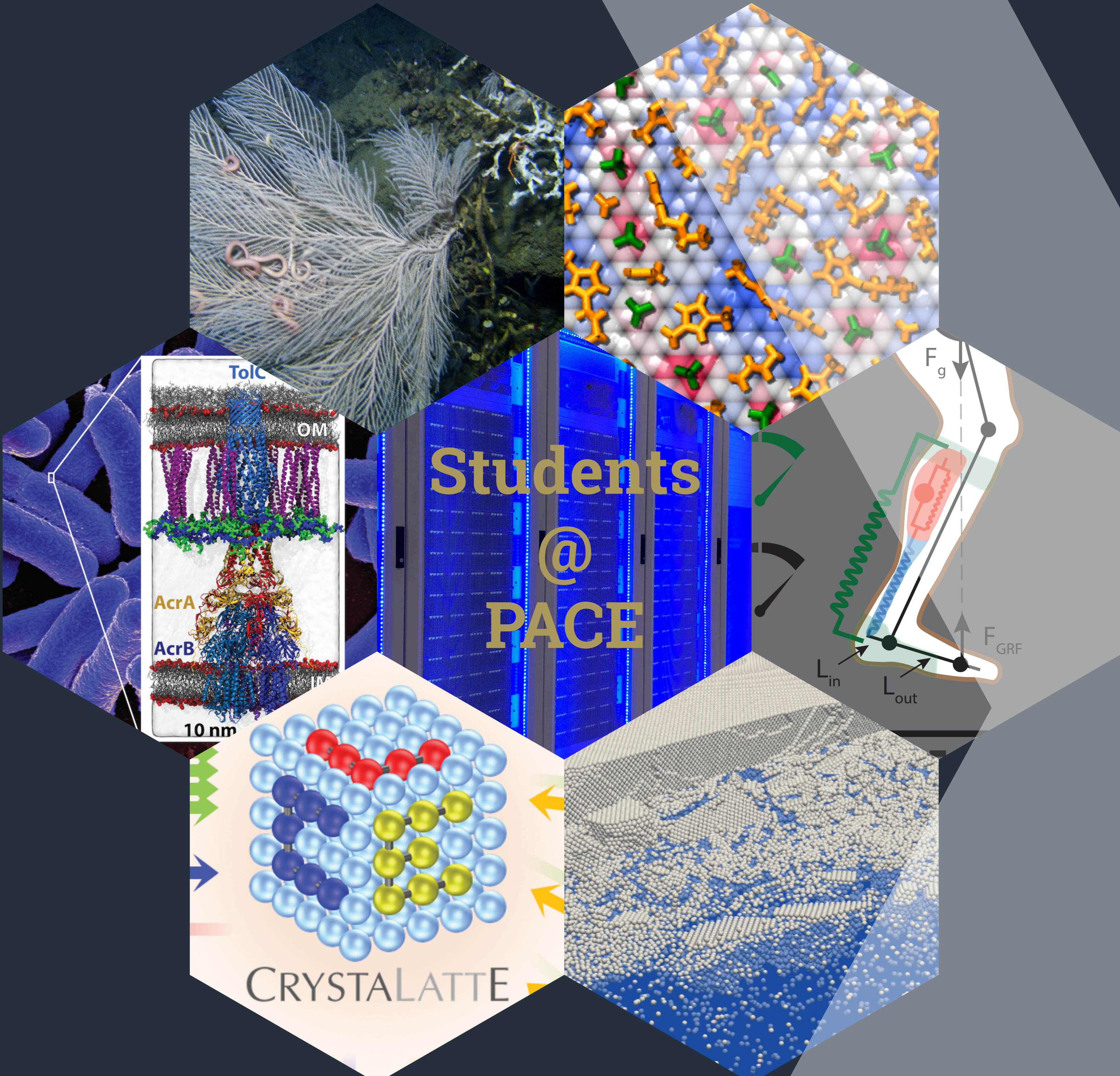


PARTNERSHIP FOR AN ADVANCED COMPUTING ENVIRONMENT (PACE)

WINTER 2020 NEWSLETTER

Georgia Tech
CREATING THE NEXT

FROM THE DIRECTOR'S DESK

Neil Bright
Credit: Malynda Dorsey/OIT

As you are aware from our prior communications and recent issues of our PACE Newsletter, we have been busy. We've deployed the Hive cluster, a state of the art resource funded by NSF; we continue to expand our team to provide an even higher level of service to our community; and we are preparing the Coda data center to receive research workloads migrated from the Rich data center. We will be following up with you on this latest point very soon. This time, I want to inform you about the PACE purchasing schedule for the remainder of FY20 and provide an update on how the recent changes in procurement requirements have impacted our timelines, as I'm sure you have seen in your departments as well.

First, the situation with procurement. The sizable orders we are placing on behalf of the faculty have come under increased scrutiny. This added complexity has resulted in much more time devoted to compliance, and the flexibility that we once enjoyed is no longer achievable. More significantly, each order we place is now requiring a competitive bid process. As a result, our first order of the year, FY20-Phase1, has been considerably delayed and is still in the midst of a bid process. We have started a second order, FY20-Phase2, in parallel to address situations of urgent need and expiring funds. We are making preparations to begin the bid process for this order shortly. Please refer to page 6 for the FY20-Phase 3 milestone schedule.

Please know that we are doing everything possible to advocate for the research community and navigate the best way through the rapid growth at PACE.

Neil Bright
Associate Director for Research Cyberinfrastructure

Editor-in-Chief:
Semir Sarajlic, MS, GISP

Managing Editor:
Michael Weiner, PhD

Contributors:
Srinivas Aluru, PhD
Mehmet Belgin, PhD
Christopher Blanton, PhD
Carlos Borca, PhD
Annalisa Bracco, PhD
Neil Bright, MS
Malynda Dorsey, MPA, MBA
JC Gumbart, PhD
Sebastian Hollister
Aaron Jezghani, PhD
Paul Manno
Jesse McDaniel, PhD
Laksh Kumar Punith
Ziad Rashed
Alexander Robel, PhD
Semir Sarajlic, MS, GISP
Gregory Sawicki, PhD
Jordyn Schroeder
C. David Sherrill, PhD
Harsh Shrivastava
Ken Suda
Sid Vemuri
Michael Weiner, PhD
Nuyun Zhang, PhD

PACE • Coda 10th and 11th floors
756 West Peachtree Street NW • Atlanta, GA 30332-0700

In the past quarter...

