Annual Report FY2023

PURDUE
UNIVERSITY

Rosen Center for Advanced Computing

HIGH PERFORMANCE COMPUTING
AT THE HIGHEST PROVEN VALUE

Director's Welcome

As we complete the 2023 fiscal year, we can all look back and see the pace of change in everything around us – new leadership around the university, growth in our world-class facilities, the creation of Purdue Indianapolis, exciting new research initiatives, and the reimagining of IT across the Purdue system, creating the new Purdue IT organization, bringing what was formerly "Central" IT, regional campuses, and academic units all into one single organization.

I hope that you find the information in this report informative as to all of the truly transformative things happening at Purdue and across the nation with the help of our campus cyberinfrastructure. Purdue's investment in research computing continues to provide value to our competitiveness in securing new and executing on research awards, conducting ground-breaking research, and training future scientists.

Some highlights of this impact:

- RCAC-using faculty partners received 60% of Purdue's new research awards in FY23, for a total of $395M.
- RCAC-using faculty accounted for 64% of sponsor expenditures in FY23, a total of $362.7M.
- In FY 2023, Purdue faculty submitted 335 proposals for $340.9M that specified the need for research computing facilities in the pre-award worksheet.
- 234 of AY 2021-22's 851 earned Doctorates (27%)have been an HPC user during their career.
- During the 2022-23 Academic Year (Fall 2022 – Spring 2023 semesters), 167 courses used Scholar intheir curriculum, impacting a total of over 3000 students
- And Research Scientists from RCAC were awarded $3.49M of their own new sponsored research, and $3.7M of expenditures in FY 2023.

Other facts and figures about RCAC's usage and reach can be found later in this report.

More specifically, during FY 2023, Purdue IT continues to support RCAC's mission to enable research at Purdue, as we've delivered another top-notch community cluster supercomputer "Negishi", and invest in GPU computing resources to support the ever-growing use of AI in all sorts of domains.

The NSF-funded "Anvil" system has provided tremendous impact to research across the country, supporting over 500 projects among 231 institutions. Several projects using Anvil, and its workforce development activities, are highlighted in the report.

You will see how RCAC scientists and research software engineers (RSE) partner with Purdue faculty to provide expertise, develop training programs, develop research software, and provide solutions in AR and VR.

Also this year, RCAC outreach programs have continued to grow, hosting events, trainings, and symposia on emerging topics and techniques, building connections between Purdue researchers.

I'm always glad to have an opportunity to highlight our student training programs, where we expose students both to the application of advanced computing for their sciences, but also provide experiential learning to ones who may be interested in careers in high performance computing or research software engineering.

And as always, we remain grateful for your continued support and partnership – we strive to be the computing center of choice and a trusted partner at Purdue, providing high performance computing at the highest proven value.

Best Wishes,

Preston M Smith

Preston M Smith, PhD
Executive Director, Rosen Center for Advanced Computing
Purdue IT

MISSION & VISION

Mission

Purdue IT provides the technology infrastructure, services, solutions, and information security that support teaching and learning, enhance research, and enable faculty and staff to achieve their objectives while providing a positive student experience.

Vision

To empower giant leaps across Purdue by providing safe, efficient and reliable services in our pursuit to become the benchmark for IT in higher education.

TABLE OF CONTENTS

Research Highlights

Education and Outreach

Staff News

External Funding

NEGISHI SUPERCOMPUTER

now available for faculty use

The Purdue Rosen Center for Advanced Computing (RCAC)'s latest research supercomputer, Negishi, is now available for faculty use.

Negishi, which is optimized for traditional, tightly-coupled science and engineering applications, was built through a partnership with Dell and AMD in 2022. Negishi consists of 460 Dell PowerEdge compute nodes with two 64-core 3rd Generation AMD EPYC "Milan" processors (128 cores per node) and 256 GB of memory, six large memory nodes with 1 TB of memory, and 15 AMD Instinct M1210 GPUs.

Negishi is named in honor of the late Ei-ichi Negishi, the Herbert C. Brown distinguished professor of chemistry and the winner of the 2010 Nobel Prize in chemistry. The Negishi supercomputer was dedicated in a February 2023 ceremony featuring campus leaders including Purdue President Mung Chiang and members of the Negishi family.

President Chiang spoke about how Purdue faculty will benefit from the Negishi community cluster, which is the largest and most powerful cluster Purdue has built other than the NSF-funded Anvil system. Provost Wolfe added that just as Ei-ichi Negishi helped set a new standard for Purdue, his namesake cluster will help set Purdue apart from other universities and attract top talent.

The ceremony was followed by a reception and the unveiling of the Ei-ichi Negishi Archives in the Wetherill Laboratory of Chemistry. The Archives contain Negishi's personal memorabilia, including family photographs and awards collected throughout his esteemed career.

There are 68 research groups who have already been onboarded to Negishi cluster through an Early User Program (EUP) and have received individualized assistance from RCAC consultants with corresponding domain knowledge. An accelerated weekly maintenance cycle during the EUP greatly helped in deploying new features quickly and reliably.

Tonglei Li, the Allen Chao Chair and Professor of Industrial and Physical Pharmacy, has used Negishi to develop a novel deep-learning concept, Manifold Embedding of Molecular Surface (MEMS), to predict molecular properties and functions and to design molecular and supermolecular structures by machine intelligence.

"The state-of-the-art capabilities of Negishi have allowed me to calculate the quantum information of tens of thousands of molecules within a short period of two months and predict drug developability including solubility, drug-induced liver toxicity, and liver enzymatic activities against drug molecules," says Li.

Researchers at Purdue can now purchase capacity in Negishi through the RCAC cluster orders website and it has a six-year lifespan. Negishi delivers the lowest cost per FLOP of any community cluster system to date.

ENVISION CENTER COLLABORATION AIMS TO MAKE CONSTRUCTION WORK SAFER

New technology developed by the Rosen Center for Advanced Computing's Envision Center in collaboration with a Purdue professor aims to make construction environments safer by helping workers better assess risk and avoid judgment errors.

"IT WAS A WONDERFUL EXPERIENCE WORKING WITH THE ENVISION CENTER"

Sogand Hasanzadeh, assistant professor of construction engineering and management and civil engineering, studies how construction workers change their risk-taking behavior under different task stressor such as time pressure, and environmental stressors such as heat stress.

“We can't ask people to go on the roof of a two-story building without protection,” explains Hasanzadeh. “We were looking for a visualization that would allow us to provide the same sense of presence of being on the roof and allow them to safely install shingles.”

Enter the Envision Center, which worked with Hasanzadeh's team to use extended reality (using Varjo XR) to develop a partially virtual environment that allows the user to stand on a sloped piece of the actual roof, which is located on the ground and believe they're working two stories up. The worker can selectively see their own body and the shingles on the roof they're installing, but not the rest of the room, where environmental modalities like wind, heat and humidity of a hot summer day, and noise from the suburban area were simulated.

Hasanzadeh, her doctoral student, Shiva Pooladvand, and undergraduate research assistants then use devices such as eye trackers embedded in the Varjo XR headset, medical wristbands collecting heart rate and emotional responses, and fNIRS neuroimaging caps for monitoring the brain's oygen concentration to examine how workers are psychophysiologically reacting to the environment.

The project's ultimate goal is to save thousands of lives and billions of dollars of costs by avoiding falls, near misses, and other resulting incidents/injuries in the construction industries.

Hasanzadeh was familiar with the Envision Center after collaborating on a previous proposal about electrical safety, and she reached out to them during the proposal phase of this project. After she was awarded funding from the National Science Foundation (award # 2049711), the extended reality environment came to life.

