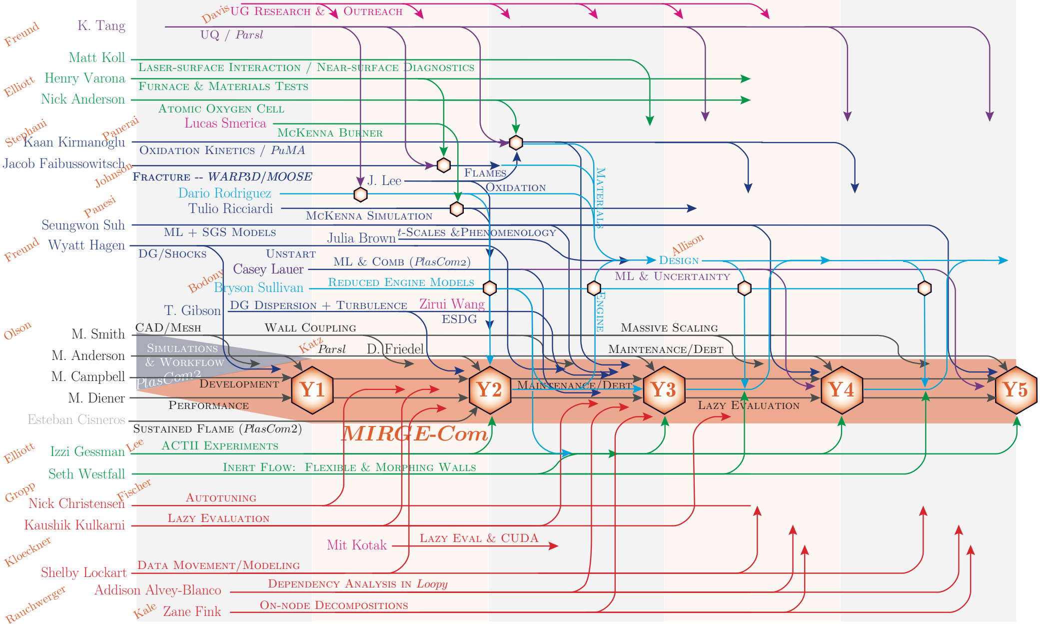

Davis
UG Research & Outreach
Freund
K. Tang
UQ / Parsl
Matt Koll
Laser-surface Interaction / Near-surface Diagnostics
Elliott
Henry Varona
Furnace & Materials Tests
Nick Anderson
Atomic Oxygen Cell
Lucas Smerica
McKenna Burner
Stephani
Panerai
Kaan Kirmanoglu
Oxidation Kinetics / PuMA
Jacob Faibussowitsch
Johnson
Fracture -- WARP3D/MOOSE
J. Lee
Flames
Oxidation
Materials
Dario Rodriguez
Tulio Ricciardi
McKenna Simulation
Panesi
Seungwon Suh
ML + SGS Models
Julia Brown
t-Scales & Phenomenology
Freund
Wyatt Hagen
DG/Shocks
Unstart
Allison
Design
Casey Lauer
ML & Comb (PlasCom2)
ML & Uncertainty
Bodony
Bryson Sullivan
Reduced Engine Models
Engine
T. Gibson
DG Dispersion + Turbulence
Zirui Wang
ESDG
Olson
M. Smith
CAD/Mesh
Wall Coupling
Massive Scaling
M. Anderson
Simulations & Workflow
Katz
Parsl
D. Friedel
Maintenance/Debt
M. Campbell
PlasCom2
Development
Y1
Y2
Maintenance/Debt
Y3
Y4
Y5
M. Diener
Performance
Lazy Evaluation
Esteban Cisneros
Sustained Flame (PlasCom2)
MIRGE-Com
Elliott
Izzi Gessman
Lee
ACTII Experiments
Seth Westfall
Inert Flow: Flexible & Morphing Walls
Gropp
Fischer
Nick Christensen
Autotuning
Kaushik Kulkarni
Lazy Evaluation
Mit Kotak
Lazy Eval & CUDA
Kloeckner
Shelby Lockart
Data Movement/Modeling
Addison Alvey-Blanco
Dependency Analysis in Loopy
Rauchwerger
Kale
Zane Fink
On-node Decompositions

Task Graph Parallelism on GPUs via CUDAGraphs

Mit Kotak, Kaushik Kulkarni

University of Illinois at Urbana-Champaign

October 27, 2022

Motivation

MIRGE-Com: Overview

- DG-FEM solver using array-valued data flow graphs.
- For our experiments, we use an unfused version of the operator.

Problem

- Target the concurrency across the launched GPU kernels.
  - Would lead to better device usage at lower problem sizes.
  - Hypothesize: Cost savings from:
    - Lower launch overhead
    - Exploiting overlap

Approach

- Develop a new `ArrayContext`.
- Map *Numpy*-like operations to graph-based IR (*Pytato*).
- Generate *CUDAGraph* source code by mapping *Pytato* IR onto *PyCUDA*.

Code Transformation

```
actx = PytatoCUDAGraphArrayContext()
def f():
    return actx.zeros(100, dtype="float") + 1
f_compiled = actx.compile(f)
f_compiled()
```

Figure: `Arraycontext` Program

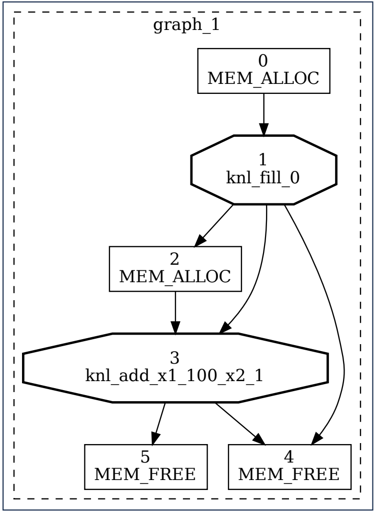


Figure: Generated *CUDAGraph*

Code Transformation

```
actx = PytatoCUDAGraphArrayContext()
def f():
    return actx.zeros(100, dtype="float") + 1
f_compiled = actx.compile(f)
f_compiled()
```

Figure: `Arraycontext` Program

```
import pycuda.driver as _pt_drv
import numpy as np
from pycuda.compiler import SourceModule as _pt_SourceModule
from pycuda import gpuarray as _pt_gpuarray
from functools import cache
_pt_mod_1 = _pt_SourceModule('#define bIdx(N) ((int) blockIdx.N)\n#define tIdx(N) ((int) threadIdx.N)\n\
      nextern "C" __global__ void __launch_bounds__(32) knl_fill_0(double *__restrict__ out)\n{\n  if (99 +
       -1 * tIdx(x) + -32 * bIdx(x) >= 0)\n  {\n    int const ibatch = 0;\n\n    out[tIdx(x) + 32 * bIdx(x)
      ] = 0.0;\n  }\n}')
_pt_mod_2 = _pt_SourceModule('#define bIdx(N) ((int) blockIdx.N)\n#define tIdx(N) ((int) threadIdx.N)\n\
      nextern "C" __global__ void __launch_bounds__(32) knl_add_x1_100_x2_1(double *__restrict__ out,
      double const *__restrict__ _in0)\n{\n  if (99 + -1 * tIdx(x) + -32 * bIdx(x) >= 0)\n  {\n    int
      const ibatch = 0;\n\n    out[tIdx(x) + 32 * bIdx(x)] = _in0[tIdx(x) + 32 * bIdx(x)] + 11;\n  }\n}')

@cache
def exec_graph_builder():
    _pt_g = _pt_drv.Graph()
    _pt_buffer_acc = {}
    _pt_node_acc = {}
    (_pt_memalloc_0, _pt_array_0) = _pt_g.add_memalloc_node(size=800, dependencies=[])
    _pt_kernel_0 = _pt_g.add_kernel_node(_pt_array_0, func=_pt_mod_1.get_function('knl_fill_0'), block=(32,
       1, 1), grid=(4, 1, 1), dependencies=[_pt_memalloc_0])
    _pt_buffer_acc['_pt_array_0'] = _pt_array_0
    _pt_node_acc['_pt_kernel_0'] = _pt_kernel_0
    (_pt_memalloc, _pt_array) = _pt_g.add_memalloc_node(size=800, dependencies=[_pt_kernel_0])
    _pt_kernel = _pt_g.add_kernel_node(_pt_array, _pt_array_0, func=_pt_mod_2.get_function('
      knl_add_x1_100_x2_1'), block=(32, 1, 1), grid=(4, 1, 1), dependencies=[_pt_memalloc, _pt_kernel_0])
    _pt_buffer_acc['_pt_array'] = _pt_array
    _pt_buffer_acc['_pt_array_0'] = _pt_array_0
    _pt_node_acc['_pt_kernel'] = _pt_kernel
    _pt_g.add_memfree_node(_pt_array_0, [_pt_kernel, _pt_kernel_0])
    _pt_g.add_memfree_node(_pt_array, [_pt_kernel])
    return (_pt_g.get_exec_graph(), _pt_g, _pt_node_acc, _pt_buffer_acc)

def f(allocator=_pt_drv.mem_alloc):
    _pt_result = _pt_gpuarray.GPUArray((100,), dtype='float64', allocator=allocator)
    (_pt_exec_g, _pt_g, _pt_node_acc, _pt_buffer_acc) = exec_graph_builder()
    _pt_exec_g.set_kernel_node_params(_pt_buffer_acc['_pt_array_0'], kernel_node=_pt_node_acc['_pt_kernel_0
      '])
    _pt_exec_g.set_kernel_node_params(_pt_result.gpudata, _pt_buffer_acc['_pt_array_0'], kernel_node=
      _pt_node_acc['_pt_kernel'])
    _pt_exec_g.launch()
    _pt_tmp = {'_pt_out_': _pt_result}
    return _pt_tmp
```

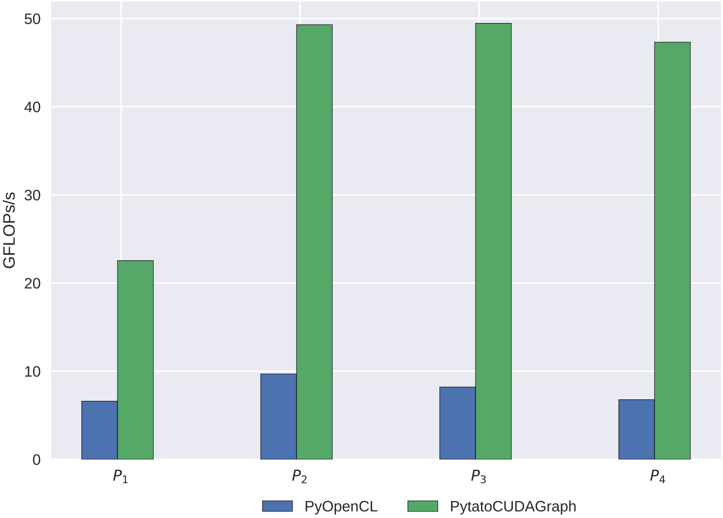
Graph dependencies

Figure: Generated *PyCUDA* code

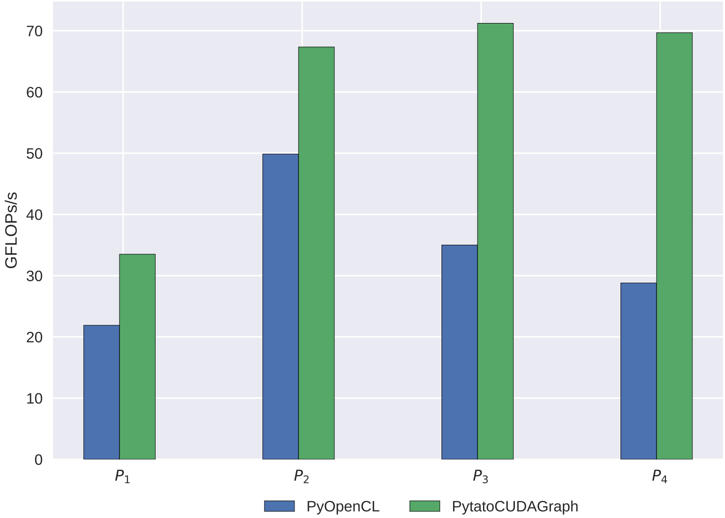
Experimental Setup

- Nvidia Titan V
  - Peak Double prec. FlOps: 6144 GFlOps/s
  - Peak bandwidth: 652.8 GB/s
- 3D Wave and Euler Operators
- $p \in \{1, 2, 3, 4\}$
- #Tetrahedrons in mesh: 10K (for lower orders)- 27K (for higher orders)
- OpenCL Implementation: PoCL-CUDA (3.0)
- Benchmarks, run instructions at github.com/mitkotak/dg_benchmarks

Wave Operator



Euler Operator



Key Takeaways

- `PytatoCUDAGraphArraycontext` abstraction can compile real-world DG-FEM operators.
  - as a drop-in replacement for array program backend.
- Observed a speedup of 3-6x for Wave and 1.5-3x for Euler.

Open Questions

- Is CUDAGraph+FusionActx profitable for MIRGE-Com?
  - Develop a performance model for *CUDAGraphs*.
  - Model peak memory usage for *CUDAGraphs*.

Future Work

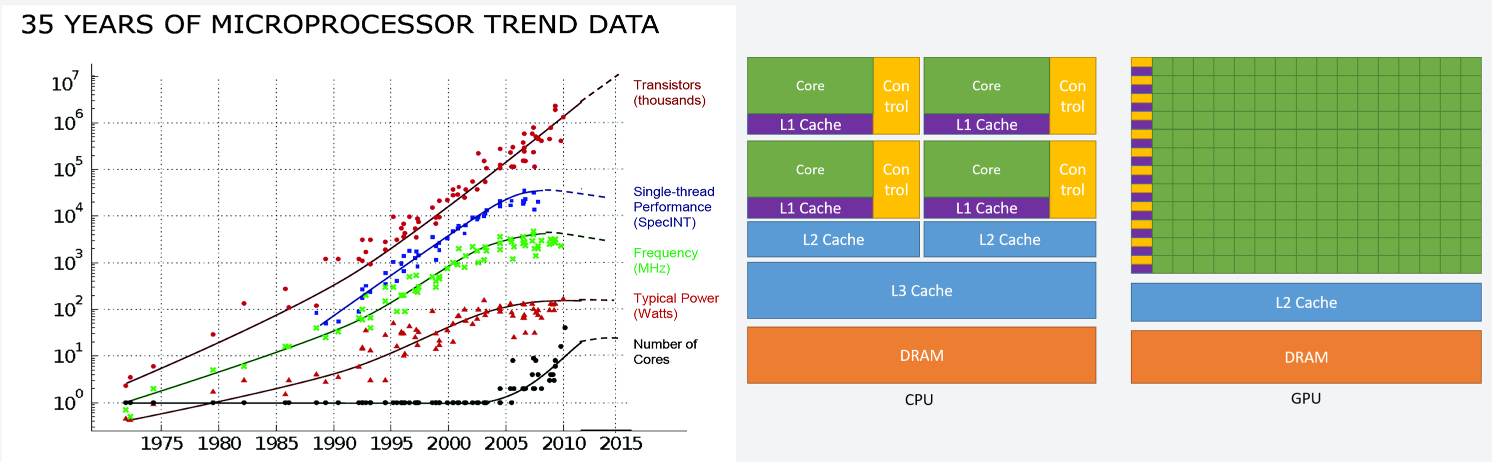
- Upstream the work: Integrate with *PyCUDA*, *Pytato*, `Arraycontext`.