178 Georgia Tech PIs used PACE for their research

961 Georgia Tech researchers ran jobs on PACE

430 users employed high-throughput computing (HTC) workflows

50M CPU hours were used on PACE

2.6M jobs ran on PACE clusters

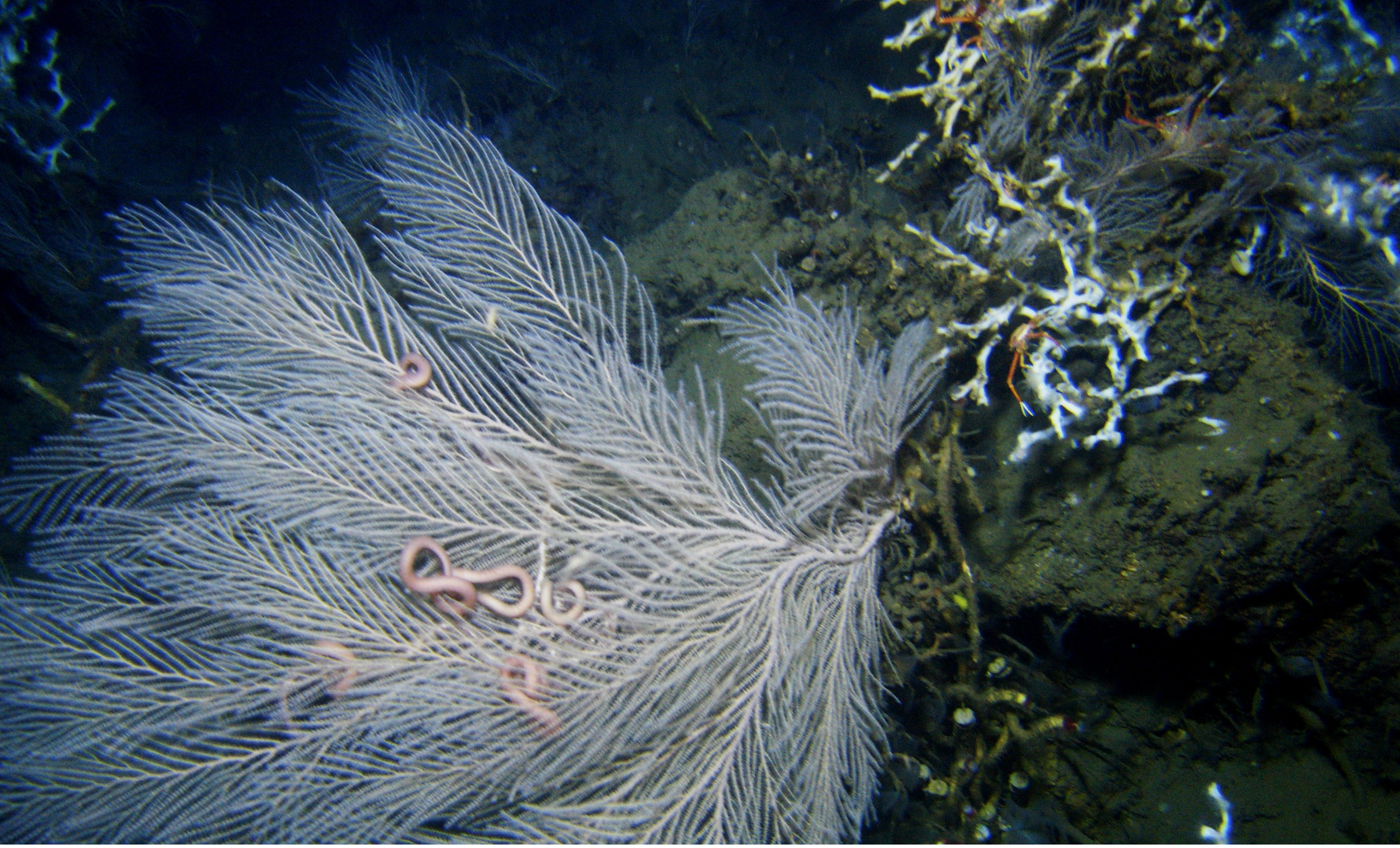
24M CPU hours were used via XSEDE by GT researchers

130k core hours were provided via OSG connect by GT

1.5M core hours were used via OSG by GT researchers

Data for October 1, 2019 - December 31, 2019

IN THIS ISSUE



CORAL MIGRATION IN OCEAN CURRENTS

With continued anthropogenic threats in the marine environment, it is urgent to make decisions that will lead to the effective management and conservation of vulnerable coral ecosystems. This is particularly true in the Gulf of Mexico. In this basin, the establishment of marine protected areas (MPAs) not only is essential to protect and conserve coral ecosystems, but also has been identified as one of the key restoration strategies for benthic communities impacted by the Deepwater Horizon oil spill. The Flower Garden Banks National Marine Sanctuary has proposed to expand the boundaries of its current protected areas to encompass additional coral sites. A group of scientists including marine ecologists, biologists and physical oceanographer **Dr. Annalisa Bracco** are working hard to provide ecosystem connectivity information and tools to resource managers to effectively manage this MPA.

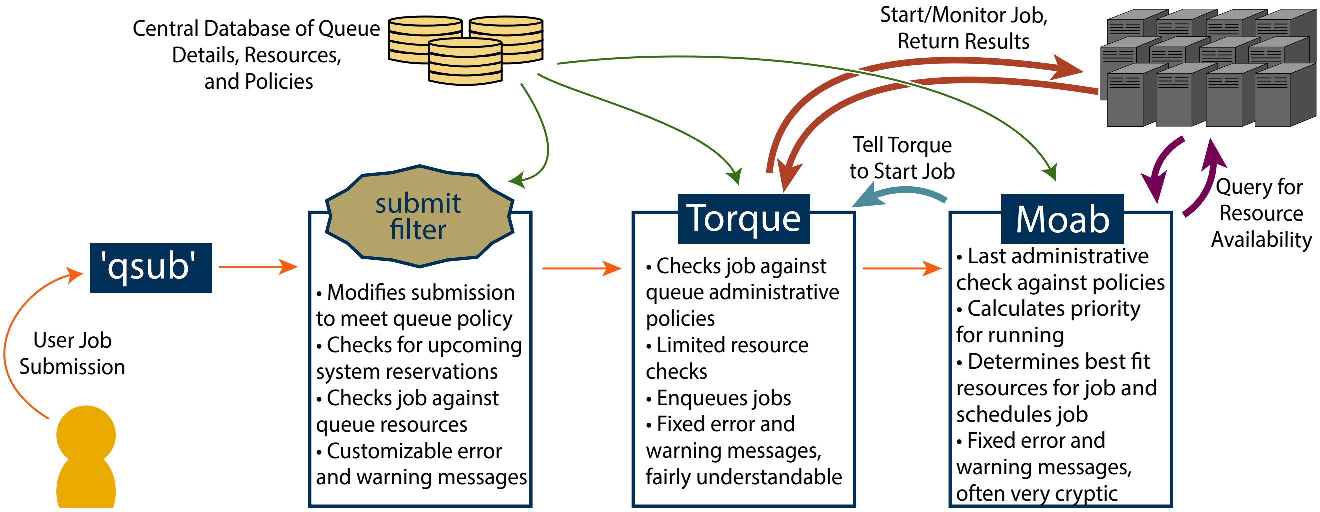
Marine animals that live on the reefs and banks of the MPA have propagules (i.e., eggs, larvae, or juveniles) that disperse to other reefs and banks transported by ocean currents. Because it is difficult to track a single egg or larva, the researchers are modeling the oceanic currents that affect the reefs and banks located on the outer continental shelf in the northwestern Gulf of Mexico using PACE facilities. These reefs are located at depths of a few tens to a few hundreds of meters and little is understood about the various physical processes that influence the current direction and velocity at these depths (such as the role of wind gusts and how tides affect these areas).

