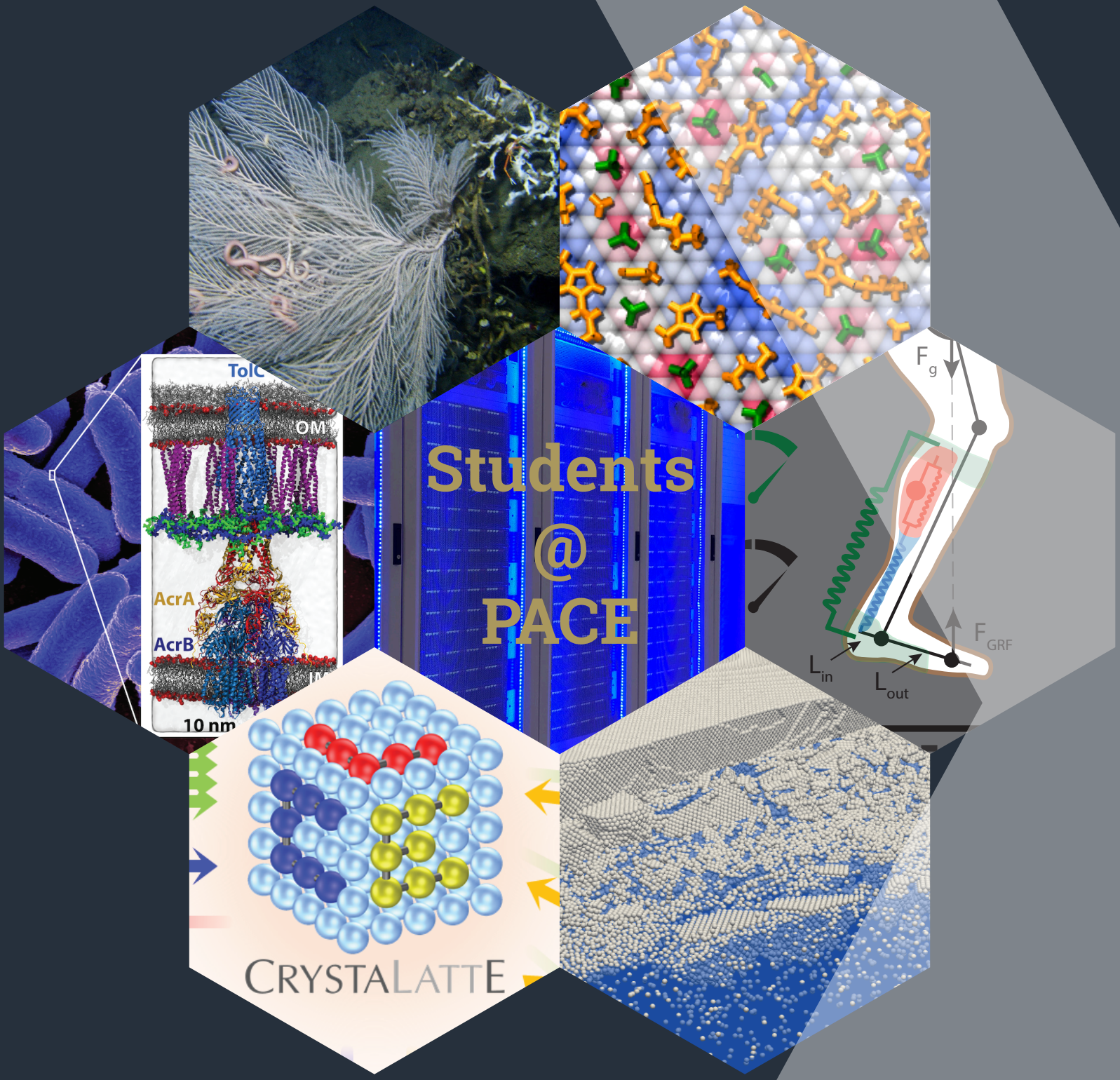


PARTNERSHIP FOR AN ADVANCED COMPUTING ENVIRONMENT (PACE) **WINTER 2020 NEWSLETTER**

Georgia Tech

CREATING THE NEXT

PARTNERSHIP FOR AN ADVANCED COMPUTING ENVIRONMENT (PACE) WINTER 2020 NEWSLETTER

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

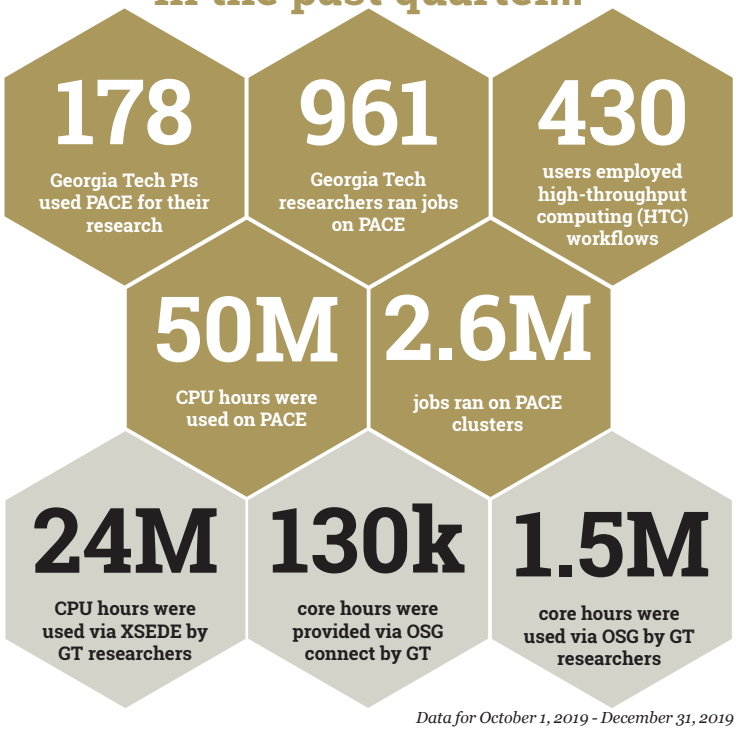
Editor-in-Chief:
Semir Sarajlic, MS, GISP

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...