Efficiently Executing *NumPy* on GPUs via the *CUDAGraph* API

Mit Kotak · Kaushik Kulkarni · Andreas Klöckner

National Center for Supercomputing Applications · Urbana, IL

Why GPUs?



- CPU performance is flatlining
- GPUs devote more transistors to data processing

Abstract

- *NumPy* is the defacto standard for array-based numerical computations in scientific workloads (for ex. PDE solving, Image Processing, etc.)
- However, its ability to harness GPU acceleration is limited by its single-node, occasionally multi-threaded, CPU- only execution model
- Goal: Should be able to efficiently execute array operations on GPUs
- Challenges:
  - Saturating all available execution units to efficiently use the entire GPU
  - *NumPy* is a high-level programming interface, does not address performance directives such as:
    - Execution grid-size which affects data partitioning
    - User-specified local memory management
    - Kernel launch overhead costs
  - **Realizing concurrency across array operations**
- Our Approach: Use *CUDAGraph* API to concurrently launch array operations using tuned kernels that are accessed through *PyCUDA*
- Past work:
  - *Legate*: A runtime system for scheduling operations in a task graph
    - Representation lacks a global view of the program limiting the program optmizing space
  - Lazy evaluation: *Theano* | *JAX* | *PyTorch*
    - Requires expensive algorithms (ex. kernel/loop fusion)
    - Ex. *Theano*[3] claim to have super-linear codegeneration algorithm
  - Single Stream: *cuPy* | *GPUArrays.jl*

Why *CUDAGraphs*?

- *CUDA Streams*
  - Operations are enqueued in-order into the stream object
- *CUDAGraph API*
  - Takes in a task dependency graph where each node corresponds to a *CUDA* kernel
  - Scheduler realizes concurrency across nodes in the graph through multiple streams

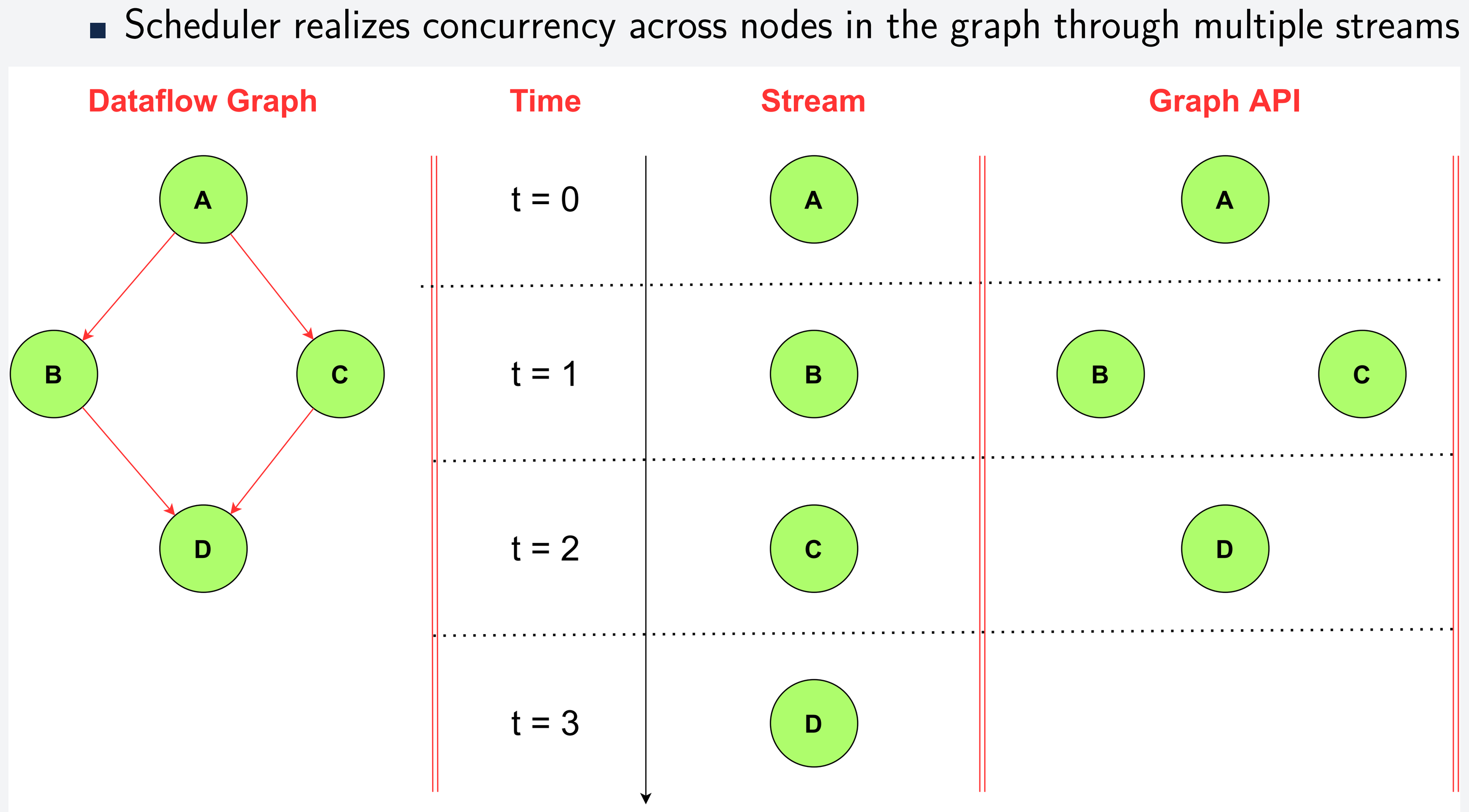


Figure: Single Stream vs CUDAGraphs scheduling

Interface Implementation

```
actx=PyCUDAGraphArrayContext()
x = actx.zeros((100,1))
tmp=x+1
result=actx.freeze(tmp)
```

```
cuGraphCreate(&m_graph, flags)
cuGraphAddMemAllocNode(&memalloc_x, m_graph)
cuGraphAddMemAllocNode(&memalloc_tmp, m_graph)
cuGraphAddMemSetNode(&memset_node, m_graph, [memalloc_x])
cuGraphAddKernelNode(&k_node, [memset_x, memalloc_tmp], 2)
cuGraphInstantiate(&m_exec, m_graph)
cuGraphLaunch(m_exec)
cuGraphExecDestroy(&m_exec)
cuGraphDestroy(&m_graph)
```

Figure: User Input Program **Figure: Driver C code**

- Array operations realized as a composition of *CUDA* calls (memcpy, kernel_launches, memalloc) that are added onto a task dependency graph with precise edges.
- Object cleanup tied to lifetime of objects in array operations

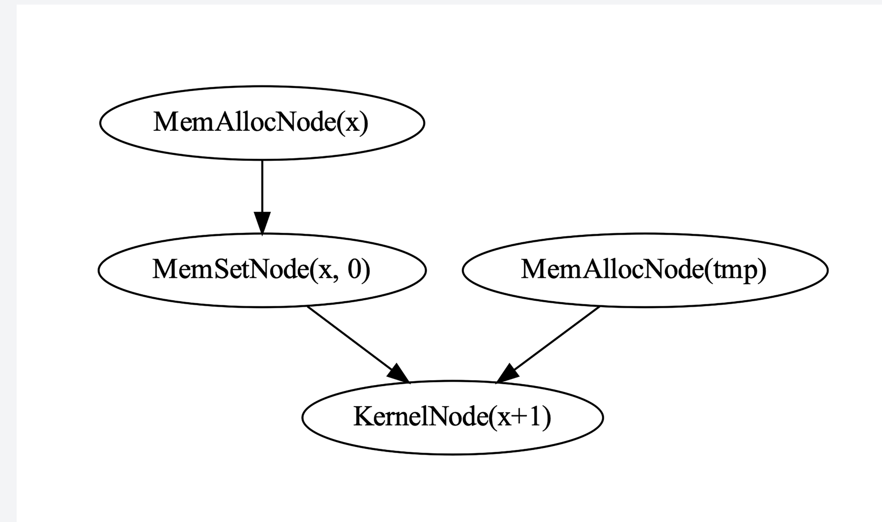
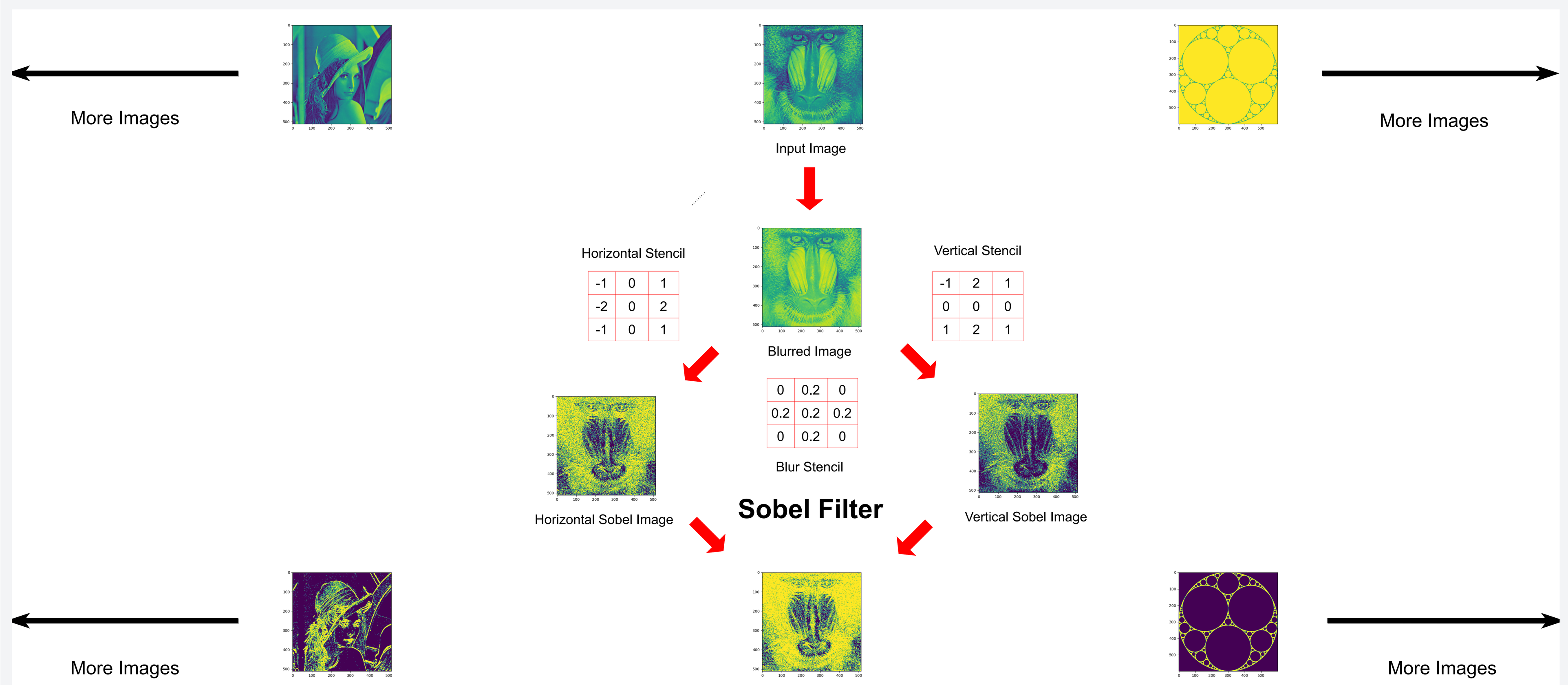


Figure: Generated CUDAGraph

Experimental Setup

We ran a set of image processing algorithms with and without *CUDAGraph* API on different image batches



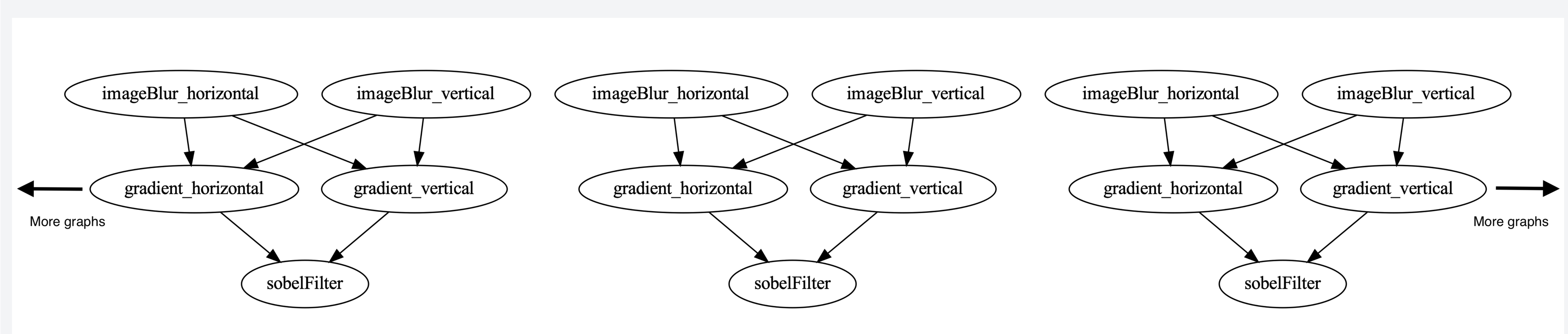


Figure: Sobel Filter Graph

Results

- We observed a speed up of upto 15% on NVIDIA Titan V for smaller problems which was attributed to high task parallelism
- We observed overlapping kernels across multiple streams for the *CUDAGraph* API program

