



FRONTERA USER'S MEETING WELCOME AND INTRO

Dan Stanzione
PI, Frontera project
January 2021

WELCOME AND THANK YOU!

PURPOSE

- ▶ Given the pandemic, and everyone spending far too much time on zoom, we've tried to cut the agenda down to just the most important things:
 - ▶ We will update you on the project, and the future (this morning and tomorrow).
 - ▶ You will hear from your colleagues about exciting ways Frontera is being used (11 application talks and one technology talk).
 - ▶ You can tell us what we are doing wrong! (Two Feedback Panels)

WHAT IS FRONTERA? A PROJECT AND AN INSTRUMENT

▶ The Instrument

- ▶ The NSF's Leadership-class Computer System.
- ▶ Awarded in September 2018, began production operations in August 2019 (on budget).
- ▶ **Debuted as the #5 Computer in the world.**
 - ▶ Only one in Top 10 under \$100M.
- ▶ For some problems, was the fastest in the world (because of our focus on programmability, general purpose architecture).
- ▶ Over a year old, still #9 in the world
 - ▶ New #1 is Japan's \$1B+ Fugaku
 - ▶ Followed Frontera in the "All CPU" path.

▶ The Project:

- ▶ ~30 FTEs spread over ~50 people at TACC + 10 partner institutions
- ▶ Sysadmin, operators, networking, security, user support, allocations
- ▶ Scientific code support, performance optimization and tuning, data management and curation, AI support, etc.
- ▶ Education and Outreach (K-12, undergrad, grad, MSI engagement, training, public)
- ▶ Project Leadership and Management



FRONTERA VIRTUAL TOUR



WHAT'S FRONTERA DONE IN A YEAR (AND A HALF)?

- ▶ About 30% of the time was diverted to emergency COVID response work between March-August (still a smaller fraction).
- ▶ Number of Jobs run: **Over 2.4 Million (Job 2,461,862 started this morning)**
- ▶ Node hours delivered: **80M (4 Billion+ Core Hours)**
- ▶ Uptime is over 96%; the CPU and 2 GPU systems all have utilizations of 80% or higher.
- ▶ Requests in the last round were about 3x available capacity.
 - ▶ All "Highly Recommended" requests got at least 90% of the requested time.
- ▶ User Tickets: **1,043**
- ▶ Frontera Community Slack: **12,777** messages (on average, 80 unique users each week)
- ▶ Cybersecurity incidents: **ZERO**
- ▶ 9 academic courses, 3 summer institutes, conference tutorials, lots of online training.

A FEW MORE ACCOMPLISHMENTS. . .

- ▶ Top 500, still #8. I/O 500, #3
- ▶ Used for the Gordon Bell Special Prize for COVID-19
- ▶ Cover of Science first week of January!
- ▶ Portal is up and running
- ▶ Commercial cloud integration is functional with Azure, AWS
- ▶ Founding partner in the COVID-19 HPC Consortium
- ▶ First class of Fellowships underway.

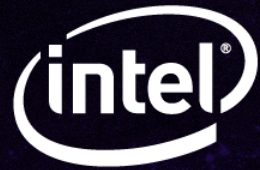


THE TEAM - INSTITUTIONS

- ▶ Operations: TACC, Oden Institute, Ohio State University (MPI/Network support), Cornell (Online Training), Texas A&M (Campus Bridging)
- ▶ Science and Technology Drivers and Phase 2 Planning: Cal Tech, University of Chicago, Cornell, UC-Davis, Georgia Tech, Princeton, Stanford, Utah



THE TEAM - VENDORS



THE HARDWARE (ORIGINAL) – COMPUTE

- ▶ The “main” compute system: 8,008 Dual-Socket Intel Xeon 8280 (Cascade Lake) nodes, 192GB of RAM, 56 cores, HDR Infiniband
 - ▶ Peak (39PF) and HPL (23PF) performance considers only these nodes
- ▶ FronteraRTX
- ▶ Frontera “Longhorn” subsystem
- ▶ Frontera NVDIMM – 16 nodes with 6TB (each) of Intel Optane NVDIMM
 - ▶ DAOS testbed using these nodes basis of I/O 500 submission

EXPANSION

- ▶ Since we dedicated so much time last year to COVID work (that we can't get back).
- ▶ With gracious support from NSF and Dell, we are expanding the system to provide some extra cycles over the next few years.
 - ▶ Availability starting this week!
 - ▶ 396 additional compute nodes installed (now 8,404 total!)
 - ▶ Additional GPU time.

THE HARDWARE – STORAGE

- ▶ First Tier (for certain users): All-flash
 - ▶ DDN hardware
 - ▶ Currently mostly used in production with DDN IME software layer – that will likely evolve to a pure flash filesystem.
 - ▶ 3PB capacity, 1.2TB/sec peak speed
- ▶ Second Tier: Home + 3 scratch filesystems
 - ▶ DDN Disk, ~50PB, ~300TB/sec
 - ▶ Users distributed among scratch1, scratch2, scratch3 (3 has highest BW).
- ▶ Third Tier: Sitewide “/work” filesystem (Stockyard)
- ▶ Fourth Tier(s):
 - ▶ Data publication/collaboration system (Corral)
 - ▶ Archive System (Ranch) – encrypted and unencrypted options now available.