Bracco's colleagues collect samples from a number of sea fans, black corals, sponges, and fish and analyze them to determine their genetic and chemical information. This information is then combined with circulation models of larval dispersal that Bracco's lab is developing using PACE supercomputers to better characterize the key habitats for target species and determine how physical processes are shaping how the various populations found on different reefs and banks are related.

For example, the researchers used this integrated approach to investigate the connectivity of a deepwater coral, Callogorgia delta. The circulation in the basin has been simulated by a regional ocean model at 1 km horizontal resolution, which is sufficiently detailed to allow for the generation and evolution of small vortices and vorticity filaments, and for a reliable representation of major bathymetric features. Building upon data collected at four sites spanning about 250 km of distance and 400 m of depth, they were able to conclude that depth differences on scales of tens to at most a few hundreds of meters are sufficient to limit Callogorgia delta connectivity among sites (Bracco et al,. *Journal of Marine Systems,* 2019). These results have important implications in the development of restoration and preservation strategies of deepwater corals in the Gulf of Mexico because they call for carefully accounting for the depth dimension in these efforts. ◇

Above: Callogorgia Delta. Image credit: NOAA Photo Library.

PACE EXTENDS SCHEDULER FUNCTIONALITY TO IMPROVE RESEARCHER EXPERIENCE



Above: With the new Torque submit filter in place, job submissions will be fully checked against queue policy and resource availability by the submit filter, then Torque, and then Moab (orange arrows), and then, after acceptance into the queue, assigned resources by Moab (purple arrows), which tells Torque when and where to start the job (blue arrow), followed finally by Torque running the job and returning the results (red arrows). As with Moab and Torque, the submit filter will draw its configuration from a central repository (green arrows) so that queue details can be efficiently administered. Image credit: Aaron Jezghani/PACE.

The PACE batch scheduler is comprised of a tandem effort between the resource manager, Torque, and the workload manager, Moab. Torque addresses the tasks of job management, including enqueueing, starting, cancelling, and monitoring jobs. Once Torque routes the job to the appropriate queue, Moab enqueues it and determines the optimal resources on which it can run before instructing Torque to start the job. Each of these applications checks that submissions correctly meet administrative policy and resource availability criteria before enqueueing and starting them. Due to the nature of each tool, they perform a slightly different role. Torque, as the user-facing job manager, tends to enforce policies from an administrative perspective and provides specific details in the event of an error. Moab, the behind-the-scenes scheduler, mostly checks for resource availability and simply tells Torque to let the user know the job was cancelled, but not why. As such, each utility is configured to provide meaningful job validation checks and user communication to ensure an efficient batch system.

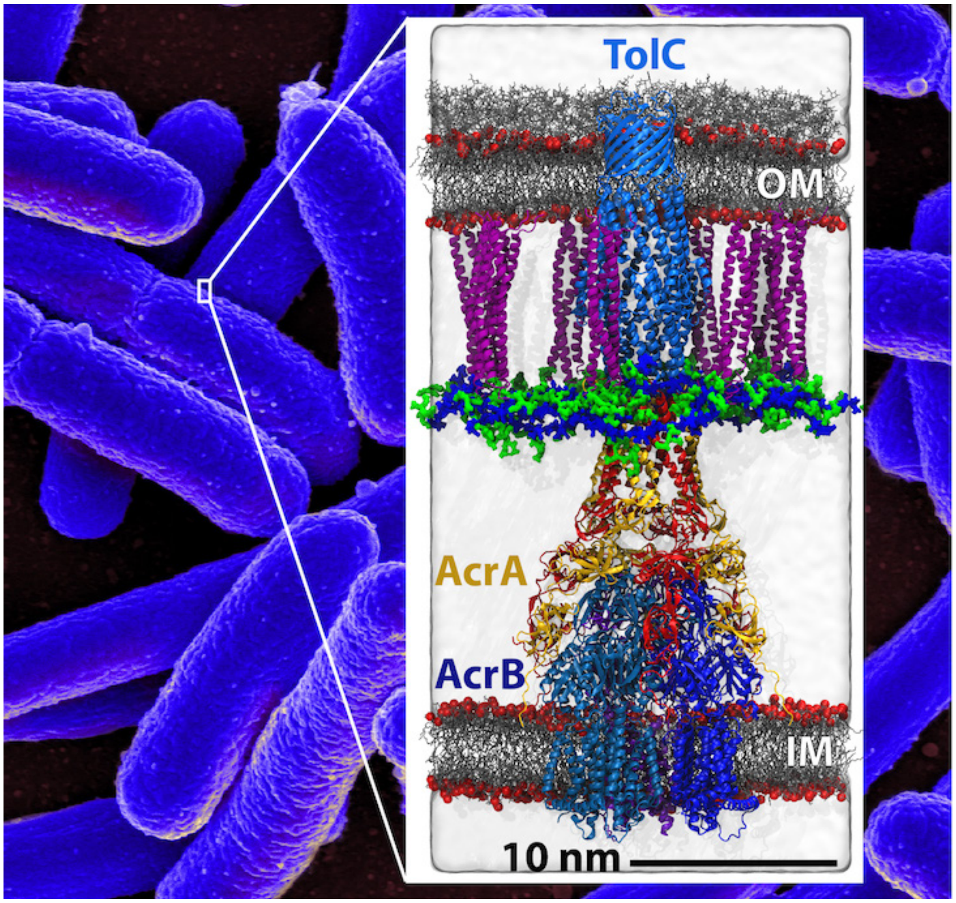
Unfortunately, Torque's lack of resource knowledge and Moab's inability to check against certain resource requests, such as GPUs and node features, leads to successfully queued jobs that are waiting for the scheduler to find valid compute nodes. Furthermore, certain policies such as the hive-gpu queue CPU:GPU ratio (see story on page 8) can only be enforced retroactively with the default tools; i.e., PACE administrators must manually cancel such jobs. And, as many users have experienced, jobs that overlap with the PACE quarterly maintenance will readily be enqueued but held without warning until after the maintenance period has passed. Consequently, users have been left with jobs in the queue indefinitely and idle resources because the scheduler lacked the ability or knowledge to proactively reject or warn the user at submission time. Fortunately, a utility known as a job submit filter can be written to complement the scheduler and provide the missing job checks at submission time.