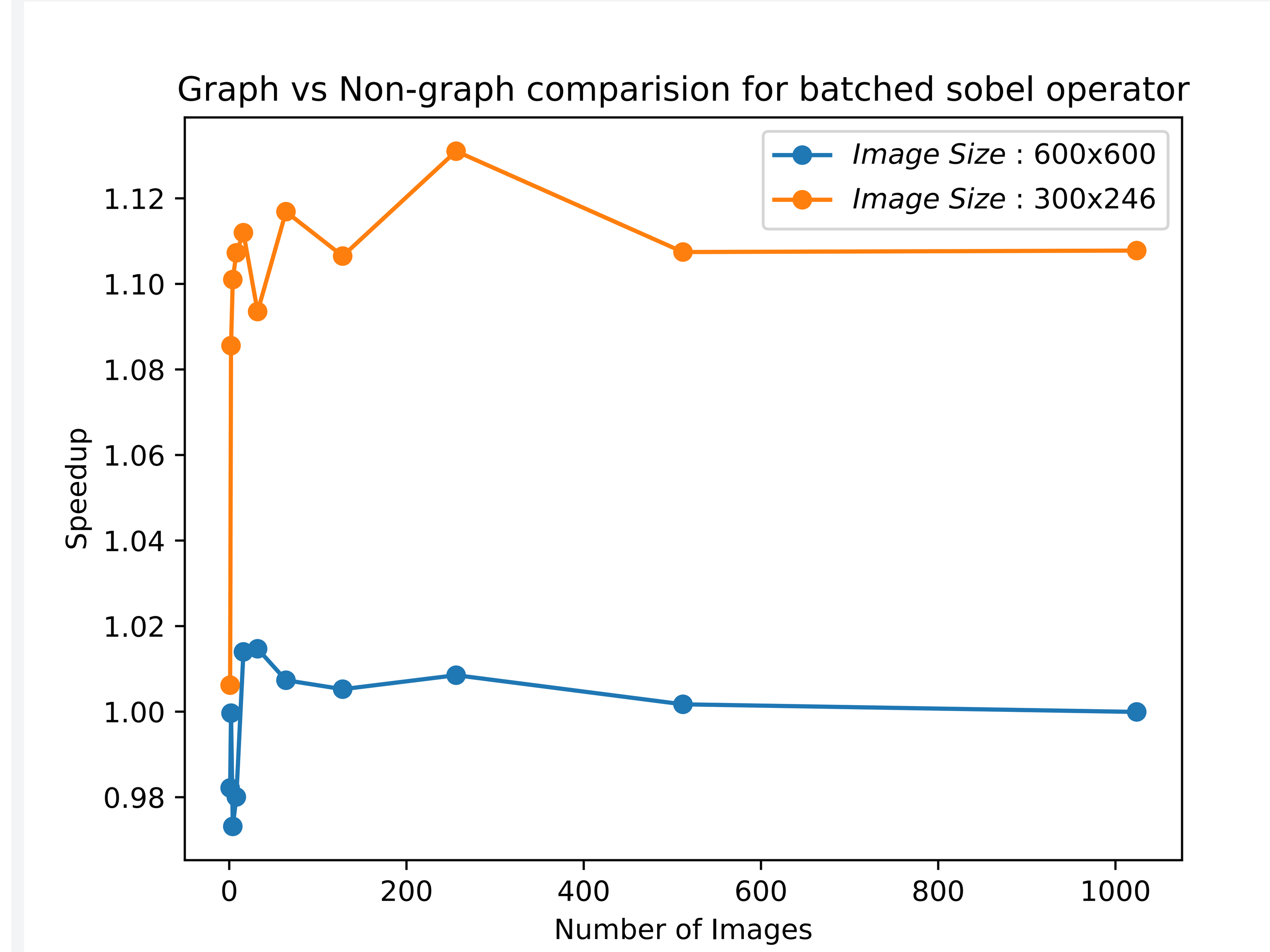


Figure: Sobel Filter on Image Batches

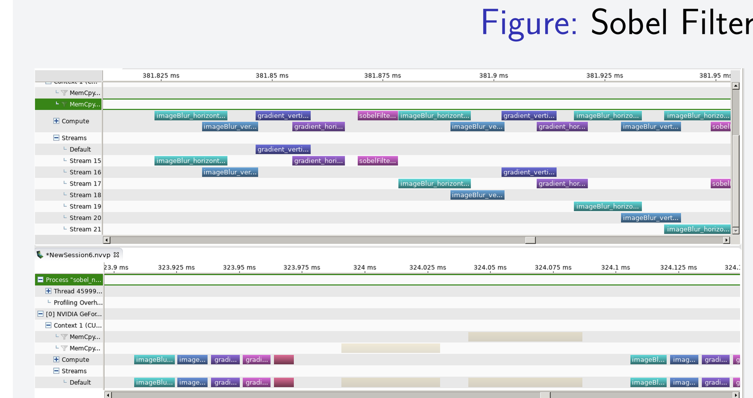


Figure: Kernel Execution timeline with (top) and without (bottom) *CUDAGraph* API Image size 300 x 246

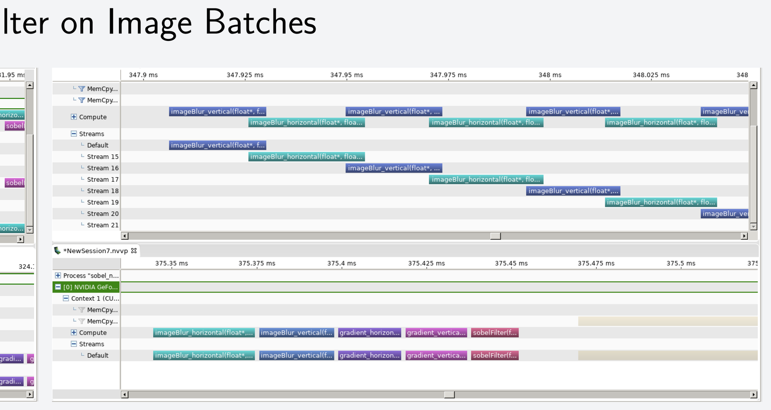


Figure: Kernel Execution timeline with (top) and without (bottom) *CUDAGraph* API Image size 600 x 600

Results, run instructions at
`https://github.com/mitkotak/sobel`

Future Work

- Upstream work: Integrate with *PyCUDA* and *Arraycontext*
- Evalulate this approach on large scale scientific simulations

Acknowledgements

This material is based in part upon work supported by the Department of Energy, National Nuclear Security Administration, under Award Number DE-NA0003963

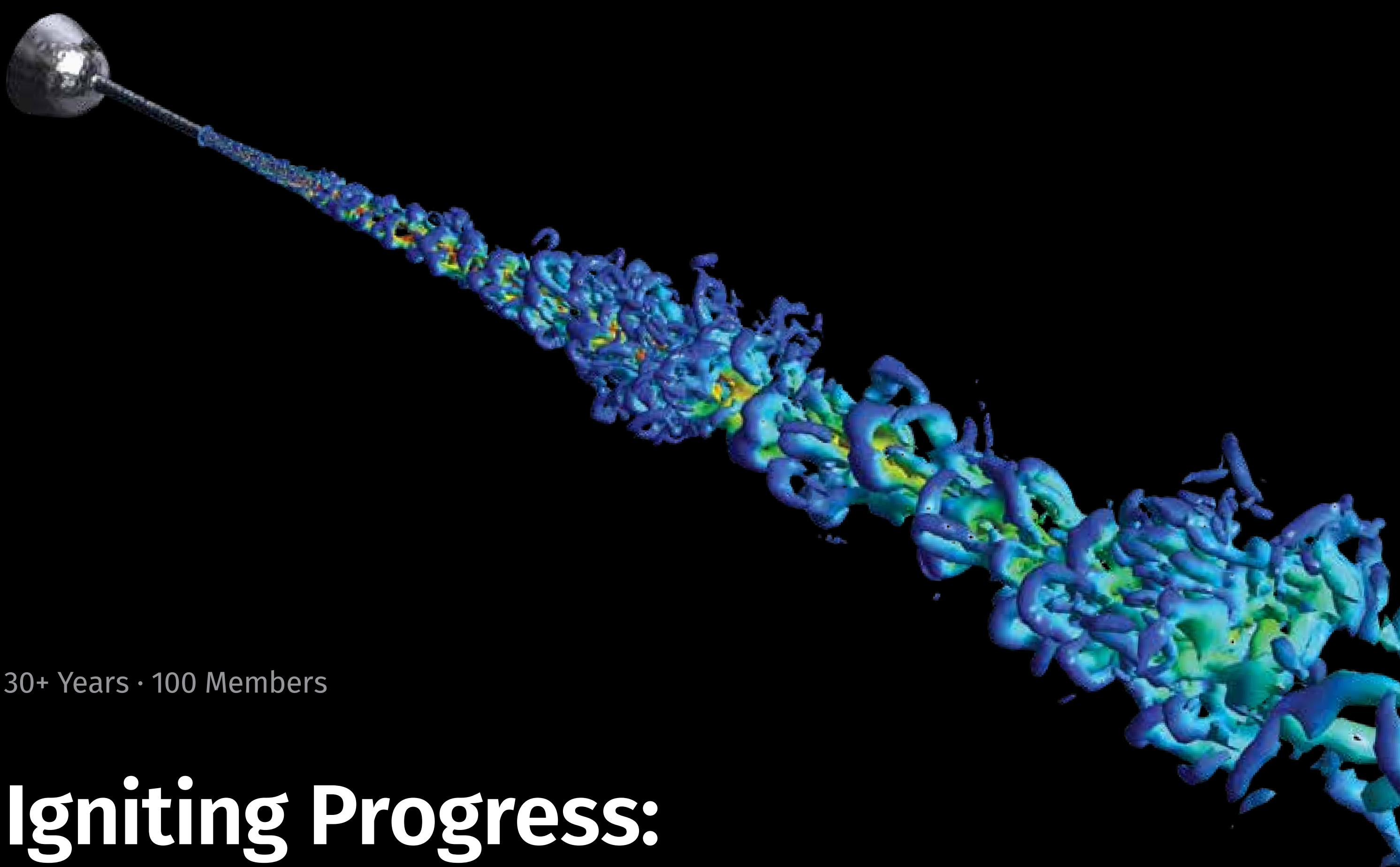
References

- CUDA Programming: A Developer's Guide to Parallel Computing with GPUs (https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html)
- Chapter 27 - GPU Scripting and Code Generation with PyCUDA (https://doi.org/10.1016/B978-0-12-385963-1.00027-7)
- Theano: A Python framework for fast computation of mathematical expressions (https://doi.org/10.48550/arXiv.1605.02688)

mitak2@illinois.edu

NCSA ILLINOIS

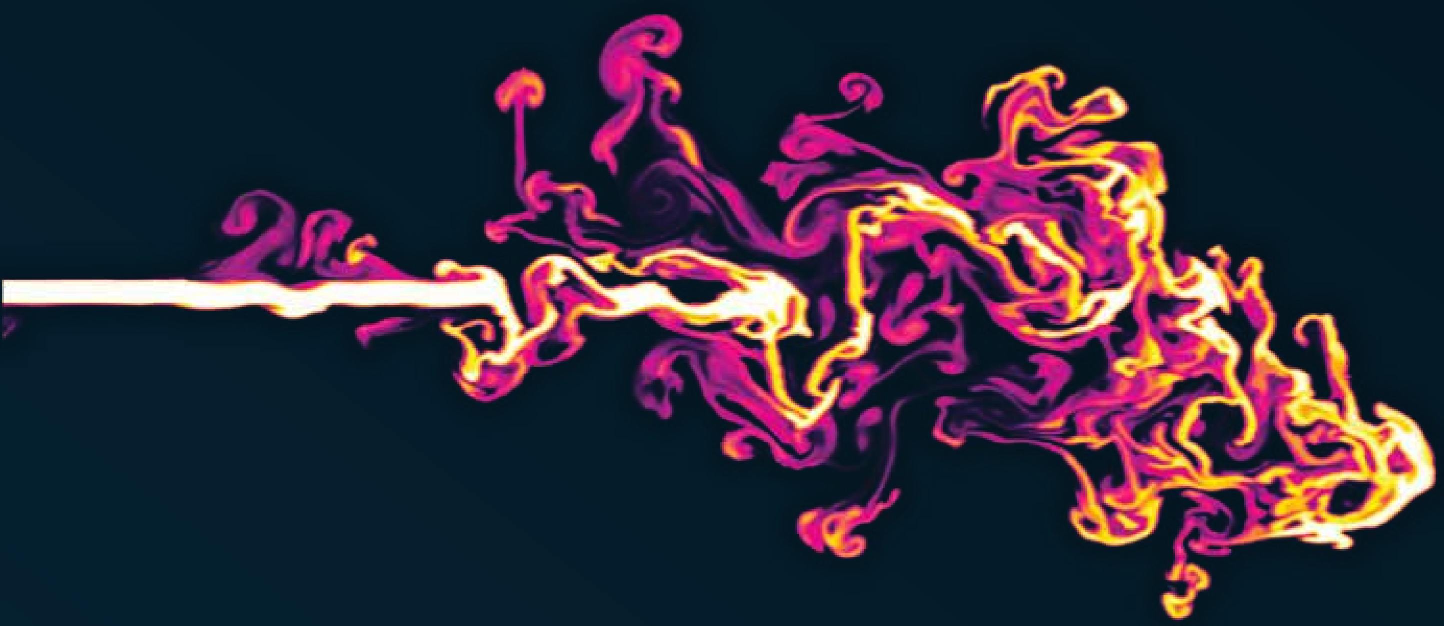
CASC



30+ Years · 100 Members

Igniting Progress: From Big Data to Bold Solutions

CASC Coalition for Academic Scientific Computation



About the Cover

Even as we look toward a future with many more electric vehicles on the road, increasing the efficiency of traditional internal combustion engines remains vital to curbing harmful emissions. To aid in this task, Joonsik Hwang and Charles Michael Gibson from the Mississippi State University Center for Advanced Vehicular Systems are working with researchers at Sandia National Laboratories to capture the complexities of fuel combustion in detailed computer simulations.

With computation performed at Mississippi State University's Malcolm A. Portera High Performance Computing Center, the researchers use their models to generate data and visualizations probing the fluid dynamics within combustion engines, allowing them to study how different fuels break down and mix during the combustion process. The insights derived from these simulations can be used to predict the behavior and efficiency of alternative fuel mixtures, as well as to help researchers confirm these predictions with physical experiments. Ultimately, the work aims to help industrial partners improve fuel efficiency, reduce CO_2 emissions, and supplement gasoline with renewable fuels.

Contents

About CASC

The Coalition for Academic Scientific Computation (CASC) is a nonprofit organization dedicated to the use of advanced computing technology to accelerate discovery.

For over 30 years, CASC has represented many of the nation's most forward-thinking research computing and data services organizations — a vibrant community of excellence that today totals 100 member organizations. By providing a voice for this broad coalition of members, facilitating the exchange of ideas and resources, and advocating for funding and policies to enable the field to reach its full potential, CASC advances our community's vision for a robust, sustainable research ecosystem to support national competitiveness, global security, economic success, and a diverse and well-prepared 21st century workforce.

CASC Mission

- To advocate for the importance of and need for public and private investment in research computing and data services to support academic research.
- To serve as a trusted advisor to federal agencies on the direction of relevant funding programs.
- To actively engage in discussions of policies related to research computing and data services.
- To foster advancement of a robust and diverse community of current and emerging leaders in this field.
- To provide a forum for the community to share strategic ideas and best practices.