FRONTERA SYSTEM --- INFRASTRUCTURE

- ▶ Frontera consumes almost 6 Megawatts of Power at Peak
- ▶ Direct water cooling of primary compute racks (CoolIT/DellEMC)
- ▶ Oil immersion Cooling (GRC)
- ▶ Solar, Wind inputs (~30% today)



TACC Machine Room Chilled Water Plant

FRONTERA – RTX SUBSYSTEM

- ▶ 90 nodes
- ▶ 4 NVIDIA Quadro 5000 RTX cards per node
- ▶ 16 GB per card
- ▶ 11 TFlop single-precision per card
- ▶ Mineral oil-cooled immersion
- ▶ Additional GPU modules
- ▶ Accessed via rtx queues



LONGHORN

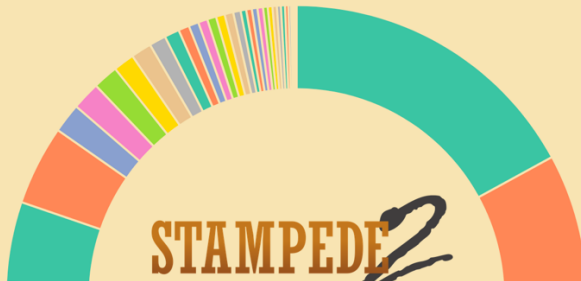
If you have used DOE's *Sierra* system at Livermore, it's that.

- ▶ 104 IBM Power9 nodes
- ▶ 4 NVIDIA V100 GPUs per node
- ▶ V100s connected with NVLINK on node
- ▶ 7 TFlop double-precision performance per card
- ▶ 14 TFlop single-precision performance per card
- ▶ 16 GB per card

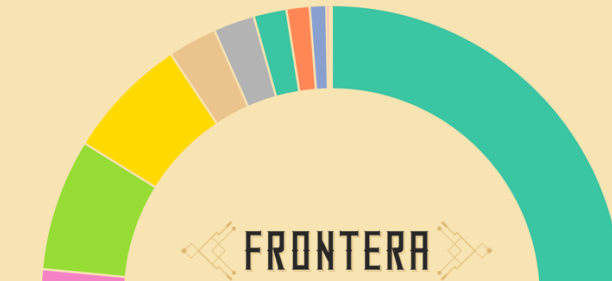


A FOCUS ON LARGE SCALE

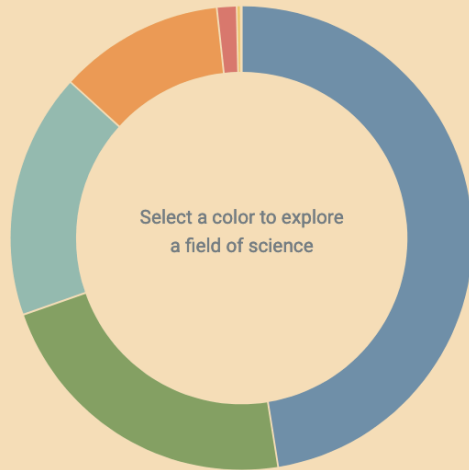
- ▶ Frontera is not *that* much bigger than Stampede2 (the physical footprint is actually smaller!).
- ▶ But it was built with the *largest* computational problems in mind.
- ▶ And we try to use it that way. . .



[Explore More stamped2 >](#)



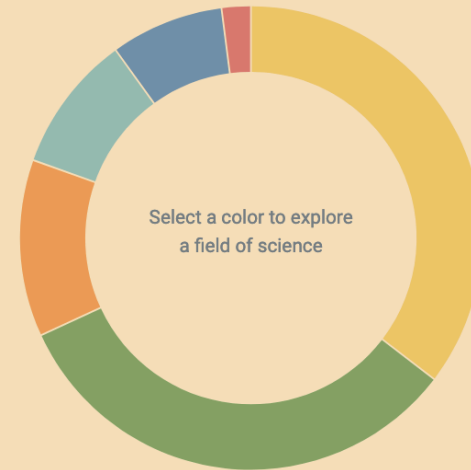
[Explore More frontera >](#)



timestamp : Jul 17, 2020, 4:28:36 PM

All Stampede2 Research Projects (207)

Institution	Domain Science	Jobs	Nodes	Project Description
Texas Advanced Computing Center	COMPUTER AND INFORMATION SCIENCE AND ENGINEERING (CISE)	4	4	View Description
University of California, Santa Barbara	COMPUTER AND INFORMATION SCIENCE AND ENGINEERING (CISE)	1	16	View Description
Texas Advanced Computing Center	COMPUTER AND INFORMATION SCIENCE AND ENGINEERING (CISE)	1	1	View Description
SUNY - Farmingdale State College	COMPUTER AND INFORMATION SCIENCE AND ENGINEERING (CISE)	1	5	View Description
Department of Electrical and Computer Engineering	COMPUTER AND INFORMATION SCIENCE AND ENGINEERING (CISE)	1	1	View Description



timestamp : Jul 17, 2020, 4:27:33 PM

All Frontera Research Projects (42)

Institution	Domain Science	Jobs	Nodes	Project Description
California Institute of Technology	MATHEMATICAL AND PHYSICAL SCIENCES (MPS)	7	49	View Description
University of Utah	MATHEMATICAL AND PHYSICAL SCIENCES (MPS)	2	256	View Description
Los Alamos National Laboratory	MATHEMATICAL AND PHYSICAL SCIENCES (MPS)	1	128	View Description
University of Chicago	MATHEMATICAL AND PHYSICAL SCIENCES (MPS)	1	1	View Description
Northwestern University	MATHEMATICAL AND PHYSICAL SCIENCES (MPS)	2	2	View Description

Project Name

A-ccsc

TG-ASC160051

TG-ASC160041



A FOCUS ON LARGE SCALE

- ▶ Running the largest jobs is still something that takes some work.
- ▶ The **Texascale Days** are periods of time where we let teams push to new scales.
 - ▶ A focus on 4k/8k node jobs (200k/400k+ cores).
 - ▶ A period of exclusive access to debug and restart at large scale without having to re-queue.
 - ▶ Using the whole system for the very largest problems.
- ▶ 11 teams have requested access for next week's period.
- ▶ (with the new expansion nodes, we hope to allow some small jobs to continue to run when the system is otherwise occupied with Texascale jobs).