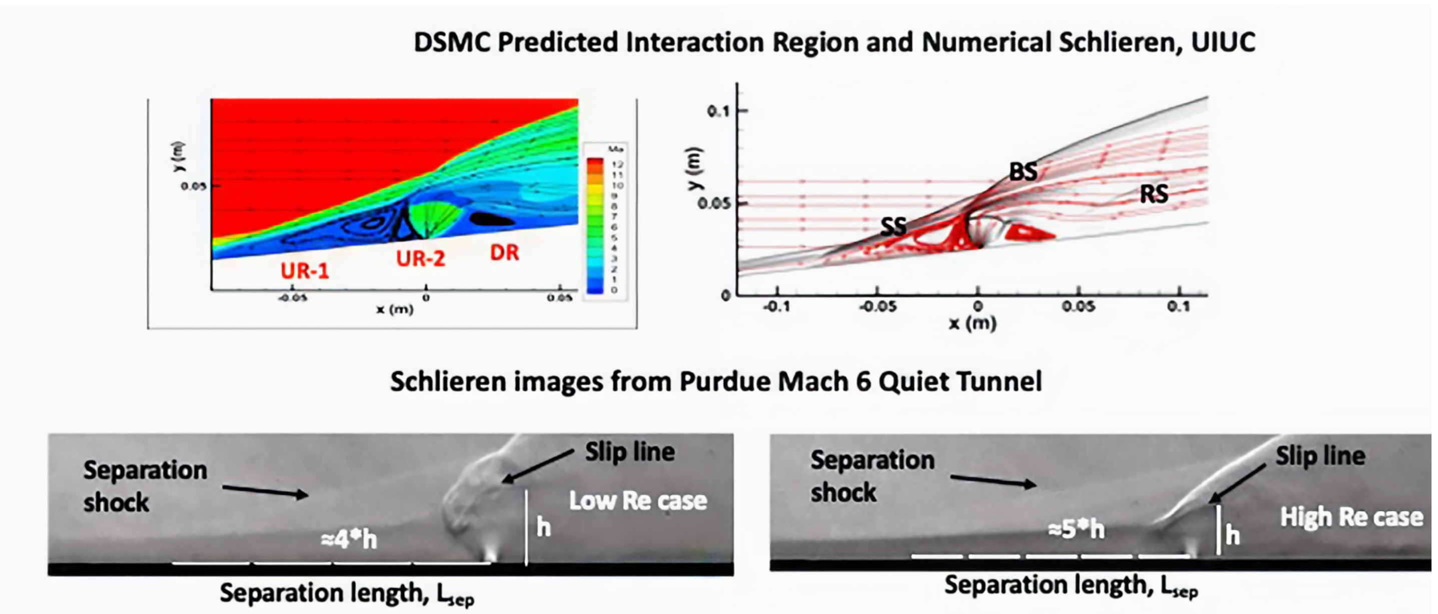
“It was a wonderful experience working with the Envision Center,” says Hasanzadeh, who plans to continue her collaboration with the Center. “It was also a great learning opportunity for our students.”

“WE CAN'T ASK PEOPLE TO GO ON THE ROOF OF A TWO-STORY BUILDING WITHOUT PROTECTION,

We were looking for a visualization that would allow us to provide the same sense of presence of being on the roof and allow them to safely install shingles.”

Proposal Submissions: In FY 2023, Purdue faculty submitted 335 proposals for $340.9M that specified the need for research computing facilities in the pre-award worksheet. (Vs 269 proposals for $391.5M in FY22)

RESEARCH TEAM USING ANVIL IN HYPERSONICS RESEARCH



A team of researchers has been using the Rosen Center for Advanced Computing (RCAC)'s powerful Anvil supercomputer to study gaseous jets in hypersonic crossflow, which is useful for understanding the operations of the systems that are used to steer aircraft and spacecraft in the upper atmosphere.

Joseph Jewell, an assistant professor of aeronautics and astronautics, collaborated with Deborah Levin, a professor of aerospace engineering at the University of Illinois at Urbana-Champaign, and her team of students, to compare their experimental results from Purdue's hypersonic wind tunnel with computational simulations done on Anvil.

The simulation methods they use are computationally intensive but easily parallelizable, meaning they can make good use of Anvil's many cores and massive computing power.

"This Anvil allocation has really helped us to do some work that we can put in proposals in the future," says Jewell.

Anvil has helped the team make progress on answering questions about the degree to which jet interactions scale across different parameters, such as whether the penetration depth of the jet will change with pressure if the Mach number is consistent.

Access to Anvil is provided through the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), previously the Extreme Science and Engineering Discovery Environment (XSEDE), in which Purdue has long been a partner. Purdue's participation in these national cyberinfrastructure programs has made it easier for faculty to benefit from their resources.

Anvil is Purdue University's most powerful supercomputer yet, providing researchers from diverse backgrounds with advanced computing capabilities. Built through a $10 million system acquisition grant from the National Science Foundation (NSF), Anvil went into production in February of 2022 and will continue serving the research community through 2027.

Over the coming years, the Anvil supercomputer will continue to enable significant discoveries across many areas of science and engineering. Anvil will also serve as an experiential learning laboratory for students to gain real-world experience using computing for their science, and for student interns to work with the Anvil team to innovate and enhance Anvil's capabilities.

Anvil, which was built in partnership with Dell and AMD, consists of 1,000 nodes with two 64-core AMD Epyc "Milan" processors, each with a peak performance of 5.3 petaflops. The supercomputer ecosystem also includes 32 large memory nodes, each with 1 TB of RAM, and 16 nodes each with four NVIDIA A100 Tensor Core GPUs providing 1.5 PF of single-precision performance to support machine learning and artificial intelligence applications.

RCAC, COLLEGE OF PHARMACY PARTNER ON AI TOOL TO IDENTIFY ADVERSE DRUG EFFECTS FROM SOCIAL MEDIA POSTS

Adverse drug effects are thought to be vastly underreported, but a team from Purdue's College of Pharmacy is working with the Rosen Center for Advanced Computing (RCAC) to use artificial intelligence to change that.

RCAC senior research data scientist Sarah Rodenbeck partnered with Kyle Hultgren, the director of Purdue's Center for Medication Safety Advancement and a clinical assistant professor of pharmacy practice, to design and implement a natural language processing (NLP) tool that can be used by a computer to identify social media posts regarding medication use and recognize the ones mentioning harm or side effects associated with them.

The tool identified potential adverse drug events (ADEs) in social media posts and gave researchers a way to interact with both the obtained data and the Food and Drug Administration (FDA)'s adverse event reporting system (FAERS) data. The model was successful in achieving 93% accuracy in classifying ADE tweets from non-ADE tweets.

ADEs are thought to be underreported because not everyone reports them to their doctors, for example, if a side effect is mild, or people do not think to report anything back to medical professionals.

Social media users seldom use precise medical terminology. Instead, the text is filled with slang, pop culture references, and jokes or sarcasm. Many NLP models face challenges to correctly interpret this type of text. However, the tool RCAC developed made it possible to interpret such language.

"By identifying ADEs in social media we hope to augment the view researchers have about how medications are used and experienced," said Rodenbeck. "Combined with other data, this could eventually be used to identify additional side effects not already included on a drug label or to discover potential manufacturing problems. This could result in anything from additional safety studies all the way to partnering with manufactures to improve drug design."

The NLP model was part of a larger effort to improve and enhance the HubZero SafeRX tool, with Rajesh Kalyanam and Shawn Rice from RCAC also playing vital roles. Rice worked on updating the HubZero SafeRX database ingestion tool and Kalyanam was responsible for deploying the model.

This study illustrates how technology like GPUs and high-performance computing can be used to benefit a previously underrepresented domain such as pharmacy. In addition to the interdisciplinary partnership between the College of Pharmacy and RCAC, this work makes use of Anvil, Purdue's recently built $10 million NSF-funded supercomputer.

Anvil was used to train the NLP model and perform binary classification inference on tweets collected each month. Anvil GPUs were used to train the model, and Anvil CPUs were used to perform inference with the trained ML model. A startup allocation of 100,000 service units on Anvil is being used to collect Twitter data and run the inference on a regular basis by manual job submission via Anvil's Slurm scheduler.

"I really want to emphasize how important this collaboration has been for my team," says Hultgren. "Working with RCAC has truly made a powerful impact on our work and our ability to understand more about how the US population is interacting with prescription medications. Ultimately, it is collaborations like this that will help us make a clinical impact in our communities as we strive to reduce the harm associated with medication use."

This project was supported by funding from the FDA.

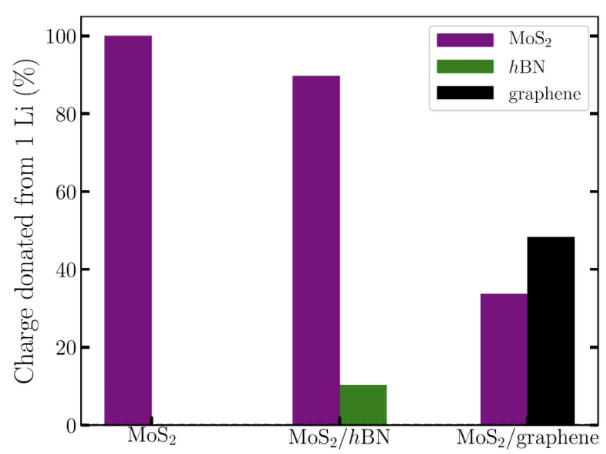
ANVIL HELPS YALE RESEARCHERS IN THE SEARCH FOR BATTERIES OF THE FUTURE

Purdue's Anvil supercomputer has been helping researchers forge a path in developing the batteries of the future.