CASC Executive Committee*

- Chair: Jim Wilgenbusch
- Vice Chair: Rich Knepper
- Secretary: James Barr von Oehsen
- Treasurer: Dave Hart

Communications Committee

- Melyssa Fratkin, Texas Advanced Computing Center *(Chair)*
- Andrew Bell, University of Virginia
- Marisa Brazil, Arizona State University
- Brian Connelly, University of Pittsburgh
- Ben Tolo, San Diego Supercomputer Center
- Kristin Lepping, Rutgers University
- Paul Redfern, Cornell University Center for Advanced Computing
- Stephanie Suber, Renaissance Computing Institute

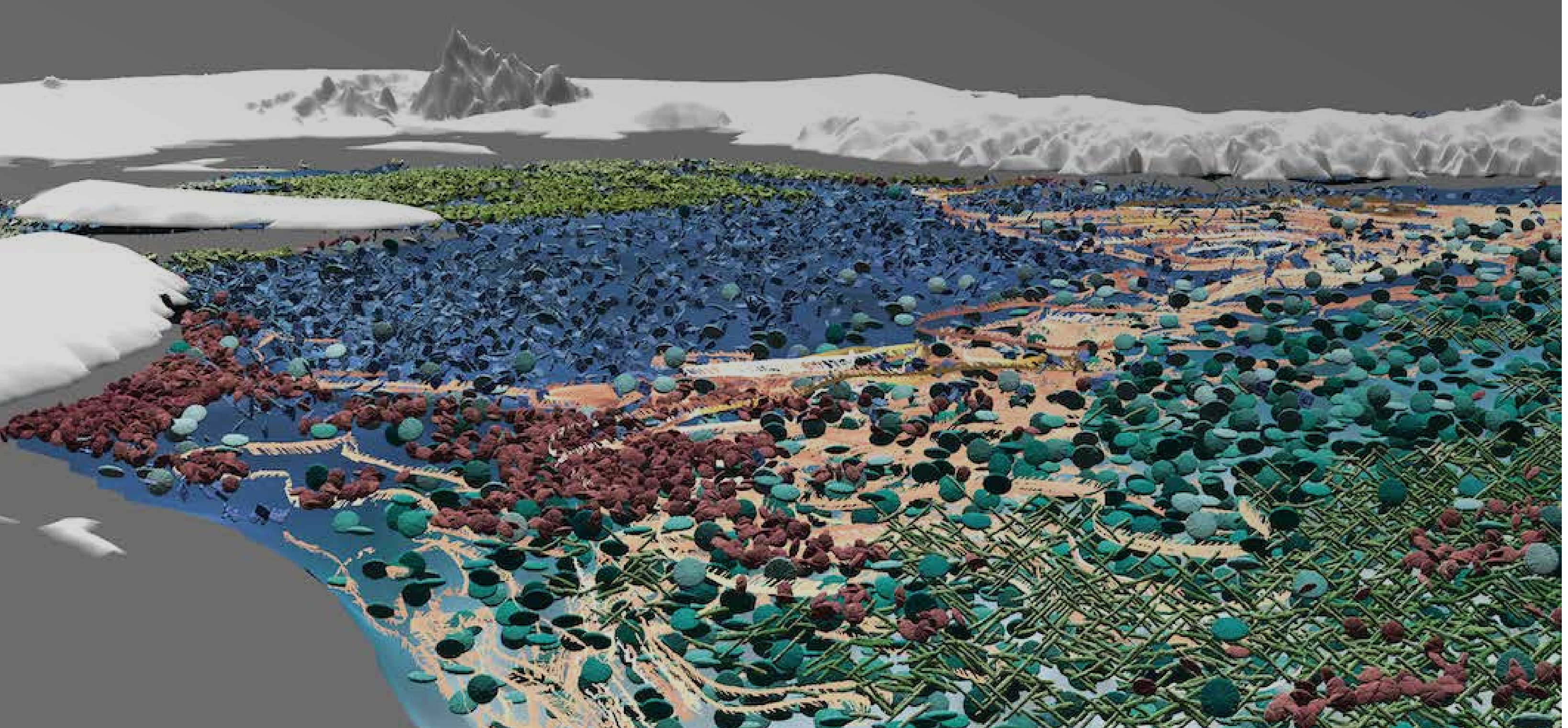
Capturing Climate Complexities

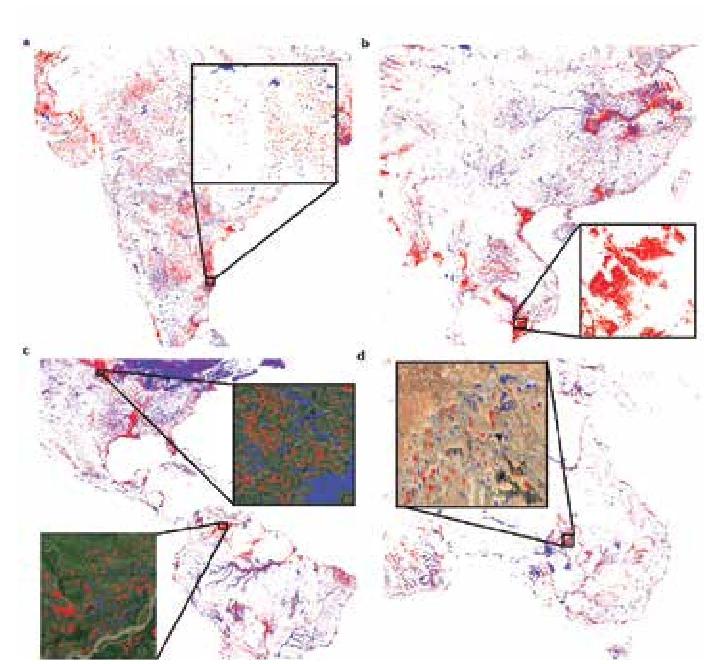
Visualizing Change Beneath the Ice

To predict how increasing temperatures might affect sea-level rise and coastal communities in the coming century, it is critical to study how ice is melting across Earth's polar regions, especially in the ice shelves that form where a glacier or ice sheet flows down to the coastline, extending over the (warming) sea. However, most of this melting occurs on the bottom of the shelves where it is difficult to track.

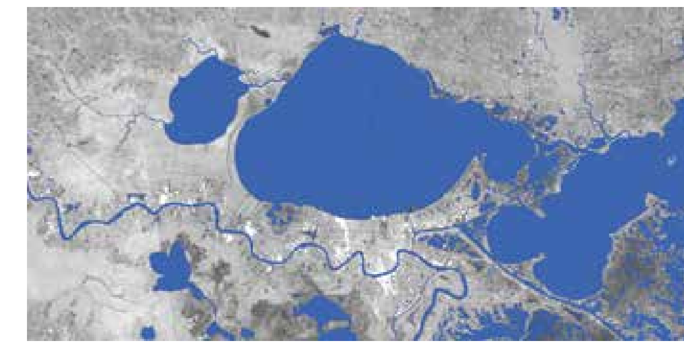
An interdisciplinary group of experts — including ocean modelers, computer scientists, visualization professionals, and an artist — came together to help scientists visualize the complex processes taking place under ice shelves. The team, which included Francesca Samsel and Gregory Abram of the University of Texas at Austin and the Texas Advanced Computing Center, Daniel Keefe from the University of Minnesota, and Mark Petersen from Los Alamos National Laboratory, developed a new process to show various types of water masses in three dimensions as well as the movement of these masses over the course of a year, interactions with the ocean floor, and mixing caused by turbulent flows. The team created this image by applying their process to Antarctica's Filchner-Ronne Ice Shelf, on the edge of the Weddell Sea, showing how four types of water masses and relevant ocean currents flow and interact beneath an ice shelf that is — for now — hundreds of feet thick. These visualizations will help scientists assess Antarctica's ice shelves at a time when some of the continent's vast bodies of ice appear to be disintegrating faster than previously thought, with alarming implications for future sea-level rise.

Copyright Francesca Samsel, TACC UT Austin





Copyright Vipin Kumar, University of Minnesota



Copyright University of Alaska Fairbanks

Where's the Water?

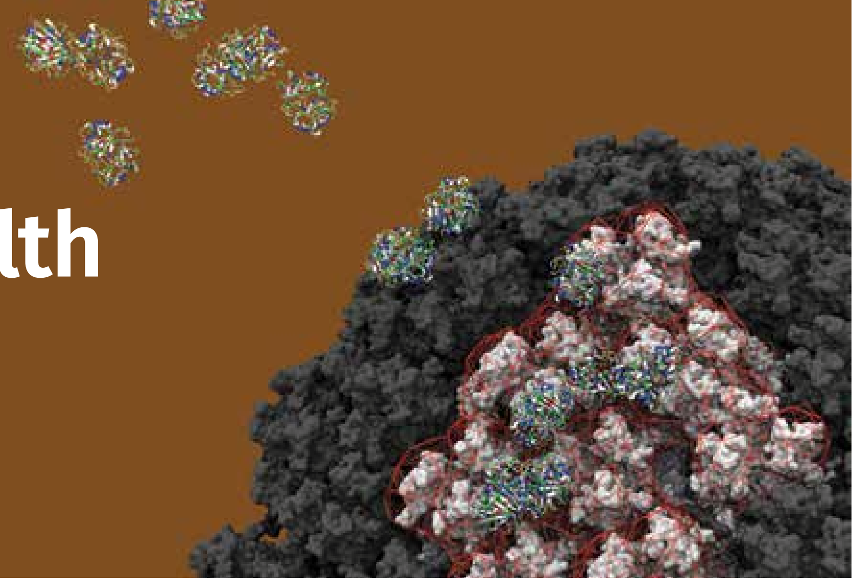
Fresh water is a critical resource around the world — and one that is dwindling rapidly in many places. Tracking changes in the sizes of lakes and reservoirs can reveal how human activities affect water availability and help communities plan for future climate change. Ankush Khandelwal from the University of Minnesota Supercomputing Institute for Advanced Computational Research led the development of a comprehensive new dataset reflecting how lakes and reservoirs have varied month-by-month over more than 30 years. Called RealSAT, the dataset contains information for almost 650,000 lakes and reservoirs around the world, a significant expansion over the 250,000 water bodies within the scope of the ReaLSAT dataset that are also captured in the widely used HydroLAKES database.

Developing the dataset took the interdisciplinary team over eight years and required a new class of machine learning algorithms that combine satellite imagery with existing knowledge on the physical dynamics of water bodies. This image shows lakes identified by RealSAT and HydroLAKES (blue) as well as lakes present in RealSAT but not HydroLAKES (red). The expanded dataset offers a new view on a wide variety of water bodies, such as small reservoirs in India (a), water-intensive agriculture in Vietnam (b), natural lakes in the United States and wetlands in Venezuela (c), and shallow lakes in Australia (d). An early version of this dataset already helped scientists to automatically identify many man-made reservoirs, and researchers are now using it to study how irrigation practices affect the relationship between surface water and groundwater.

Ready to Respond

When Hurricane Harvey brought record-shattering rainfall to Texas and Louisiana in 2017, emergency responders were desperate to know which areas were flooded or likely to flood in order to guide rescue efforts. Scientist Franz Meyer from the University of Alaska Fairbanks Geophysical Institute and collaborators at NASA used remote sensing technology known as synthetic aperture radar to peer through the storm's thick cloud cover and measure water levels on the ground. But they had to scramble to manually process the satellite data, limiting their ability to provide detailed information in time.

Four years later, as Hurricane Ida barreled toward the Louisiana coastline in 2021, the team was ready. This time, Meyer and colleagues had a new tool in their arsenal: a computer algorithm, implemented using modern cloud computing technology, to automatically extract surface water measurements from satellite images. As the storm neared, the scientists produced hundreds of images using this system and met with FEMA officials daily to interpret the latest imagery, helping to inform on-the-ground decisions and save lives. This image shows surface water levels (blue) around New Orleans on Aug 29, 2021, as Hurricane Ida approached.



Pioneering Health Innovations

Racing for Answers on Vaccine-Linked Clots

Despite the overwhelming success of COVID-19 vaccines, some rare side effects have emerged, triggering intensive studies to understand the causes and reduce the risk. When some vaccines were found to be linked to life-threatening blood clots — an ultra-rare adverse event called vaccine-induced thrombotic thrombocytopenia (VITT) — some countries paused vaccination campaigns as scientists raced to learn more.

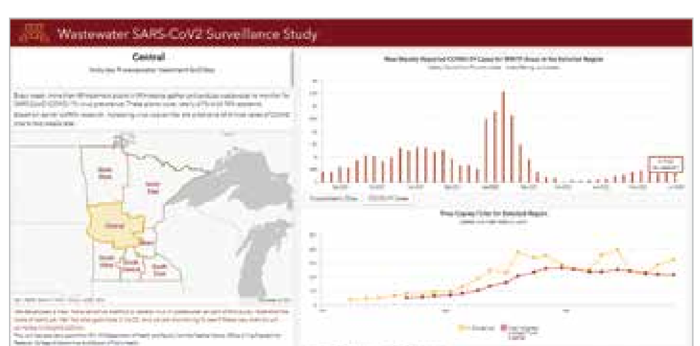
Abhishek Singharoy from Arizona State University is working with scientists from the Mayo Clinic, AstraZeneca — which makes one of the vaccines linked with VITT — and other institutions to better understand what causes this dangerous condition. They found that one of the key interactions that occurs when the AstraZeneca vaccine enters a cell involves platelet factor 4 (PF4), a protein that is also involved in blood clotting. To better understand this interaction on a molecular level, the researchers ran 400 parallel simulations on Bridges-2 at the Pittsburgh Supercomputing Center. This revealed that the outer surface of the adenovirus vector that carries the vaccine is negatively charged — information that wasn't available from laboratory studies alone. This negative charge could cause PF4 to bind to vaccine components in a way that starts a signaling cascade that leads to VITT. The discovery could eventually make it possible to identify people at risk for VITT or help researchers reengineer the vaccine to prevent the adverse effect. This image shows the adenovirus vaccine vector (gray), PF4 proteins (multicolored), and the negatively charged area where PF4 binds to the vector (white/red).