Data for October 1, 2019 - December 31, 2019

IN THIS ISSUE

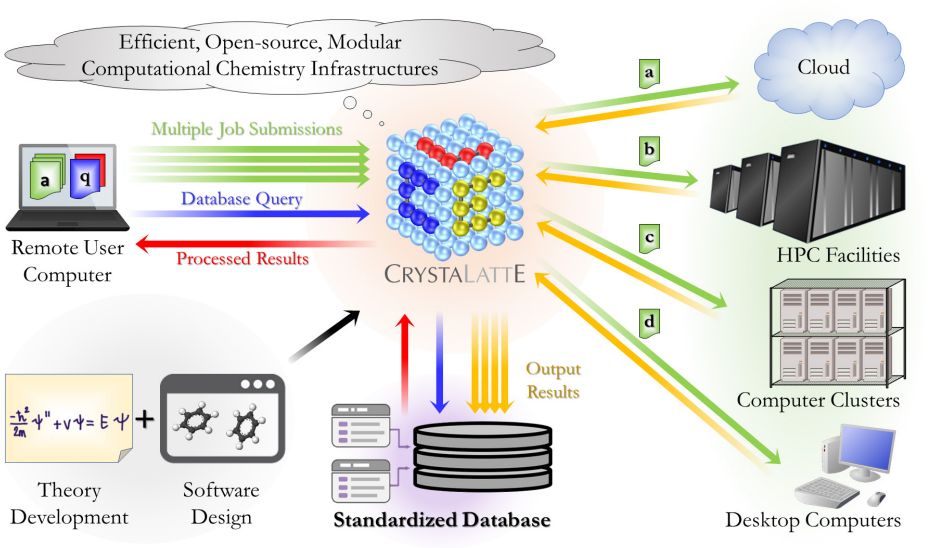
SOFTWARE SPEEDILY FINDS LATTICE ENERGIES

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 multi-scale 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. ◇



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.

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.

CREATING THE NEXT®

Georgia Tech

pace.gatech.edu