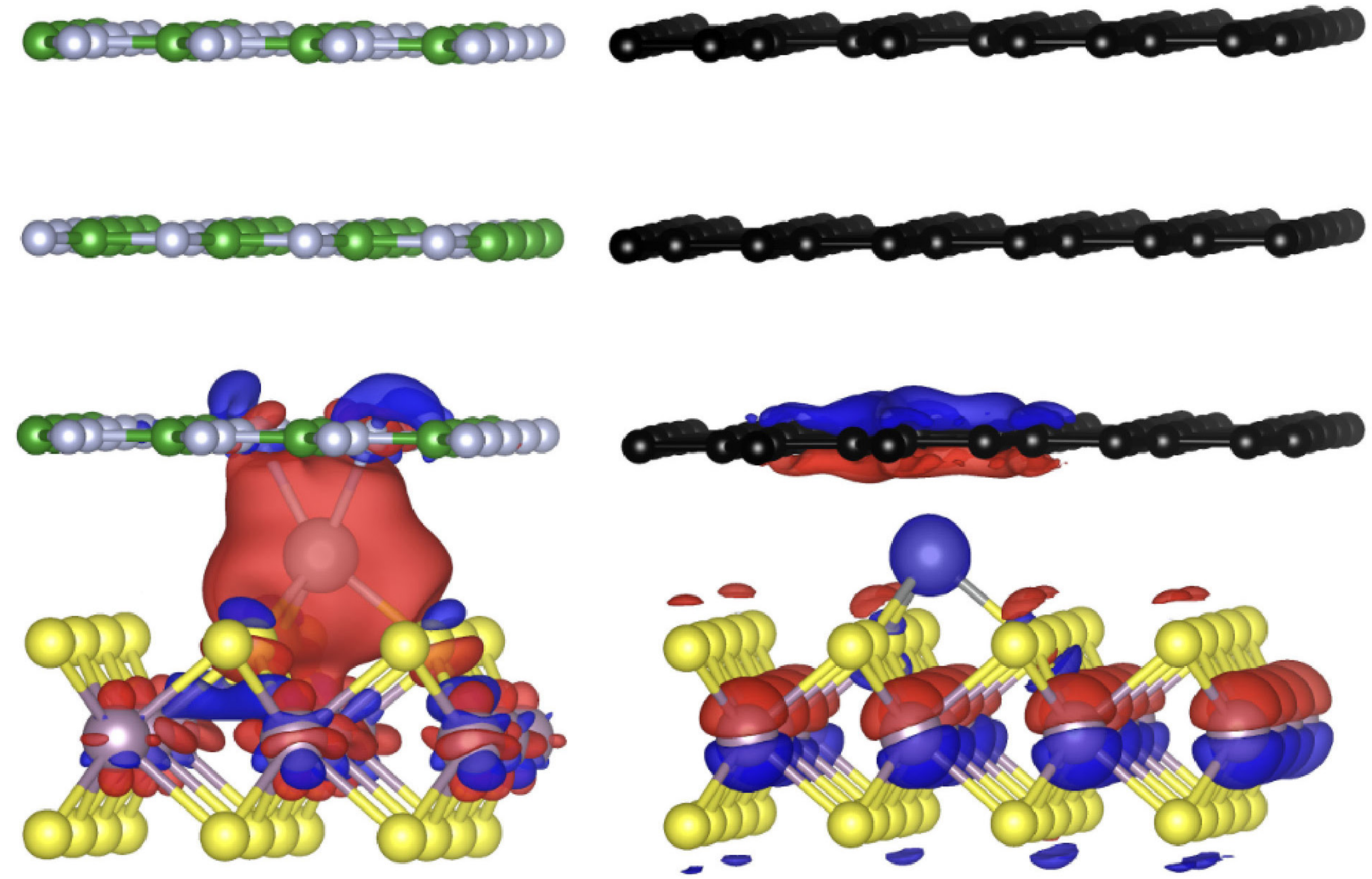
Dr. Aakash Kumar, a postdoctoral associate in the Department of Mechanical Engineering and Materials Science as well as the Energy Sciences Institute at Yale University, utilized the powerful computational capabilities of the Anvil supercomputer to help his experimental collaborators explain a novel finding in the search for potential new materials for energy storage. Their research involved a combined experimental-computational investigation of lithium intercalation in heterostructures of quasi-2D materials. In this collaborative work, the experimental work was led by Dr. Josh Pondick and Prof. Judy Cha, who carried out laboratory experiments, while the modeling team carried out the computational work on Anvil to elucidate the results of those same experiments. The goal of their project was to understand how heterostructures—a single, layered structure created by stacking two different materials—could influence the phase transition caused by lithium intercalation (incorporation of Li within a host material, such as in the interlayer gaps of a 2D material), and whether this process could be controlled. Researchers can then use this knowledge to determine what combinations of heterostructures would work best for energy storage and how to optimize their construction, potentially leading to the creation of more stable and longer-lasting batteries.

For Kumar's research, the team looked specifically at two-dimensional (2D) nanomaterials. 2D materials are essentially ultrathin "sheets" with a high surface-to-volume ratio, meaning they have large surface areas exposed where other atoms, such as lithium, can be intercalated. For purposes related to energy storage, researchers inject lithium ions into the layers of these 2D heterostructures and observe the associated phase change—specifically, they must know precisely where the charge (electron donation) from the lithium goes and how much lithium is required to initiate the phase transition. Ab initio quantum-mechanical calculations that run on supercomputers can be extremely useful in understanding such phenomena. This information is crucial for determining if certain materials are conducive to energy storage applications and how they could be later tailored for use in batteries. To quote Kumar and team's research publication (Pondick et al. 2021), "To effectively exploit these phase transitions for these applications [batteries], understanding the thermodynamics and kinetics of the nucleation pathways is essential." And so that's exactly what the team set out to do.

Kumar used the Anvil supercomputer to compute the thermodynamics and electronic structure of MoS2/hBN and MoS2/graphene, utilizing a software called Quantum Espresso. MoS2/hBN is a heterostructure that combines molybdenum disulfide (MoS2) and hexagonal boron nitride (hBN), while MoS2/graphene is a heterostructure combining molybdenum disulfide and graphene. MoS2, hBN, and graphene have all shown to be promising materials for a wide range of electronics, including batteries. Comparing these two heterostructures with bare MoS2, the research team found that manipulating the heterostructure design of 2D materials is an effective strategy to control lithium intercalation, which could greatly aid the development of more stable, longer-lasting batteries.



After running his calculations, Kumar found that graphene took roughly 50% of the charge donated from lithium, impeding the intercalation process, whereas hBN did not interfere, meaning that you must inject more lithium into the MoS2/graphene heterostructure in order to see the same phase transition. This information is crucial for researchers who want to develop 2D nanosheet batteries. In the end, Kumar and his team used Anvil to provide the first experimental evidence that different 2D nanamaterials can be used to control the lithium-induced phase transition in MoS2 nanosheets.



Charge density difference isosurfaces of MoS2 heterostructures with hBN (left) and graphene (right)

Anvil was extremely helpful in this discovery. Thanks to Anvil's 128 cores per node, and up to 1TB of memory per node, Kumar and his team could perform their calculations much faster and more reliably than otherwise possible.

"In my mind, running calculations on Anvil is literally like striking a hammer," says Kumar. "There is so much power."

In this research, Kumar uses ab initio quantum mechanical calculations to compute the changes in electronic structure of materials due to lithium intercalation to determine their suitability for solid-state battery applications. He found that on Anvil, he could complete his medium-sized calculations (~50-100 atoms) using one or two nodes, whereas on HPC systems he has used in the past, he needed to use many more. "There [on other HPC systems], in my experience, you tend to use more nodes, which is fine, but what happens is, when you use more nodes, there is a communication issue between the machines." These communications issues led to what Kumar described as a bottleneck, slowing down the calculations and the research. But because Anvil has such a large number of powerful cores per node, he could avoid this communication bottleneck and complete all of his calculations using a smaller number of nodes, which saved him a lot of time.