Image Above: *Copyright Chun-Kit Chan and Abhishek Singharoy, Arizona State University*

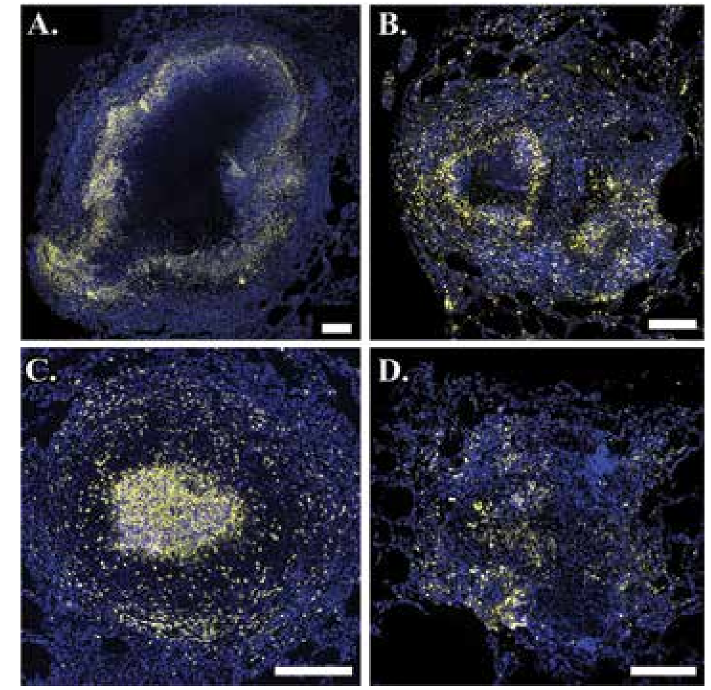
Staying Ahead of COVID Cases

Scientists have monitored wastewater to detect the spread of diseases for some time, but it wasn't until the COVID-19 pandemic that the full usefulness of this approach became apparent. Studies show that COVID-19 wastewater data can provide about two weeks' notice of viral trends, information that lets hospitals know when to prepare for a spike in cases and helps community members figure out when protective precautions are most needed. But to get these benefits, people need wastewater monitoring data to be presented in a way that's easy to understand.

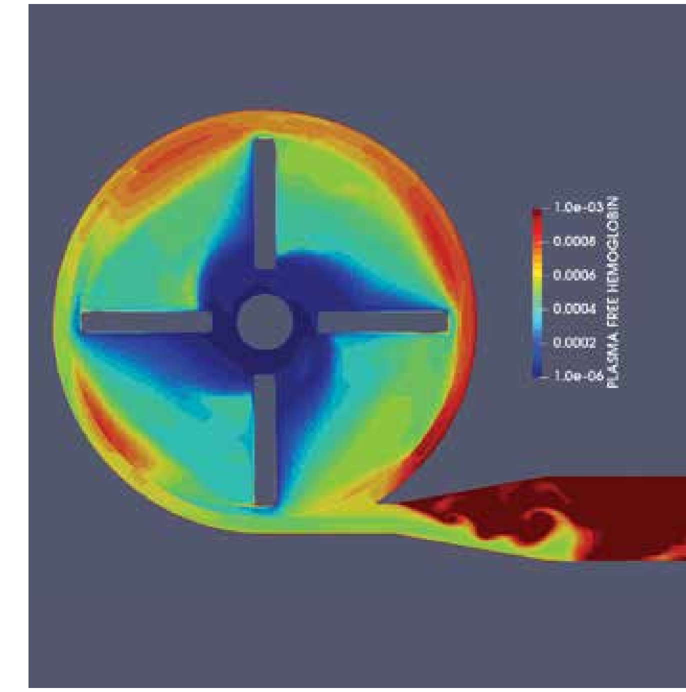
To this end, Stacey Stark from the University of Minnesota U-Spatial, part of Research Computing in the Office of the Vice President for Research, created a dashboard that compares COVID-19 positivity and hospitalization rates from county public health departments with wastewater data in each of seven regions of the state. Pictured is the public-facing dashboard, which draws weekly data from more than 40 wastewater treatment plants across the state and has had over 50,000 views. Providing public access to this dashboard demonstrates public accountability for the work and has spurred the media's interest in following the data. Scientists are now working to test wastewater for additional pathogens such as polio and monkeypox, which would help expand the predictive value of this statewide network.



Copyright 2022 Regents of the University of Minnesota



Copyright Russ Butler, AdventHealth University and Josh Mattila, University of Pittsburgh



Copyright Greg W. Burgreen, Mississippi State University

A Modern Take on an Ancient Disease

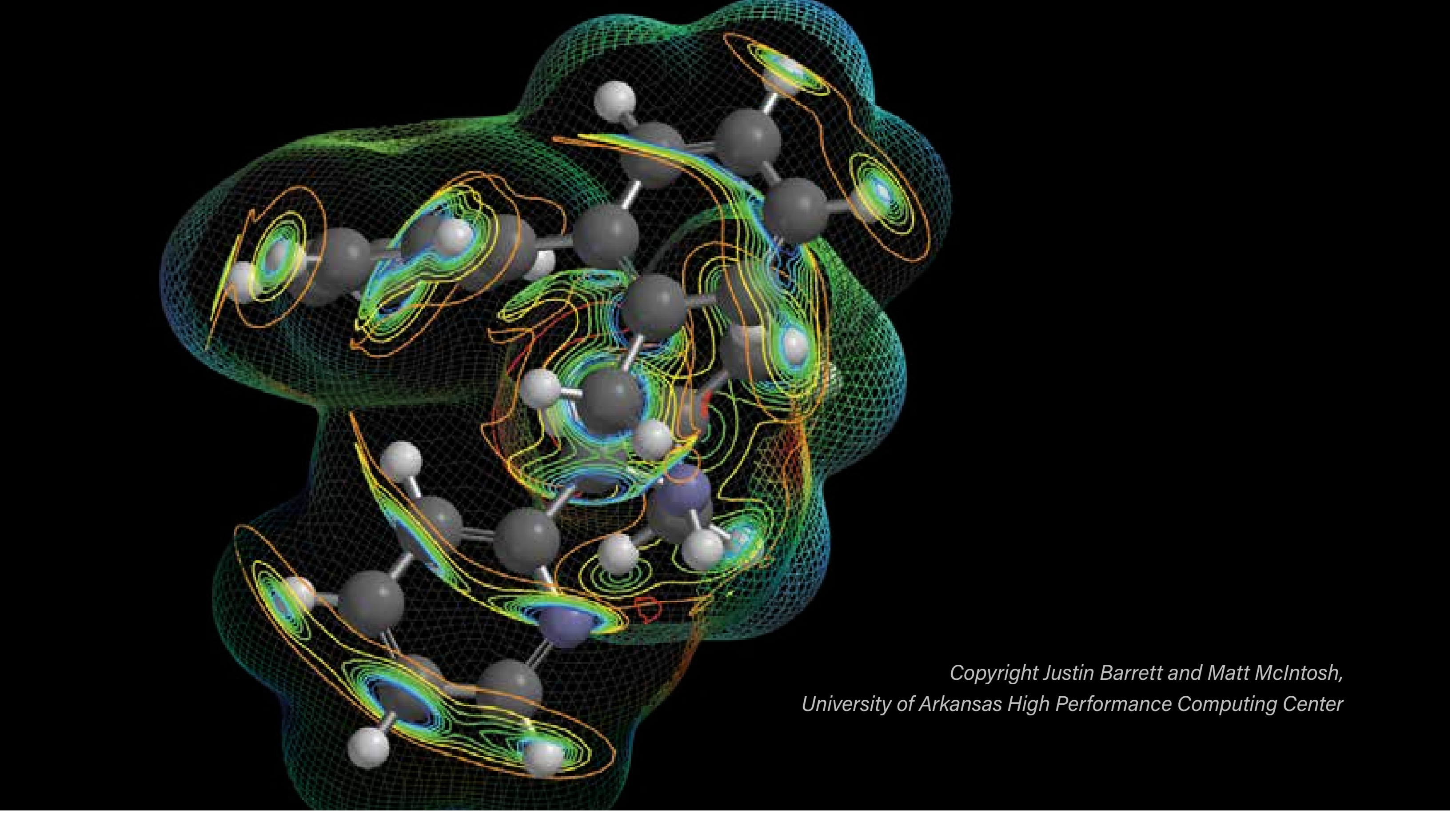
Tuberculosis (TB) has plagued humanity for thousands of years and still ranks among the leading causes of death worldwide, causing an estimated 10 million new cases of active disease and over a million deaths annually. This staggering toll points to a critical need to better understand the biological underpinnings of the disease so that more effective treatments and vaccines can be developed.

Denise Kirschner and colleagues from the University of Michigan Medical School are using supercomputers to study the role that immune cells known as neutrophils play in the disease. Because of their instability, neutrophils are hard to grow and maintain in the lab. The researchers decided to see if predictive simulations performed on the Expanse supercomputer at the University of California's San Diego Supercomputer Center could help with this challenge. They created high-resolution computer models to explore how neutrophils behave when inside spherical masses of infected TB tissues called granulomas. As researchers continue to use advanced computational approaches to explore these immunohistochemistry images further, the models can be useful for making predictions that could lead to better ways to treat and prevent TB. Pictured are granulomas containing neutrophils (yellow) at various stages of TB infection, with cell nuclei shown in blue.

Modeling Fluid Dynamics for Safer Medical Devices

For people with severe heart failure, ventricular assist devices (VADs) — small pumps implanted in the heart to improve blood flow — are a lifesaving innovation. They can provide temporary treatment while waiting for a heart transplant or extend life for those who are not eligible for a transplant. But these devices also can damage blood cells and increase the formation of dangerous blood clots, side effects scientists are working hard to minimize.

To find ways to make VADs and other types of blood pumps safer for patients, Greg Burgreen from Mississippi State University and James F. Antaki from Cornell University are using simulation to better understand how artificial heart pumps can damage blood. By modeling the fluid dynamics of blood flow within and around these medical devices, scientists can study how various factors could work together to damage blood or create clots. These images show results from applying the model to a centrifugal blood pump recently developed by the U.S. Food and Drug Administration (FDA). The researchers quantified the pump's effects on blood dynamics by running several combinations of models and types of blood damage on the Orion supercomputer at the Mississippi State University High Performance Computing Center. The predictions corroborate published experimental findings and offer additional reassurance that the FDA pump should not cause extensive damage to blood cells under typical operating conditions.



Copyright Justin Barrett and Matt McIntosh, University of Arkansas High Performance Computing Center

Big Insights from Tiny Molecules

A Recipe for New Drugs

In the search for new medicines, computer simulations can save researchers both time and money by reducing the need to synthesize and test chemicals that might not work. The quantum-mechanical atomistic simulation method known as Density Functional Theory (DFT) has proven especially valuable for rapidly and accurately modeling chemical phenomena and providing insights into why chemicals react the way they do.

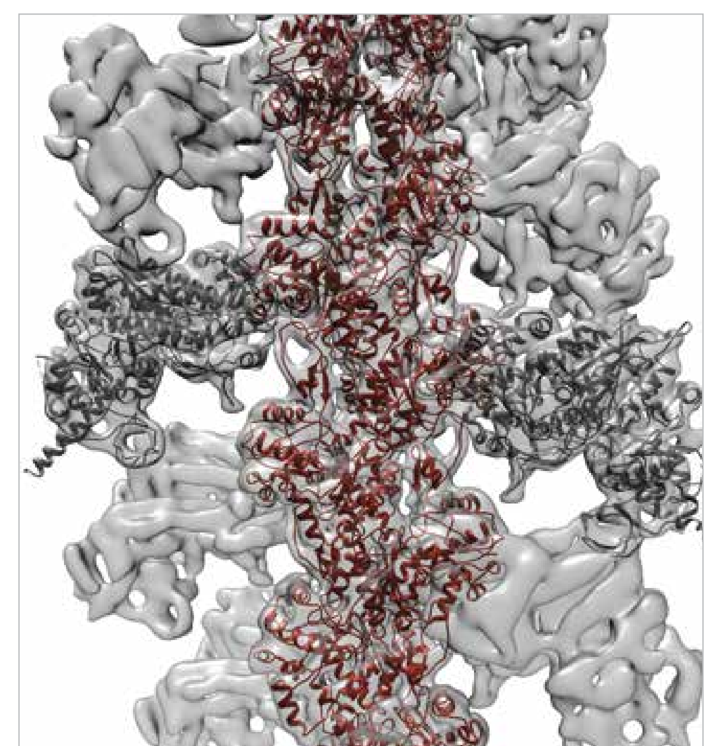
Researcher Justin Michael Barrett, a Ph.D. candidate working under Matt McIntosh at the University of Arkansas, is using DFT simulation to identify pharmaceutically useful molecules and to develop recipes to synthesize these molecules. So far, he and his team have used their findings from DFT to guide the synthesis of pyridine derivatives. Related compounds are currently being screened for HIV-1 treatment and antifungal medication.

This image shows a DFT simulation of one of the pyridine-derived compounds Barrett synthesized. The molecule's electrostatic potential field lines are colorized in a gradient from blue, for a positive charge, to red, indicating a negative charge. The mesh outline represents the electronic density of the molecule and can be interpreted as the molecule's volume. Knowing the shape of the molecule and what parts of it are positively and negatively charged can help predict how and where the compound will interact with a pharmaceutical target. The simulation was performed at the University of Arkansas High Performance Computing Center.

Natural Nanomachines

Machines may have made the modern world, but as with most things, nature was there first. Recent research has revealed the intricate workings of molecular motors that operate inside every living cell. Bridging physics and biology, the study of these tiny machines can uncover new insights into the fundamental processes that sustain life and provide new tools for medical research and more.

Krishna Chinthalapudi of The Ohio State University College of Medicine and colleagues are using the Pitzer cluster at the Ohio Supercomputer Center to analyze the molecular motors involved in muscle contraction and cell division, among other functions. Shown here is an image, captured using cryo-electron microscopy, that illustrates how myosin motors (dark gray) walk on biological tracks of actin filaments (red). Using computational methods to process images from cryo-electron microscopy, X-ray crystallography, and high-resolution fluorescence microscopy enables the scientists to magnify tiny proteins and study these enzymes at work. Based on their findings, the scientists aim to identify ways to leverage molecular machines to discover new drugs or precisely deliver therapies within cells.

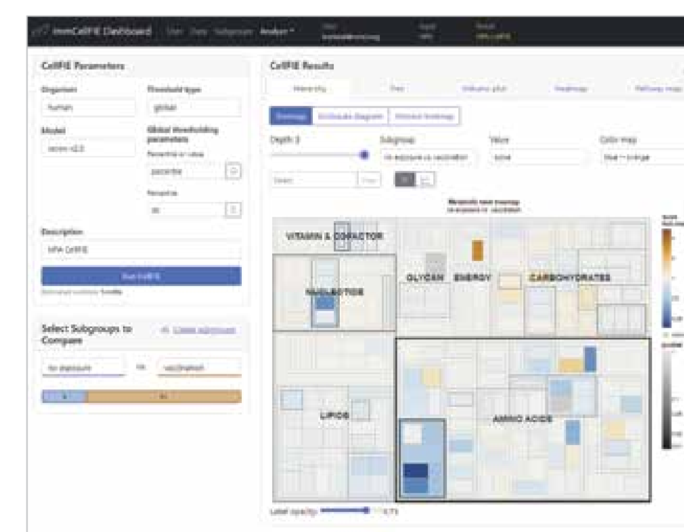


Copyright Krishna Chinthalapudi, The Ohio State University

Broadening Access to the Power of Computational Biology

Recent years have seen an explosion in the types and amount of data capturing the inner workings of our cells. Today's scientific techniques tell us not only what genes are there but how they are expressed; not only which proteins and chemicals are present, but which ones are made and consumed as cells accomplish intricate tasks. To make sense of all this information, researchers can use genome-scale models that simulate and predict biological pathways, producing detailed insights into the individual drivers behind complex cellular functions. However, the reach of these approaches has been largely limited to those with specialized training in computational systems biology.