A job submit filter, also known as a "qsub wrapper", provides additional checks at the time of submission to ensure that the job meets policy and resource requirements. Jobs that should not or will not run for these reasons can be rejected by the filter, preventing user frustration. The script can be written in any language, but there are several design challenges that must be considered for a successful filter. First, the script must be fast, as it examines every job individually as they are submitted. Second, it must preserve the integrity of the original job submission, preserving batch script formatting and command line options. Next, the script must be available from any node where jobs can be submitted, so it must be scalable across a diverse, distributed system. And finally, the submit filter must be robust enough to accommodate the many options for job submission supported by Torque.

Drawing inspiration from a script shared by Lev Gorenstein of Purdue University, PACE Research Scientist **Dr. Aaron Jezghani** recently wrote and deployed such a filter across the PACE clusters. By utilizing cached database queries, the filter can impose additional checks for job submissions and prevent the addition of bad requests to the scheduler queue, while allowing for efficient administration of queue policies and details. Furthermore, when a job is rejected, the messages are designed to be instructive to the user, providing key feedback to help the user

Continued on next page.

FIGHTING RESISTANT BACTERIA VIA SIMULATION



Above: Gram-negative bacteria such as E. coli contain efflux pumps in their cell envelope, spanning both inner and outer membranes. The computational model reveals the fully assembled AcrAB-TolC multidrug efflux pump in its native environment. Image credit: JC Gumbart (Background image from National Institute of Allergy and Infectious Diseases).

The growth of antibiotic resistance is outpacing our ability to develop new drugs to counter it. Resistant bacteria are now responsible for at least 2.8 million infections and 35,000 deaths annually.

One of the innate antibiotic resistance mechanisms of Gram-negative bacteria such as Escherichia coli and Pseudomonas aeruginosa is the expression of so-called multidrug efflux pumps. These protein pumps, which span the approximately 25-nm space between the inner and outer membranes (IM and OM) in Gram-negative bacteria, expel most drugs out of the cell as fast as they can come in. One example is E. coli's AcrAB-TolC, a tripartite complex made up of multiple proteins: AcrB in the IM, TolC in the OM, and AcrA, which fuses the other two together, in the middle. If a way to prevent the assembly of these pumps could be found, resistance would go down dramatically, making "old" antibiotics new again.

The research group of **Dr. JC Gumbart** is currently using computationally-intensive atomistic molecular dynamics (MD) simulations to develop a more precise understanding of how AcrAB-TolC, as well as other efflux pumps, assemble. In collaboration with Dr. Helen Zgurskaya (U. Oklahoma), Dr. Jeremy Smith (Oak Ridge National Lab), and Dr. Jerry Parks (Oak Ridge), they will use these simulations to then design small-molecule drugs that can bind to AcrA and inhibit formation of the efflux pump. PACE resources, including both high performance CPU clusters and GPUs, have been instrumental in reaching the necessary lengths and number of simulations to take some of the first steps in this direction (Darzynkiewicz et al., *Biophys. J*, 2019, and Hazel et al., *ACS Infect. Dis.*, 2019). Ongoing efforts continue on PACE resources, including on the new NSF-funded Hive cluster. ◇

SUBMIT FILTER

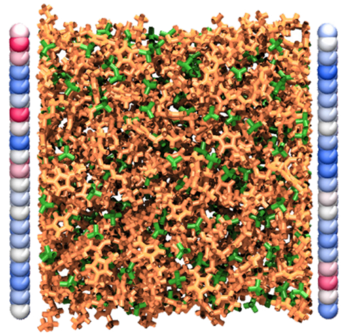
Continued from previous page.

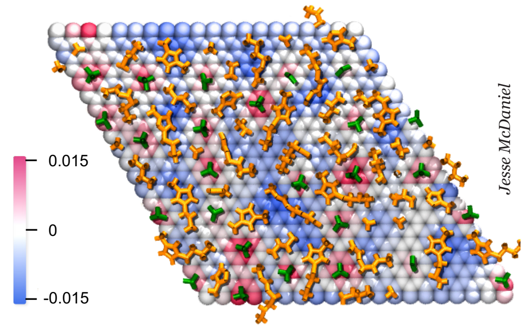
correct their job submission to ensure it runs correctly. "In a batch job system, such as the PACE clusters, there is nothing more frustrating than to submit a job and watch it sit in the queue while the compute nodes remain idle," Jezghani explains. "This filter is meant to improve the user experience and workflow efficiency, as problems can be addressed at submit time, rather than in days or weeks later when results are expected to be in hand."

The new submit filter performs a number of functional checks that Torque and Moab are unable to do, including checking per node memory, CPU, and GPU requests against the available resources within the queue; if a job request cannot be accommodated within the queue, the job will be rejected with an error message indicating the problem. Also included in the submit filter are policy checks to ensure that jobs submitted to shared GPU queues such as force-gpu, testflight-gpu, and hive-gpu request GPUs and that jobs submitted to hive-gpu adhere to a 6:1 CPU to GPU ratio. Finally, the submit filter will check the job submission time against upcoming system reservations and warn users if their job overlaps with the maintenance cycle, so that they know why their job sits in the queue. As Research Facilitation Lead **Dr. Mehmet Belgin** says, "[This is] a functionality we desired for a very long time."

Jezghani is continuing to work on advancing the scheduler functionality to provide the friendliest and most efficient HPC workspace possible for Georgia Tech researchers. Ultimately, he is working towards a scheduler configuration that enables users to do more each and every day, and as Team Lead for Outreach and Faculty Interaction **Semir Sarajlic** says, "at PACE, we aim to accelerate research in engeneering and science for our user community." ◇

SUPERCAPACITORS FOR ENERGY STORAGE





Jesse McDaniel

Designing better energy storage devices such as batteries, supercapacitors, and fuel cells requires fundamental characterization of electrochemical interfaces—i.e., the interface formed between an electrode and electrolyte. Molecular dynamics (MD) is a type of computer simulation serving as a powerful tool for such characterization, as it provides atomic-scale spatial detail and femtosecond (10^{-15} s) temporal resolution. These simulations are very computationally demanding, however, and require new algorithms and software optimized for graphics processing unit (GPU) architectures. The group of **Dr. Jesse McDaniel** utilizes a 32-GPU-card computer cluster on PACE to perform MD simulations of supercapacitor devices to understand fundamental energy storage mechanisms. This cluster, designed with the help of PACE, has a custom 1:1 GPU:CPU architecture, enabling higher computational efficiency for these simulations.

Current work within the McDaniel group seeks to optimize the performance of supercapacitor devices that store energy within high-capacitance, electrical double layers. Top-performing devices often utilize highly nanoporous electrodes in combination with redox-active chemical species that either coat the electrode surface and/or are dissolved within the electrolyte. However, it is often not well-understood how the chemical redox activity is influenced by particular electrode nanostructure within these systems, and this lack of fundamental understanding hinders device improvement. The McDaniel group utilizes MD simulations to directly characterize changes in the physical and chemical properties of electrochemical interfaces with external voltage, replicating processes during charge and discharge cycles of supercapacitors. In addition, the group develops new algorithms and software to enable predictions of redox potentials and provide novel characterization of how redox chemistry is altered within these complex environments. ◇

APPLICATIONS OF MACHINE LEARNING

You have learned some theory, but how you do your first Machine Learning project can be challenging: How to handle the data? How to identify what is essential in the data? How to visualize correlations? How to evaluate your model? What are some common pitfalls and how can you avoid falling into them?