Another benefit Anvil provided for Kumar and the research team was the large amount of memory per node. "Because there is so much memory associated with each core, I can let them work on more memory-intensive problems."

Kumar went on to note that the Anvil technical staff were very helpful throughout his allocation. "I have to say that the support staff was amazing," says Kumar. "One person [Ryan DeRue] helped me with this one particular thing I was trying for a few weeks, and he would reach out to me every few days to check in." Ryan DeRue is a Senior Computational Scientist at the Rosen Center for Advanced Computing (RCAC). He works on the scientific applications team, building, testing, and optimizing the software that researchers use for their scientific workloads. In this instance, DeRue worked with Kumar to customize the Quantum Espresso software and to ensure that the installation of this custom software was well-optimized for Anvil. Once this was completed, Kumar was able to use an alternative solver to enhance his calculations.

Anvil performed beyond expectations on this project, and Kumar and his team were thrilled with the system overall. They applied for a new allocation on Anvil and have been approved for another year.

References:

1. Pondick, Joshua V., Aakash Kumar, Mengjing Wang, Sajad Yazdani, John M. Woods, Diana Y. Qiu, and Judy J. Cha. 2021. "Heterointerface Control over Lithium-Induced Phase Transitions in MoS2 Nanosheets: Implications for Nanoscaled Energy Materials." ACS Applied Nano Materials 4 (12): 14105–14. https://doi.org/10.1021/acsanm.1c03402.

RCAC PARTNERS WITH PURDUE FACULTY ON NSF-FUNDED CYBERTRAINING PROGRAM

Purdue's Rosen Center for Advanced Computing (RCAC) is a partner on an $850,000 grant awarded to a team led by Venkatesh Merwade, professor of civil engineering, to create a cybertraining curriculum for climate, water and environmental (CWE) sustainability.

Chief scientist Carol Song and senior research scientist Lan Zhao are the co-PI and senior personnel, respectively, on the award, which will address a gap in training on how to integrate and work with data from different CWE sources by creating a set of cybertraining modules that will incorporate rapidly developing technology and community best practices.

The cyberinstrastructure development team also includes RCAC's Shawn Rice, lead computational scientist, and Jaewoo Shin, senior computational scientist. They are developing an open, scalable online platform that will support the interactive cybertraining modules and give students access to additional tools and resources RCAC has developed over the years.

The learning modules are organized at three levels of depth: foundation, expert, and developer, with badges and certifications available at each level. While the foundational modules teach essential skills such as coding, the expert modules impart in-depth knowledge and skills such as data processing, simulation, and visualization for specific CWE subareas. The developer modules aim at training students to be able to contribute by expanding existing and creating new tools.

"Our traditional curriculum includes the science of Climate, Water and Environmental Sustainability, which does not equip the students to tackle new challenges using computational tools and cyberinfrastructure," says Merwade. "Our cybertraining project is filling this critical gap by providing students these critical skills."

In June 2023, the Purdue team hosted a workshop on campus to educate researchers from around the country about how to use the cybertraining platform to develop and disseminate training modules. On the second day of the workshop, attendees divided into groups and started to develop their own modules.

The recent NSF award is the second round of funding for this project. A previous NSF award in 2018 supported climate and water cyber training courses at Purdue and a series of virtual workshops for participants around the United States.

To advance the advanced training and learning activities in this round, the RCAC team is migrating to a new cybertraining platform that will deliver a seamless connection to large scale computing systems such as Anvil.

Says Song, "we are in an excellent place right now: leveraging the new on-prem composable system Geddes, the large capacity Anvil HPC resource, and the Halcyon web framework that powers the campus research computing management system, the new cybertraining platform will provide a powerful but smooth learning experience to the project's audience." The new platform will debut later in 2023.

Another goal of the project is to help researchers make their science "FAIR" – findable, accessible, interoperable and reusable. To that end, the funding has previously supported fellowships to help students turn software applications in development into modules that follow FAIR principles.

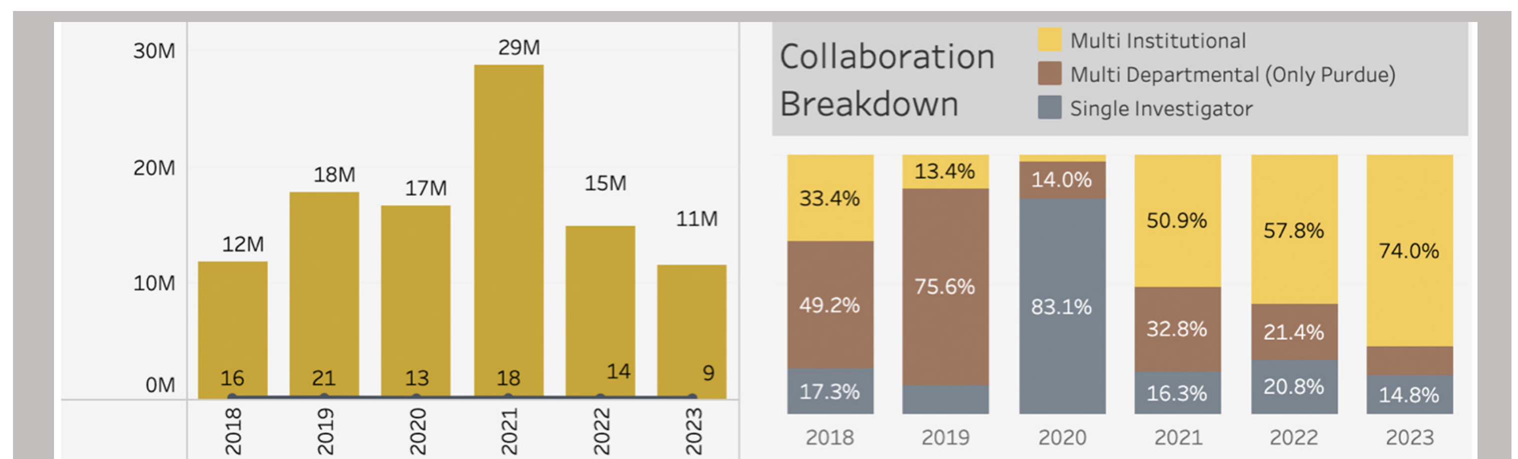
This project is an interdisciplinary collaboration at Purdue that includes not only Merwade's expertise in water resources and RCAC's cyberinfrastructure development, but

Researchers from around the country gathered at Purdue University to attend a Cybertraining Workshop focused on teaching them how to use the new cybertraining platform.

also Matt Huber, professor of earth, atmospheric and planetary sciences, and Jake Hosen, assistant professor of forestry and natural resources, who provide climate and environmental expertise, respectively. Nicole Kong, associate dean for research and associate professor of library science, and Wanju Huang, clinical associate professor of learning design and technology, are also part of the team.

The Purdue team is also working with a group led by Adnan Rajib at the University of Texas at Arlington to make the cybertraining modules available to minority institutions at the high school and college levels.

This work is supported by NSF Award Nos. 2230092 and 1829764.



9 RCAC PIs submitted 22 proposals for $17.8M in FY23, with 5 awarded for $2.7M, and 11 not funded ($8.7M). 6 proposals totaling $6.3M to RCAC are currently pending.

RCAC ADDS GPUS TO CLUSTER AS PART OF BROADER INVESTMENT IN AI RESEARCH

After significantly expanding GPU computing capabilities in fall 2022, the Rosen Center for Advanced Computing (RCAC) has added 104 new NVIDIA A100 GPUs to the Gilbreth community cluster. Based on Dell PowerEdge R7525 compute nodes with .5 TB of RAM, two Nvidia A100 Tensor Core GPUs, and 100 Gbps HDR Infiniband, this expansion increases the GPU capacity in the cluster by 46%.

As part of a push to expand GPU computing capabilities to meet the needs of researchers working in AI, machine learning, and accelerated modeling and simulation, the Rosen Center for Advanced Computing (RCAC) recently went through two major expansions of the Gilbreth community cluster to include more GPUs.

In fall 2022, RCAC added 93 new NVIDIA A10, A30, and A100 GPUs to Gilbreth, increasing the number of available GPUs on Gilbreth by 40%. In spring 2023, 104 new NVIDIA A100 GPUs were added to Gilbreth, further increasing GPU capacity by 46%.