OTHER MODES TO THINK ABOUT...

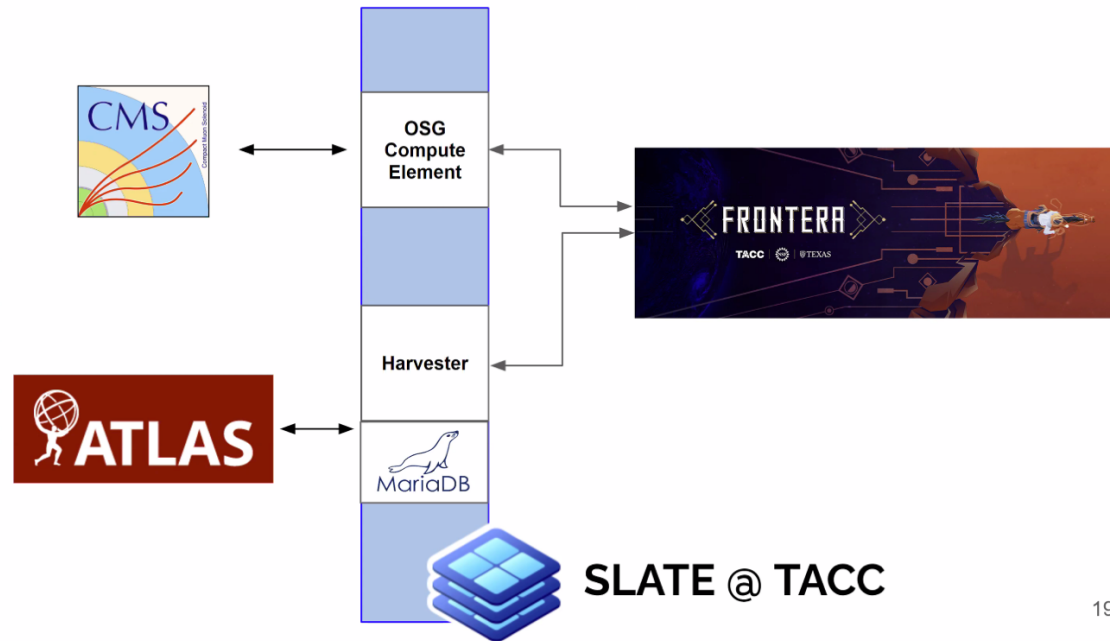
- ▶ Everyone here knows how to run batch jobs for the “normal system”, but in addition to Texascale, remember some of our other “special” capabilities:
- ▶ Hardware:
 - ▶ Single precision and double precision GPU systems
 - ▶ Large memory/fast storage NVDIMM nodes (on their own, or in conjunction with other nodes).
 - ▶ Flash filesystem for the highest I/O demand jobs.
- ▶ Software:
 - ▶ API support for automated workflows
 - ▶ Cloud storage integration
 - ▶ Interactive jobs for the portal (including Jupyter, Visualization, etc.).
 - ▶ Reservations where QoS matters.

FRONTERA AS-A-SERVICE: EMBEDDING IN AUTOMATED WORKFLOWS: WORKFLOW SUPPORT IN HIGH ENERGY PHYSICS

Examples of Federated Operations Gateways



U.S. ATLAS and U.S. CMS are both investigating or developing SLATE-based solutions for sending workloads to the Frontera supercomputer at TACC.



The **Frontera Portal** is a web-based research platform that provides computational tools to manage, analyze, and understand critical data for research.

The Portal has cloud-based tools that support the analysis and integration of diverse data types.

FRONTERA Training User Guide Allocations Fellowships News About Help [Log in](#)

FRONTERA
TACC | TEXAS

COVID-19 Updates From TACC: Plans are in place at TACC to continue providing normal support for computing operations. We don't anticipate any interruption in the availability of computing resources or user support staff due to the COVID-19 situation.

Upcoming & Archived Training

Frontera leverages TACC and its partners' innovative approach to education, outreach, and training to encourage, educate, and develop the next generation of leadership-class computational science researchers.

10/29 High Performance Computing on Frontera Day 2	10/22 High Performance Computing on Frontera Day 1	02/06 Introduction to PETSc
---	---	---------------------------------------

UPCOMING ALLOCATIONS OPPORTUNITIES:

- [Pathways & LSCP Quarterly Allocations Submissions](#)
Deadline: 8/15/2020
- [LRAC Allocations Submissions](#)
Deadline: 1/15/2021

STATUS	LOAD	RUNNING	QUEUED
Operational	84%	950	249

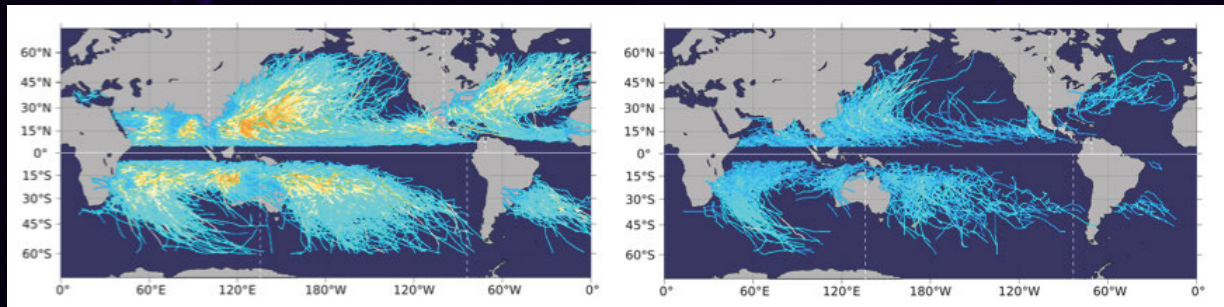
Getting Started

A brief overview of system plans and the broader project that surrounds it, the architectural design choices, and a discussion of application community that will run on it.