To bring the power of computational models to more scientists, Kimberly Robasky and David Borland of the University of North Carolina at Chapel Hill Renaissance Computing Institute along with Nathan E. Lewis of the University of California, San Diego, created CellFIE. Connecting data on gene expression and metabolic processes, CellFIE generates a set of precomputed biological pathways that scientists can use as a scaffolding to help interpret their experimental results. CellFie has been used for studies such as identifying metabolic dysregulation in Alzheimer's disease and investigating the effect of certain drugs on kidney function. The ImmCellFIE portal, pictured here, provides data and CellFIE visualization tools specially curated for immunologists, helping to provide insights into the activities of the immune system and support research on vaccines, cancer, autoimmune diseases, and more.



Copyright David Borland 2022

Seeing Patterns in the Past

Using ‘Eyes in the Sky’ for Heritage Preservation

People have inhabited the part of the world now known as Afghanistan longer than almost anywhere else on the planet, with evidence of human communities dating back as far as 50,000 years. We have much to learn from the traces of our early ancestors, yet modern agriculture, development, and wars have destroyed many irreplaceable artifacts. To find and protect sites of cultural and historic importance before it’s too late, Gil Stein of the University of Chicago Oriental Institute is using state-of-the-art remote sensing and computer vision technology to uncover promising archaeological sites like the one pictured here.

For the project, called the Afghan Heritage Mapping Project, Stein and colleagues trained an AI algorithm with images from declassified Cold War-era spy plane photographs, hand-drawn topographic maps, and satellite and LIDAR scans, using the computational resources of the University of Chicago Research Computing Center to process vast amounts of data about Afghanistan’s vast terrain. Trained on 3,200 sites labeled by human researchers, the AI model scans the entirety of Afghanistan, cut up into 70 million individual tiles, in roughly a week, finding promising patterns that are then manually validated by researchers to further hone the model. The approach has already used LIDAR data to recognize evidence of ancient sites that wouldn’t have been visible to the naked eye and helped researchers focus on areas faced with imminent threat — helping to forge connections across cultures and centuries and preserve our shared past.

Copyright Gil Stein, University of Chicago Oriental Institute

Copyright Mitsui Bunko

Codes of Commerce

Relationships are at the heart of any business. Over the years, people have used a variety of methods to document identities and relationships when making business commitments such as financial arrangements and legal contracts. In East Asia, personalized stamps such as the ones pictured have often taken the place of signatures in this role. University of Pittsburgh researcher Raja Adal studies these stamps for insights into corporate histories and the nature of communication, trust, and certification.

When Adal began working with documents of the Mitsui Mi’ike Mine, a Japanese coal mine that operated throughout much of the 20th century, it quickly became apparent that the company’s trove of tens of thousands of pages of documents sporting close to 100,000 stamps would be an exceptionally rich dataset. However, a large number of the stamps only appear a few times each, creating a “long tail” distribution that makes it challenging to analyze the enormous dataset with conventional machine learning methods. Led by Paola Buitrago at the Pittsburgh Supercomputing Center (PSC), scientists from PSC, the University of Pittsburgh, Carnegie Mellon University, and DeepMap Inc. took up this challenge, using PSC’s Bridges and Bridges-2 systems to build a new two-step machine learning method that allows for a more flexible AI algorithm capable of recognizing rare data points. In addition to the historical research value of helping Adal explore a unique window into the workings of a large Japanese corporation, the work represents a microcosm of a much larger issue in deep learning. By finding better ways to handle imperfect, long-tail datasets, the team’s new AI approach can be used as a benchmark for other image classification algorithms with a similarly open-ended and highly unbalanced class set.

Tweets by the Billion

Social media has fundamentally altered political communication, delivering messages from candidates and elected officials directly into the hands of the people — and giving people a ready means to talk back. While the power of surveys and focus groups to provide useful insights on Americans’ political views may be waning, the potential to tap into the dynamic conversations happening on social media is on the rise.

University of Virginia (UVA) political scientist Paul Freedman and colleagues from UVA’s School of Data Science and Research Computing teamed up to document Twitter conversations involving America’s leaders, producing a unique database for political analysis. Their system captures and stores each tweet to and from every major presidential candidate and elected U.S. President since the 2016 primaries — a collection of about 1 billion tweets to date. Rather than simply documenting what messages candidates and presidents send out, the database reflects what messages people send to these leaders, a rich resource to understand the interplay of online conversations with real-world endorsements, scandals, and polling.

Revealing the Hidden Environment

Copyright Kyla Dahlin, Michigan State University

A Full Color View of Carbon Capture

Anticipating the future effects of climate change requires a nuanced understanding of where carbon dioxide ends up on the planet. Since trees use carbon dioxide to grow, they help to keep this greenhouse gas out of the atmosphere and some have proposed large-scale tree-planting programs as a way to curb climate change. However, it has been difficult for scientists to calculate just how much carbon dioxide trees take in because it varies by species, location, and season.

Kyla Dahlin from Michigan State University (MSU) is using a technique called hyperspectral imaging to get a better picture of how the trees in various places take up carbon. By using both the visible and shortwave infrared portions of the spectrum, hyperspectral imagery allows scientists to map the locations of particular tree species and create global predictions of how much carbon trees can absorb. Pictured is a composite image derived from a hyperspectral image of the University of Michigan Biological Station in August 2019. Differences in color within the forest represent differences in leaf chemistry and forest structure, which in turn influence how much carbon a forest can absorb. The Airborne Observation Platform, a light aircraft operated as part of the National Ecological Observatory Network, collected the hyperspectral images, which scientists processed at the MSU High Performance Computing Center in collaboration with Scott Stark from MSU, Shawn Serbin from Brookhaven National Laboratory, and the MSU Institute for Cyber-Enabled Research. This work, funded by the National Science Foundation, will help forest managers better manage forests for carbon uptake and monitor forests as they change in the future.

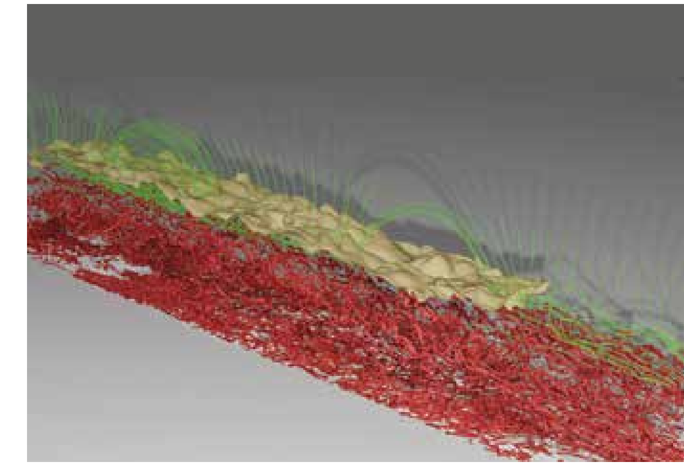
Copyright R.C. Gatti, P.B. Reich, J. Liang, and GFBI Consortium 2022

Searching for Undiscovered Trees —Before They Disappear

Even though scientists and foresters have studied trees for centuries, it is still difficult to get an accurate picture of how many tree species exist in the world. Most tree diversity estimates rely primarily on published lists of species known to exist in particular areas, but some areas are much more heavily studied than others, creating significant gaps in our knowledge of species diversity at a global scale.

Jingjing Liang from Purdue University led a group of researchers in using new data from The Global Forest Biodiversity Initiative (GFBI) and TREECHANGE to estimate the number of tree species at the biome, continental, and global scales. GFBI consists of abundance-based records for over 28,000 species of trees, and TREECHANGE contains occurrence-based data on distributions, traits, and phylogeny for around 65,000 identified tree species. The researchers used these datasets to compare species diversity across different spatial scales. Their calculations, performed using Purdue University Research Computing resources, showed an estimated 73,300 species of trees around the world, about 14% higher than previously estimated. The results suggest that most of the yet-to-be identified species are rare and only found on a single continent. They also tend to grow in tropical or subtropical climates where many forests are facing significant risks from logging, mining, and other activities.

This image shows estimates of the number of tree species and number of trees on each continent, with images of some of the most common tree species surrounding the map. Green areas represent global tree cover. The researchers merged the GFBI dataset (blue) with the TREECHANGE data (purple) to estimate tree cover and species diversity. Blue represents sample areas where species abundance data were collected, and purple represents sample areas where only species incidence/occurrence information is available.



Copyright S. Balachandar, University of Florida

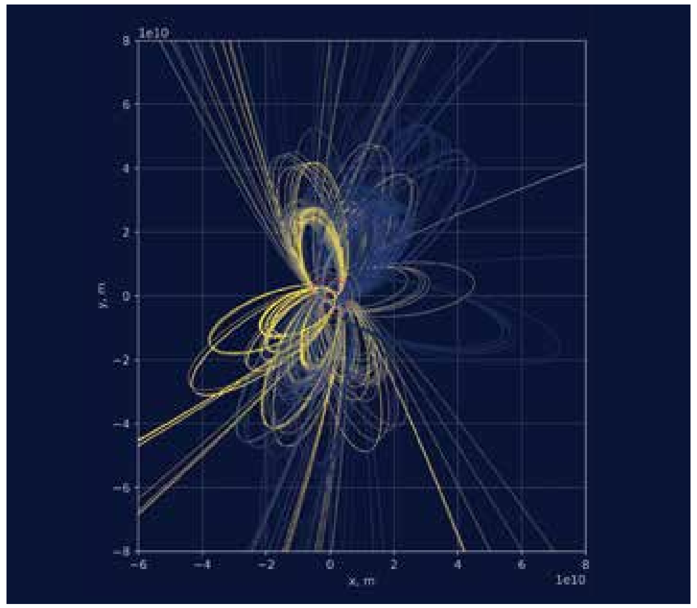
Cracking the Secrets of Underwater Oases

As rivers run into the ocean, they don't just vanish — rather, they continue to flow in underwater channels known as turbidity currents. These currents carry large amounts of sediments and often create biodiverse oases along the sea floor. Some are as deep as the Grand Canyon and travel for hundreds to thousands of miles, transporting large amounts of carbon, nutrients, and fresh water through the world's oceans.

How these sediment-laden currents can travel so far without mixing with ocean water has been something of a scientific mystery. S. Balachandar and colleagues at the University of Florida developed mathematical simulations to better understand how these currents remain self-contained. The key is a middle layer that lies between the current's smooth, mostly turbidity-free top layer — closest to the ocean water — and the churning, sediment-laden layer near the ocean floor. This middle layer prevents the near-floor turbulence from penetrating into the layer near the ocean waters by transferring energy back from turbulence to average flow.

This image shows the complex internal structure of an under-ocean turbidity current. The middle layer (brown surface) buffers the clear ambient water (green lines) from the turbid, sediment-rich part of the current (red). The simulation was performed using HiPerGator, the University of Florida's supercomputer. The unprecedented insights from these simulations can inform ocean-related science and engineering in a variety of areas, including by helping experts to more accurately estimate the extent and nature of offshore oil reservoirs.

Beyond the Horizon

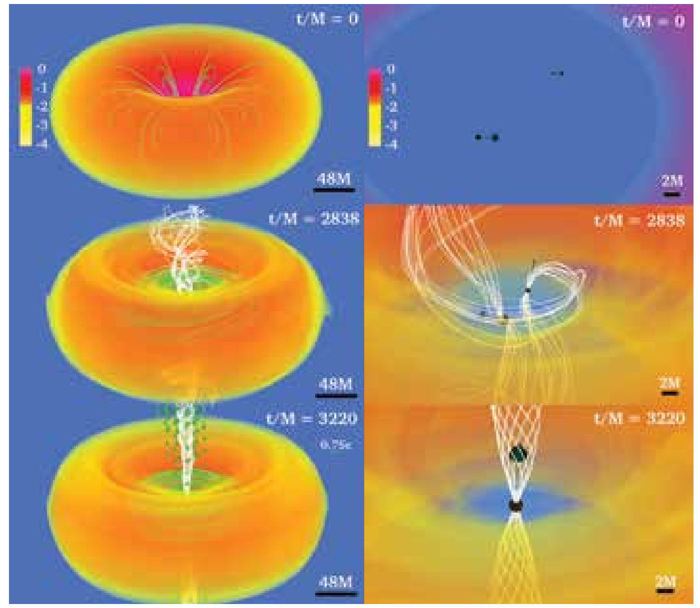


Copyright Yuri Shimane, Georgia Institute of Technology

The Hunt for Habitable Planets

People have long dreamed of finding habitable worlds beyond our solar system. But even if we achieved the technology necessary for interstellar space travel, many other challenges would need to be overcome in order to identify the most promising planets and sites for human settlement. To begin to grapple with some of these issues, the European Space Agency's Advanced Concepts Team and the Genetic and Evolutionary Computation Conference challenged experts around the world to compete to solve three complex optimization problems in an imagined space exploration scenario.

Yuri Shimane from the Georgia Institute of Technology tackled one of the problems: mapping the optimal trajectories for automated probes to survey the planets of the Trappist-1 star system. This real-life star system is thought to host as many as seven terrestrial planets, four of which show evidence of liquid water. For the challenge, Shimane calculated the paths a hypothetical probe could travel to tour the system in order to determine the best planets for habitation and identify suitable human settlement sites while minimizing fuel consumption and flight time. This image represents the 100 most optimal trajectories, according to the analyses. Contradicting objectives within the challenge created trade-offs that resulted in multiple families of trajectories with a wide variety of incoming directions and ellipse sizes. The work was done at the Georgia Institute of Technology Partnership for an Advanced Computing Environment.



Copyright Mit Kotak, University of Illinois at Urbana-Champaign

What Happens When Black Holes Collide?

Black holes represent one of the deepest mysteries of space. The gravity in these regions of space-time is so strong that nothing can escape — not even light. New telescope observations have revealed a dramatic magnetic swirl action occurring when two black holes come together, but scientists do not fully understand the complex physical processes that occur during this type of merger.

To take a closer look at merging black holes, Mit Kotak along with the Illinois Relativity Group led by Stuart Shapiro from the University of Illinois at Urbana-Champaign simulated the process on a supercomputer at the National Center for Supercomputing Applications. The simulation required new mathematical models that combined Einstein's equations describing the gravitational field around a black hole with equations governing the motion of matter moving close to the speed of light in a magnetic field. It took about six months to perform all the calculations necessary to produce these visualizations of the process. The simulation time is indicated by t/M, and the scale bar at the bottom indicates the black hole mass. Models like these help astronomers and astrophysicists measure and interpret observations of far-away cosmic events; for example, the jets coming from the black holes in this visualization could be measured by electromagnetic emission detectors, and the colors contain light intensity fluctuations that can be detected by large telescopes such as Pan-STARRS in Hawaii.