The new GPUs bring Gilbreth to an aggregate peak performance of 32 single-precision PetaFLOPs (one quadrillion operations per second) – doubling Gilbreth's AI performance.

The added resources are just a part of RCAC's ongoing investment in supporting researchers performing AI and machine learning work. Along with the additional hardware, RCAC has full-time research scientists with AI and machine learning expertise, who offer training opportunities and are available to partner with faculty on proposals.

"Since the summer of 2022, continued investment by Purdue IT has allowed RCAC to grow our GPU capacity available to Purdue researchers and students from 104 to 331 after this latest expansion — a two-and-a-half times increase," said Preston Smith, Executive Director for RCAC in Purdue IT.

"This investment in resources, and pricing structures that are better than do-it-yourself options will ensure that Purdue faculty have access to cutting-edge resources to ensure their competitiveness, while also benefitting from the community cluster program's professional maintenance and scientific support to ensure that their research is appropriately safeguarded, and they can focus on scientific outcomes instead of technology problems."

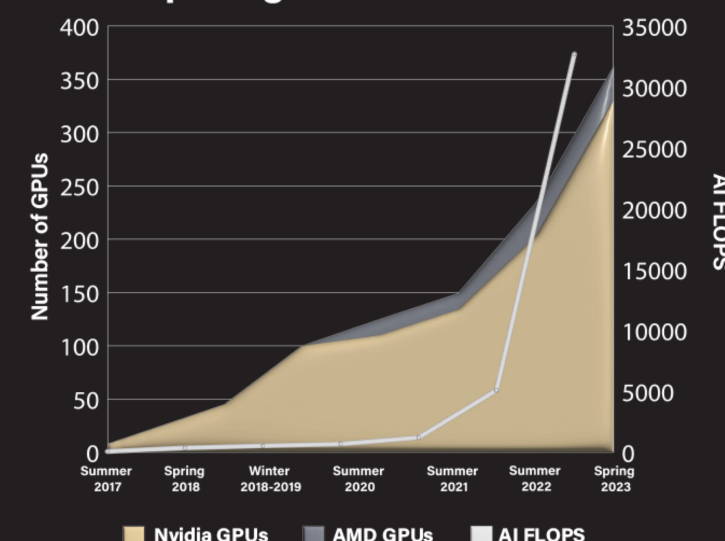
Daniel Aliaga, associate professor of computer science, led a team that used 100 of the Gilbreth GPUs to create urban representations for 330 cities across the United States, with a significant emphasis on sustainability and the cities of the future.

"The availability of the GPUs enabled us to improve algorithm development as well as run the model to enable creation of the results," said Aliaga.

Aliaga expressed his excitement over the accomplishments of the team, stating, "Our research has pushed the boundaries of what is possible in generating urban-related content. The utilization of the RCAC GPUs has allowed us to achieve unprecedented milestones and significantly contribute to various fields, from computer science to ecological sciences." "This project demonstrates a unique capacity at Purdue to push the boundaries of science via computing.

The Rosen Center is proud to have contributed to this exciting research project and grateful for the investments in our infrastructure that have enabled us to support projects such as this," adds Arman Pazouki, director of scientific applications for RCAC."We are excited that this project offers a model for how to enable future research in the fields of AI and machine learning."

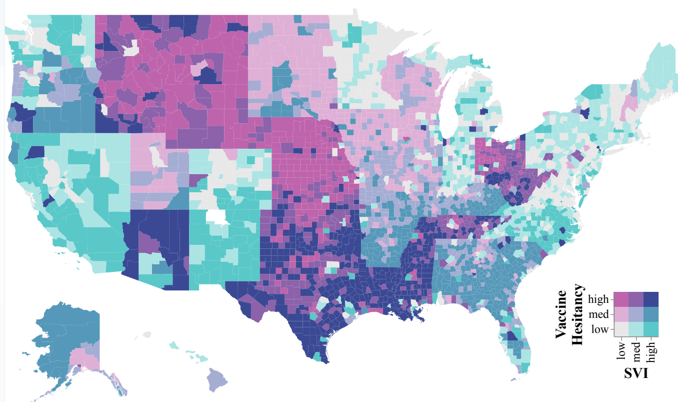
Centrally Provided GPU Computing Resources at Purdue



> **Since the summer of 2022, continued investment by Purdue IT has allowed RCAC to grow our GPU capacity available to Purdue researchers and students from 104 to 331 after this latest expansion a two-and-a-half times increase,**
> **- Preston Smith, Executive Director for RCAC in Purdue IT.**

PURDUE'S ANVIL SUPERCOMPUTER ASSISTS WITH COVID-19 PANDEMIC RESPONSE

A research group from the University of Virginia (UVA) utilized Purdue's Anvil supercomputing cluster to help provide COVID-19 scenario modeling for local, state, federal, and university officials and departments.



Social Vulnerability Index vs. Vaccine Hesitancy across US

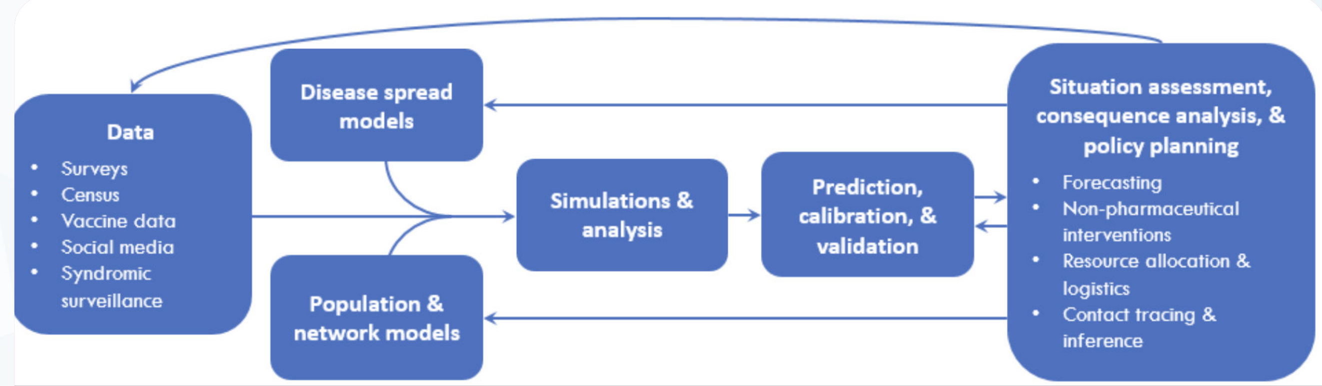
The team used computer models to predict what might happen with the virus, such as how it might spread, how many people could get sick and need to go to the hospital, and how vaccines might help over time. The computational power of Anvil was used to run these scenario models, which produced data that helped drive pandemic response efforts in real-time at varying levels of government.

Dr. Madhav Marathe, the Director of the Network Systems Science and Advanced Computing (NSSAC) division of the Biocomplexity Institute and Initiative at the University of Virginia, works in the field of computational epidemiology. He and his team use advanced modeling techniques and simulations to investigate large-scale biological, social, and technological systems. In this instance, the NSSAC team was tasked with simulating scenarios surrounding the COVID-19 pandemic. The scenario models produced by the team were not only used by local, state, and federal officials to help make informed decisions regarding pandemic response, but were also used to bolster the CDC's COVID-19 Scenario Modeling Hub, a consortium combining long-term COVID-19 projections from different research teams. The importance of computational epidemiology during a pandemic cannot be overstated—more data equates to more accurate projections, and high-performance computing (HPC) is needed to run the simulations and analyze that data quickly. To quote a paper co-authored by Dr. Marathe (Marathe & Vullikanti, 2013):

"Computational models help in understanding the space-time dynamics of epidemics. Models and algorithms can be used to evaluate various intervention strategies, including pharmaceutical interventions such as vaccinations and anti-virals, as well as non-pharmaceutical interventions such as social distancing and school closures."

Without computational epidemiology, pandemic-response teams and decision-makers would be flying blindfolded. But even with the help of supercomputers, not all datasets are created equal—the type of model you use to obtain the data matters.

In the realm of epidemiology, three main types of modeling are used, each with its strengths and limitations: statistical-based models, compartmental mass-action models, and agent-based models. Agent-based modeling is a relative newcomer to the party. Pioneered by Marathe and his group in 2004, agent-based models take into account real-world factors, such as geography, social relationships, and individual differences—things that can have a major impact on how a disease spreads through a community but are lost in traditional



Elements of computational real-time epidemiology
Diagram courtesy of Lily Li, a Postdoctoral Research Associate with the NSSAC team.

epidemiology models.