[Welcome to Frontera](#)

FRONTERA BRIDGES TO THE COMMERCIAL CLOUD: NEW HIGH-RESOLUTION GLOBAL CLIMATE MODEL PREDICTS CHANGES TO EXTREME WEATHER, OCEAN CURRENTS

- ▶ Ping Chang, Texas A&M (collaborators at NCAR and QNLM).
- ▶ CESM, 25KM resolution
- ▶ One of the UN-supported HighResMIP simulations – the official climate change forecasts.
- ▶ In addition to large simulations, also published the data results to Azure, where they do online analytics on ~550GB of data with the “PanGeo” Cloud platform.



ALLOCATIONS

- ▶ We've tried to streamline this process for you.
- ▶ Three Main Tracks:
 - ▶ **Leadership Resource Allocations** – Ready to run at large scale, 250k-5M node-hours per year. (Currently 49 projects active)
 - ▶ **Pathways Allocations** – Not yet at that scale, but scientific potential to get there, up to 200k node hours.
 - ▶ **Large Scale Community Partnerships** - For Gateways, Community Codes, or large scientific collaborations, up to 3 years, 25k-1M hours per year.
- ▶ Also Startup, Educational, and Discretionary allocations

ALLOCATIONS

	Leadership Resource Allocation (LRAC)	Pathways	Large-scale Community Partnerships (LSCP)	Totals
Requests	62	38	13	113
Unique Pis	62	37	12	103
Unique Orgs	45	31	12	69
CPU SUs Requested	93,914,427	5,634,333	12,377,780	111,926,540
CPU SUs Awarded	54,076,692	3,433,463	6,086,100	63,596,255
GPU SUs Requested	2,545,684	439,665	1,385,600	4,370,949
GPU SUs Awarded	613,684	146,975	160,000	920,659

▶ Three Main Tracks:

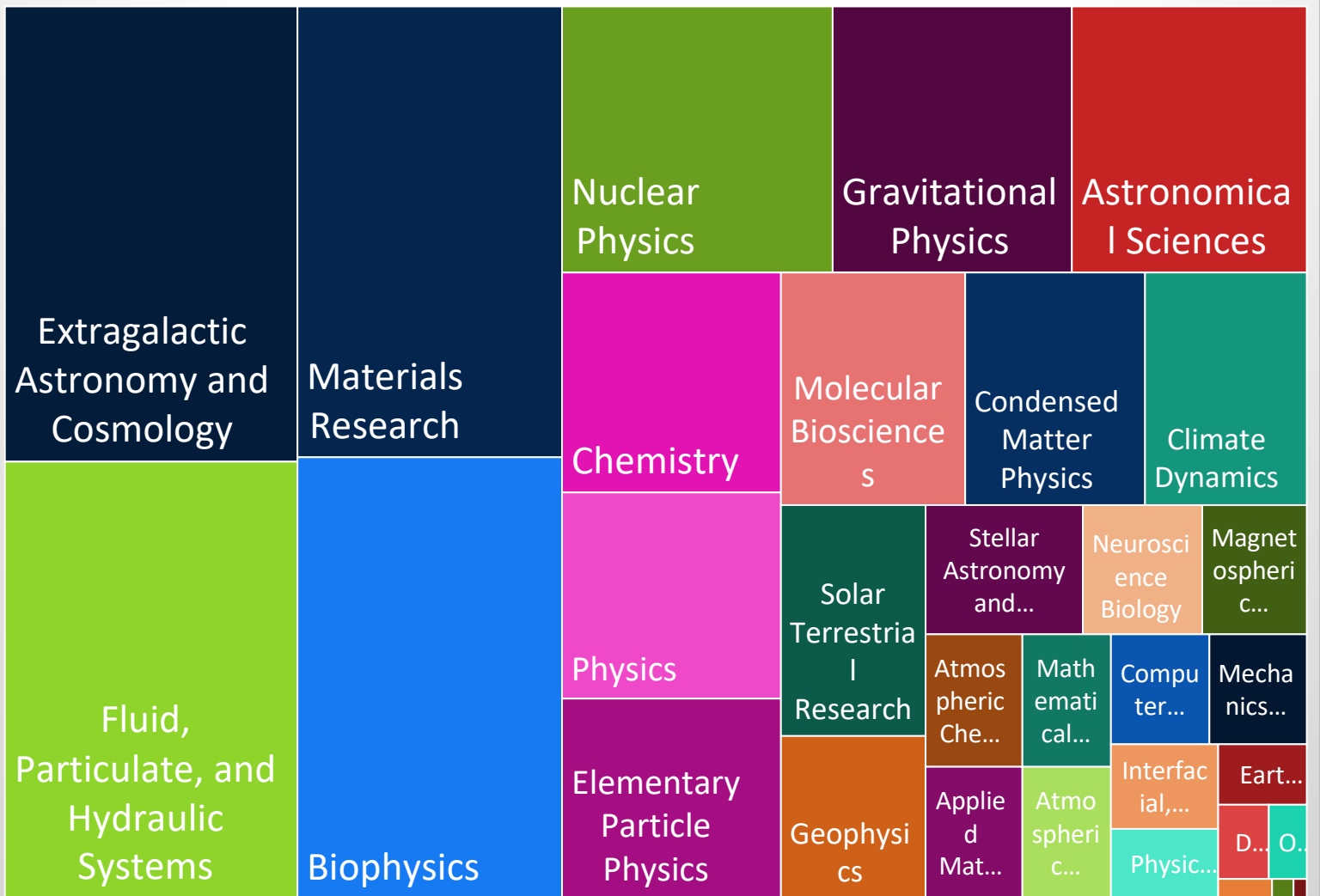
- ▶ **Leadership Resource Allocations** – Ready to run at large scale, 250k-5M node-hours per year. (Currently 49 projects active)
- ▶ **Pathways Allocations** – Not yet at that scale, but scientific potential to get there, up to 200k node hours.
- ▶ **Large Scale Community Partnerships** - For Gateways, Community Codes, or large scientific collaborations, up to 3 years, 25k-1M hours per year.

