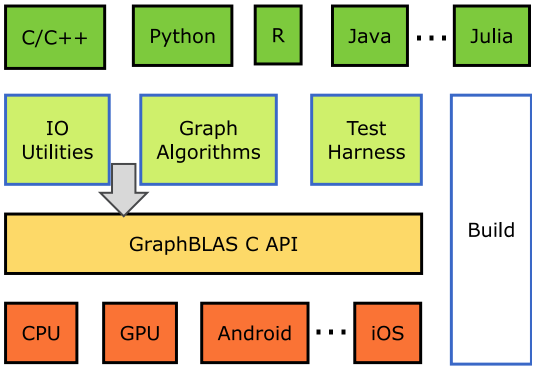
NERSC

BERKELEY LAB

LAGraph: A Community Effort to Collect Graph Algorithms

Project Aims

– Bring together the full range of known graph algorithms that can be constructed with the GraphBLAS.
– GraphBLAST is the first high-performance GPU implementation of GraphBLAS
– Systematically assess the coverage of graph algorithms based on linear algebra.
– Provide raw material in ongoing studies of the fundamental design patterns exploited by linear algebra-based graph algorithms



LAGraph Project Overview

Algorithms Implemented to Date

- Breadth-First Search, Shortest Paths
- Triangle and k-truss enumerations
- Connected Components, graph clustering
- Graph coloring, graph matching
- Collaborative filtering via Stochastic Gradient Descent

New Avenues for GraphBLAS

- Graph neural network training/inference
- Graph kernels for supervises learning
- A* search

Tim Mattson et al. *LAGraph: A Community Effort to Collect Graph Algorithms Built on Top of the GraphBLAS, IPDPSW, 2019*

Anticipatory Data Delivery in Extreme Scale In-situ Workflows

Scientific Achievement

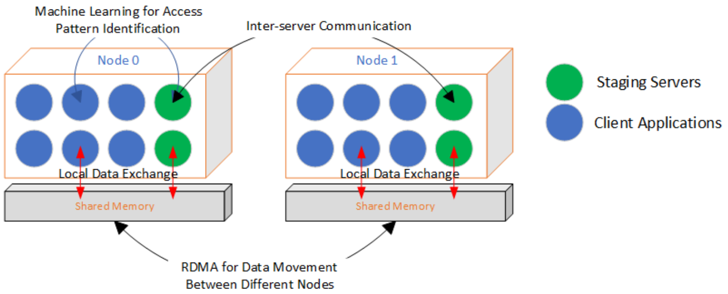
Enables machine learning guided, anticipatory data delivery for extreme-scale in-situ workflows to significantly reduce in data access costs.

Significance and Impact

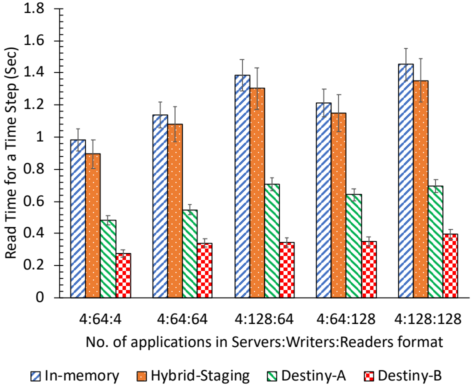
DESTINY can achieve a reduction of up to 75% and 53% in read response time, for collocated application processes and processes residing in separate nodes, as compared to in-memory staging service for production scientific workflows.

Research Details

- Uses n-gram machine learning model to anticipate future data accesses, proactively packages and delivers the data necessary to satisfy these requests as close to the consumer as possible.
- Amortizes expensive data discovery and assembly operations in data staging.



Overview of DESTINY. The data is exposed to applications via shared memory abstractions and delivered to the closest staging server before read-request is issued.



Hybrid-Staging represents just exposing data through shared memory. Destiny-A represents applications residing on separate nodes than staging and Destiny-B represents collocated staging and application processes.