Our new PACE workshop will include an example first project in machine learning. We will begin with a concept of a project, ingest the data, visualize and view potential correlations, select, train our model, and evaluate the model.

This new PACE workshop will debut in February 2020. Visit pace.gatech.edu for more information. Prerequisite: Some familiarity with Python.

PACE PROCUREMENT UPDATE AND SCHEDULE

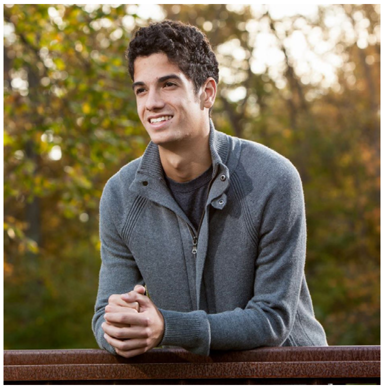
Given the extended time that is required to process orders, we have time for only one more order before the year-end deadlines are upon us. We will accept letters of intent to participate in FY20-Phase3 from now through February 20, 2020. We will need complete specifications, budgets, account numbers, etc. by February 27, 2020. Please see the schedule below for further milestones. This rapidly approaching deadline is necessary for us to have sufficient ability to process this order in time to use FY20 funds. Due to the bidding process, we will have reduced ability for configuration changes once the bid is released to vendors. By extension, we have reduced ability to precisely communicate costs in advance. We will continue to provide budgetary estimates, and final costs will be communicated after bids are awarded.

**Note: Purchases of PACE storage are not affected. Storage can be added as needed via request to pace-support@oit.gatech.edu.

Deadline	Description	Notes
02/20/2020	intent period	Intent to participate in FY20-Phase3 due to pace-support@oit.gatech.edu
02/27/2020	actionable requests	All details due to PACE (configuration, quantity, not-to-exceed budget, account number, financial contact, queue name)
04/22/2020	bid award	Anticipated date to award bid
04/29/2020	final bid quote	Anticipated date to finalize quote with selected vendor
05/06/2020	final quote, faculty approvals	Exact pricing communicated to faculty; all formal approvals received
05/08/2020	enter requisition	Requisition entered into Workday
05/22/2020	release PO	GT-Procurement issues purchase order to vendor
07/31/2020	installation	Estimated completion of physical installation by the vendor
08/28/2020	acceptance testing	PACE will perform acceptance testing to ensure proper operation. Upon completion, resources will be made "ready for research"

WHAT'S IT LIKE TO BE A STUDENT ASSISTANT AT PACE?: A WORD FROM TWO PACE STUDENTS

Georgia Tech's Partnership for an Advanced Computing Environment (PACE) provides incredible learning opportunities for students and a chance for them to do meaningful work that contributes to the Georgia Tech research community. We are **Sid Vemuri** and **Sebastian Hollister**, undergraduate Computer Science Students from the College of Computing, PACE Student Assistants, and XSEDE Student Champions mentored and managed by PACE's **Semir Sarajlic** and **Paul Manno**. There are a variety of different projects that students can work on that are valuable for both the students and the department. The best way to showcase what this means is to talk about our personal experiences as Student Assistants at PACE. We were lucky to receive, as our first assignment, a large and open-ended project: rewrite and redesign the PACE documentation. It was incredibly interesting to be able to come in every day and learn enough about a new piece of software or a new part of this complex system to enable us to write guides that are used to set up important research on PACE Clusters. As we approached the end of building the base set of pages for the documentation, we were able to take on new responsibilities. We went from changing the design of the documentation, to improving its aesthetics and ADA compliance, to building a framework for the internal staff documentation, to writing documentation for the new Hive Cluster, to doing a complete redesign of the PACE Blog. Throughout this process we learned a lot about Linux and Bash, running software on distributed systems, communicating technical ideas, some web development, and even data analytics using PowerBI. It's fulfilling for us to see that the projects we worked on have a real impact and continue to be used by the Georgia Tech community as well as within PACE. The projects that we currently work on at PACE remain challenging and technical, and we enjoy the freedom to creatively think of solutions. The freedom we get with projects is balanced by the support of the PACE team, who are always willing to provide assistance or offer useful insights and are a pleasure to work with. The PACE team has been the most important part of working here, and we can say for certain that this team is a group of people who are genuinely kind, smart, interesting, and above all, passionate about their work. Overall, we're happy to be part of the PACE team, and we can say unequivocally that PACE provides a unique and valuable experience for its students. Plus, we get to say we work with supercomputers! ◇



"PACE has been an amazing team of people to work with. I've picked up so many unique skills, and had unbelievable experiences and opportunities at PACE I wouldn't have had anywhere else."

Sebastian Hollister

Computer Science major, College of Computing

XSEDE Student Champion

PACE IS HIRING STUDENTS!

Join Sebastian and Sid as a PACE Student Assistant! Interested students should visit pace.gatech.edu/careers for more information.

"I've had a great time working with PACE's incredible team and learning new skills in the process. It's really fulfilling to know that I was able to work on projects that are truly valuable to PACE and its user community."

Sid Vemuri

Computer Science major, College of Computing

XSEDE Student Champion

POLICIES ON HIVE TO RAISE PRODUCTIVITY

After deploying Georgia Tech's new $5.3 million Hive cluster in the fall of 2019, PACE was pleased to watch the rapid growth of the Hive user community and of the cluster's utilization. PACE monitored resource usage and collected feedback from initial users throughout the first few months of operation. In consultation with the Hive PIs, **Drs. Srinivas Aluru, Rich Vuduc, Surya Kalidindi, C. David Sherrill,** and **Deirdre Shoemaker,** PACE made several policy adjustments in January 2020 to increase productivity for all users of Hive:

1. *Hive-gpu*: The maximum walltime for jobs on *hive-gpu* was decreased to 3 days from the previous 5 day limit, in order to address the longer job wait times that users have experienced on the *hive-gpu* queue.

2. *Hive-gpu*: To ensure that GPUs do not sit idle, jobs are not permitted to use a CPU:GPU ratio higher than 6:1 (i.e., 6 cores per GPU). Each *hive-gpu* node contains 24 CPUs and 4 GPUs.

3. *Hive-nvme-sas*: A new queue, *hive-nvme-sas*, that combines the compute nodes of the *hive-nvme* and *hive-sas* queues and shares the nodes with those queues, was created.

4. *Hive-nvme-sas, hive-nvme, hive-sas*: The maximum walltime for jobs on the *hive-nvme*, *hive-sas*, and *hive-nvme-sas* queues was increased to 30 days from the previous 5 day limit.