▶ Also Startup, Educational, and Discretionary allocations

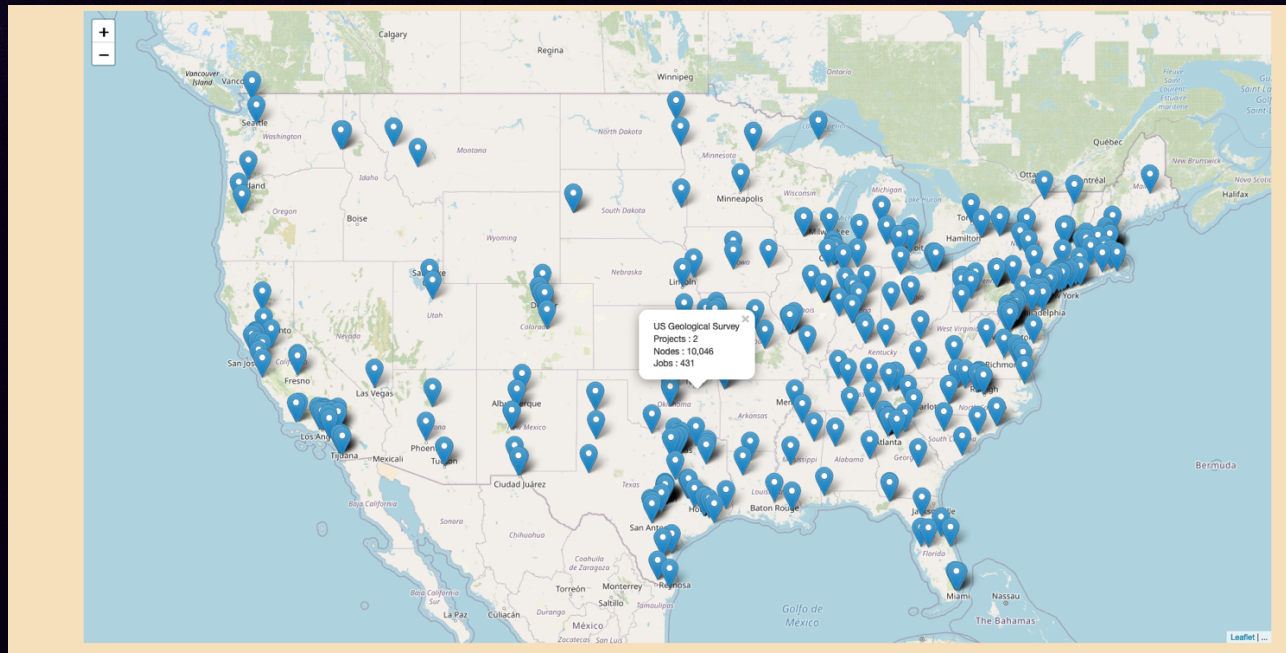
- ▶ Allocations are widely advertised to the community through multiple channels, including an NSF DCL (NSF 20-018 was the most recent: <https://www.nsf.gov/pubs/2020/nsf20018/nsf20018.jsp>)

