SSI: Removing Bottlenecks in High Performance Computational Science

Mark Gordon (PI), Theresa Windus, Daniel Crawford, David Sherrill, Lyudmila Slipchenko, Todd Martinez

Iowa State University · Virginia Tech · Georgia Institute of Technology · Purdue University · Stanford University

Interoperability Among Quantum Chemistry Programs

- Psi4
- GAMESS
- NWChem
- Cfour

}

- New library-level interfaces
- Light-weight data-sharing
- Driver/control codes
- Well-defined data standards

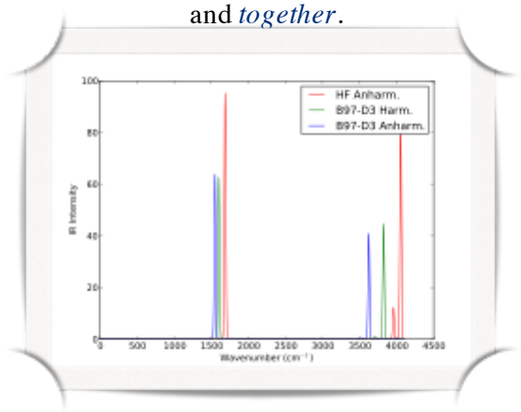
	Psi4	Cfour	Psi4 + Cfour
DFT 1st Deriv.	Yes	No	Yes
Anharmonic 2nd Deriv.	No	Yes	Yes
DFT Anharmonic 2nd Deriv.	No	No	Yes

best vibrational spectrum from *each* and *together*.

⟨CC|CC⟩ Psi4 +

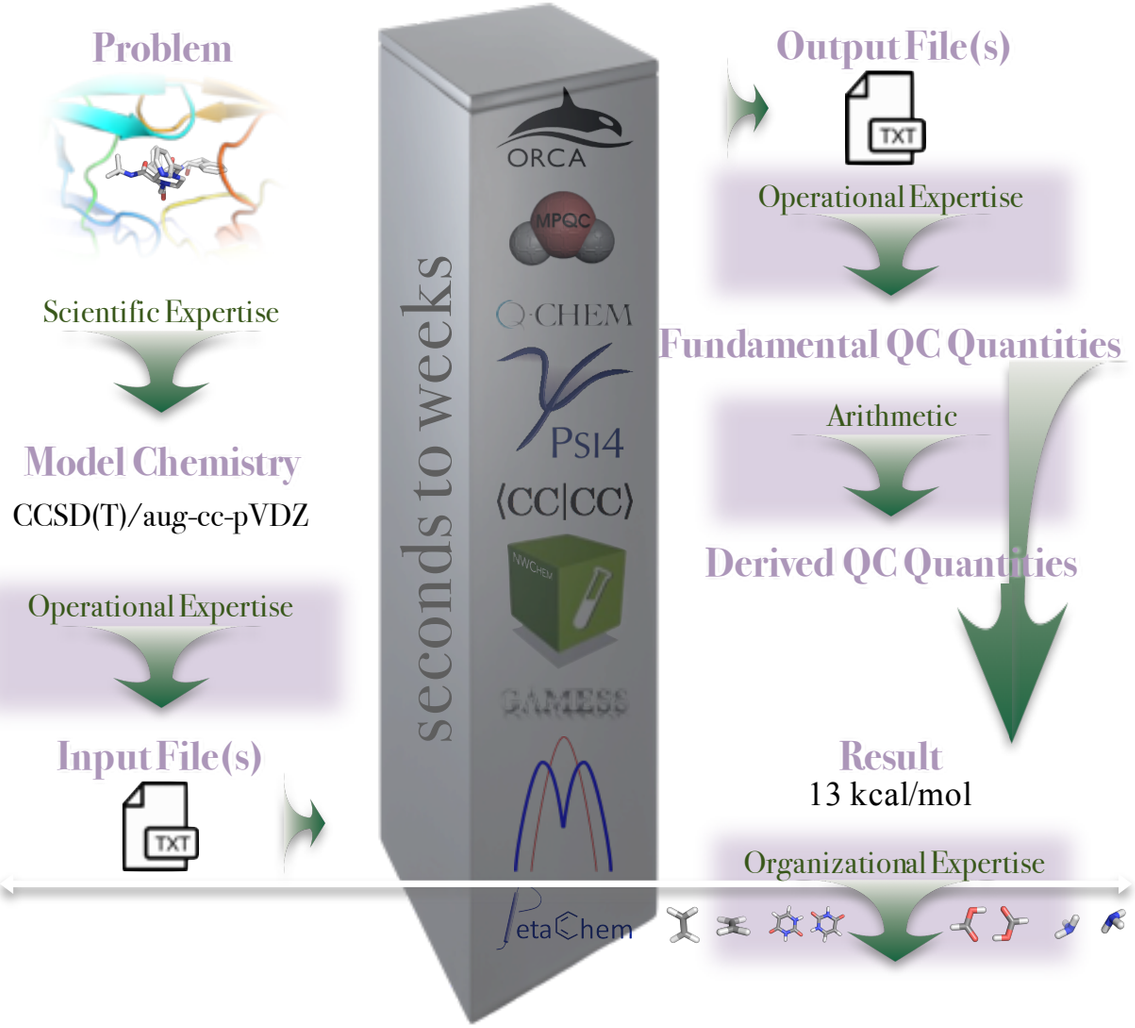
```
memory 4 gb

molecule h2o {
O  -1.51  -0.13   0.00
H  -1.95   0.73   0.00
H  -0.56   0.09   0.00
}

set basis aug-cc-pvdz

vpt2('b97-d3')
```

=



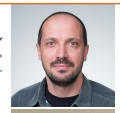
thanks to Prof. John Stanton & Dr. Devin Matthews for help with interface

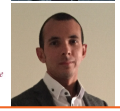
Targets of QCDb in the Research Workflow



WRENCH: A Simulation Workbench for Scientific Workflow Users, Developers, and Researchers

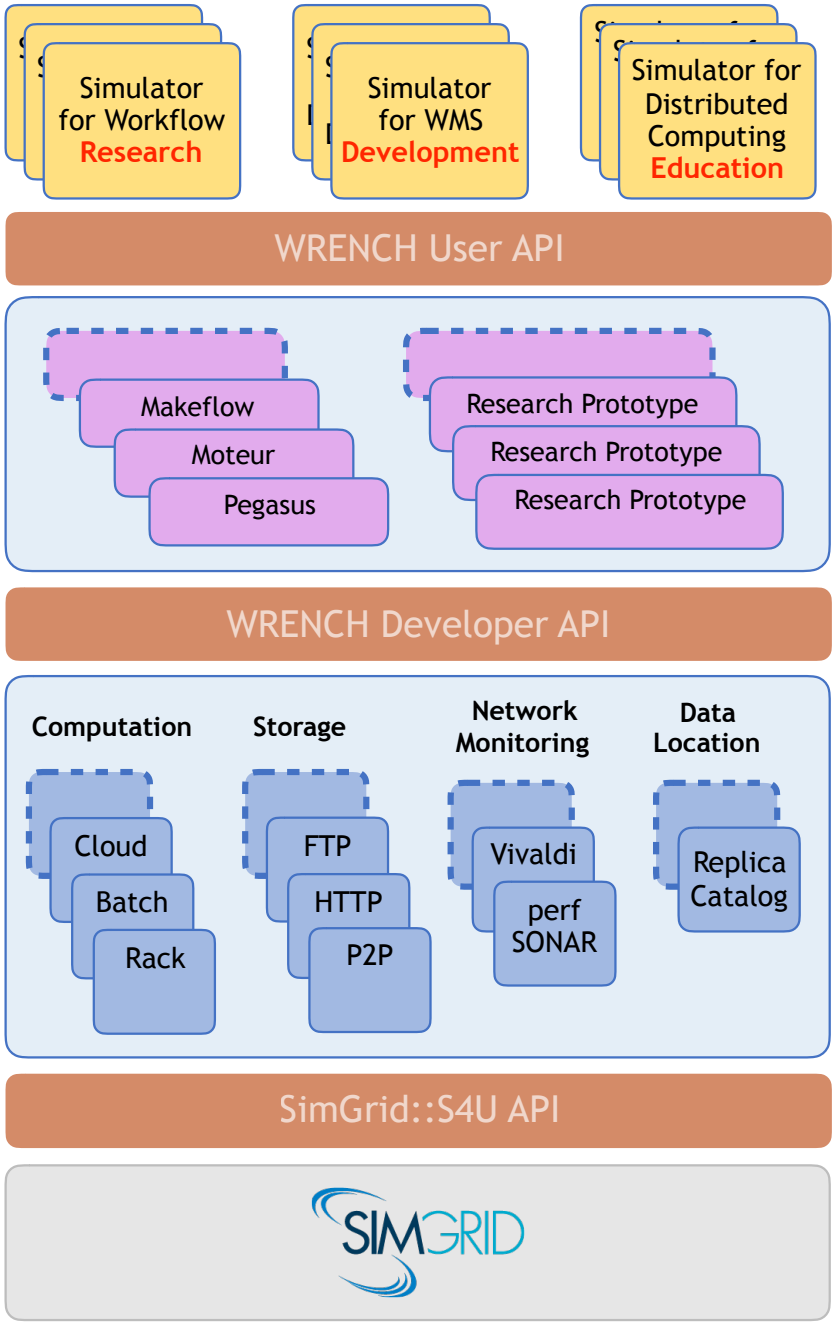
https://wrench-project.org


Henri Casanova


Rafael Ferreira da Silva

University of Hawai'i Mānoa

USC Viterbi School of Engineering Information Sciences Institute



MOTIVATION

- ✦ **Scientific Workflow** applications are important
- ✦ But applications, software infrastructures, hardware platforms are complex
- ✦ We need a solid **experimental science** approach to understand all this
- ✦ Yet real-world experiments have limited scope

OBJECTIVE

- ✦ Realize workflow execution **simulation** that is **accurate, fast,** and **scalable**
- ✦ Useful to
  - ✦ Domain (workflow) scientists
  - ✦ Workflow Manager Systems (WMS) developers
  - ✦ Educators

APPROACH

- ✦ Build on the decades of research and development in the SimGrid project
- ✦ Implement the software stack on the left

Jet Energy-Loss Tomography with a Statistically and Computationally Advanced Program Envelope (JETSCAPE)

Mohammad Ebrahim Khalaj

The JETSCAPE collaboration is tasked with the construction of a modular simulation tool and statistical data comparison package for high energy heavy-ion collisions. These collisions at Brookhaven National Lab. and at CERN create exploding droplets of matter which reach temperatures over a trillion degrees. At these temperatures neutrons and protons melt into a liquid plasma of quarks and gluons: the Quark Gluon Plasma (QGP). The modular JETSCAPE simulator, simulates all aspects of the collision of the ions: from the initial overlap, to the explosive expansion and evaporation into conventional matter. It allows theorists to assume a factorized approach by focusing on only one or two aspects of the evolution within the simulator, and modifying only those portions of the code base. It allows experimentalists access to a state-of-the-art event simulation tool to compare with experimental data, and to simulate their detector response.

OpenQBMM – www.opeqbmm.org

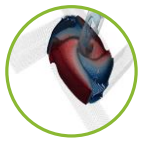
What is OpenQBMM?

A suite of libraries and solvers for OpenFOAM® to implement quadrature-based moment methods.

Features

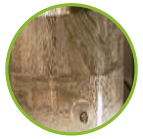
- Robust
  - Automatic enforcement of moment realizability
  - Moment-preserving advection schemes
  - Realizable integration of stiff source terms
- Validated
  - Test-case provided for each core component
  - Validation cases provided as example application for solvers

What problems can it solve?



Nanoparticle formation

(St << 1)



Gas-liquid systems with coalescence and breakup

St = O(1)

Polydisperse particulate systems (dense and dilute flows)

St >> 1

Full range of Stokes numbers

Broader impact and community

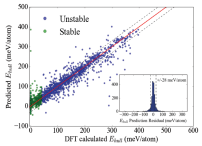
- 3 research groups at other institutions actively collaborate
  - University of Sherbrooke, Canada; Ecole Centrale Paris, France; University of Warwick, UK; Politecnico di Torino, Italy
- 6 external contributors
- Five graduate students supported (4 Ph.D., 1 M.Sc)
- 2 published journal articles (+ 2 in preparation)
- 24 invited talks
- Two training courses
- 3065 code builds (21 countries, non-unique IPs, ISU developers excluded)

SI2-SSI Collaborative Research: A Computational Materials Data and Design Environment

Dane Morgan (*Univ. of Wisconsin*), **ACI-Award 1148011**

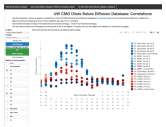
Developed tools/data for computational materials design:

The Materials Simulation Toolkit (MAST) for high-throughput defect and diffusion modeling[1]

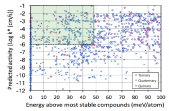


A Machine Learning extension (MAST-ML) to rapidly generate machine learning models from materials data[2].

Online defect and diffusion analysis apps on MaterialsHub[3].

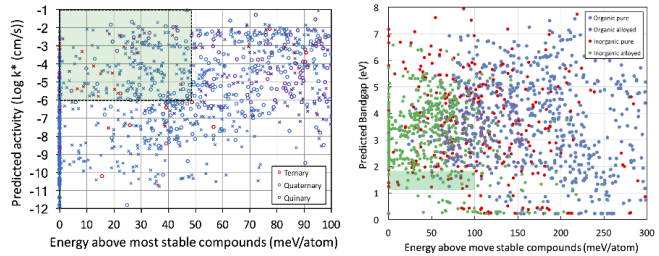


The world's largest computed and machine learning enhanced diffusion database with easy online search[4].