5. *Hive-interact*: A new interactive queue, *hive-interact*, was created. This queue provides access to 32 Hive compute nodes (192 GB RAM and 24 cores). It is provided for quick access to resources for testing and development. The walltime limit is 1 hour.

6. *Hive-priority*: A new *hive-priority* queue was created. This queue is reserved for researchers with time-sensitive research deadlines. For access to this queue, please communicate the appropriate dates/upcoming deadlines to the PACE team in order to get the necessary approvals for PACE to provide you access to the high-priority queue. Please note that PACE may not be able to provide access to priority queue for requests made less than 14 days in advance of the time when the resource is needed, due to running jobs at the time of the request. ◇

PACE IS HIRING!

» XSEDE/OSG Architect (Research Technologist)
» Scheduler Architect (Research Technologist)
» Student Assistants (see page 7)

Visit pace.gatech.edu/careers for more information.

XSEDE SAVES THE DAY DURING MAINTENANCE

At PACE, we remember the long hours of work from our days in graduate school, and recognize the challenges our student users face. When one of **Dr. Srinivas Aluru**'s students, PhD Candidate **Harsh Shrivastava**, reached out to us with an urgent need to use PACE resources due to a rapidly approaching paper deadline, we could not tell him that we could not help due to our scheduled maintenance being in progress. Instead, PACE Research Scientist **Semir Sarajlic** quickly leveraged his XSEDE Campus Champion allocation to provide Harsh access to GPU resources from Pittsburgh Supercomputing Center's Bridges cluster and guided Harsh on transitioning his workflow to the SLURM scheduler. Within hours, Harsh had access to the Bridges GPU resources, and he was able to complete the experiments required for his paper submission. Now, with the paper having been accepted in the International Conference on Learning Representations, we can celebrate and take a quick moment to reflect on the value that XSEDE and the Campus Champion program provide us!

At the conference, Harsh's colleagues, fellow students and many more will be able to hear about his research. Harsh shared about his paper's focus: "Recovering sparse conditional independence graphs from data is a fundamental problem in machine learning with wide applications. A popular formulation of the problem is an l1 regularized maximum likelihood estimation. Many convex optimization algorithms have been designed to solve this formulation to recover the graph structure. Recently, there has been a surge of interest to learn algorithms directly based on data, and in this case, learn to map empirical covariance to the sparse precision matrix. However, it is a challenging task in this case, since the symmetric positive definiteness (SPD) and sparsity of the matrix are not easy to enforce in learned algorithms, and a direct mapping from data to precision matrix may contain many parameters. We proposed a deep learning architecture, GLAD, which uses an Alternating Minimization (AM) algorithm as a model inductive bias and learns the model parameters via supervised learning. In a poster at the ICLR 2020 Conference, we will show that GLAD learns a very compact and effective model for recovering sparse graphs from data. We ran our experiments partially on the PACE clusters, and with the support from PACE in the 11th hour, we ran part of our experiments on PSC's Bridges via XSEDE that helped us get our paper submitted prior to the deadline!"

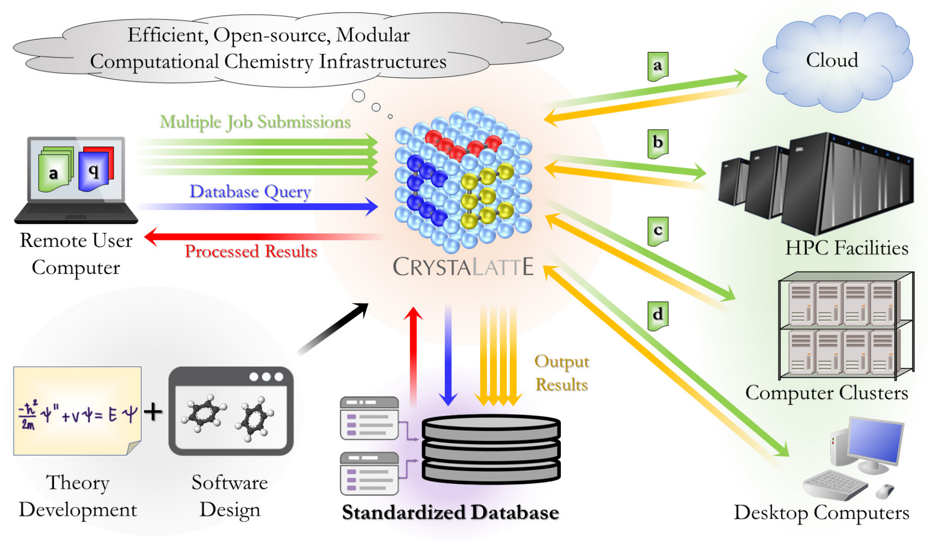
PACE GROWS

PACE is pleased to welcome **Dr. Christopher Stone** to the PACE team. Chris, a Georgia Tech graduate, joined the Software and Collaboration Support team as a Research Scientist in January 2020.

PACE, SCHELLER PARTNER FOR ML

PACE is partnering with the Quantitative & Computational Finance Program of the Scheller College of Business to offer a series of hands-on workshops for machine learning and big data, taught by PACE's **Dr. Nuyun Zhang**. The workshops, specifically designed for the students' research and job interview needs, include 18 hours of instruction covering programming skills with Scikit-learn, TensorFlow 2.1, PySpark and MLlib. Students utilize PACE's instructional cluster (ICE) with GPUs and Google Cloud with TPUs. In this way, students can learn practical skills and a pipeline to solve real-world problems with some of the most recent algorithms, such as CNN, LSTM, RNN and attention.

SOFTWARE SPEEDILY FINDS LATTICE ENERGIES



Above: CrystaLattE automates the calculation of lattice energies of molecular crystals by employing fragmentation approaches that have been developed to enable modular, scalable software to drive large numbers of parallel computational chemistry jobs. Image credit: Carlos Borca/Sherrill Group.

Recent advances in efficient software to sample lattice structures has drastically improved the prediction of probable crystal polymorphs. However, there are many cases where the lattice energies of several polymorphs differ by a few kJ mol^{-1}, and the prediction of lattice energies at that level of accuracy is beyond that of all but the most expensive computational chemistry methods. However, such predictions are important for determining the stability of crystalline forms of pharmaceuticals, among other applications. Therefore, **Drs. Carlos H. Borca** and **C. David Sherrill** are working to develop software that couples highly-accurate wave-function methods with distributed computation techniques, which could have a significant impact in overcoming this challenge.

In a recent publication (*The Journal of Chemical Physics*, 2019), they presented CrystaLattE, an algorithm to compute lattice energies of molecular crystals. CrystaLattE exploits the many-body cluster expansion, a widely-applied fragmentation approach for intermolecular interactions based on the decomposition of the full interaction of an N-particle aggregate as an expansion of two-, three-, four-, ..., N-body interactions. In this case, a single molecule constitutes a body, and the required computations on two-bodies, three-bodies, etc., within the crystal are independent of each other, leading to a naturally parallel approach. The algorithm also exploits the long-range periodic order of crystals to automatically detect and avoid unnecessary computations, which are performed with the computational chemistry package Psi4.