FIELDS OF SCIENCE

► From last allocation request



GEOGRAPHIC DISTRIBUTION



2020-2021 FRONTERA FELLOWS



Gabriele Bozzola
University of Arizona

Field of Research:
Physics & Astrophysics



Maureen Kitheka
State University of NY
at Binghamton

Field of Research:
Computational
Chemistry



Danat Issa
Northwestern
University

Field of Research:
Astrophysics



Jeane Camelo
University of Central
Florida

Field of Research: Coastal
Engineering

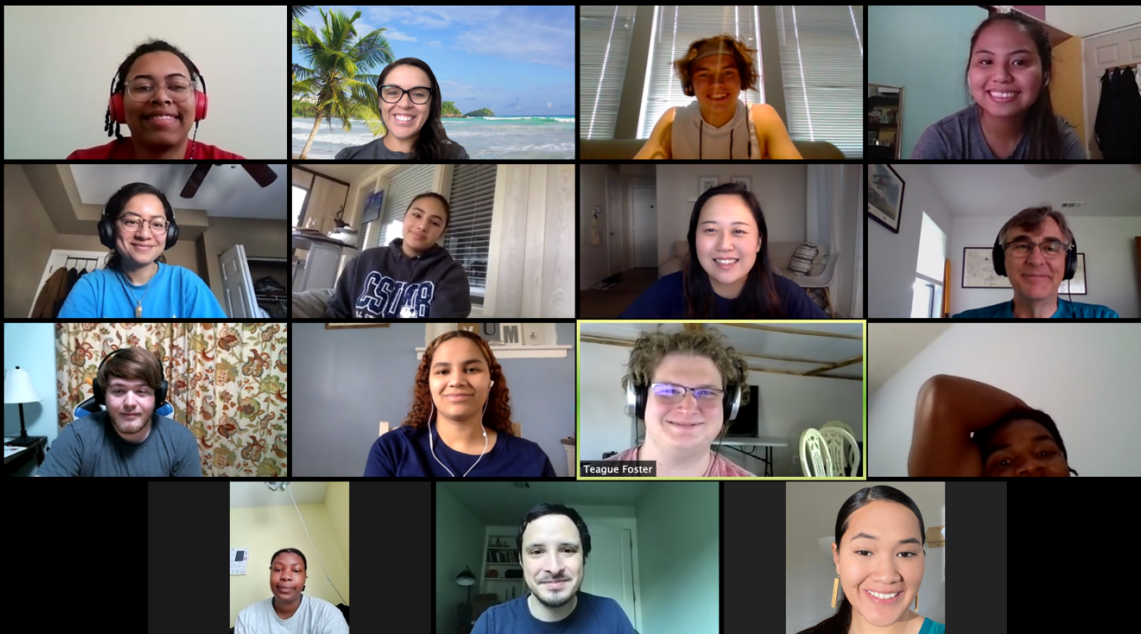


Dubem Okafor
University of Iowa

Field of Research:
Elastoplastic wave
propagation in
impacting media

2020 REU COHORT

JUNE 8-AUGUST 7



Syema Ailia
CS, Northeastern Illinois University

Teague Foster
CS, Central Texas College

Carolyn Gonzalez
CP & IS, Farmingdale State College

Zach Steed
CS, Hendrix College

Tyshawn Ferrell
BIO, Albany State University

Regina-Mae Dominguez
Math/CS, University of Guam

Jakerra Lance
CIS/History, Washington & Jefferson
College

Erin Searcy
Math, Fort Valley State University

Ricky Brown
CIT, Broward College

Mariana Duarte
CS, CA State University Monterey Bay

Courtesy Rick
Stevens, U. of
Chicago/ANL

Cryo-EM

FFEA

all-atom

AI-driven multi-resolution simulations of SARS-CoV-2

- All-atom viral envelope simulations are limited by timescales that are accessible
- Experimental data still challenging to integrate
- AI is used as a “glue” to drive adaptive sampling and improve accessible experimental observables
- Multi-resolution techniques implemented using FFEA:
 - intermediate representation from experimental data
 - AI builds automatic conversion across multiple scales

EM

Continuum scale

all-atom

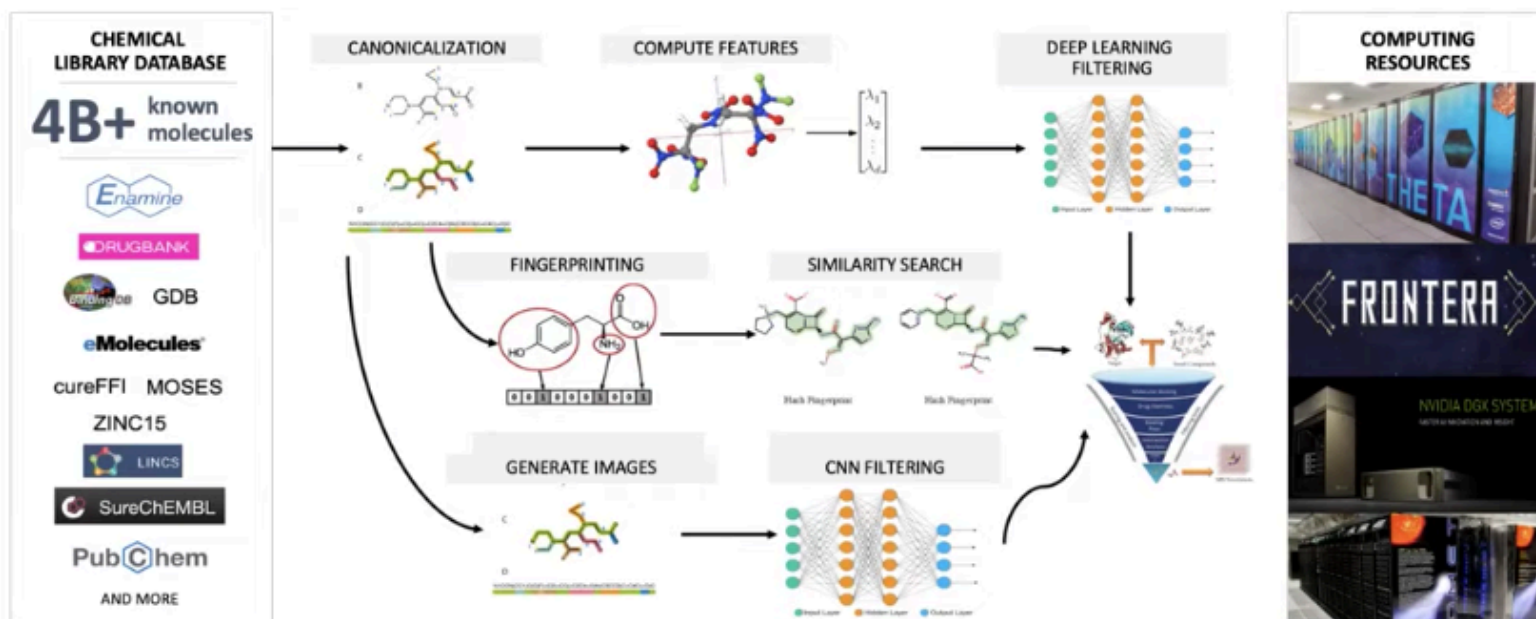
R. Amaro (UCSD), S. Harris (Leeds),
C. Simmerling (Stony Brook), S. Jha
(Rutgers/ Brookhaven)