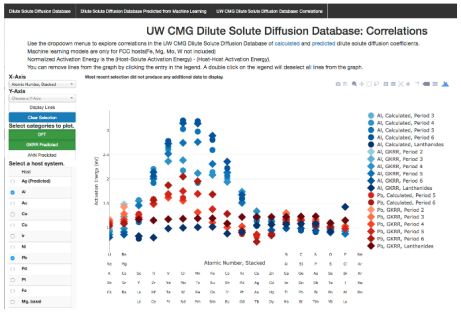


Valuable research results using these tools and data, e.g. new fuel cell materials[5].

Workforce training through the *Informatics Skunkworks*, an undergraduate materials informatics group[6].



New materials for fuel and solar cells.



Database of of diffusion coefficients.

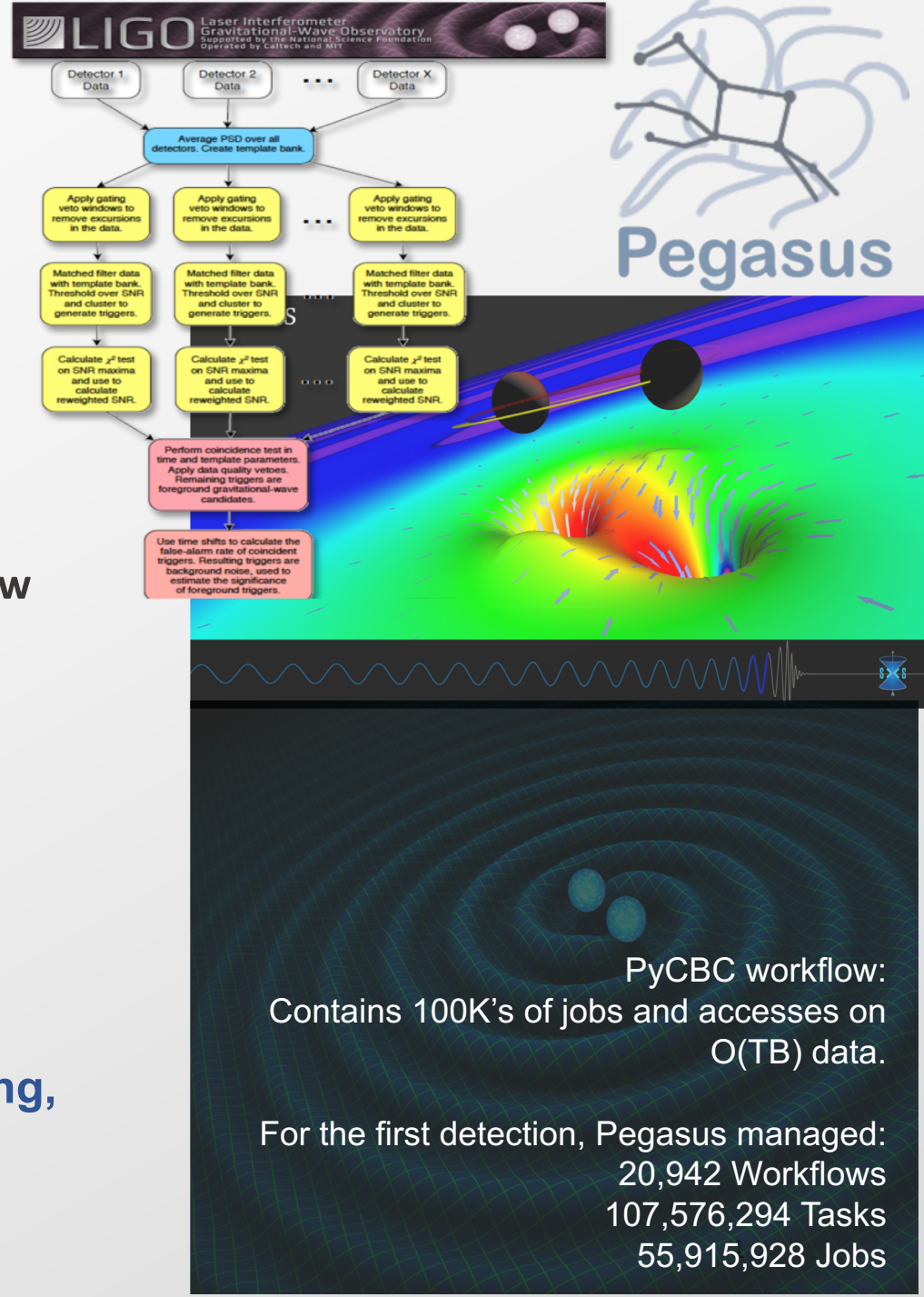
50+ undergraduates in Skunkworks

WISCONSIN innovation awards Finalist

[1] T. Mayeshiba, et al. Comp. Mat. Sci. '17; pypi.python.org/pypi/MAST ; [2] github.com/uw-cmg/MAST-ML ; [3] materialshub.org ; [4] diffusiondata.materialshub.org ; Wu, et al., *Scientific Data* '16 and *Comp. Mat. Sci.* '17; [5] E.g., Jacobs, et al., *Advanced Energy Materials* '18 ; [6] skunkworks.engr.wisc.edu/

Pegasus: Automate, Recover, Debug

- **Collaboration with Miron Livny, HTCondor**
- **Construct workflows in your favorite language (also Jupyter)**
- **Portability across heterogeneous infrastructure**
  - Separation of workflow description and execution
  - Support for campus and leadership class clusters, OSG, XSEDE, academic and commercial clouds
  - Can interact with a number of different storage systems (with different protocols)
  - Supports containers
- **Supports data reuse– useful in collaborations and ensemble workflow runs**
- **Reliability:** Recovers from failures, retry, workflow-level checkpointing
- **Scalability:** O(million) task, O(TB) data in a workflow
- **Restructures workflow for performance**
- **Supports reproducibility**
- **Web-based monitoring and debugging tools**
- **Can be included in various user-facing infrastructures**
- **Open source, available on Github**
- **Since 2001 used in astronomy, bioinformatics, climate modeling, earthquake science, molecular dynamics, helioseismology**
- **Funded by NSF under grant #1664162**



http://pegasus.isi.edu

Ewa Deelman, University of Southern California, PI

INNOVATIVE COMPUTING LABORATORY

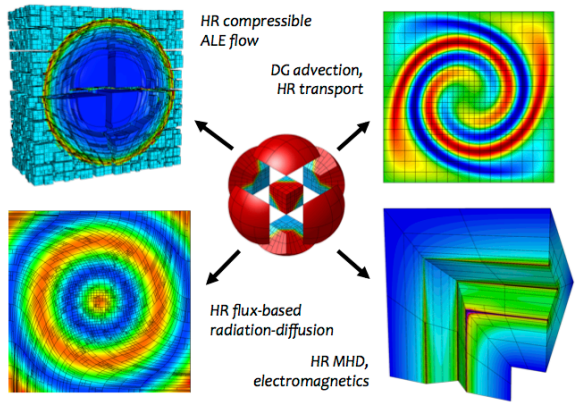
SI2-SSE: MATEDOR

MAtrix, TEnsor, and Deep-learning Optimized Routines

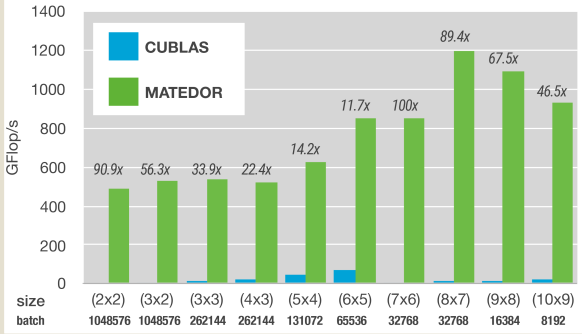
Azzam Haidar (PI)
Stanimire Tomov (Co-PI)
Ahmad Abdelfattah
Ichitaro Yamazaki
Jack Dongarra
UNIVERSITY OF TENNESSEE

Breadth of MATEDOR's impact on Application Domains

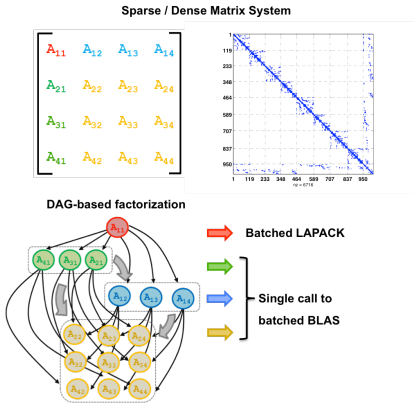
Tensor Contractions
High Order FEM & Applications



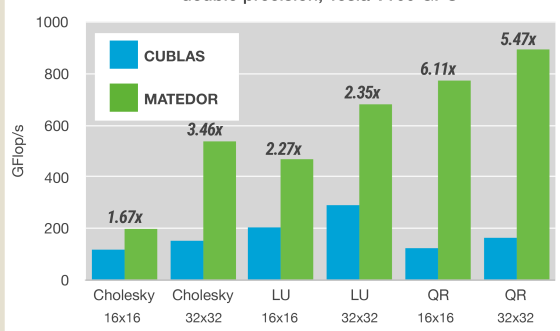
Tensor Contractions: computing $B^T D (BAB^T) B$, double precision, Tesla V100 GPU



Sparse/Dense
Solvers & Preconditioners



Batch Matrix Factorization, 100k matrices, double precision, Tesla V100 GPU



MATEDOR SCOPE

The project seeks to develop

- **Software Technologies and Standard Interface for Batched Routines,**
- **Sustainable and Performance-Portable Software Library**

for large-scale computations, but whose individual parts are **many very small matrices or tensor computations**.

The main target is the acceleration of applications from important fields that fit this profile, including deep learning, data mining, astrophysics, image and signal processing, hydrodynamics, and more.

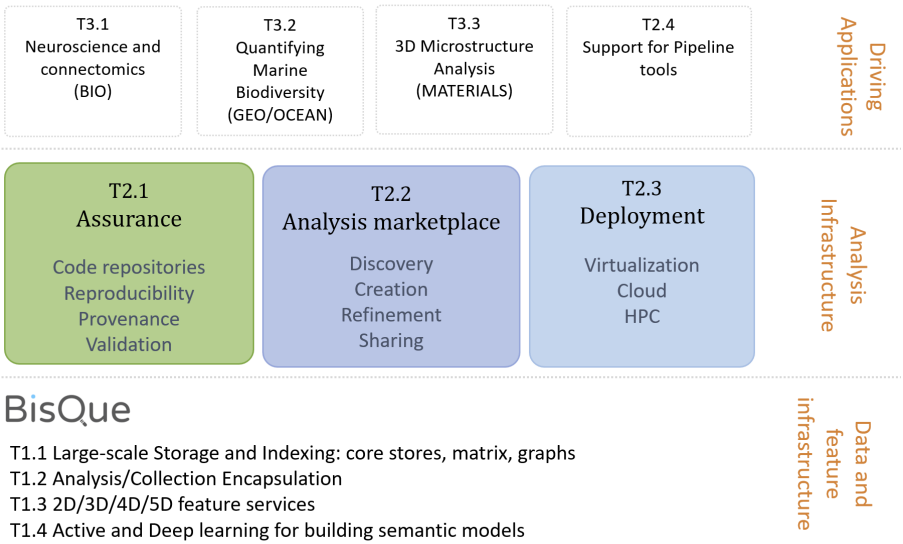
We really appreciate your input and would encourage collaboration with your applications

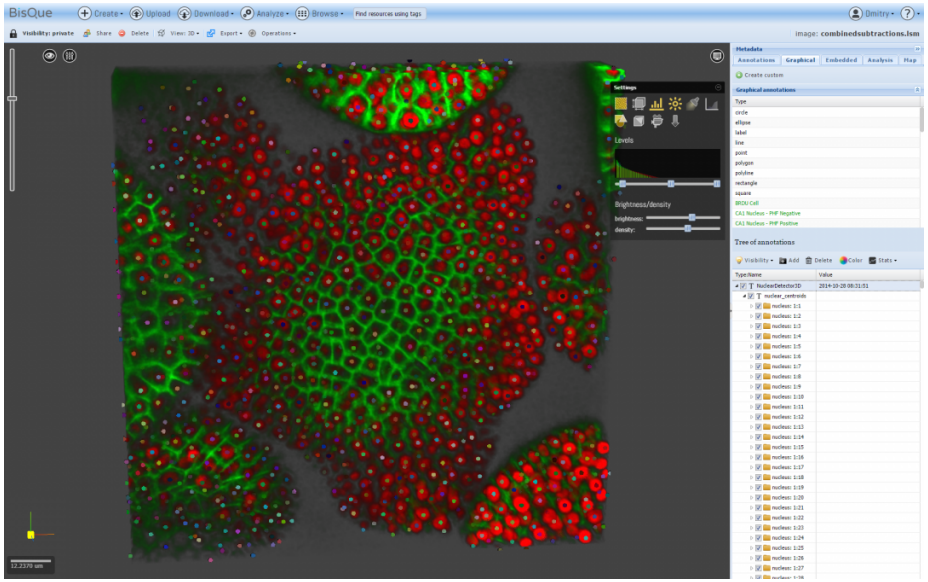
Contact: haidar@icl.utk.edu, tomov@icl.utk.edu

LIMPID/BisQue Overview

- LIMPID is built on cloud-based analysis platform BisQue
- Management, analysis, and sharing of images and metadata for large-scale data science
- Flexible and scalable query system across network of multimodal data items
- Module system for scalable integration of analysis tasks over images and metadata
- 200+ life science image and video formats
- Analysis marketplace: easy sharing and discovery of analysis modules

- *More information on BisQue:* http://bioimage.ucsb.edu/bisque





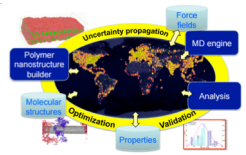
SI2-SSE Collaborative Research: Molecular simulations for polymeric systems in the cloud

PIs: Alejandro Strachan (Purdue) Coray Colina (Florida)
CoPIs: Benjamin Haley, Chunyu Li
Graduate students: Michael Fortunato, Lorena Alzate

Awards: 1440727
1613155

Approach and Goals

GOAL: enable pervasive, high-quality molecular simulations of polymers and their nanostructures



Develop a framework for molecular simulations of polymers and their nanostructures, universally accessible and useful to the community for cloud computing via NSF's nanoHUB.