Dancing With the Stars

Stars emerge and perish over almost unfathomable scales of time and space, yet we are all connected with even the most distant cosmic phenomena through the physics of the universe and our innate human curiosity. In a stunning blending of science and art led by University of Arizona (UA) researcher Kay He (Yuanyuan), dancers took to the stage against an immersive backdrop of astrophysical simulation to tell the story of a star from birth to death. The performance, called Stellarscape, was developed as an interdisciplinary collaboration involving faculty from the UA departments of Dance, Music, Astronomy, and School of Information, with visualization support from Devin Bayly and the computational resources of UA Research Computing. The dancers' movements, tracked with a series of cameras and a virtual reality device, were used as inputs to render astronomical simulation snapshots in real time, creating an interactive experience. Linking the grace of human movement with the rhythms of the universe, the science-inspired multimedia work offered an inspiring new way for audiences to connect with the cosmos and their own curiosity.

Copyright Yuanyuan (Kay) He, University of Arizona

The Art of Inspiration

Copyright Robert Twomey, University of Nebraska-Lincoln

Putting the Art in Artificial Intelligence

Machine learning (ML) and artificial intelligence (AI) may be best known for powering robots and informing data-driven decisions in fields from finance to medicine, but these techniques are also ushering in a new era of computer-assisted creativity in the arts and humanities. Robert Twomey of the University of Nebraska-Lincoln is using the Crane system at the University of Nebraska Holland Computing Center to experiment with new forms of media arts and explore how ML and AI can extend human perception and imagination.

Twomey's experiments have included a neural network trained on children's artwork (pictured here), a ML-generated visual essay reflecting on the historical concept of the sublime, and a live performance of an AI-written radio play. Incorporating creative applications of these technologies into classes, Twomey trains students to be "computation artists," developing fluency with ML and compute-intensive workflows via hands-on exercises to creatively apply different models, compare them to human creations, and gain a deeper understanding of AI and ML techniques.

Visualization framework by Nicholas Polys and Srijith Rajamohan; photo copyright Alex Parrish for Virginia Tech

Fuel for Thought

Upheavals in energy markets have brought increased attention to nuclear power as a long-term, low-carbon domestic energy source, but the ability to safely store spent nuclear fuel is a key priority if nuclear power is to play a larger role in our energy future. Achieving this requires meticulous engineering as well as a workforce with the expertise necessary to support tomorrow's nuclear technologies.

Virginia Tech researchers Alireza Haghighat and Valerio Mascolino are advancing both of these goals with the Real-time Analysis for Particle-transport and In-situ Detection (RAPID) Code System, a simulation platform that allows users to interact with nuclear technology and collaborate with colleagues through a high-resolution multi-user virtual reality (VR) interface. The system is designed to help engineers optimize the safety of spent-fuel storage across a variety of containers, assemblies, and configurations, and it has also proven valuable as a teaching tool. This image shows nuclear engineering students working with the simulation to examine how different designs affect the fate and security of spent fuel inside the high-resolution virtual reality room managed by Virginia Tech Advanced Research Computing.

Copyright Kimberly Mann Bruch, San Diego Supercomputer Center at the University of California San Diego

Learning to Love Data

Before Amara Sanchez ever heard of data science, she already had a hand in the field — literally. The middle schooler developed a habit of recording her interactions with others in a series of symbols drawn on her hand, demonstrating an instinct for data collection that would serve her well in the year she spent analyzing the pH of a local river. That project ultimately earned Sanchez and her partner Maniya Zwicker, both from the Pala Band of Mission Indians in southern California, recognition as "Best New Team" in a national competition.

As important as the experience was to the students, their participation in the program also reflected an important development for the competition itself. DataJam, launched in 2013 as an initiative of Pittsburgh DataWorks' Cheryl Begandy, also a staff member of the Pittsburgh Supercomputing Center (PSC), challenges teachers and students to apply data analytical skills through an annual competition. Unlike most student competitions, DataJam lasts an entire school year, giving middle and high school teams time to plan and complete an independent research project while developing data analysis skills. When the COVID-19 pandemic hit, organizers wondered if the program would have to shut down, but soon realized that virtual collaboration could offer an opportunity to expand DataJam's geographic scope beyond Pittsburgh. Thanks to connections from a long-standing relationship between PSC and the San Diego Supercomputer Center, the program welcomed teams from across the country, including a team from New Jersey, in addition to Sanchez and Zwicker from California. As DataJam enters its tenth year, organizers aim to continue to expand the program into a truly national contest, bringing students, teachers, and mentors together from across the United States to foster the next generation of data science leaders.

CASC Membership

30+ Years · 100 Members

Albert Einstein College of Medicine Department of Information Technology

Arizona State University Research Computing

Boston University

Brown University Center for Computation and Visualization

Carnegie-Mellon University & University of Pittsburgh Pittsburgh Supercomputing Center

Case Western Reserve University Core Facility Advanced Research Computing

Chan Zuckerberg Biohub CZ BioHub Scientific Computing

City University of New York High Performance Computing Center

Clemson University Computing and Information Technology (CCIT)

Columbia University

Cornell University Center for Advanced Computing

Dartmouth College

George Mason University

Georgetown University UIS

Georgia Institute of Technology PACE

Harvard University

Icahn School of Medicine at Mount Sinai

Indiana University Pervasive Technology Institute

Johns Hopkins University

Lawrence Berkeley National Laboratory

Louisiana State University Center for Computation & Technology (CCT)

Massachusetts Green High Performance Computing Center

Michigan State University High Performance Computing Center

Mississippi State University High Performance Computing Collaboratory (HPC2)

National Center for Atmospheric Research (NCAR)

New York Genome Center

New York Structural Biology Center Simons Electron Microscopy Center

New York University

North Dakota University System

Northwestern University

NYU Langone Hospitals

Oak Ridge National Laboratory (ORNL) Center for Computational Sciences

Ohio Supercomputer Center (OSC)

Old Dominion University Information Technology Services

Princeton University

Purdue University Research Computing

Rensselaer Polytechnic Institute

Rice University Ken Kennedy Institute for Information Technology (K21)

Roswell Park Comprehensive Cancer Center

Rutgers University

Southern Methodist University

Stanford University

Stony Brook University Research Technologies

Texas A&M University High Performance Research Computing

Texas Tech University High Performance Computing Center

The George Washington University

The Pennsylvania State University

The University of Alabama at Birmingham IT-Research Computing

The University of Texas at Austin Texas Advanced Computing Center (TACC)

University at Buffalo, State University of New York Center for Computational Research

University of Alaska Fairbanks Research Computing Systems

University of Arizona Research Computing

University of Arkansas High Performance Computing Center

University of California, Berkeley Berkeley Research Computing

University of California, Davis HPC Core Facility

University of California, Irvine Research Cyberinfrastructure Center

University of California, Los Angeles Institute for Digital Research and Education

University of California, Merced Cyberinfrastructure and Research Technologies

University of California, San Diego San Diego Supercomputer Center (SDSC)

University of Chicago & Argonne National Laboratory Research Computing Center

University of Cincinnati Research Technologies

University of Colorado Boulder

University of Connecticut

University of Florida

University of Georgia Advanced Computing Resource Center (GACRC)

University of Hawaii Information Technology Services

University of Illinois at Chicago Advanced Cyberinfrastructure for Education and Research

University of Illinois at Urbana-Champaign National Center for Supercomputing Applications (NCSA)

University of Iowa

University of Kentucky Center for Computational Sciences

University of Louisville

University of Maryland Division of Information Technology

University of Massachusetts

University of Miami Institute for Data Science and Computing

University of Michigan Office of Research

University of Minnesota Minnesota Supercomputing Institute for Advanced Computational Research

University of Nebraska Holland Computing Center

University of Nevada, Las Vegas National Supercomputing Institute (NSI)

University of Nevada, Reno Research Computing

University of New Hampshire Research Computing Center

University of New Mexico Center for Advanced Research Computing

University of North Carolina at Chapel Hill

University of North Carolina at Chapel Hill Renaissance Computing Institute (RENCI)

University of Notre Dame Center for Research Computing

University of Oklahoma Supercomputing Center for Education and Research

University of Oregon Research Advanced Computing Services (RACS)

University of Pittsburgh Center for Research Computing

University of Rhode Island

University of Southern California Information Sciences Institute

University of Tennessee at Chattanooga SimCenter

University of Tennessee National Institute for Computational Sciences (NICS)

University of Texas at San Antonio Research Computing

University of Utah Center for High Performance Computing

University of Virginia Research Computing (RC)

University of Wyoming Advanced Research Computing Center (ARCC)

Vanderbilt University Advanced Computing Center for Research and Education

Virginia Tech Advanced Research Computing

Washington University in St. Louis

West Virginia University

Yale University Yale Center for Research Computing (YCRC)

CASC

Coalition for Academic Scientific Computation

1717 Pennsylvania Ave., NW, Ste. 1025
Washington, DC 20006

casc.org

Progress Report: 3D Visualizations of Black Hole Binaries and Disks in Full General Relativity

Mit Kotak

Introduction

Accreting black holes are interesting sources of multi-messenger signals and are crucial in explaining a range of high-energy astrophysical phenomena that we observe in our Universe. In order to understand the wealth of observations black holes can produce, it is crucial to compare them to predictions from theoretical modeling. Illinois Relativity Group's research focuses on simulations of relativistic astrophysical processes in order to understand the physics of compact objects and the basis of physics of matter under extreme conditions and the role of general relativity. These simulations provide the necessary waveforms needed for the detection of gravitational waves (GWs) by the LIGO/Virgo scientific collaboration (LVSC), KAGRA, NanoGrav, Pulsar Timing Arrays, LISA, the Einstein telescope, and other current and future detectors. They also provide descriptions of the electromagnetic (EM) signals, such as Gamma Ray Bursts (GRB) and the outgoing EM luminosity, as well as the fraction of the escaping mass that may give rise to kilonovae phenomena that can be detected by current space observatories, including the NICER, CHANDRA, and XMM-Newton X-ray instruments.

Over the summer, we focused on wrapping up two main projects, that we had been working on throughout the academic year, and getting them ready for publication.The two projects are as follows:

Single Black Hole Binaries

Most galaxies are believed to host a central supermassive black hole (SMBH) and as a result of galaxy mergers, SMBH binaries (SMBHBHs) are expected to form (see, e.g., Colpi et al. 2014 [3] for a review). These systems form unique sources for multimessenger signals. SMBHBH evolution is particularly promising as potentially strong sources of gravitational and various forms of electromagnetic (EM) emission. SMBHBHs are understood to be the loudest sources of low frequency GWs, whose detection will be one of the main scientific goals of future spaceborne interferometers such as LISA (Amaro-Seoane et al. 2017 [1]).

Black Hole Disks

Black holes (BHs) immersed in gaseous environments are ubiquitous in the Universe. Black hole-disks (BHDs) appear on a great variety of scales, reflecting their diverse birth chan- nels and sites. From the core collapse of massive stars[22, 14] and the cores of active galactic nuclei[13, 18, 16], to asymmetric supernova explosions in binary systems[5], and the merger of compact binaries where at least one of the companions is not a BH, BHDs may be formed and serve as prime candidates for multimessenger astronomy.

While the grant proposal only mentioned Single Black Hole Binaries, we were also able to complete the Black Hole Disks project during the allocation period.

Methods

Modeling these systems realistically is a central problem in theoretical astrophysics, but has proven extremely challenging, requiring the development of numerical relativity codes that solve Einstein's equations for the space time, coupled to the equations of general relativistic magnetohydrodynamics (GRHMD) for the magnetized fluids. Over the past decade, the Illinois Relativity Group's dynamical space time GRHMD code has proven itself as one of the most robust and reliable tools for theoretical modeling of such astrophysical phenomena[4, 17, 19]. The code evolves the spacetime metric by solving Einstein's field equations in the BSSN formulation (Baumgarte & Shapiro 1999 [2]; Shibata & Nakamura 1995 [20]) coupled to the equations of ideal GRMHD in a conservative scheme via high-resolution shock capturing methods for the evolution of matter and magnetic field. The methods used by the group for solving Einstein's field equations for the gravitational field, and the equations of ideal magnetohydrodynamics in curved spacetime have been described previously in (Khan et al. 2018[7, 8]), which can be refered for more details and further references.

We have also developed a Python-Visit code that allows us to visualize these simulations on large-scale supercomputing clusters. These visualizations are critical in determining if thee is a jet confined by the helical B-field lines emanating from the black-hole poles, if fluid elements in the accretion disk end up into the funnel, the spatial distribution of EM and GW signatures from our systems, etc.

Summary of Findings

Single Black Hole Binaries

We were successfully able to complete stationary stills and movies that will be displayed on our group web page[6] and referenced in our upcoming research article. The project consisted of three simulations, each having a unique spin configuration. The links for these movies are listed below:

- 00 degree Case (Spin arrows are pointed in the plane)[9]
- 45 degree Case (Spin arrows are pointed 45 degree from the plane)[10]

- 90 degree Case (Spin arrows are pointed perpendicular to the plane)[11]

These movies will help researchers understand the formation of X-shaped galaxies. The popular hypothesis is that they are formed due to SBHBH mergers. However, there are also models that show that these galaxies can be formed due to "feeding" of a supermassive black hole[15]. Our research article will test our simulations against predictions of all of these theories.

Black Hole Disks

We were able to complete stationary stills for our upcoming research article[21] that will be published within the next few weeks. In this article we analyze the spin of BHdisks and how they influence the multi messenger signals that we can detect for these systems.

Impact

This grant was crucial in funding my living expenses throughout the grant allocation period. All of the visualizations and publications, that were produced during this period wouldn't have been possible without the grant. This grant advanced the field of astronomy by helping researchers create models that can be used by astronomers to detech BHs.

This also indirectly helped advance society since detecting BHs requires breakthrough technology which in turn has immense practical benefits[12]. The difficult technical challenges in detecting gravitational waves have led to spinoff technologies that have made improvements in lasers, glass technology, semiconductor manufacturing and information processing among other applications. For these detectors to succeed, technology has been developed by LIGO scientists and engineers to measure displacements less than 1/10,000 the diameter of an atomic nucleus. Innovations in areas as diverse as lasers, optics, metrology, vacuum technology, chemical bonding and software algorithm development have resulted directly from this pioneering work.