CANDLE Team has been using CANDLE infrastructure to support large-scale docking and machine learning training and inference

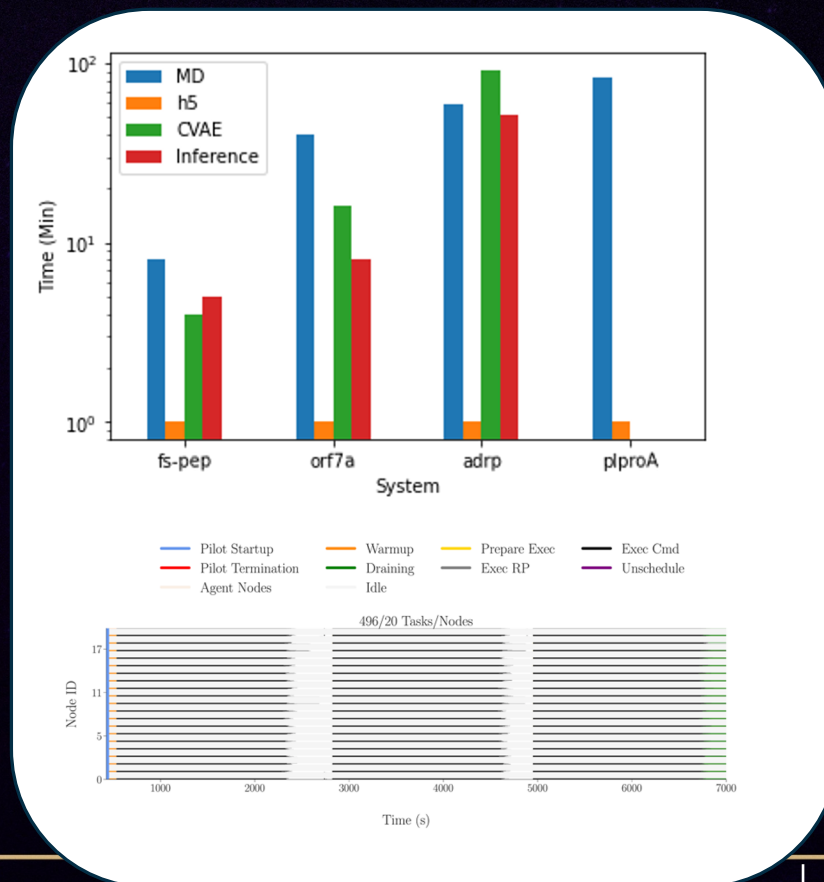
Courtesy Rick Stevens, U. of Chicago/ANL