1. Powerful simulation tools for polymer nanostructures (molecular builders, a parallel MD engine for property characterization and post-processing);
2. A UQ framework to orchestrate the molecular simulations and propagate uncertainties in input parameters to predictions and compare the predictions to experimental values;
3. Databases of force fields and molecular structures as well as predicted and experimental properties.

Impact: tool usage

PolymerModeler

1200+ users
42,000+ simulation runs

nuSIMM

320+ users
47,000+ simulation runs

Struc2LAMMPS: 100+ users

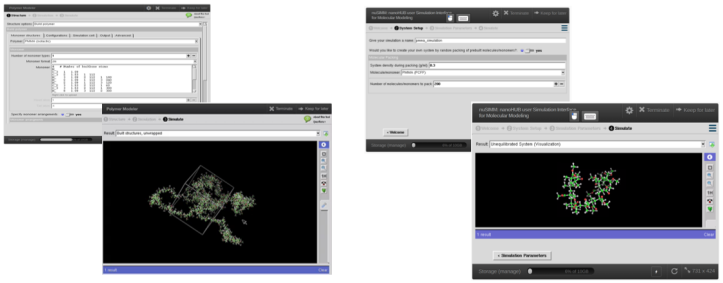
Use in research by outside groups

- Sahputra IH, Alexiadis A, Adams MJ. Molecular Simulation. 2018 Mar 24:1-7.
- Rzeznik L, Fleming Y, Wirtz T, Philipp P. Beilstein Journal of Nanotechnology. 2016 Aug 2;7(1):1113-28.
- Sebeck K, Shao C, Kieffer J. ACS Applied Materials & Interfaces. 2016 Jun 10.
- Rashidi V, Coyle EJ, Sebeck K, Kieffer J, Pipe KP. The Journal of Physical Chemistry B. 2017 Apr 24;121(17):4600-9.
- Sundarram SS, Li W. Polymer Engineering & Science. 2013 Sep 1;53(9):1901-9.
- Ingvason GA, Rollin V. In MRS Proceedings 2014 (Vol. 1700, pp. 61-66). Cambridge University Press.

Tool set for polymer simulations

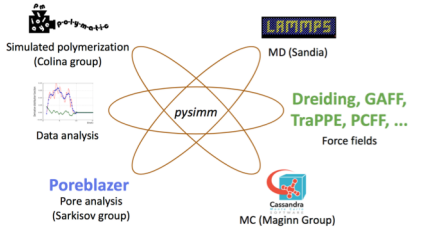
PolymerModel – https://nanohub.org/tools/polymod

nuSIMM – https://nanohub.org/tools/nusimm

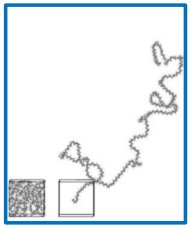


pySIMM – http://github.com/polysimtools/pysimm

Fortunato ME, Colina CM. pysimm: A python package for simulation of molecular systems. SoftwareX. 2017 Dec 31;6:7-12.



Self-avoiding random walk



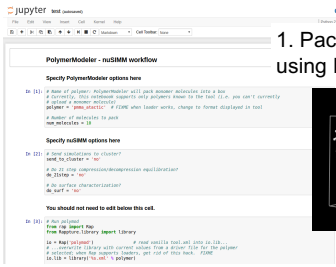
Setting up MD simulations

https://nanohub.org/tools/struc2lammpsdf

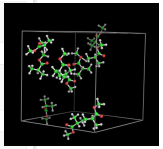
- Determine bond connectivity and topology
- Perform atom typing
- Create energy expression for LAMMPS
- Create required LAMMPS input files

Scientific workflows

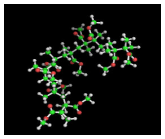
Jupyter notebooks running in nanoHUB



1. Pack monomers using PolymerModeler



2. Polymerize with nuSIMM

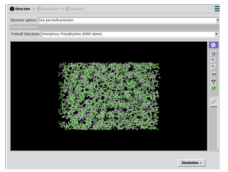


Jupyter notebooks:
- Combine text, graphics and live code
- nanoHUB notebooks seamlessly connect to simulation tools

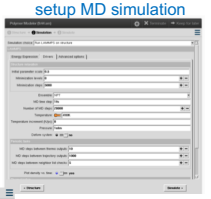
Impact: classroom use

Learning module: heat of fusion of PE

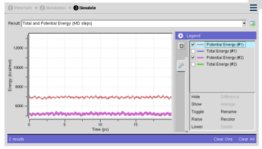
STEP 1: select initial structure



STEP 2: setup MD simulation



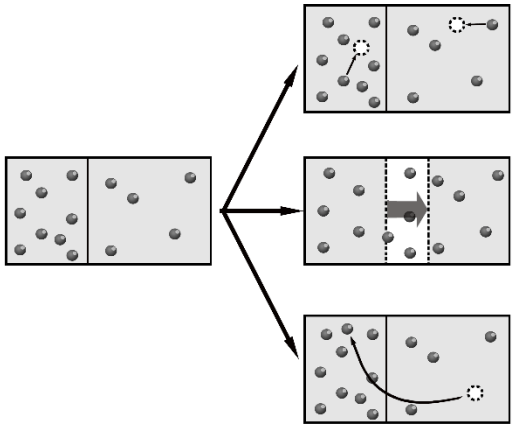
STEP 3: data analysis



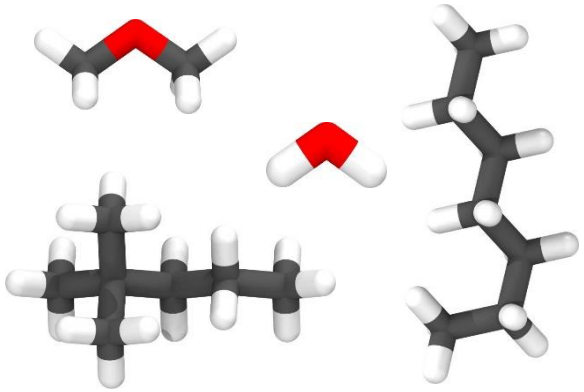
Students use MD simulations of heat of fusion together with DSC experiments to learn about crystallization dynamics

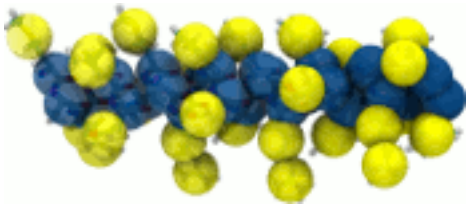
Classroom use:
- U of Florida CHM6586: Computational Chemistry. Total of 23 students, with 2 being undergrads.
- Purdue University MSE 697: Atomistic view of materials: Modeling and Simulations. Total of 18 graduate students.
- Purdue University MSE 235 Materials Properties Laboratory. Instructor: Prof. Michael Titus.

NVT, NPT, GCMC, and GEMC



Linear, Branched, Polar

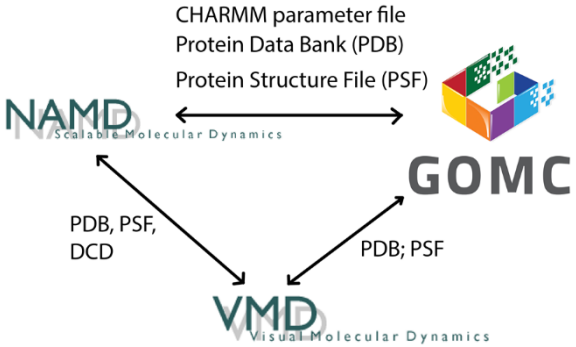




Martini
Charmm
OPLS
Mie
Custom Potential

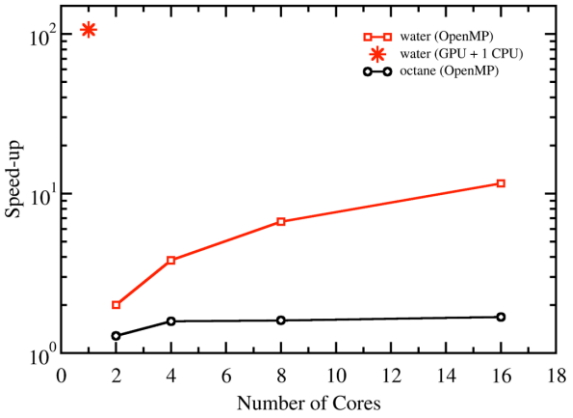
Moves Supported:

1. Displacement
2. Rotation
3. Swap
4. ID Exchange
5. Volume Exchange
    1. Isotropic
    2. Anisotropic



High Performance

- OpenMP+GPU



Award #1450459 :
Collaborative Research:
SS2-SSI: The Agave Platform:
An Open
Science-As-A-Service Cloud
Platform for Reproducible
Science

Agave Platform

https://agaveplatform.org

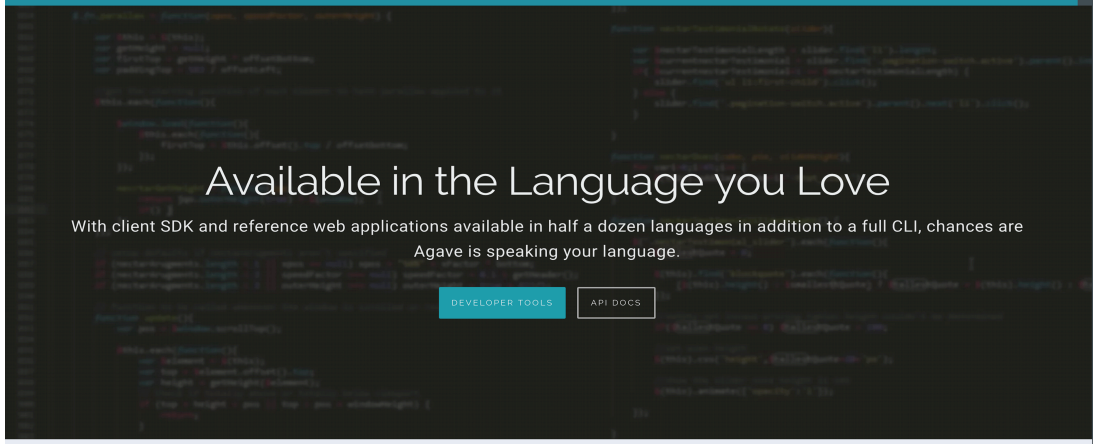
THE LEADING ALL-IN-ONE SCIENCE-AS-A-SERVICE PLATFORM FOR THE OPEN SCIENCE COMMUNITY

- Manage Data
- Run Code
- Collaborate Meaningfully
- Integrate Anywhere

Works with the academic and commercial research infrastructure you already use

XSEDE Extreme Science and Engineering Discovery Environment
Jetstream
amazon web services
Open Science Grid
CloudLab
Azure

ZERO INSTALLATION FOR DEVELOPERS AND END USERS

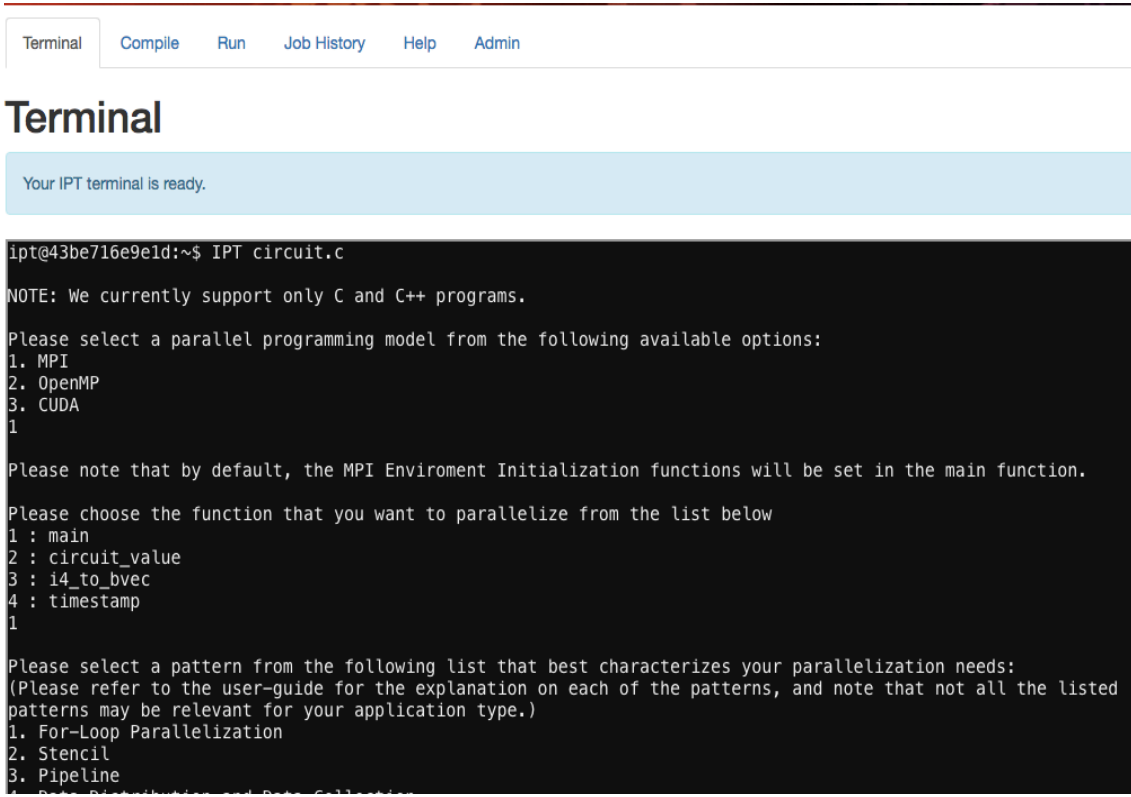


FULLY OPEN SOURCE AND FREE TO USE FOR THE OPEN SCIENCE COMMUNITY

IPT: INTERACTIVE PARALLELIZATION TOOL

https://ipt.tacc.cloud *NSF SI2:SSE Award # 1642396*

- IPT is a high-productivity tool for (1) semi-automatically parallelizing C/C++ code, (2) teaching parallel programming via demonstration
- Lowers the effort involved in parallel programming by more than 90% without significant loss in performance
- Deployed in the cloud – brings the parallel programming environment to a web browser
- Being used for workforce development in HPC – used in parallel programming trainings for TACC/XSEDE users