In a first application, using the research cyberinfrastructure resources and services provided by PACE, the lattice energy of crystalline benzene was computed using the fast, dispersion-corrected Hartree–Fock method, HF-3c, in CrystaLattE. This application involved the distributed execution of 7203 multithreaded Psi4 jobs, each one running independently on a full node. Those jobs corresponded to each of the 73 symmetry-unique two-body and 7130 symmetry-unique three-body interactions that can be formed from molecules within a 15 Å cutoff from a central reference molecule in the benzene crystal. Once all outputs were gathered, CrystaLattE reconstructed the lattice energy from the contributions of each interaction. HF-3c plus an Axilrod-Teller-Muto estimate of three-body dispersion exhibits an error of only −1.0 kJ mol^{-1} compared to the estimated 0 K experimental lattice energy of −55.3 ± 2.2 kJ mol^{-1}.

In another state-of-the art application using high-accuracy methods and multiscale techniques, the lattice energy of the triazine crystal was computed at benchmark-level accuracy in under 5 hours by distributing 41 two-body and 3618 three-body interactions as 4-thread jobs executed nearly-simultaneously on the shared queue at PACE. Unlike these examples, competing methodologies such as periodic density-functional theory, which are often less accurate, do not allow for independent distribution of calculations. In turn, they demand more execution time when running on similar resources. Thus, the CrystaLattE approach promises to be more computationally efficient. Current work on CrystaLattE involves tests on many different types of molecular crystals and development of screening rules to eliminate computations on unimportant configurations with long contact distances. ◇

SPRING 2020 PACE WORKSHOPS

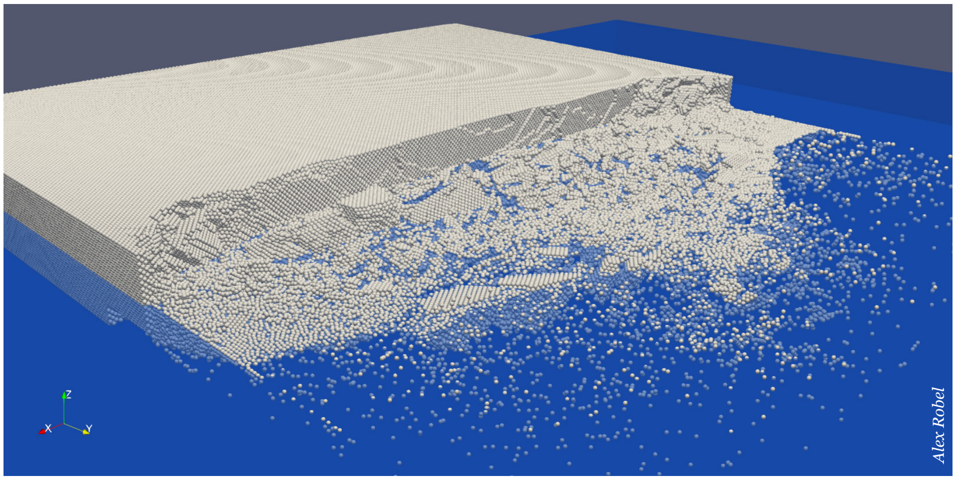
PACE continues to offer regular hands-on training workshops during the Spring 2020 semester. Each workshop is developed and taught by PACE Research Scientists and is designed to help PACE users make fuller use of PACE resources.

- Linux 101 by **Dr. Aaron Jezghani**: January 7, February 5, March 3, April 1
- Linux 102 by **Dr. Aaron Jezghani**: February 19, March 17, April 15
- Python 101: Intro to Data Analysis with NumPy by **Dr. Michael Weiner**: January 8, February 4, March 4, March 31
- Optimization 101 by by **Dr. Chris Blanton**: January 21, March 18
- Applications of Machine Learning: Your First ML Project by by **Dr. Chris Blanton**: February 18, April 14
- PACE Orientation by **Semir Sarajlic**: alternate Wednesdays, beginning January 15

Please note that registration will be closing 2 days prior to each class, to ensure that we may provision all the necessary training accounts and materials needed for the hands-on portion of the class. For more information, visit pace.gatech.edu.

We also continue to offer PACE Consulting Sessions on alternate Tuesdays, beginning January 14. Consulting sessions are a great opportunity to drop in and get face-to-face help from the PACE Resarch Facilitation team with any problems impeding your research on PACE. No registration is needed for PACE Consulting Sessions.

HOW DO FJORDS PREVENT GLACIER CALVING?



Glaciers in contact with the ocean are prone to fracture, leading to detachment of icebergs, which fall into the ocean in a process known as calving. This calving process exerts a strong control on the amount of ice lost from ice sheets and contributes to sea level rise. However, the multi-scale complexity of the calving process makes it a challenge to understand and an even greater challenge to simulate on time scales relevant to projections of future sea level rise. Research in the Georgia Tech Ice & Climate Group, led by **Dr. Alex Robel**, is seeking to bridge these scales of the iceberg calving process using new state-of-the-art modeling techniques running on PACE.

Some glaciers end in long and narrow sunken valleys, known as fjords, which may fill up with a mélange of fragmented ice particles, sea ice, and snow, ranging in size from ones to hundreds of meters. Strongly seasonal calving behavior at glaciers in Greenland indicate that mélange within fjords can potentially shut down calving activity by applying an opposing pressure along the glacier front, acting to reduce tensile stresses responsible for creating faults within glacier ice. However, the current understanding of the dynamics controlling the calving-mélange feedback is poor due to the difficulty of obtaining environmental observations and the inherent complexity of ice fracture. Processes governing ice fracture occur at spatial scales of meters and temporal scales of seconds to minutes which make explicit calving simulations especially difficult when applied to glaciers which can range in scale from ones to hundreds of kilometers.

To understand the calving process and the feedbacks at play when mélange is present, the researchers use the Helsinki Discrete Element Model (HiDEM), an open-source discrete element model designed specifically to study large-scale ice fracture. Unlike continuum models, which represent important calving variables such as velocity, pressure and density as continuum fields, discrete element models represent ice as a lattice of interconnected particles which can break apart and interact discontinuously. Ice particles are connected by elastic beams, which can deform and break. Simulating realistic glacier geometries require millions of particles, resulting in extreme computational costs for relatively short time frames (seconds to minutes). Computing clusters accessible through PACE allow Georgia Tech researchers to rapidly deploy parallelized HiDEM jobs and prototype new versions of HiDEM with more advanced physics. ◇

STUDENTS GAIN HPC EXPERIENCE ON ICE

The Instructional Cluster Environments (ICE) at PACE are a powerful teaching tool that give students the advantage of working directly in a high-performance computing environment for educational experiences. PACE welcomes requests to use these resources for classes from all schools, at no cost to faculty and students.