AI and Supercomputers used to accelerate drug development

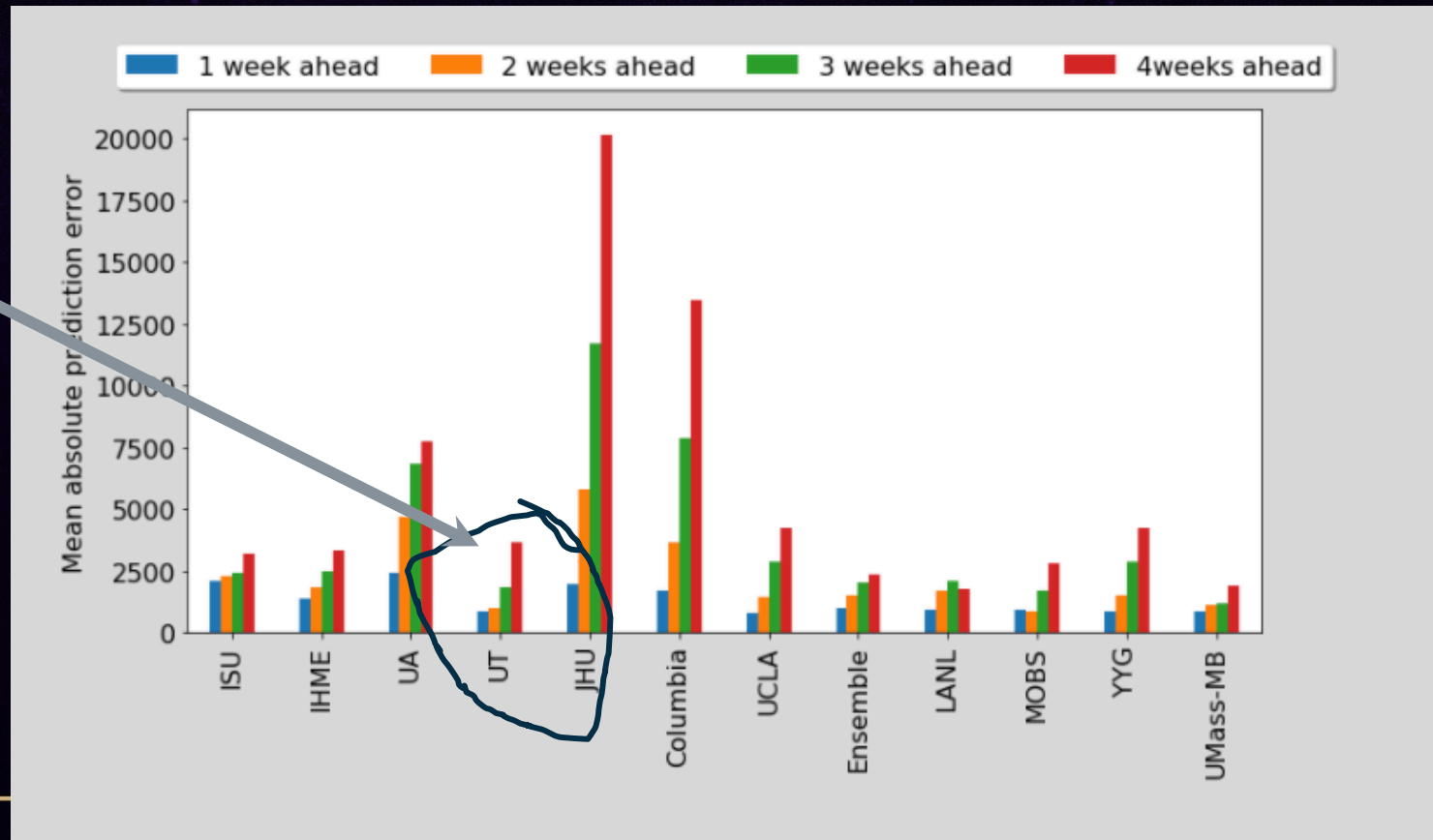


AI-HPC INFRASTRUCTURE PERFORMANCE (COURTESY SHANTENU JHA)

- ▶ **Stage-1:** ~20M SMILES/hour (Frontera)
 - ▶ Docking tasks have 2x fluctuations RCT load balances!
 - ▶ Scales to more 100,000 cores (i.e., 2000 nodes) on Frontera
- ▶ **Stage-2:** DeepDriveMD (Summit)
 - ▶ Summit: 20 nodes, 120 GPUs
 - ▶ Training time becomes increasingly challenging
 - ▶ 96% resource utilization although heterogeneous tasks are swapped
- ▶ **Stages-3 and 4:** ESMACS and TIES protocols use RCT individually but also are hybridized to efficiently use Summit nodes (6 GPUs + 42 CPU cores)
 - ▶ Efficiently executes >16K concurrent tasks on ≈ 2000 nodes (half of the entire Summit machine)



EPIDEMIOLOGY

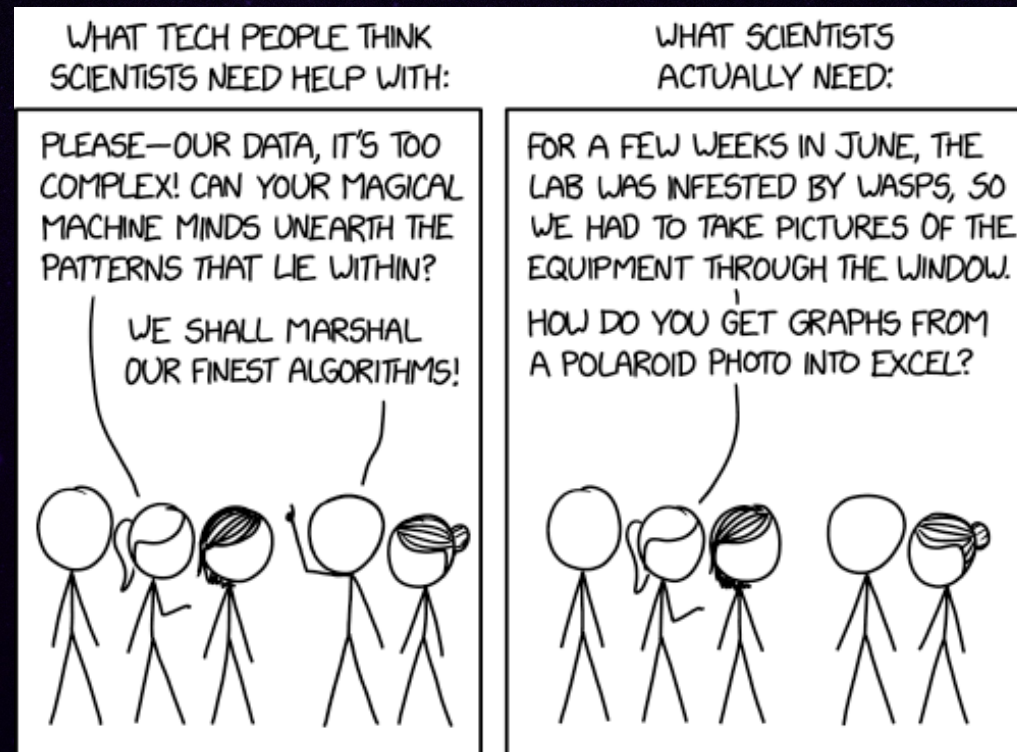


That one is us
(smaller means less error)

ALSO A GREAT EXAMPLE OF OUR EXTENDED SUPPORT

► Although:

Re-writing a lot of bad Matlab, cleaning a lot of data, building web sites, making graphs.



PLEASE KEEP IN MIND, WE NEED TO HEAR ABOUT YOUR SCIENCE ACCOMPLISHMENTS!

- ▶ We (intentionally) set the bar low for any *mandatory* reporting.
- ▶ But we still love to hear about work you do with the system!
- ▶ In addition to all the slides I am about to steal from your talks, please send us slides, pictures, pointers to papers, news about new grants or findings whenever you can!

LEADERSHIP CLASS COMPUTING FACILITY

- ▶ In essence, our plan to transform the NSF and Open Science Computing Landscape from 2025-2035
- ▶ Construction set to begin 2023, Congress permitting.
- ▶ Datacenter expansion to 30MW, new and bigger systems, new approaches to storage and interactive, new buildings, staff expansion, more direct engagement in moving codes to new architectures
 - ▶ Relying on 50% software improvements to meet our targets.



LEADERSHIP-CLASS COMPUTING FACILITY

SEE THE CALL FOR APPLICATION PARTNERS. . .

▶ <https://lccf.tacc.utexas.edu/application-partners/>

THANKS!!



Mostly virtual, but still here!



FRONTERA

TACC



TEXAS

dan@tacc.utexas.edu