Leverages NSF investments in other projects:

Agave Platform

TACC / abaco

SGCI Science Gateways Community Institute

XSEDE

TACC

Contact: Ritu Arora, Email: rauta@tacc.utexas.edu

THE UNIVERSITY OF TEXAS AT AUSTIN

SSI: Inquiry-Focused Volumetric Data Analysis Across Scientific Domains: Sustaining and Expanding the yt Community

Nathan Goldbaum
NCSA, University of Illinois
ngoldbau@illinois.edu
yt-project.org
data-exp-lab.github.io

Enzo

Athena

FLASH

Ramses

MOOSE

Gasoline

ChaNGa

Gadget

Ingest Data

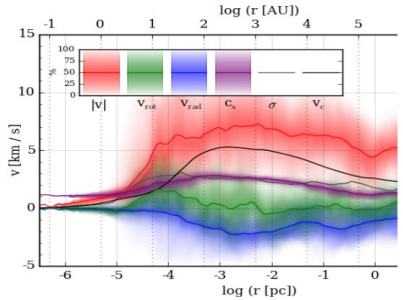
SciPy Ecosystem

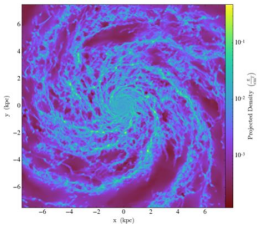
NumPy SciPy python pandas SymPy matplotlib

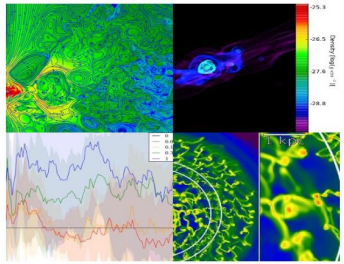
yt

NUMFOCUS
OPEN CODE = BETTER SCIENCE

Analysis/ Visualization

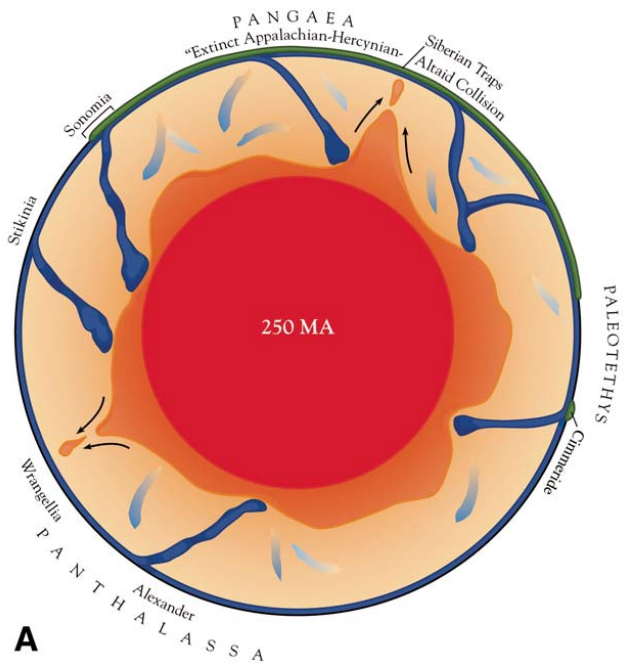






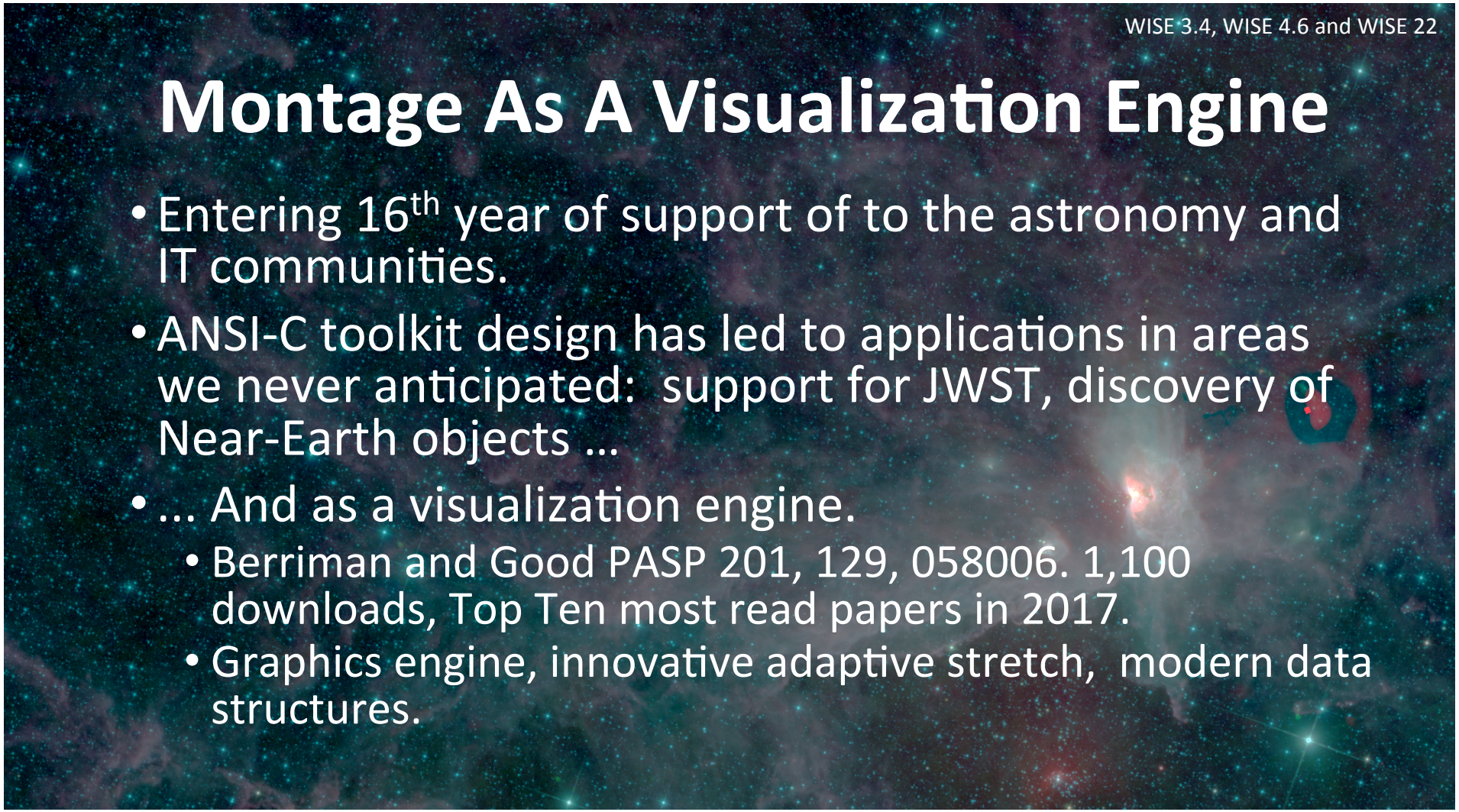
SSE's FOR MODELING PROCESSES IN THE MANTLE

Elbridge Gerry Puckett, University of California, Davis



Schematic cross section of Earth 250 million years ago

- Discontinuous Galerkin Method with Bound Preserving limiters
- Volume-of-Fluid interface tracking algorithm
- Active tracer particles
- First two SSEs are new in the computational mantle convection community
- New mathematical demonstrations that the particle algorithms converge to the *exact* solution of the Stokes equations
- All SSEs are implemented in the open-source, community driven code ASPECT



Montage As A Visualization Engine

- Entering 16th year of support of to the astronomy and IT communities.
- ANSI-C toolkit design has led to applications in areas we never anticipated: support for JWST, discovery of Near-Earth objects …
- … And as a visualization engine.
  - Berriman and Good PASP 201, 129, 058006. 1,100 downloads, Top Ten most read papers in 2017.
  - Graphics engine, innovative adaptive stretch, modern data structures.

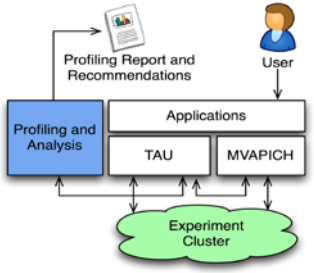
SI2-SSI (2018): Collaborative Research: A Software Infrastructure for MPI Performance Engineering: Integrating MVAPICH and TAU via the MPI Tools Interface

H. Subramoni, P. Kousha, A. Ruhela, S. Chakraborty, and D.K. Panda
The Ohio State University

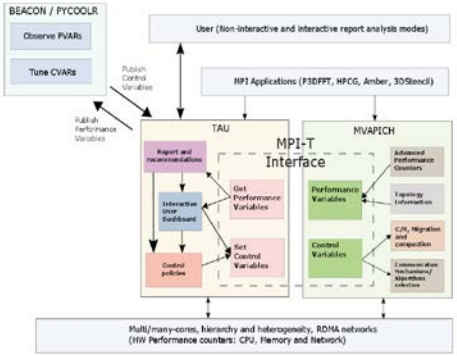
S. Shende, A. D. Malony, A. Maheo, and S. Ramesh
University of Oregon

Research Challenges

Creating an MPI programming infrastructure that can integrate performance analysis capabilities more directly, through the MPI Tools Information Interface, monitor Performance metrics during run time, and deliver greater optimization opportunities for scientific applications.



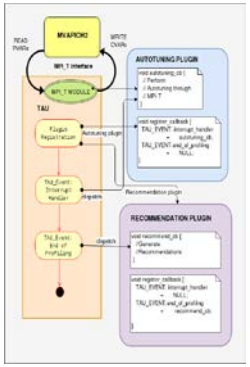
Proposed Approach



Overall Architecture

Design of the TAU Plugin Infrastructure and Associated Plugins

- Plugins are written in C/C++ and register callbacks for desired events
  - e.g. FUNCTION REGISTRATION, ATOMIC EVENT REGISTRATION, ATOMIC EVENT TRIGGER, INTERRUPT TRIGGER, END OF EXECUTION
- Design of Auto tuning Plugin
  - Registers callback for INTERRUPT TRIGGER
    - Triggered when TAU configured to collect PVARs at regular intervals
  - Use Case: Works in conjunction with MVAPICH2 to track use of VBUFs using PVARs and triggers CVARs to reduce VBUF usage on exceeding pre specified threshold
  - Benefit: Reduces additional memory consumed by unused VBUFs



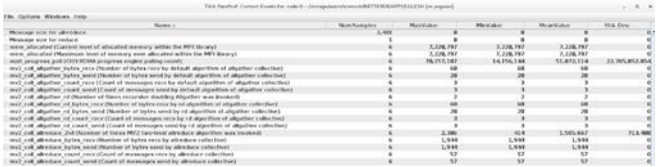
TAU Plugin Architecture

Design of Associated Plugins (Cont.)

- Design of Recommendation Plugin
  - Registers callback for END OF PROFILING
    - Triggered when TAU configured to collect PVARs at regular intervals
  - Use Case: Works in conjunction with MVAPICH2 to track use of Allreduce collective using PVARs
  - Benefit: Recommends the use of appropriate collective algorithms using CVARs based on profiling
    - e.g. Recommends use of SHARP-based allreduce through the CVAR MPIR_CVAR_ENABLE_SHARP if small message Allreduce is used frequently

Enhanced MPI_T Support in MVAPICH2 and Utilizing it in TAU

- PVARs to measure various aspects of collective algorithm usage has been added to MVAPICH2
  - e.g. bytes/messages sent/received
- Introduced support for new MPI_T based CVARs to MVAPICH2
  - MPIR_CVAR_USE_BLOCKING
  - MPIR_CVAR_RNDV_PROTOCOL
  - MPIR_CVAR_USE_SHARED_MEM



Listing of new PVARs added to MVAPICH2 as Displayed by ParaProf

Future Work & Research Dissemination

- Multiple releases of MVAPICH2 and TAU have been made that include support for PVARs, CVARs and associated interactions
- Future plans include:
  - Enhancing the MPI_T support in MVAPICH2 and co-designing TAU to take advantage of it
  - Study benefits of utilizing CVARs exposed by MVAPICH2 at application level at scale
  - Study challenges in providing an interactive performance engineering functionality for end users

Supported by
ACI-1450440, ACI-1450471, ACI-1053575, TG-ASC090010 & TG-NCR130002

THE OHIO STATE UNIVERSITY
MVAPICH
UNIVERSITY OF OREGON

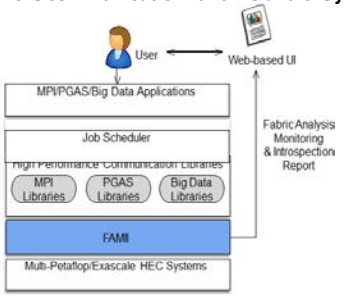
SI2-SSI (2018): FAMII: High-Performance and Scalable Fabric Analysis, Monitoring and Introspection Infrastructure for HPC and Big Data

Dhabaleswar K. (DK) Panda, Hari Subramoni, and Xiaoyi Lu
The Ohio State University
{panda, subramon, luxi}@cse.ohio-state.edu

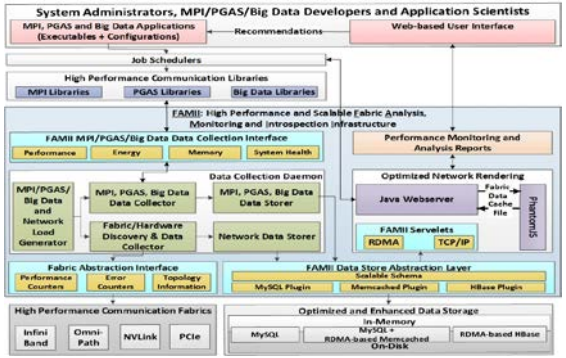
Karen Tomko
Ohio Technology Consortium
ktomko@osc.edu

Vision

Can a high performance and scalable tool be designed which is capable of analysing and correlating the communication on the fabric with behaviour of HPC/Big Data applications through tight integration with the communication runtime and the job scheduler?