Since the deployment of instructional clusters (COC-ICE and PACE-ICE) in 2018, PACE has supported educational learning for nearly 9,000 students. Some of the classes taught included Digital Image Processing, Advanced Programming Techniques, and Engineering Software Design.

For the Spring 2020 semester, PACE is supporting the computational resource needs of 1,359 students. Multiple sections of the Engineering Software Design course, taught by **Dr. Jeffery Hurley** and **Dr. Jared Ivey**, are training future engineers.

There are multiple Vertically Integrated Project courses hosted on ICE:

» **Dr. A.J. Medford**'s Big Data and Quantum Mechanics explores applications of machine learning to quantum mechanical calculations, particularly using Density Functional Theory (DFT).

» The Team Phoenix - Cluster Competition Team allows students to devise competitive strategies for the Supercomputing 2020 Student Cluster Competition, advised by **Dr. Richard Vuduc, Dr. Jeffrey Young, Will Powell,** and PACE's **Dr. Aaron Jezghani**.

» Automated Algorithm Design, guided by **Dr. Greg Rohling** and **Dr. Jason Zutty**, trains the team to use automated methods to generate efficient algorithms that teach machines to solve problems faster and with better solutions.

Spring 2020 is shaping up to be a great opportunity for the Instructional Cluster Environments at PACE. Please see the PACE website for information about gaining access to the ICE clusters for future terms.

To learn more, visit https://pace.gatech.edu/pace-ice-instructional-cluster-environment-education. ◇

WEARABLE ROBOTS OFFER MOBILITY SOLUTIONS

Mobility is a key to happiness and longevity. Indeed, the goal of the Physiology of Wearable Robotics (PoWeR) Lab, led by **Dr. Gregory S. Sawicki**, is to enable people to do the things they enjoy for as long as possible. Unfortunately, millions around the world have limited mobility due to deteriorating musculoskeletal health that comes with age and conditions like stroke, Parkinson's disease and cerebral palsy. To help overcome these limitations, the PoWeR lab develops exoskeletons. These are wearable robots that can be worn by people as an accessory to normal clothing to provide external support to muscles, tendons, ligaments and bones. Effective exoskeleton technology could augment musculoskeletal function, negate mobility deficits and expand human capabilities beyond current limits.

Understanding and optimizing human-robot interaction is a challenging problem. Unlike other fields of robotics where the engineer has total control over what the robot does, wearable robots have to work in tandem with the biological limbs—often with unpredictable results. Thus, it is critical to understand the fundamental principles that govern how the underlying physiology of the human user responds to external assistance, in order to develop symbiotic wearable robotic technology.

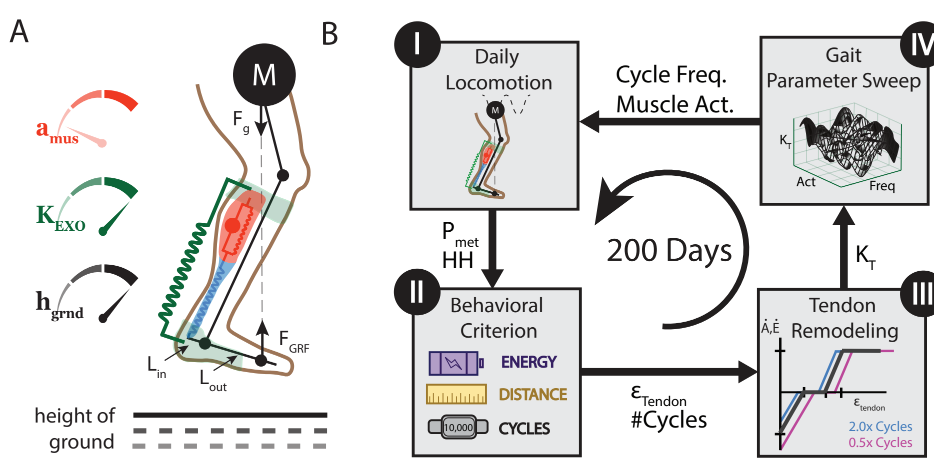
Exploring the physiological interaction between biological and engineered systems in humans is limited to an overhead view without using invasive procedures. To look 'under the skin', the PoWeR Lab uses computer simulations as a tool to develop a more complete theoretical understanding of the underlying musculotendon neuromechanics that drives the response to external assistance during human movement. Insights from modeling and simulation studies provide the critical insight that drives targeted human and animal experiments to ultimately broaden our knowledge of physiological interactions with wearable robotics.

Laksh Kumar Punith, a PhD student in the PoWeR Lab, is interested in whether wearable robots can help people navigate in natural environments with variable terrain. To study this, Laksh developed a model of human locomotion with an exoskeleton moving over ground with uneven surface height. With the help of PACE, Laksh ran numerous simulations over the large parameter space of human neural control, wearable robot stiffness, and ground height variability to develop deep mechanistic insight on the question. He discovered that wearable robots can enable fast recovery from unexpected changes in the height of the ground, while also protecting the underlying biological tissues from large stretches.

Jordyn Schroeder, another PhD student in the PoWeR Lab, is interested in how wearable robots alter musculotendon structural properties over weeks or months of use. She hypothesized that with exoskeletons unloading the limb, Achilles' tendon properties would change, and people would alter their locomotion patterns to maximize their efficiency. To study this, Jordyn developed a model of human locomotion where Achilles' tendon properties can change as a function of load on the tissue and number of steps taken over time. PACE enabled Jordyn to run large parameter sweeps across many locomotion patterns to find the most efficient movements for a given tendon stiffness – a key piece of the puzzle that allowed her to conduct her study. Using these computer simulations of human movement with exoskeletons over long time scales, Jordyn discovered that exoskeleton use can cause tendon atrophy by unloading the biological limbs, but this atrophy can be mitigated if people choose to move more due to the reduced effort of movement.

Computer simulations of human movement with exoskeletons are computationally expensive, and without PACE the PoWeR Lab would not be able to answer hard and innovative questions that cannot be addressed using a standard desktop or laptop computer. ◊

Below: Two examples of modeling and simulation studies conducted by the PoWeR Lab using PACE clusters: (A) This study examines how exoskeleton properties affect stability of human locomotion on uneven terrain. (B) This study examines how long-term exoskeleton use may change mechanical properties of the tendon due to chronic unloading of the tissue. Both studies used neuromuscular models of locomotion simulated in MATLAB/Simulink. Image credit: Sawicki Group.



NEW LINUX 102 TRAINING

PACE is pleased to offer a new course for advanced command line utilization so that users can more efficiently address computational research in a Linux environment. After taking this class, users will be able to demonstrate their ability to utilize regular expressions for stream manipulation with tools like sed and awk, compile code from source with version control and build utilities such as git and make, implement advanced I/O manipulation for in-situ processing, and integrate advanced compression with system tools for better data management. The new Linux 102 course will debut in February 2020. Visit pace.gatech.edu for more information.

CREATING THE NEXT®

pace.gatech.edu