"The idea is," says Dr. Marathe, "if I want to understand how a disease spreads in a population, then I want to understand the social network that underlies this population because I want to capture the underlying heterogeneity that exists. Compartmental mass-action models treat everyone as identical. They lose all aspects of heterogeneity and asymmetry that exist in the real world."

Agent-based modeling eliminates this oversight by representing individual people moving throughout their specific environment, behaving as normal humans would behave. These models simulate realistic populations, complete with a proper distribution of age, gender, ethnicity, job type, and geographical location. They also factor in individual behavior, such as vaccine usage or mask-wearing, which is traditionally difficult to account for, yet is a huge factor in disease spread.

"These models tell you something that is quite unique in their form. So, for instance, you can start understanding the role of network structure on the outcome—whether a dense part of the town would have more diseases than a more sparse part of the town. Or, would it affect people from some racial or ethnic background more than others because they stay together in close quarters? Or is it the case that folks who are resistant to taking vaccines are likely to be impacted somewhat disproportionately to people who have taken vaccines? These are decisions and issues where individuals and individual behavior and their situatedness in the network play a role. And this aspect is lost in aggregate models."

Agent-based models are excellent at providing accurate, real-world data that can be used to inform policy and decision-makers on how best to respond to an evolving situation, but there is one downside—they require a massive amount of computing power.

For agent-based modeling to work, each individual accounted for within the model has to be its own data point. So when the NSSAC team set out to provide "what-if" COVID-19 scenarios for the entire population of the United States, they needed to replicate more than 300 million people with-individual level resolution, a task that required an astounding amount of computations to complete. According to Dr. Marathe, "This is not a computation that you can do on a single laptop or even a small cluster. Certainly not multiple runs. These are very, very large computations. And that's where Anvil-like systems come in very, very handy, and are extremely useful in the process."

To manage such heavy computations, NSSAC not only needs access to multiple high-performance computing systems, but they also need an efficient interdisciplinary team to manage the workflow.

Jiangzhuo Chen is a Research Associate Professor and a computer scientist working on epidemiology at NSSAC. Chen's role in the project involves building the agent-based models for COVID-19, setting up the experiments, and interpreting the results. Once the models are built, Chen coordinates with Stefan Hoops, a Research Associate Professor at NSSAC, and Dustin Machi, the Senior Software Architect at NSSAC. Hoops is the author of EpiHiper, the simulation software used by the team. "EpiHiper," says Hoops, "is a simulation tool that is highly adaptable to the policy question without any alteration. This improves the quality of the results and thus is creating trust."

Machi takes the models and experiments that have been designed by Chen and—using the EpiHiper software created by Hoops—runs the simulations on various HPC resources, including Anvil. Machi, Hoops, and Chen, along with the rest of the team, then work to optimize and calibrate the parameters needed for successful simulations, ensuring that the simulations are set to provide the answers that they need.

Another crucial member of the team is Parantapa Bhattacharya, a research scientist at NSSAC. Due to the volume of scenarios the team needs to model, they cannot rely on one supercomputing system to complete all their simulations. Outside of the Anvil cluster, the team utilized a UVA local system named Rivanna, as well as Bridges, a supercomputer run by the Pittsburgh Supercomputing Center. Bhattacharya developed an efficient pipeline to manage the team's resource usage, ensuring that the individual simulations get access to the resources they need. "My job," says Bhattacharya, "is to make sure the simulation runs fast, is scalable, and we are using the resources we get as efficiently as possible."

Here is a snapshot of how the process typically works:

1. The CDC's Scenario Modeling Hub sends the team a "What-If" scenario that needs to be simulated.
2. The team creates the agent-based model. These models need to be optimized and calibrated.
3. The team runs calibration and optimization simulations on their local UVA HPC resource, Rivanna. This allows them to figure out what they need for the full simulations.
4. The full simulation, depending on the computational needs based on the parameters of the scenario, is then run on a more powerful HPC cluster, e.g., Anvil.
5. The results from the full simulation are then gathered, analyzed, and compiled in a report, which is sent back to the CDC's Scenario Modeling Hub.

Normally, experiments such as this could take a significant amount of time to complete. On the one hand, if you do not have access to a powerful enough HPC resource, it could take days or weeks to simply run a simulation, not to mention the amount of time it would take to interpret the terabytes of data obtained from agent-based models. The power and speed of the Anvil supercomputer proved to be extremely useful in this regard. On the other hand, even if you have access to the fastest supercomputer in the world, getting the actual timeslot on the machine to run your experiment could take way longer than is required for pandemic response situations. Oftentimes the information needed by the authorities has a very short turnaround time, and waiting in a queue for such projects is simply not an option. This is where Anvil truly shined.

The Anvil team made it a point to prioritize computing time for the NSSAC team so that the information needed to optimize vaccination strategies or minimize infections and hospitalizations was quickly delivered to the appropriate officials. This expedited process allowed the decision-makers to develop science-based responses based on real-time data, which is crucial during an evolving pandemic.

According to Machi, "Anvil was just nice to use—it was easy to get access to, and the team was great to work with, so we were able to get a configuration set that worked for our style of pipeline really well. A lot of jobs that get put onto the clusters will be just single jobs that can run, and however long they take to make it into the queue and complete is fine, but in the case of these large-scale simulations, we are doing coordinated jobs, often with a relatively small window before things are due. So if we leave ourselves to the mercy of the scheduler, then we may never get them completed in time. We got into Anvil because of this need to have multiple clusters and stuck with it because it was convenient to use and worked well for our work."

It's clear that having quick access to powerful HPC resources is of the utmost importance to pandemic response efforts, not only for addressing COVID-19, but also for any future pandemics that may arise. As Dr. Marathe succinctly put it:

"We need high-performance computing systems. We were here to support an evolving pandemic, in terms of operational response. This is not just a science exercise where we can wait for a while, run some experiments, do some analysis, make some interesting scientific findings, etc. We do that too, but that's not the primary purpose. The primary purpose was to support our state, local, and federal officials as they were responding to a pandemic. And time is absolutely critical. If they need an answer in three days, then they need an answer in three days. After seven days, the results may not be that useful. So what we needed was access to [HPC] machines that were relatively easy to use, flexible in their usage, and could support this operational real-time requirement. I think this is a central question that we all have to solve should we want to prepare for the next pandemic, and I think that you folks [at Anvil] did a tremendous job in supporting our desire, and I would say, the desire of federal agencies."

References:

1. Marathe M, Vullikanti AKS. Computational epidemiology. Communications of the ACM. 2013;56(7):88-96. doi:https://doi.org/10.1145/2483852.2483871

2. https://nssac.github.io/

3. Bhattacharya, et al. "AI-Driven Agent-Based Models to Study the Role of Vaccine Acceptance in Controlling COVID-19 Spread in the US." 2021IEEE International Conference on Big Data (Big Data). 2021

ANVIL

FORGING THE FUTURE OF COMPUTING

867 Million Allocated CPU hours

897 Thousand Allocated GPU hours

513 Allocations

4271 Anvil users

46 Science Domains

10 Million Processor hours from On-Demand jobs

16 Students directly supported

204 Outreach events

1800+ Undergraduate student users

10 Training sessions

PURDUE UNIVERSITY | Rosen Center for Advanced Computing

Anvil is funded by NSF grant #2005632

RCAC HOSTS DIGITAL TWIN SYMPOSIUM FEATURING PURDUE RESEARCHERS, NVIDIA

The Rosen Center for Advanced Computing (RCAC) held a symposium on digital twins at the Envision Center on April 11 featuring presentations from Purdue faculty and NVIDIA.

A digital twin is a virtual representation of a physical object that is used to model or simulate what's happening to the object. The symposium was designed to explore how RCAC's campus cyberinfrastructure and the Envision Center can support and enable faculty partners to use digital twins in their science.