Framework



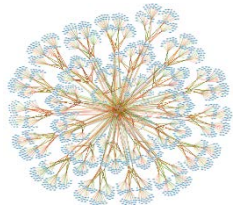
The Proposed Performance Monitoring, Analysis, and Introspection Framework

Release and Research Dissemination

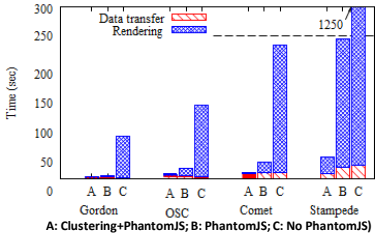
- A v0.9.3 release of the enhanced FAMII tool has been made on 03/16/18
  - http://mvapich.cse.ohio-state.edu/tools/osu-inam/
  - Over 150 downloads so far
- This release has been installed at OSC to monitor various clusters
- Installed and available on multiple clusters at OSU
- Other early adopters include NOAA and Pratt & Whitney

Optimized Network Rendering

- PhantomJS pre-renders and caches network by running the visualization module used for the network stabilization process
- Pre-rendered view updated in background when fabric is scanned to avoid stale views



Fully Expanded Visualization of the Comet Supercomputer at SDSC

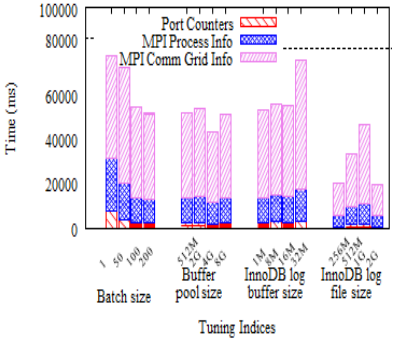


A: Clustering+PhantomJS; B: PhantomJS; C: No PhantomJS)

Performance breakup of UI load time with 100 Mbps Ethernet connection

Enhanced Data Store Abstraction Layer and Data Collection Daemon

- **Scalable Schema for Optimized Data Store**
  - Redesigned database schema to store communication grid information in single text field
- **Multithreaded Design for Data Collection Daemon**
  - Use separate database thread with separate connection to increase parallelism and reduce contention
- **Optimized MySQL with various tuning parameters**



Impact of MySQL optimization on performance of insert operations

Future Work

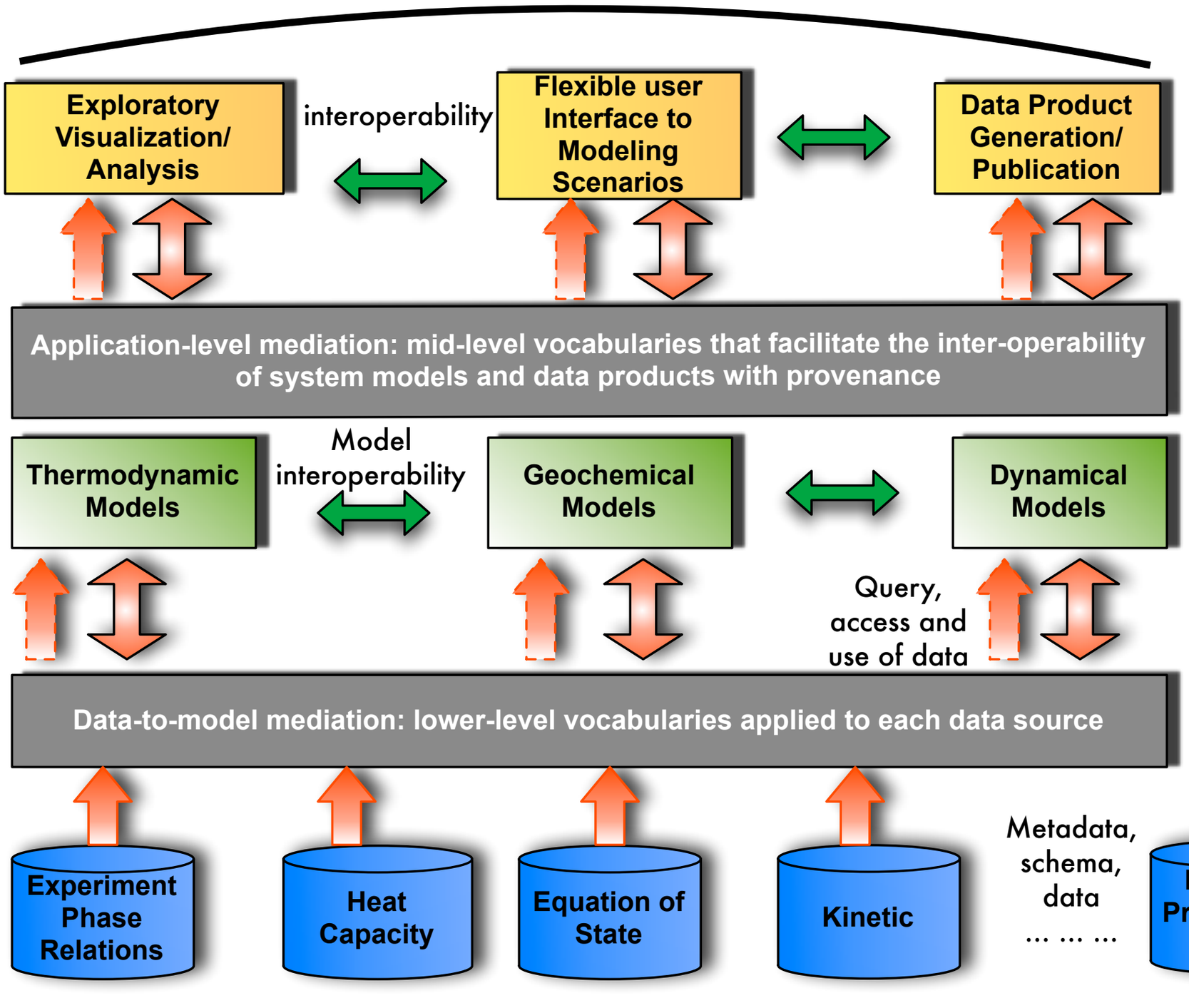
- Extend data collection daemon to gather intra-node topology, GPU-related information and update the data collection interface to gather information from the Omni-Path fabric
- Design Network and Application Load Generator
- Enhance PGAS/Big Data middleware to take advantage of FAMII

Supported by
OAC-1664137 & TG-NCR130002

ENKI: Software infrastructure that ENables Knowledge Integration for modeling coupled geochemical and geodynamical processes

PIs: MS Ghiorso, G Bergantz, P Fox, E Shock, M Spiegelman, D Sverjensky, A. Wolf

Web-based user interface and scripting environment



ENKI ecosystem:
More info at enki-portal.org

User interface built on:

- Jupyter Hub
- Jupyter Lab
- Custom lab extensions
- enki.ofm-research.org

Open source code:

- Python-based API
- Standard models
- Model calibration
- Workflow tracking
- GitLab: enki-portal

Data repositories:

- Support model calibration, expansion and testing

A Plug-and-Play Software Platform of Robotics-Inspired Algorithms for Modeling Biomolecular Structures and Motions

NSF ACI
SI2-SSE #1440581

Amarda Shehu
Computer Science

Erion Plaku
Computer Science

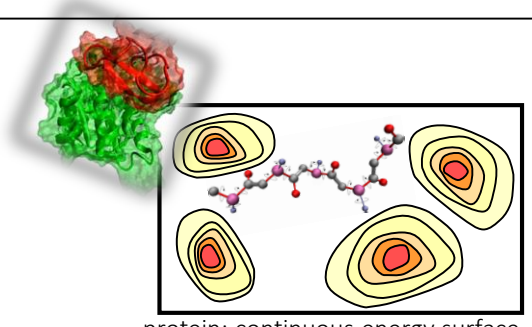
Adrian Roitberg
Chemistry

- Objective: Address algorithmic impasse on characterizing *(biological) form-to-function* relationship through a *plug-and-play platform of open-source software elements*.
- Premise: Address by integrating algorithmic efforts of AI researchers on *search and optimization* and modeling efforts of biophysics researchers on molecular mechanics and energetics.

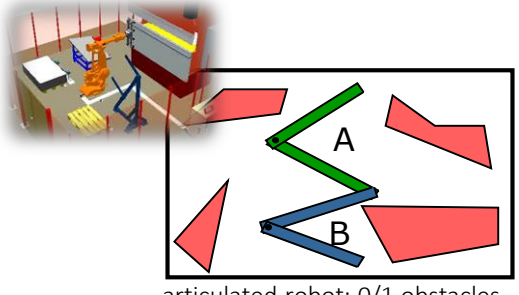
Turing, AM. (1952) Chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237(641):37-72.

Application Setting:
Protein and peptide structure and dynamics.

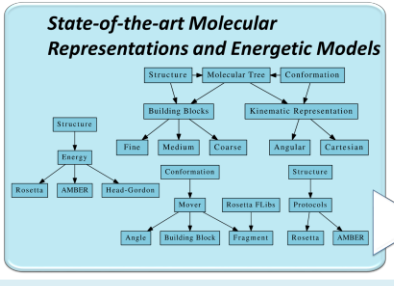
Software elements integrate search and optimization algorithms inspired from robot motion planning with sophisticated molecular models grounded in the latest understanding of protein biophysics.

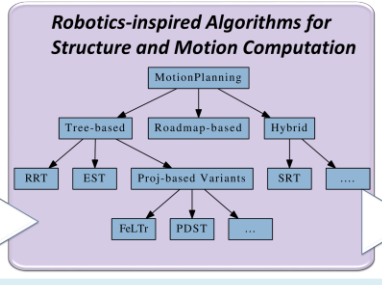


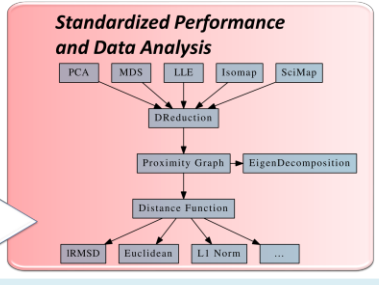
protein: continuous energy surface



articulated robot: 0/1 obstacles





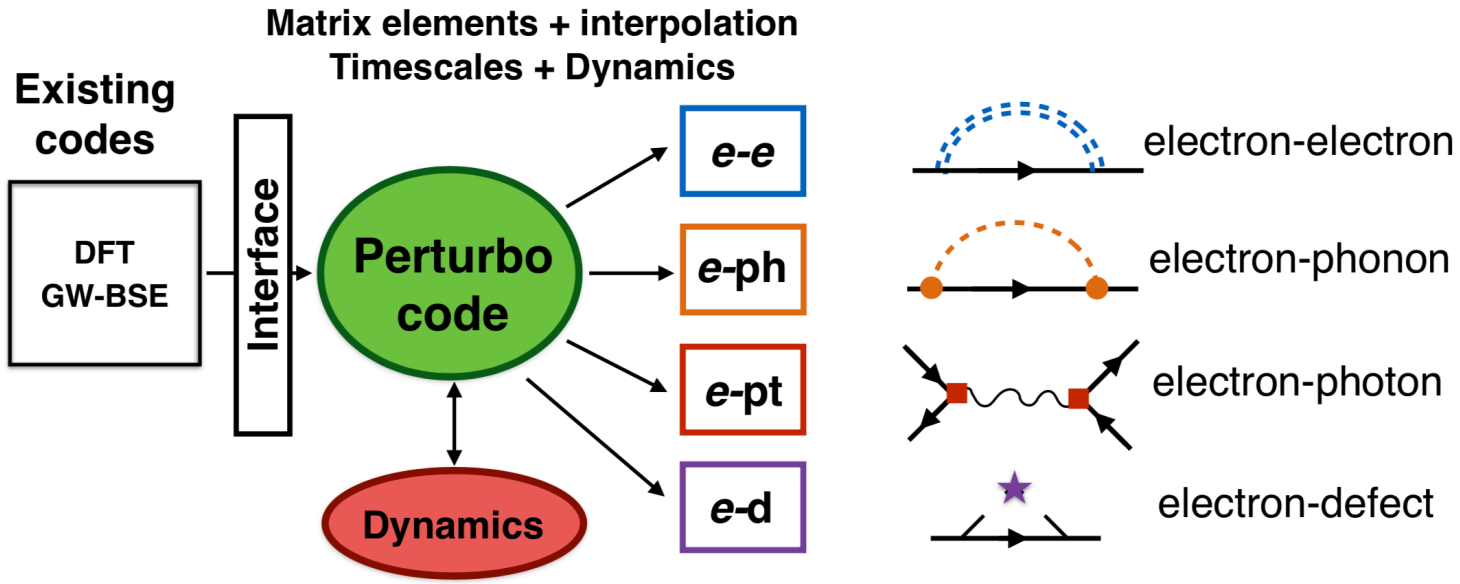


Dissemination:

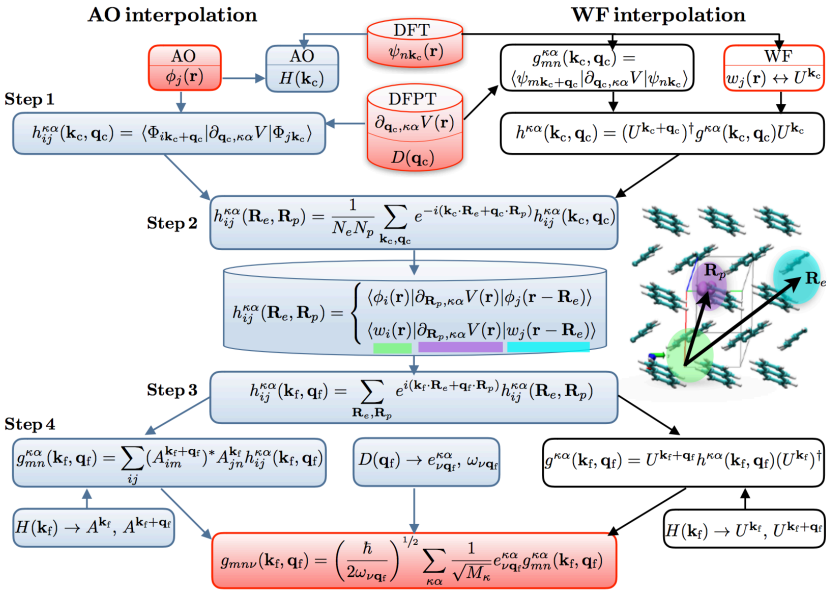
- Open-source software (ROMEO) on gitlab.
- Workshops and tutorials (on ROMEO) at ACM and IEEE conferences.
- Testing labs of research collaborators (INOVA, NCI, CUA, NEU, UPENN).
- Educational partners (UMass Boston, ODU, Lehigh, UNM).
- Research articles (BMC Genomics 2018, JCB 2017, JAIR 2016, IEEE/ACM TCBB 2016, Robotica 2016, IEEE NanoBioScience 2016, and others).

Project Page: http://cs.gmu.edu/~ashehu/?q=SI2-SSE15Project
Contact Info: amarda@gmu.edu

PERTURBO: A software platform for accelerated discovery of microscopic processes in materials

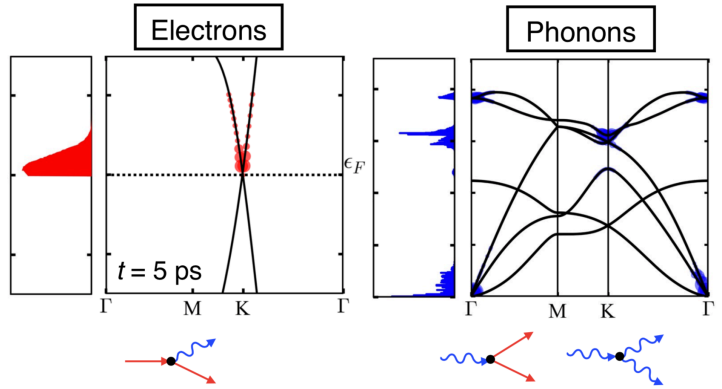


Workflow for computing the e-ph matrix



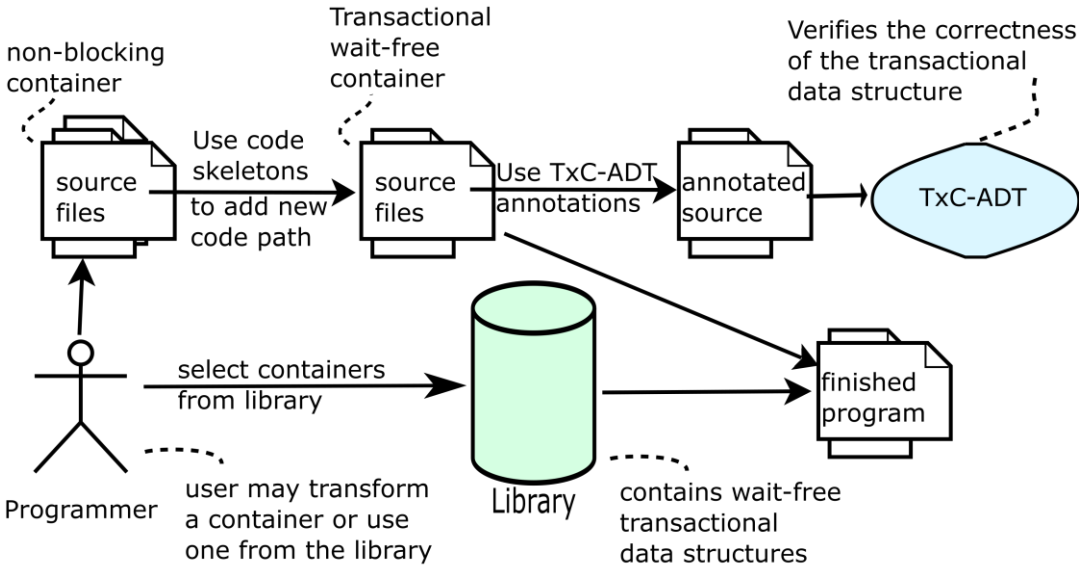
- Ab initio transport and ultrafast dynamics: hot carrier cooling, carrier conductivity, spin-flip dynamics, …
- e-ph: Atomic orbital & Wannier fn. interpolation, long-range e-ph, spin-orbit coupling, anharmonicity.
- Stand-alone code, Python interface+Fortran routines, MPI+OpenMP parallelization

Coupled dynamics of excited electrons and phonons



Transactional Lockless Data Structures

- TLDS is a framework for the development of transactional data structures
- Key Components
  - Scalable methodology for transforming non-blocking data structures into transactional data structures
  - Library of transactional data structures
  - Tool to validate their correctness
- Purpose: to enable software developers to easily construct highly scalable applications
- Potential use cases
  - In-memory databases
  - Scientific applications
  - Data analysis applications
  - Web servers

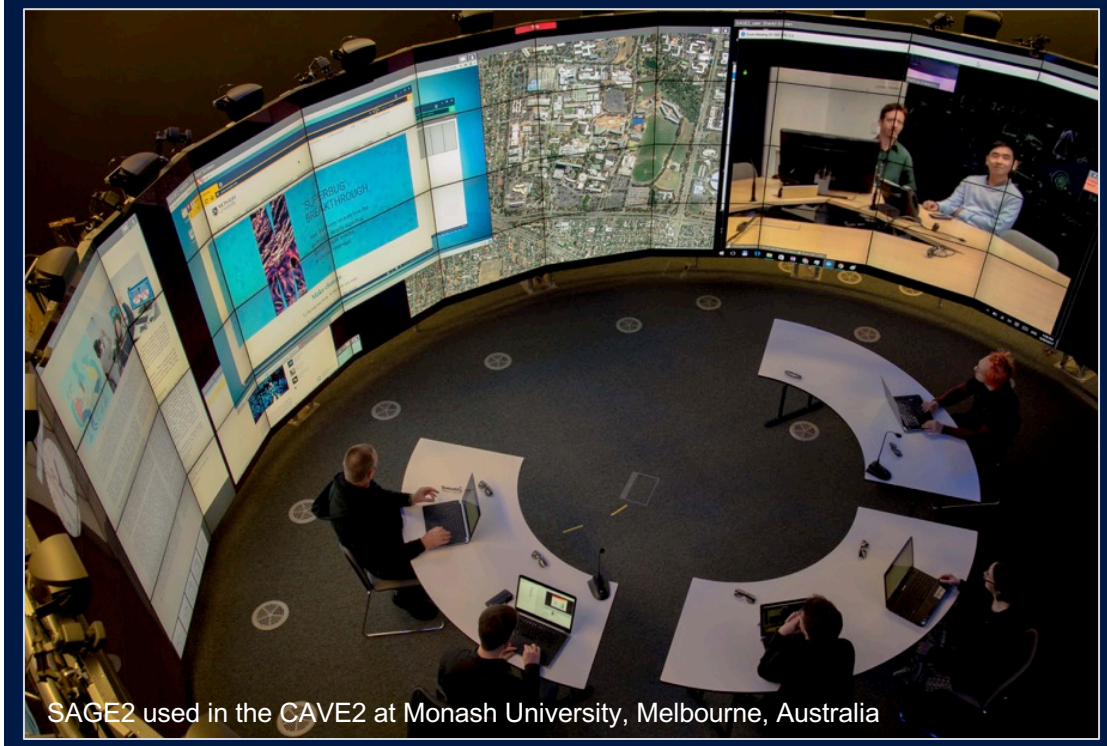


SI2-SSI: SAGE2 (Scalable Amplified Group Environment)

Jason Leigh, University of Hawai'i at Mānoa
Maxine Brown, Luc Renambot, University of Illinois at Chicago

SAGE2 is a web-based, user-centered platform for small groups or distributed teams to access digital media datasets from various sources and share and investigate content on display walls. SAGE2 builds on SAGE, previously funded by NSF.

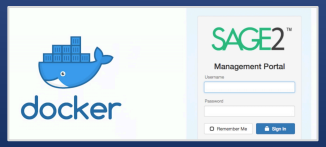
By making cyberinfrastructure more accessible to end systems and to end users, both in the lab and in the classroom, SAGE2 is transforming data visualization, data exploration and collaboration.



SAGE2 used in the CAVE2 at Monash University, Melbourne, Australia

SAGE2 features:

- SAGE2Cloud sets up multiple servers



- Jupyter Notebooks and cells (sections of notebooks) sharing

- Voice Commands

~90 SAGE2 sites worldwide since 2015 and over 160 SAGE sites (as of 2014)

www.sagecommons.org

Laboratory for Advanced Visualization & Applications
Univ. of Hawai'i at Mānoa

Electronic Visualization Laboratory
Univ. of Illinois at Chicago

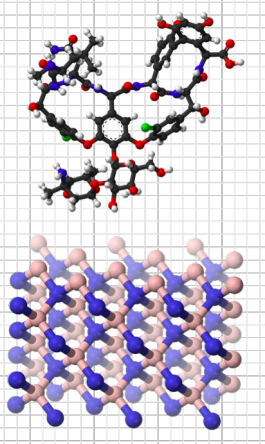
Award 1441963

Software Framework for Electronic Structure of Molecules and Solids

Garnet Kin-Lic Chan — Caltech
Toru Shiozaki — Northwestern University
Edward F. Valeev — Virginia Tech

pyscf | BAGEL | MPQC

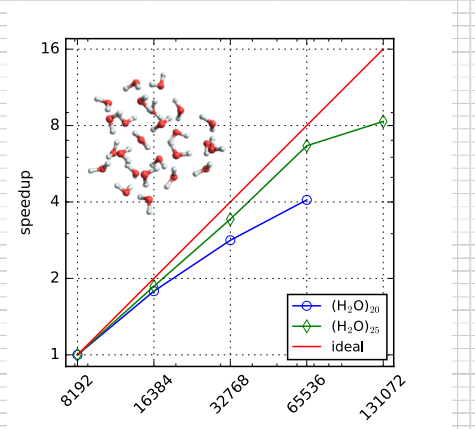
molecules and solids



reduced complexity formalisms

EMoS

high-level composition in C++/Python



high performance with 1 … 10^5 cores

Expanding Volunteer Computing

- Volunteer computing provides ~100 PetaFLOPS, to ~50 scientists
- Project goals: increase both numbers, especially the 50
- Add BOINC back ends to TACC, nanoHUB
- Create new “non-nerd” volunteer interface based on science areas rather than projects

TrajAnalytics: A Cloud-Based Visual Analytics Software System to Advance Transportation Studies Using Emerging Urban Trajectory Data.

Ye Zhao

Massively-Parallel Real-Time TDDFT Modules for Non-Equilibrium Electron Dynamics

NSF: SI2-SSE project (Oct. 2017~)

Yi Yao, Alina Kononov, Erik Draeger (Collaborator), Andre Schleife (co-PI), Yosuke Kanai (lead-PI)

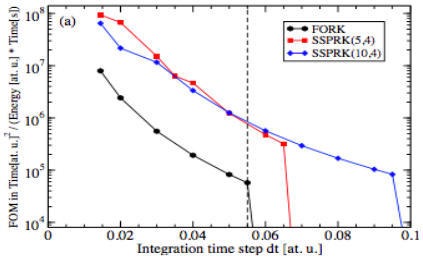
Real-Time Time-Dependent Density Functional Theory (RT-TDDFT) Simulations

$$\left\{ i\hbar \frac{d}{dt} \left| \phi_i(t) \right\rangle = \hat{H}_{KS} \left| \phi_i(t) \right\rangle \right\}_{i=1..N} \qquad \hat{H}_{KS} \equiv -\frac{\hbar^2}{2m_e}\nabla^2 + \hat{V}_{ext}(t) + \hat{V}_H[\{\phi_i(t)\}] + \hat{V}_{XC}[\{\phi_i(t)\}]$$

Coupled non-linear PDEs w/ millions of PWs for representing the single-particle states. $\phi_i(\mathbf{r},t) = \psi_{n\mathbf{k}}(\mathbf{r},t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}}^{\mathbf{G}_{max}} C_n(\mathbf{G},\mathbf{k},t) e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$

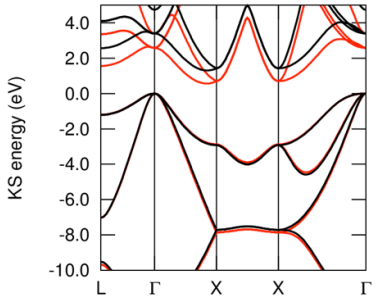
Challenges addressed through development of new modules in Qb@ll code

Highly scalable explicit integrators

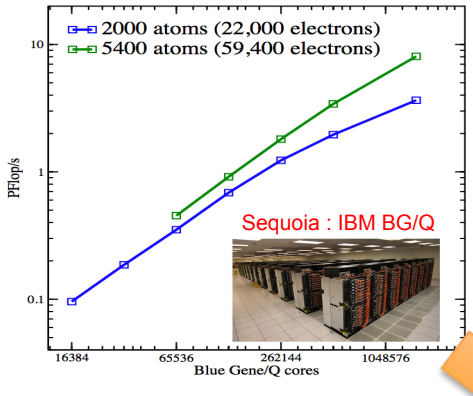


Recent *Strong Stability-Preserving Runge-Kutta (SSPRK)* methods to reduce "time-to-solution".