Rutgers
Discovery Informatics Institute

In Situ Compression Artifact Removal in Scientific Data Using Deep Transfer Learning

Scientific Achievement

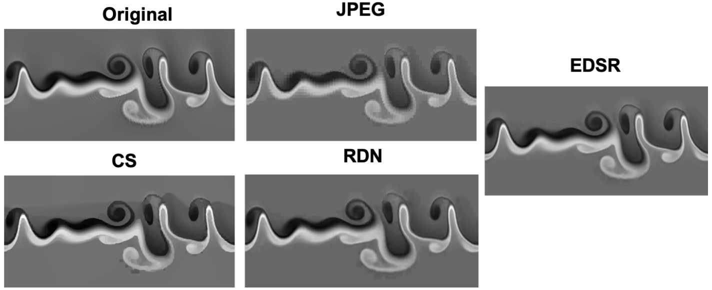
Developed a scalable in situ approach to train deep learning models that leverage knowledge from different domains and remove compression artifacts from lossy highly compressed images that correspond to streaming scientific simulated data.

Significance and Impact

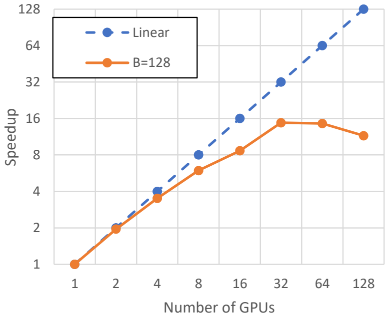
Introduces transfer learning and in situ learning paradigms into deep learning-based compression artifact removal, achieving superior accuracy and efficiency compared to standard image compression and more advanced compressed sensing methods.

Research Details

- We adopt convolutional neural network based architectures - Enhanced Deep Super-Resolution Network (EDSR) and Residual Dense Networks (RDN).
- These models are initially trained offline using the simulation data from climate domain (shallow water equations on a sphere) with JPEG compressed image as input and the original uncompressed image as the desired output.
- We adopt the discrepancy-based deep domain adaptation approach to transfer learn the knowledge from offline-trained model to an in situ setting where the model is updated at regular interval to adapt to data from different domain (Kinetic Transport).
- This approach is scaled by using data-parallel training with controlled learning rate updates, thus the CAR model will be ready as soon as the simulation is complete.



Comparison of EDSR and RDN enhanced images (using offline learning of Compression Artifact Removal (CAR) model on climate data) with JPEG compressed and compressed sensing enhancement approach result.



Scalability of EDSR training with number of GPUs (batch size is 128).

S. Madireddy, In Situ Compression Artifact Removal in Scientific Data Using Deep Transfer Learning.

Robust I/O Performance Modeling by Automated Hardware/Software Change Detection

Scientific Achievement

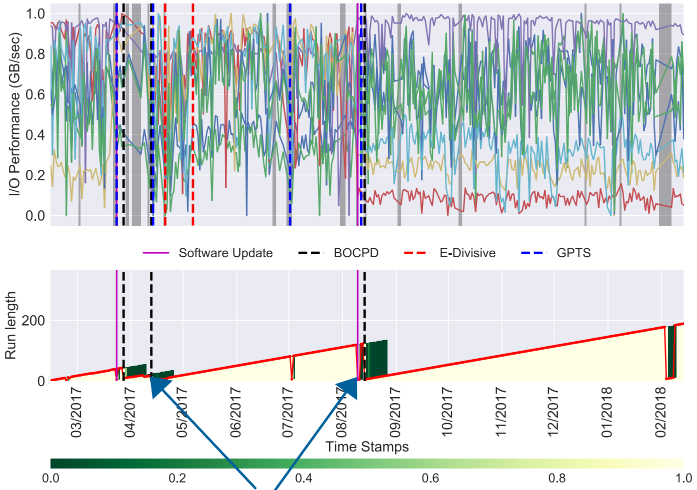
Developed a machine-learning-based I/O performance modeling approach that is robust to HPC system state changes (e.g., hardware degradation, hardware replacement, software upgrades).