References

[1] Pau Amaro-Seoane, Heather Audley, Stanislav Babak, John Baker, Enrico Barausse, Peter Bender, Emanuele Berti, Pierre Binetruy, Michael Born, Daniele Bortoluzzi, Jordan Camp, Chiara Caprini, Vitor Cardoso, Monica Colpi, John Conklin, Neil Cornish, Curt Cutler, Karsten Danzmann, Rita Dolesi, Luigi Ferraioli, Valerio Ferroni, Ewan Fitzsimons, Jonathan Gair, Lluis Gesa Bote, Domenico Giardini, Ferran Gibert, Catia Grimani, Hubert Halloin, Gerhard Heinzel, Thomas Hertog, Martin Hewitson, Kelly Holley-Bockelmann, Daniel Hollington, Mauro Hueller, Henri Inchauspe, Philippe Jetzer, Nikos Karnesis, Christian Killow, Antoine Klein, Bill Klipstein, Natalia Korsakova, Shane L Larson, Jeffrey Livas, Ivan Lloro, Nary Man, Davor Mance, Joseph Martino, Ignacio Mateos, Kirk McKenzie, Sean T McWilliams, Cole Miller, Guido Mueller, Germano Nardini, Gijs Nelemans, Miquel Nofrarias, Antoine Petiteau, Paolo

Pivato, Eric Plagnol, Ed Porter, Jens Reiche, David Robertson, Norna Robertson, Elena Rossi, Giuliana Russano, Bernard Schutz, Alberto Sesana, David Shoemaker, Jacob Slutsky, Carlos F. Sopuerta, Tim Sumner, Nicola Tamanini, Ira Thorpe, Michael Troebs, Michele Vallisneri, Alberto Vecchio, Daniele Vetrugno, Stefano Vitale, Marta Volonteri, Gudrun Wanner, Harry Ward, Peter Wass, William Weber, John Ziemer, and Peter Zweifel. Laser interferometer space antenna. 2017.

[2] T. W. Baumgarte and S. L. Shapiro. Numerical integration of Einstein's field equations. *Phys. Rev. D*, 59(2):024007, January 1999.

[3] M. Colpi. Massive Binary Black Holes in Galactic Nuclei and Their Path to Coalescence. *Space Science Rev*, 2014.

[4] Zachariah B. Etienne, Vasileios Paschalidis, and Stuart L. Shapiro. General relativistic simulations of black hole-neutron star mergers: Effects of tilted magnetic fields. *Phys.Rev.*, D86:084026, 2012.

[5] T. Fragos, M. Tremmel, E. Rantsiou, and K. Belczynski. BLACK HOLE SPIN–ORBIT MISALIGNMENT IN GALACTIC x-RAY BINARIES. *The Astrophysical Journal*, 719(1):L79–L83, jul 2010.

[6] IRG. Illinois relativity group. http://research.physics.illinois.edu/CTA/IRG/, 2022.

[7] Abid Khan, Vasileios Paschalidis, Milton Ruiz, and Stuart L. Shapiro. Disks Around Merging Binary Black Holes: From GW150914 to Supermassive Black Holes. *Phys. Rev.*, D97(4):044036, 2018.

[8] Khan A. Khan 2018 simulation. https://www.youtube.com/watch?v=kNF9No8Verw, 2018.

[9] Kotak, M. Kotak 00 2022 simulation. https://drive.google.com/file/d/1jM3bsWaOwj3u7e8Tpt5nV1iDBKDvVpWG/view?usp=sharing, 2022.

[10] Kotak, M. Kotak 45 2022 simulation. https://drive.google.com/file/d/1fPqeE_ypUUYEXfAmLqChkZw_qa1QfxNe/view?usp=sharing, 2022.

[11] Kotak, M. Kotak 90 2022 simulation. https://drive.google.com/file/d/1K1iaN_7nZO4ivFOZYNBI3GO3mqVJNlce/view?usp=sharing, 2022.

[12] LIGO. Ligo benefits. https://www.ligo.org/science/faq.php#spinoffs, 2018.

[13] D. LYNDEN-BELL. Galactic nuclei as collapsed old quasars. *Nature*, 223(5207):690–694, Aug 1969.

[14] A. I. MacFadyen and S. E. Woosley. Collapsars: Gamma-ray bursts and explosions in "failed supernovae". *The Astrophysical Journal*, 524(1):262–289, oct 1999.

[15] NW. Nw article. https://news.northwestern.edu/stories/2022/08/x-shaped-radio-galaxies-might-form-more-simply-than-expected/, 2022.

[16] B. Paczynski. A model of selfgravitating accretion disk. , 28:91–109, January 1978.

[17] Vasileios Paschalidis, Yuk Tung Liu, Zachariah Etienne, and Stuart L. Shapiro. The merger of binary white dwarf–neutron stars: Simulations in full general relativity. *Phys.Rev.*, D84:104032, 2011.

[18] N. I. Shakura and R. A. Sunyaev. Black holes in binary systems. Observational appearance. , 24:337–355, January 1973.

[19] Stuart L. Shapiro and Vasileios Paschalidis. Self-interacting dark matter cusps around massive black holes. *Phys.Rev.*, D89:023506, 2014.

[20] M. Shibata and T. Nakamura. Evolution of three-dimensional gravitational waves: Harmonic slicing case. *Phys. Rev. D*, 52:5428–5444, November 1995.

[21] Antonios Tsokaros, Milton Ruiz, Stuart L. Shapiro, and Vasileios Paschalidis. Self-gravitating disks around rapidly spinning, tilted black holes: General relativistic simulations.

[22] S. E. Woosley. Gamma-Ray Bursts from Stellar Mass Accretion Disks around Black Holes. , 405:273, March 1993.

Localizing CyberGIS-Compute through Containers

Mit Kotak
mitak2@illinois.edu
University of Illinois Urbana-Champaign
Urbana, Illinois, US

Zimo Xiao
zimox2@illinois.edu
University of Illinois Urbana-Champaign
Urbana, Illinois, US

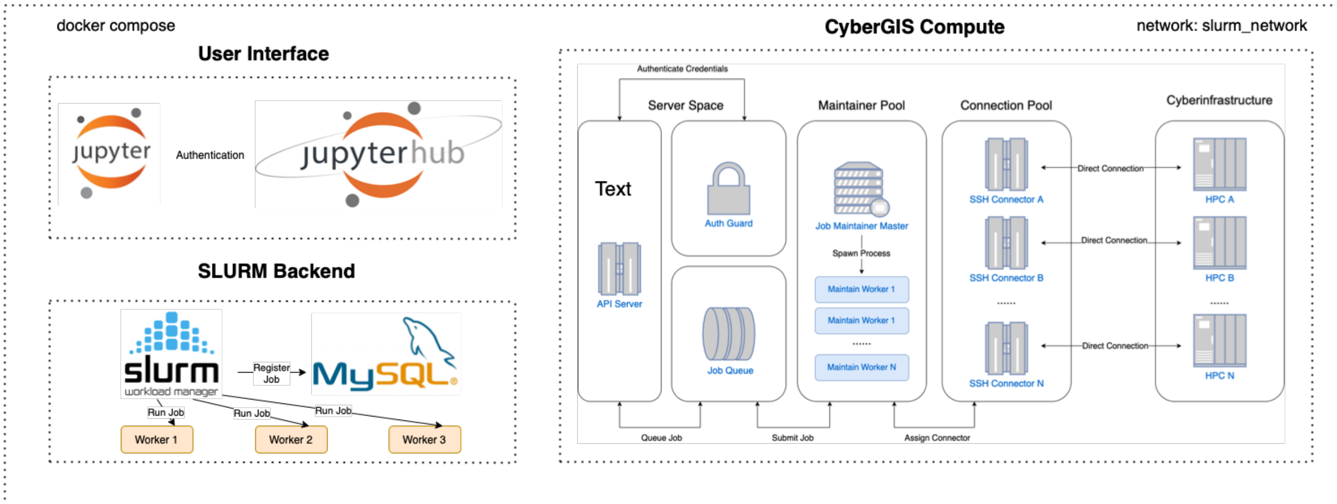


Figure 1: Overview of Local CyberGIS-Compute Container

ABSTRACT

In recent years there has been a push towards solving complex geospatial problems using advanced cyberinfrastructure (CI) systems, broadly defined as CyberGIS. While this allows the geosciences to tackler larger and more complex problems, adoption of cyberGIS techniques are slowed by the technical barriers associated with CI. Within the CyberGIS ecosystem, CyberGIS-Compute has emerged as a promising middleware for bridging access to High Performance Computing Resources (HPC) by providing a simple, easy-to-use Jupyter interface. However, its integration with external HPC resources makes it difficult for model contributors and developers to develop, debug, and test. We present a self contained Dockerized framework that can independently run the full CyberGIS-Compute stack on local machines. This streamlines the development process for model contributors and developers in the CyberGIS-Compute community.

CCS CONCEPTS

• **Applied computing → Earth and atmospheric sciences**; • **Information systems → Geographic information systems**.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

KEYWORDS

Containers, HPC, CyberGIS

ACM Reference Format:
Mit Kotak and Zimo Xiao. 2022. Localizing CyberGIS-Compute through Containers. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

CyberGIS, defined as geographic information science and systems based on advanced cyberinfrastructure (CI) (e.g., High Performance Computing (HPC) resources), has transformed how large-scale geospatial problem solving is conducted [7]. However, leveraging cybergis capabilities and CI resources is challenging, both due to the steep learning curve. CyberGIS-Compute [4] has emerged as a promising middleware that provides access to HPC resources through CyberGIS-Jupyter[5], an easy-to-use Jupyter interface. Applications of CyberGIS-Compute involve remote sensing data fusion [2], spatial accessibility [1], and hydrological modeling.

CyberGIS-Compute also provides a unique model contribution mechanism that allows community members to easily share their models with the users. This not only ensures reproducibility, but also compresses the model's complexity into a simple user interface where the model contributor can preconfigure the user-facing parameters. Model contributors package their models as Github repositories with metadata describing how CyberGIS-Compute should run the model. While this mechanism greatly simplifies the process of using these models, model contributors have to still test and debug their models on HPC resources in order to integrate them into the CyberGIS-Compute framework. This is not ideal since HPC resources cannot be securely shared across the model contributor community and often provides limited control over the model environment. The integration with HPC also complicates development and testing of CyberGIS-Compute.

We propose a novel framework that runs CyberGIS-Compute on local machines. This greatly simplifies the development process since model contributors: (a) can quickly configure the dependencies needed for the model by trying out different Singularity images, (b) can ensure that their model is able to pass through different layers of CyberGIS-Compute and successfully run on the desired back-end resource, and (c) can simulate different HPC configurations to see how their model might get parallelized across multiple nodes. We use Docker microservices [3] to create separate layers within the same network, thus providing a self-contained local deployment of all the components of CyberGIS-Compute.

2 ARCHITECTURE

The CyberGIS-Compute architecture can be broadly classified into three components: JupyterHub, CyberGIS-Compute and HPC infrastructure. These components are distributed making it tedious to deploy and debug CyberGIS-Compute, especially for community members without direct access to the infrastructure. On top of that, CyberGIS-Compute consists of multiple layers as shown in Figure 1 which makes it difficult to trace bugs through different stacks.

Our approach involves creating separate Docker containers for all of these services and then running them on the same machine and network. Instead of relying on CyberGIS-Compute logs, model contributors and developers have direct access to docker logs which helps breakdown their workflow into manageable chunks.

The various Docker containers can be grouped into three sections: user interface, Cybergis-Compute and SLURM backend. The user interface consists of a JupyterHub which spawns individual user containers. The user container can be configured to match the user's preferred Jupyter environment whereas the JupyterHub container manages user authentication for CyberGIS-Compute. These settings are be helpful to users who might be interested in setting up CyberGIS-Compute deployments for their research group or department. The Cybergis-Compute library is packaged into a separate container which provides RESTful services. The HPC backend is replaced by a set of SLURM containers [6] which can be configured to match different HPC machine configurations.

We faced two main challenges while creating this framework. The first one was networking. We had to map out all of the different SSH and REST calls that were happening across different services onto a single network without modifying any of these services since we wanted to replicate CyberGIS-Compute's behavior on HPC resources. The next challenge was to install all of the dependencies typically found on HPC machines while ensuring that the deployment was light enough to run on user machines.

3 CONCLUDING DISCUSSION

We envision this framework will empower the CyberGIS-Compute community to take a more active role in its development. Model contributors can integrate this framework into their model testing workflow to ensure long term compatibility with CyberGIS-Compute. Community developers can use this tool to locally test their contributions of new features and submitting bug reports. Lastly, this framework will greatly simplify deploying CyberGIS-Compute for those who wish to support their own instance. This will help realize CyberGIS-Compute's vision towards democratizing access to HPC resources.

ACKNOWLEDGMENTS

The authors would like to thank Alexander Michels, Anand Padmanabhan and Shaowen Wang for providing feedback on the abstract. This material is based upon work supported by the Institute for Geospatial Understanding through an Integrative Discovery Environment (I-GUIDE) that is supported by the National Science Foundation (NSF) under award No. 2118329. The material is also based in part upon work supported by NSF under grant No. 1833225. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF. Our computational work used Virtual ROGER, which is a geospatial supercomputer supported by the CyberGIS center for Advanced Digital and Spatial Studies and the School of Earth, Society and Environment at the University of Illinois Urbana-Champaign.

REFERENCES

[1] Jeon-Young Kang, Alexander Michels, Fangzheng Lyu, Shaohua Wang, Nelson Agbodo, Vincent L Freeman, and Shaowen Wang. 2020. Rapidly Measuring Spatial Accessibility of COVID-19 Healthcare Resources: A Case Study of Illinois, USA. *International journal of health geographics* 19, 1 (2020), 1–17.

[2] Fangzheng Lyu, Zijun Yang, Zimo Xiao, Chunyuan Diao, Jinwoo Park, and Shaowen Wang. 2022. CyberGIS for Scalable Remote Sensing Data Fusion. In *Practice and Experience in Advanced Research Computing (PEARC '22)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/3491418.3535145

[3] Dirk Merkel et al. 2014. Docker: lightweight linux containers for consistent development and deployment. *Linux j* 239, 2 (2014), 2.

[4] Anand Padmanabhan, Zimo Xiao, Rebecca Vandewalle, Furqan Baig, Alexander Michels, Zhiyu Li, and Shaowen Wang. 2021. CyberGIS-Compute for Enabling Computationally Intensive Geospatial Research. In *SpatialAPI'21: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on APIs and Libraries for Geospatial Data Science.* https://doi.org/10.1145/3486189.3490017

[5] Anand Padmanabhan, Zimo Xiao, Rebecca Vandewalle, Alexander Michels, and Shaowen Wang. 2021. Enabling Computationally Intensive Geospatial Research on CyberGIS-Jupyter with CyberGIS-Compute. In *Proceedings of Gateways 2021*. Zenodo. https://doi.org/10.5281/zenodo.5570056

[6] Giovanni Torres. 2013. Project Title. https://github.com/giovtorres/slurm-docker-cluster.

[7] Shaowen Wang. 2010. A CyberGIS Framework for the Synthesis of Cyberinfrastructure, GIS, and Spatial Analysis. *Annals of the Association of American Geographers* 100, 3 (June 2010), 535–557. https://doi.org/10.1080/00045601003791243