Advanced and efficient approximation to V_{xc}



Recent *Strongly Constrained and Approximately Normed (SCAN)* approximation to improve accuracy at a reasonable increase in computational cost.



HPC Software Infrastructure

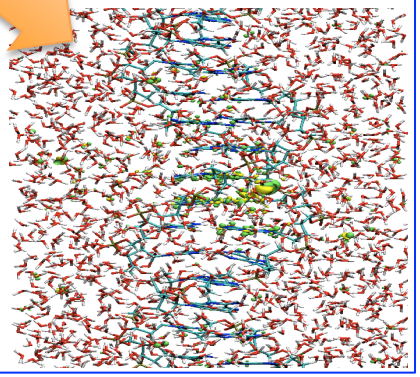
>40% peak performance
@ 1.6 million cores

- Massively parallel and tailored to modern HPCs with hybrid MPI/open-MP/SIMD.

Scientific Impacts

Simulating electronic excitation dynamics in solvated DNA under proton irradiation.

> 12,000 electrons
> 6,000,000 PWs
> 131,000 cores on BG/Q



Is your Science Hijacked by your compiler?

Ganesh Gopalakrishnan
School of Computing
University of Utah
Salt Lake City, UT 84112

PhD Student : Michael Bentley

Lab Collaborator : Dong H. Ahn (LLNL)

Enabling Multiscale and Multiphysics Applications in Fluid Dynamics, Solid Mechanics, and Fluid-Structure Interaction

Boyce E. Griffith, University of North Carolina at Chapel Hill, Matthew G. Knepley, University at Buffalo, and Neelesh A. Patankar, Northwestern University

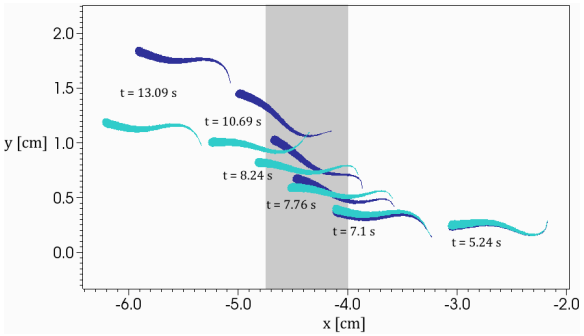
Objectives: To advance the ***IBAMR*** software infrastructure for simulating **fluid-structure interaction (FSI)** using the **immersed boundary (IB)** method with **adaptive mesh refinement (AMR)** through specific projects on:

- Numerical Methods;
- Solver Infrastructure; and
- User Interface Tools

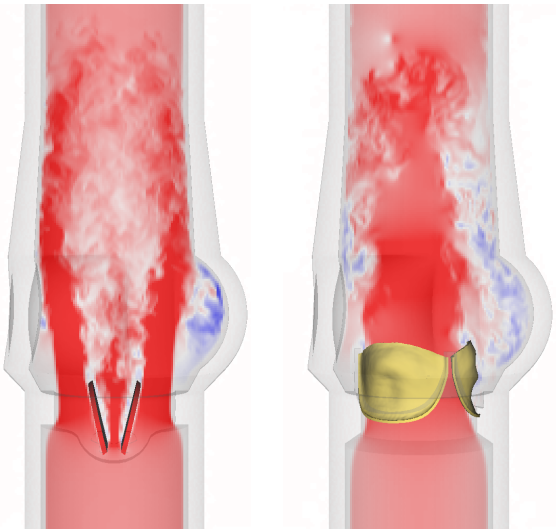
This work is coupled to applications in *aquatic locomotion*, *cardiac electromechanics*, and *esophageal transport*.

Sustained Impacts: ***IBAMR*** has been used at more than 20 colleges and universities in the U.S. and internationally and at the U.S. *Food and Drug Administration*.

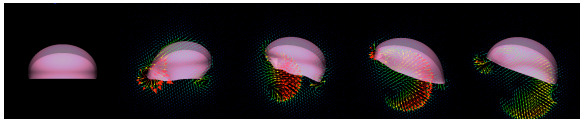
`ibamr.github.io`



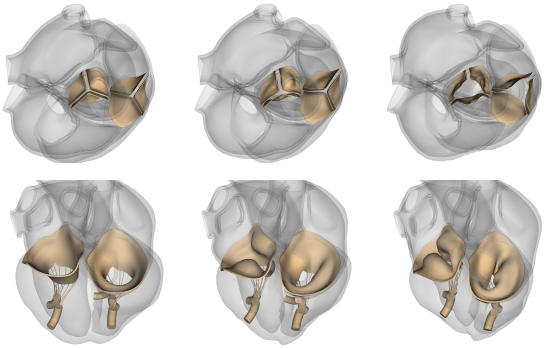
Swimming through an upward jet with (*cyan*) and without (*blue*) feedback.



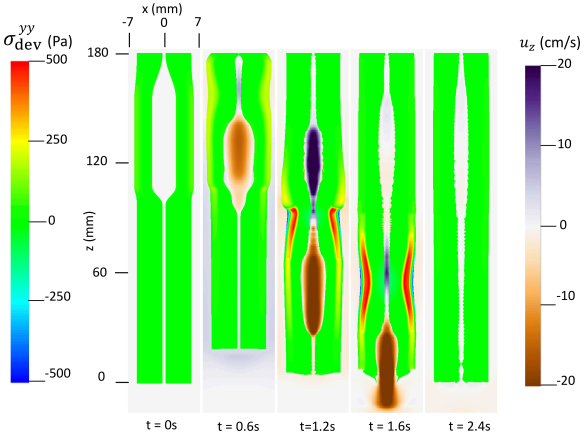
Fluid dynamics of mechanical (*left*) and bioprosthetic (*right*) heart valves.



Jellyfish turns and maneuvers.



CT-based FSI heart model.



Bolus transport and wall stress in a model of esophageal transport.

A United Theory of van der Waals forces in Non-Local Density Functional Theory

T. Thonhauser – Wake Forest University

NSF DMR 1712425

SLACKHA: Software Library for Accelerating Chemical Kinetics on Hybrid Architectures

PIs: Kyle E. Niemeyer (Oregon State Univ.) & Chih-Jen Sung (Univ. of Connecticut)

Students: Nicholas Curtis @ Univ. of Connecticut; Andrew Alferman, Parker Clayton, Morgan Mayer, Phillip Mestas @ Oregon State Univ.

Goals

accelerate the computation of chemical kinetics in simulations of reactive fluid flows, optimized for hybrid CPU/GPU processing architectures

Why?

- Cost of using detailed chemical kinetic models
- Trend towards heterogeneous computing with accelerators

Mature Components

`pyJac`: open-source Python package that generates source code used to analytically calculate Jacobian matrices for chemical kinetics

`accelerInt`: open library of vectorized solvers, usable on heterogeneous architectures (CPU, GPU, MIC, …)

Impact

- Users from combustion in academia, national labs
- Used in study of Langmuir turbulence effects on upper ocean carbonate chemistry
- Currently being used to accelerate large eddy simulation of reacting bluff-body flow

`http://slackha.github.io/`

`https://groups.io/g/slackha-users`

SI2-SSE: Open OnDemand: Transforming Computational Science through Omnidisciplinary Software Cyberinfrastructure

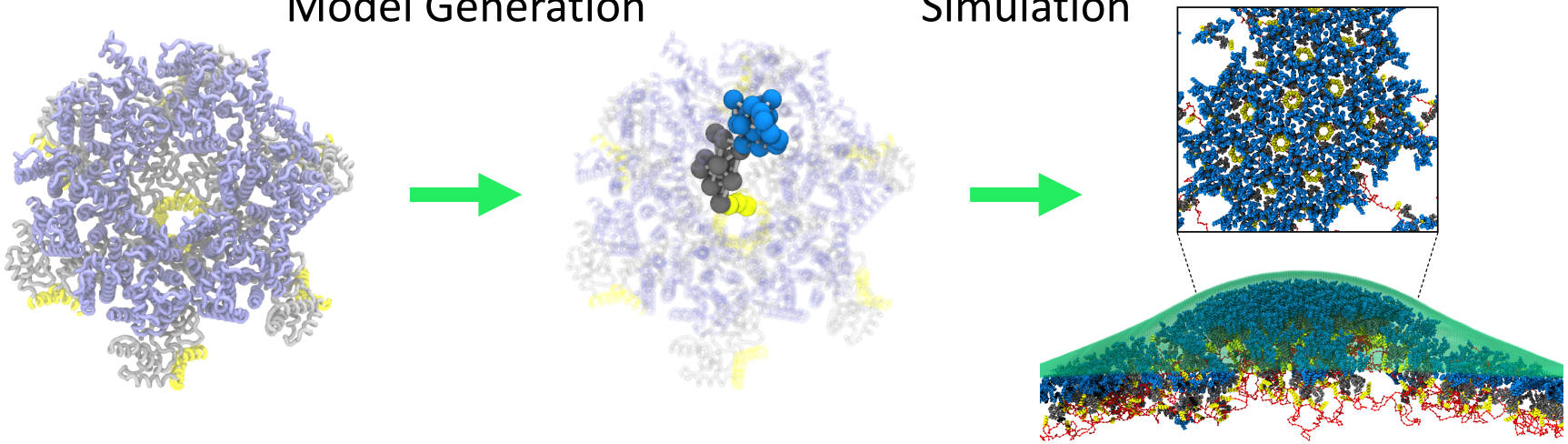
David Hudak

SI2-SSE: Highly Efficient and Scalable Software for Coarse-Grained Molecular Dynamics

PI: Gregory A. Voth

Model Generation

Simulation



Key Challenges

- Systematic CG development is difficult and fragmented
- Bottlenecks slow simulation implementation and runtime

Our Approach

- Integrated tools for bottom-up CG development and simulation
- Open-source network to enable community-driven development

THE UNIVERSITY OF CHICAGO

Grant: OAC-1740211

THE VOTH GROUP
THE UNIVERSITY OF CHICAGO | DEPARTMENT OF CHEMISTRY

Development of computational methods for the characterization of novel strongly correlated materials: from DFT to DMFT

NSF SI2 PI meeting, Washington D.C. (2018)

Hyowon Park (The University of Illinois at Chicago),
Aldo Romero (West Virginia University)

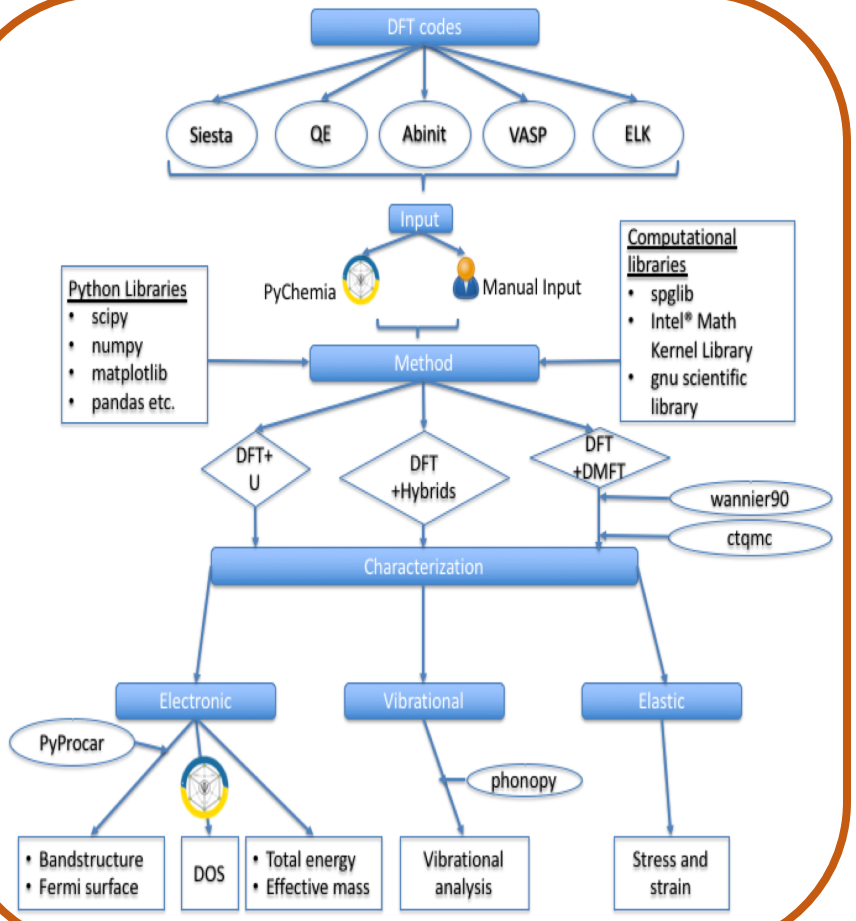
Motivations

- Dynamical mean field theory (DMFT) : a powerful method for novel strongly correlated materials.
- DMFT can be interfaced with density functional theory (DFT) for describing materials with both weakly and strongly correlated electrons.
- Popular DFT codes has different interfaces for inputs and usual DFT equation scales as $O(N^3)$.
- Combining DMFT with different DFT packages can be **problematic** due to **different choices of correlated orbitals** and **arbitrary choices of interaction parameters**.