A number of Purdue faculty researchers spoke at the symposium about how their specific domains are using and can benefit from digital twins. Speakers included:

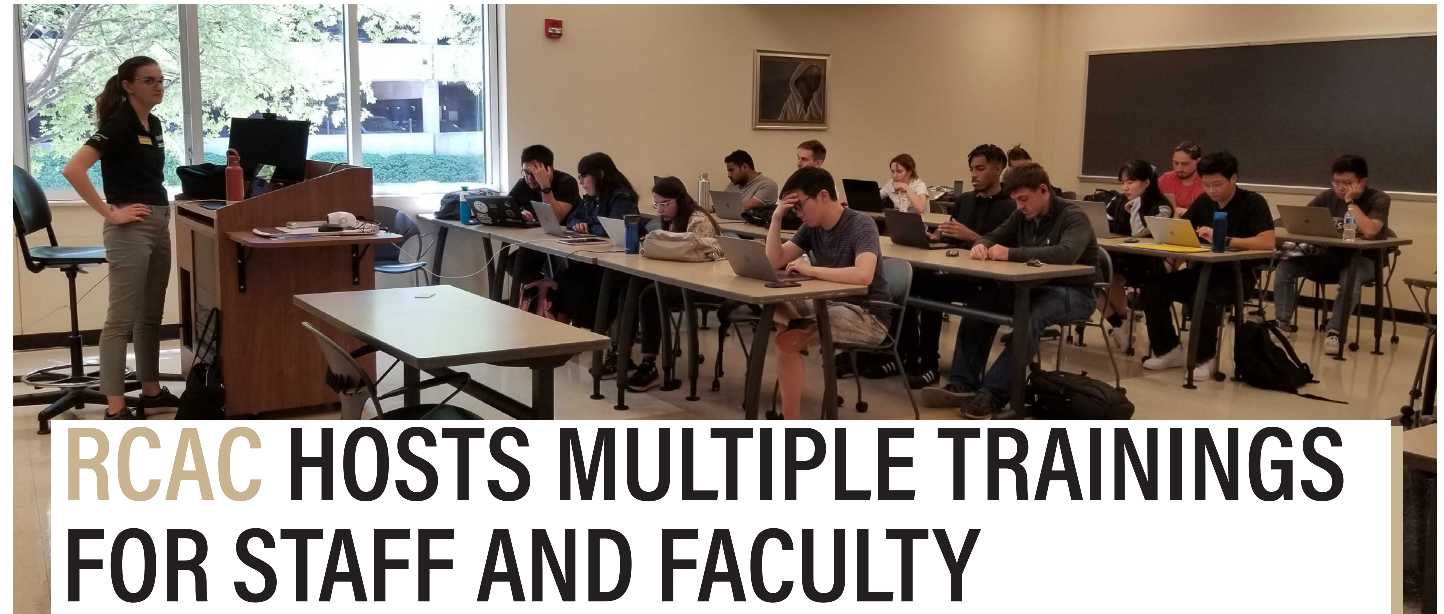
- **Nathan Hartman**, Dauch Family Professor of Advanced Manufacturing and Head of the Department of Computer Graphics Technology, who spoke about how digital twins are used in the manufacturing realm;
- **Jian Jin**, associate professor of agricultural and biological engineering, who discussed Purdue's use of crop sensors to create digital twins in agriculture;
- **Songlin Fei**, professor of forestry and natural resources, who spoke about the use of digital twins in forestry;
- **Daniel Aliaga**, associate professor of computer science, who discussed how he uses digital twins in his deep generative modeling of urban spaces; and
- **Ziran Wang**, assistant professor of civil engineering, who discussed the use of digital twins in connected and automated vehicles.
- **Zoe Ryan** from NVIDA presented about the latest developments that support research that relies on digital twins and how NVIDIA's Omniverse platform can be used to build digital twins.

Partnerships such as this alongside open communication channels with Purdue researchers have been instrumental in forming the AI vision at RCAC. Future events and trainings will be created to support this vision.

"Modeling, simulation, AI, augmented reality, and digital twins are all computational techniques that more and more Purdue researchers use on a daily basis," said RCAC Executive Director Preston Smith. "We're excited to bring together researchers, students, and our technology experts from RCAC and Envision to learn from each other and plant the seeds of future collaborations."

Smith also notes that digital twins can be a powerful tool when paired with AI. To support the growth in AI, RCAC has invested substantially in AI computing resources, adding over 100 GPUs to the Gilbreth Community Cluster.

The event concluded with a panel discussion moderated by RCAC lead research data scientist Geoffrey Lentner and featuring several faculty speakers.



RCAC HOSTS MULTIPLE TRAININGS FOR STAFF AND FACULTY

The RCAC Training program is a year-round series of educational workshops delivered in a variety of modalities (in-person, virtual, asynchronous). These targeted sessions are created to support all levels of HPC users—from novices and first-timers who are just getting started to domain experts looking to develop new skills. The purpose of these trainings is to help support the growth of the HPC and general computing communities at Purdue and nationwide. Regardless of background or prior experience levels, the RCAC Training program has what users need to accomplish more through computing.

- **Containerizing HPC applications with Singularity/Apptainer**, for researchers interested in learning about containerizing HPC applications using Singularity/Apptainer.
- **Open OnDemand 101**, This session is a beginner-level discussion focusing on how to transition from using local computing resources to using HPC cluster resources via Open OnDemand.
- **Anvil Open OnDemand 101**, for students, staff, and faculty interested in transitioning from using local computing resources to using the ACCESS Anvil cluster via Open OnDemand
- **Unix 101**, for students, staff, and faculty looking to learn how to use Linux/Unix-based high-performance computing systems like RCAC's resources. This first workshop will take you through the steps of logging into Unix systems, along with an overview of file manipulation and shells.
- **Fundamentals of Job Management**, a virtual workshop will introduce attendees to various strategies available for establishing and managing their scientific workflows. Transforming a job to a workflow, Analyzing resource usage within a job, Slurm constructs for creating/managing workflows across many jobs, Tools for creating/managing workflows within a single job.

Purdue Rosen Center for Advanced Computing supported Women in HPC Diversity at SC22.

ANVIL REU SUMMER 2023 PROGRAM IS A HUGE SUCCESS

The 2023 Anvil Research Experience for Undergraduates (REU) Summer program saw five students from across the nation gather at Purdue's campus in West Lafayette, Indiana, to learn about high-performance computing (HPC) and work on projects related to the operations of the NSF-funded Anvil supercomputer at Purdue. During the program, which is hosted and run by the Rosen Center for Advanced Computing (RCAC) with support from the National Science Foundation (NSF), the students were able to gain the knowledge and skills necessary to build and support advanced research computing systems and scientific applications.

The Anvil REU program is a paid summer internship open to undergraduate students from any background, not just those who major in Computer Science. When the application window closed in mid-February of this year, the program had received over 70 applications. The Anvil REU mentors—seven RCAC staff members who were in charge of the projects the students would work on during the summer—along with the Anvil executive team, took this list of 70+ applicants and distilled it down to five students. The five participants of the Anvil REU program were Aneesh Chakravarthula, Ved Arora, Oluwatumininu Oguntola, Nayelia Gurrola, and Henry Olson, all pictured below:

Pictured from left to right: Aneesh Chakravarthula, Computer Science major from Purdue University, **Ved Arora**, Data Science & Analytics major from Case Western Reserve University in Ohio, **Oluwatumininu Oguntola**, Computer Science major from the University of North Carolina at Chapel Hill, **Nayeli Gurrola**, Computer Science major from the University of Texas Rio Grande Valley, and **Henry Olson**, Computer Science and Cybersecurity double major from Purdue University Northwest.

The Anvil REU program consisted of four separate projects, with one or two students working on each one. These projects were chosen with real-world applicability in mind—the students would not only be gaining experience with HPC and learning new skillsets, but would simultaneously be increasing Anvil's capabilities. Each project had two mentors working with the students to help them achieve their goals. The tasks at hand were not simple. Working alongside their mentors, as well as helping each other across the different projects, it took the students the entirety of the summer to accomplish everything set before them. Completing these projects in 11 weeks was certainly a challenge, but the knowledge and skills the five students obtained—plus the camaraderie found during the program—are things that will stay with them for a lifetime.

On the final day of the program, the students presented their work to the Anvil team. As they demonstrated the results of their projects, each student discussed their accomplishments, obstacles, failures, and what they learned throughout the summer. The Anvil team then asked the students questions and gave them feedback on their work. Overall, these five students made fantastic progress: they completed their projects, learned hard and soft skills they will need when in the workforce, and gained an in-depth understanding of the world of HPC. The RCAC staff was delighted with everything the students achieved and are confident they will be successful in future endeavors.