Significance and Impact

Automatically identifies hardware and software changes that affect I/O performance in HPC systems and adapts our performance model, allowing better prediction and potentially improving the system utilization and application scheduling.

Research Details

- Online Bayesian detection to automatically identify the location of events that lead to changes in near-real time
- Moment-matching transformation that converts the training data collected before the change to be useful for retraining.
- Approach demonstrated on I/O performance data obtained on Lustre file system at NERSC.



Online method that monitors the change in the I/O performance of an application and adapt the model to these changes

We use application I/O performance data collected on Cori, a production supercomputing system at NERSC, to demonstrate the effectiveness of our approach. The results show that our robust models obtain ***significant reduction in prediction error---from 20.13% to 8.28%*** when the proposed approaches were used in I/O performance modeling.

S. Madireddy,et al. Adaptive Learning for Concept Drift in Application Performance Modeling, Preprint, ANL/MCS-P9132-0918, 2019.

Understanding How Deep Learning Models Operate

RAPIDS

Scientific Achievement

Allow developers of deep learning models to open the black box to see how and why the DNN model functions, so as to further optimize its performance

Significance and Impact

Explaining AI decision-making is a key challenge in the adoption of AI algorithms in scientific activities. Visual analytics approaches can play a crucial role in explaining modern AI models.

Research Details

- Deep Visual Interpretation and Diagnosis for Image Classifiers (**DeepVID)** is a model-agnostic approach for interpreting and diagnosing images classifiers, providing a rich user interface for understanding convolutional neural networks (CNNs).
- DeepVID is one tool in a suite of tools being developed for understanding AI models.

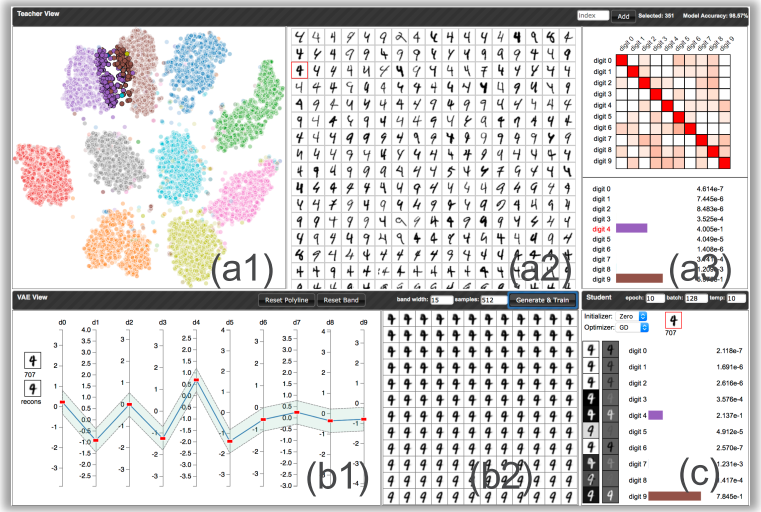


Figure: DeepVID is a visual analytics interface for understanding an image classifier based on variational autodecoder (VAE). Our goal is to understand what knowledge the neural network has acquired so as to perform the image classification tasks. **We visualize the various aspects of the neural models that will help the developer to optimize and diagnose the classification model.**

Thanks to the RAPIDS Team!

RAPIDS

… and more …

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

RAPIDS

For general questions:

Rob Ross <rross@mcs.anl.gov>

Lenny Oliker <LOliker@lbl.gov>

For engagement discussion:

Anshu Dubey <adubey@anl.gov>

Sam Williams <swwilliams@lbl.gov>

On the web:

http://www.rapids-scidac.org

… or just reach out to the RAPIDS person that you already know!

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