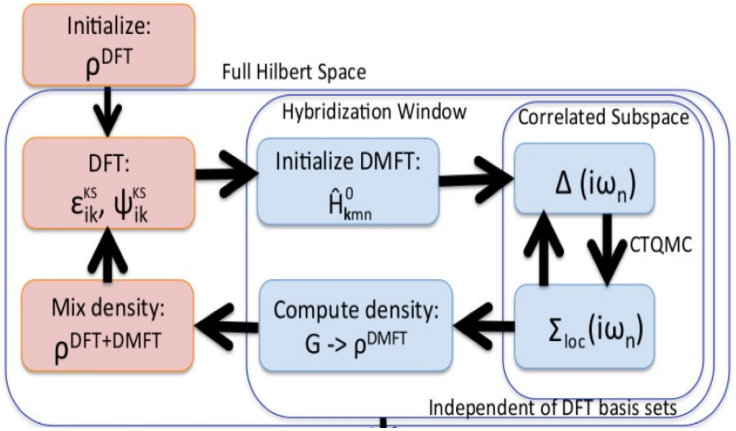
New Implementations

- We build a new software package for DFT+DMFT which can be **scaled as O(N)** for large-scale strongly correlated materials.
- Our DFT+DMFT code is **user-friendly and open-source** interfaced with efficient DFT packages.
- We study electronic, vibrational, and elastic properties of strongly correlated materials by implementing **total energy and atomic force calculations** using DFT+DMFT.
- As proof of concept, we will show the band structure and the structural phase diagram of strongly correlated nickelates computed using DFT+DMFT

Flow Chart



DFT+DMFT

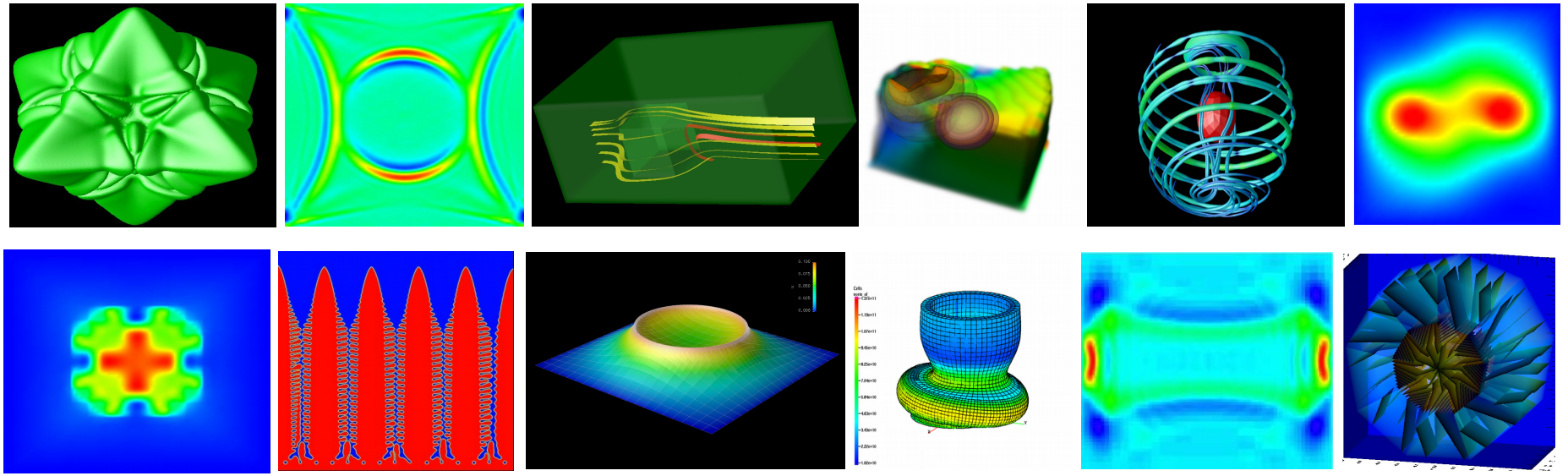


Future Works

- Implementation of the interatomic forces within DFT+DMFT and extend the functionality to calculate phonons and stress of strongly correlated materials.
- Interface of the DMFT code with free licensed DFT packages including Siesta, Abinit, and QE.
- Application of our DFT+DMFT code to the study of strongly correlated materials including Gd and oxides.
- Two-particle susceptibility calculations including the optical conductivity and the magnetic susceptibility.
- Fermi surface calculations.

deal.II

A library for finite element computations that supports...



...a large variety of PDE applications tailored to non-experts.

Fundamental premise:
Provide building blocks that can be used in many different ways, not a rigid framework.

www.dealii.org

Conceptualization of an S2I2 Institute for High Energy Physics

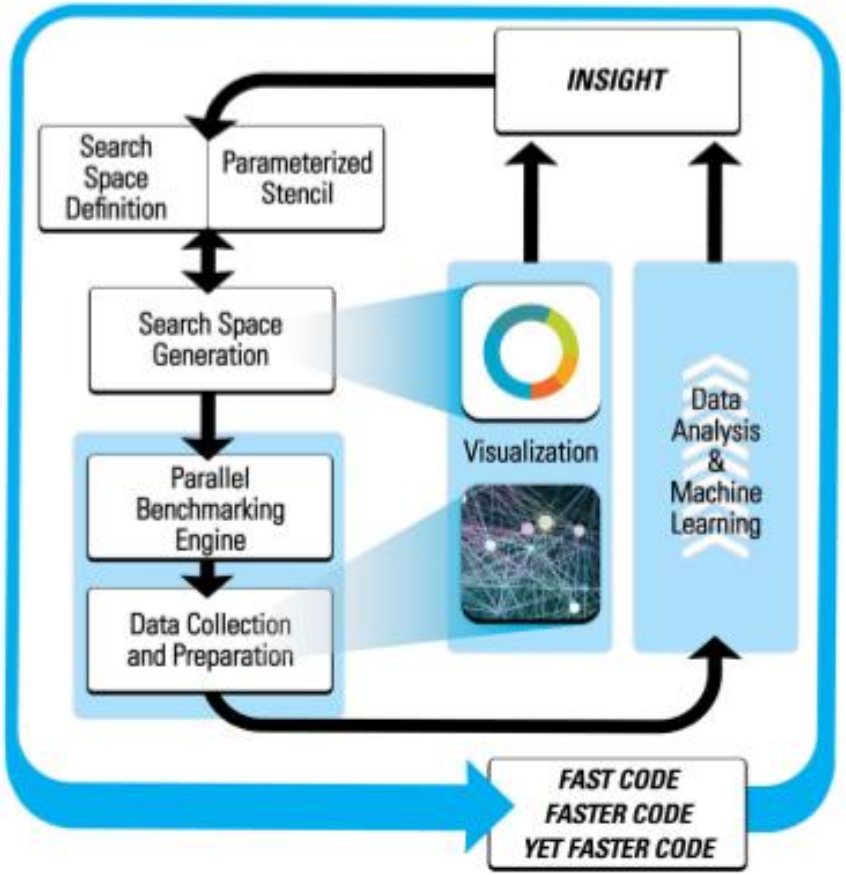
Peter Elmer

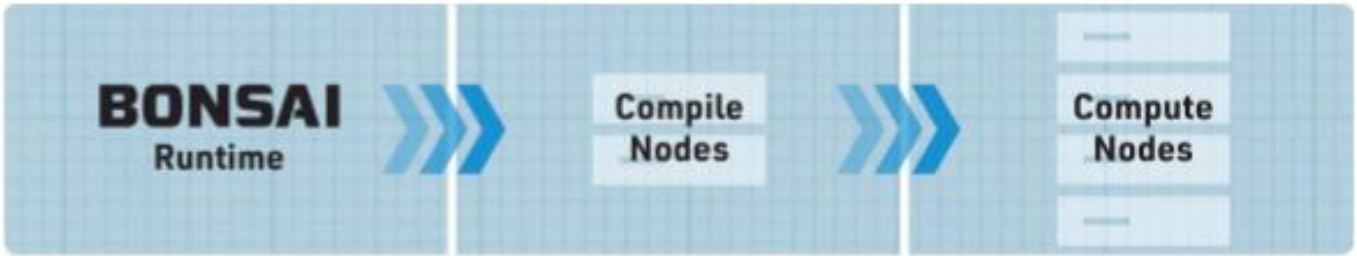
SI2-SSI: BONSAI
Benchtesting OpeN Software Autotuning Infrastructure

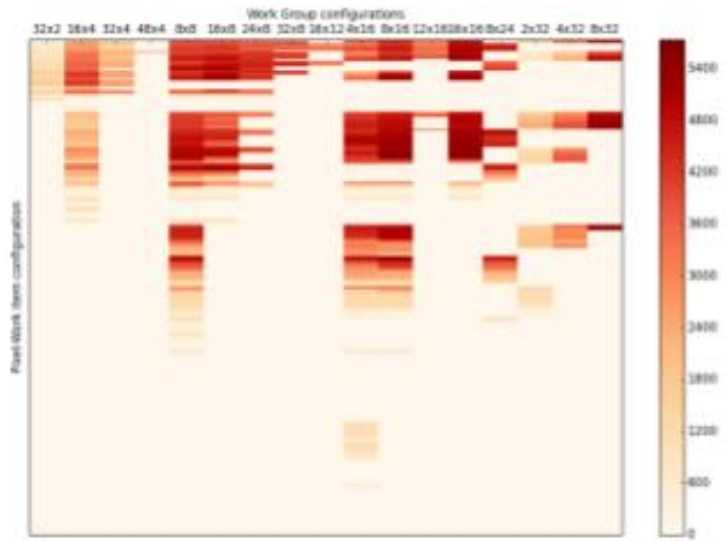
Piotr Luszczek, Jakub Kurzak, Matthew Bachstein, Yaohung Mike Tsai, Jack Dongarra

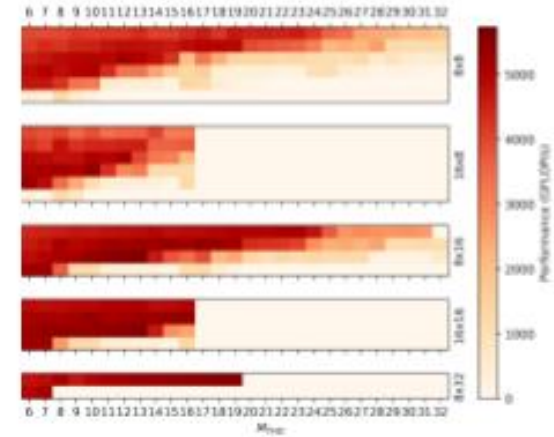
```
R1 = range(3)
R2 = range(3)

@condition
def r1_smaller_than_r2():
    return R1 < R2
```

```
R1, R2
0 , 0
1 , 0
1 , 1
2 , 0
2 , 1
2 , 2
```





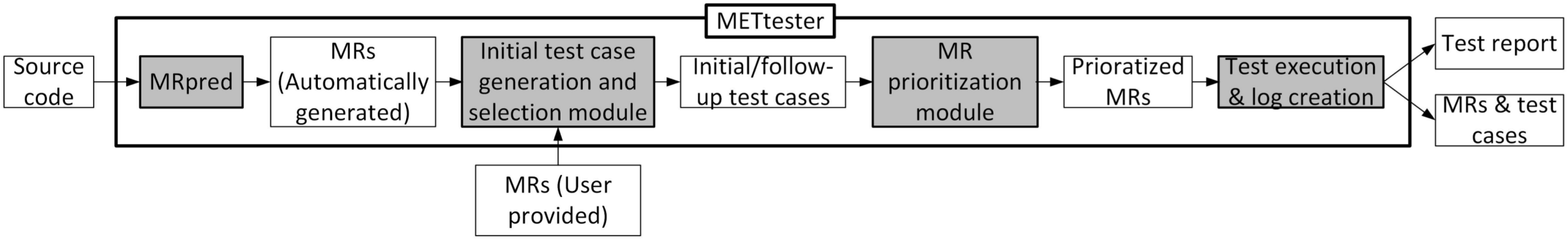




Toward Sustainable Software for Science - Implementing and Assessing Systematic Testing Approaches for Scientific Software

Upulee Kanewala, Montana State University

- Systematic testing of scientific software is challenging due to the *oracle problem*.
- **METtester:** a publicly available testing tool that can be used in the day-to-day scientific development process.



SI2-S2I2 Conceptualization: Geospatial Software Institute

Shaowen Wang

SI2-SSE: Expanding the Scope of Materials Modeling with EPW Software

PIs: E. R. Margine, M. Govindaraju, SUNY Binghamton

- EPW provides unprecedented levels of accuracy and efficiency in calculations of materials properties defined by the interaction between electrons and atomic vibrations (i.e., electron-phonon interactions)
- Electron-phonon interactions are critical for understanding and designing electronic materials and devices (e.g., semiconductors, superconductors, and thermoelectrics).
- Develop accurate and robust software elements to model spin-dependent materials properties from first principles
- Develop Gateway Interfaces to ease usage of EPW
- Apache Mesos framework, Scylla, that leverages Docker Swarm for orchestration of containerized scientific jobs