Project 1:
The first Anvil REU project for 2023 focused on Data Analytics. Two students, Ved Arora and Nayeli Gurrola, teamed up to tackle this project under the supervision of their mentors, Guangzhen Jin, Amiya Maji, as well as Anvil Executive Team member, Preston Smith. In this project, the students instrumented and analyzed scientific application workloads on the Anvil system. Specifically, the two were tasked with integrating the XALT job-level usage activity monitoring tool for a deeper analysis of workloads. On completion of their REU program, Arora and Gurrola successfully presented a visualization dashboard with a view and insights into all the scientific applications that researchers have used on Anvil. This visualization dashboard was multi-faceted, including such information as how many users per application, how many CPU/GPU hours per application, as well as the spread of usage across all of the different applications.

Project 2:
The second Anvil REU project aimed at extending the Anvil supercomputer to burst scientific workloads into the Microsoft Azure cloud. Henry Olson worked on this project under the supervision of his mentors, Erik Gough and Ryan DeRue, and Anvil Executive Team member, Arman Pazouki. Olson implemented a direct cloud burst from Anvil to Azure for HPC and accelerator workloads, creating a hybrid cloud between the two and allowing for free-flowing communication and the sharing of resources. After facing multiple obstacles throughout the summer, Olson successfully completed his project on the second-to-last day of the program.

Project 3:
The third Anvil REU project aimed to extend Anvil's Kubernetes-based composable subsystem as an on-ramp to Azure Kubernetes Service. Oluwatumininu Oguntola worked on this project with his mentors, Erik Gough and Rajesh Kalyanam. The purpose of this project was to connect the Anvil composable subsystem's Rancher management platform to Azure Kubernetes Service in order to support elastic-scaling of workloads for science gateway applications. Oguntola was successful in completing his primary objectives, and as such, has made it easy to launch workloads on the Azure cloud system from the Anvil composable web interface.

Project 4:
For the fourth Anvil REU project, Aneesh Chakravarthula worked with his mentors, Christopher Thompson, Lev Gorenstein, and Anvil Executive Team member Rajesh Kalyanam, to build and enable containers in support of education. Throughout the summer, Chakravarthula developed deployment solutions for the Jupyter Notebook interactive computing platform that simplify the integration of course materials and federated authentication on Anvil's composable subsystem. By enabling the use of these large-scale notebook deployments, Chakravarthula's work has helped to provide interactive access to Anvil for education and training activities.

PURDUE, IU TEAM UP FOR STUDENT SUPERCOMPUTING COMPETITION

For the first time ever, Purdue students have teamed up with their counterparts at Indiana University to showcase their supercomputer building skills, and work hands-on with high performance computing (HPC) and scientific computing.

The joint Purdue-IU team competed at the Student Cluster Competition (SCC) at SC22, an international supercomputing conference that took place in Dallas, Tex. from Nov. 13 to 18. The SCC challenges student teams to build their own mini-supercomputer and use it to accomplish real science during the 48-hour competition, all while staying under a specified power limit.

The Purdue-IU team, known as 'INpack,' a reference to their shared Indiana connection and the HPC benchmarking software LINPACK, was advised by Erik Gough, lead computational scientist for the Rosen Center for Advanced Computing (RCAC) in collaboration with other RCAC and IU staff, including Betsy Hillery, director of high-performance computing for RCAC, and Winona Snapp-Childs, chief operating officer for the IU Pervasive Technology Institute.

"Our advisors have been monumental in helping us set realistic goals and timelines for ourselves," said team member Karl Oversteyns, a junior in unmanned aerial systems from Purdue.

Along with Oversteyns, this year's participants from Purdue are Larkin Nickle, a senior in computer and information technology, and Dhruv Sujatha, a sophomore in computer science and data science. The participants from IU are Lucas Snyder and Nrushad Joshi, both juniors in intelligent systems engineering, and Zachary Graber, a senior in computer science.

INpack spent months preparing for the competition in which they created a small cluster, learned about scientific applications, and implemented optimization strategies for their chosen architectures.

Since this is the first time Purdue and IU have collaborated to participate in SCC, the teams have been coordinating virtually and dividing the tasks systematically.

"We took on the hardware, and the IUers took on software," said Oversteyns. "The collaboration has been a good time. Naturally, resisting the urge to take playful jabs at one another is a challenge in itself."

"Students have been practicing compiling and running applications on Anvil and IU's Jetstream cloud resource, attending webinars on the benchmarks and applications for the competition," said Gough. "Three of the team members participated in an Anvil REU program this summer which helped them get familiar with HPC concepts."

The annual SC conference welcomes more than 10,000 attendees from all over the world. Purdue has fielded a team in the Student Cluster Competition almost every year since the competition's inception in 2007. Purdue SCC team alumni have gone onto jobs at premier technology companies, national labs, and academic HPC centers, including Purdue RCAC.

RCAC STUDENT PROGRAM EVENT GIVES SENIORS A CHANCE TO SHOWCASE PROJECTS

The Envision Center hosted the RCAC Student Program senior presentation event on April 24, 2023. The event was a success and gave the program's graduating seniors a chance to showcase their hard work and accomplishments.

Attendees were able to hear from their peers about their projects, which they had been working on during their tenure at Purdue's Rosen Center for Advanced Computing (RCAC). The event provided a platform for the students to present their projects and network with each other.

"The event went well and the students seemed to enjoy being able to easily hear what their peers worked on and learned," said Amanda Warren-Glowe, program manager at the Envision Center.

Hritik Pratik Trivedi, a graduate student in computer graphics technology, spoke about his experience presenting his Unity development work: "I had a lot of fun attending the lightning talks event held at the Envision Center. All the participants had very interesting projects and it gave me a new perspective on the work being done in RCAC," said Trivedi.

"Participating in the lightning talks was a fun way to summarize the work I'd done over the semester. I enjoyed presenting, and it was a nice opportunity to share what I'd done with my coworkers," said Kate Koury, a senior in computer graphics technology.

Koury also mentioned the benefit of being able to see what other students in RCAC are working on. "I rarely get an opportunity to interact with the other students on the RCAC side of things, so this was a fantastic opportunity to see what they were doing," says Koury. "I enjoyed engaging with them and further connecting with others in my department."

"Participating in the lightning talks was a fun way to summarize the work I'd done over the semester. I enjoyed presenting, and it was a nice opportunity to share what I'd done with my coworkers," said Kate Koury, a senior in computer graphics technology.

RCAC WELCOMES NEW TEAM MEMBERS

Paul Branham, Industry Business Development Director

Paul Branham joined RCAC this year as the Industry Business Development Director. Paul is a proud two-time Purdue alumnus who earned a bachelor's degree in nuclear engineering in 2012 and a master's degree in nuclear engineering in 2015.

Before joining Purdue, Paul was the president and CEO of transportation startup Reindeer Shuttle. He's also worked at Heartland Payment Systems as a relationship manager and at APL Cargo as their logistics director.

In his role as Industry Business Development Director, Paul is always on the lookout for opportunities to connect Purdue and RCAC with outside industry partners through meaningful relationships. This includes offering RCAC services for purchase, such as computing resources on Anvil or consulting with our experienced team, as well as corporate sponsorships for events through the year, and sponsored research agreements or other mutually beneficial opportunities. He can be reached at pbranham@purdue.edu.

David Liu, Quantum Computing

Prof. David Q. Liu recently joined RCAC as a faculty affiliate in the role of Senior Research Scientist in Quantum Information Science. He is leading RCAC's partnership efforts in the Quantum Collaborative (https://quantumcollaborative.org/) at Arizona State University.

David provides faculty and researchers with training workshops in quantum computing using IBM Qiskit and NVIDIA's cuQuantum. In addition, David is also developing quantum computing use cases through working with faculty, researchers, and industry partners. He is an IBM Qiskit Advocate and developing algorithms and applications in quantum machine learning, quantum optimization, and hybrid quantum-classical computing.

As a tenured associate professor of computer science, David has taught and conducted research in wireless sensor networks, cyber-physical systems, cybersecurity, Cloud Computing, IoT and Edge AI, machine learning and deep learning, and quantum computing. He can be reached at dqliu@purdue.edu.

Arman Pazouki, Director of Scientific Applications

Arman Pazouki joined RCAC this year as the Director of Scientific Applications. Prior to joining RCAC, Arman led the research computing group at Northwestern University, a career he embarked on after serving as an Assistant Professor at California State University, Los Angeles.

Arman's research prior to joining RCAC focused on multi- and many-body dynamics and relied heavily on modeling, simulation, and high-performance computing. More recently, he has been researching the research computing and data workforce.

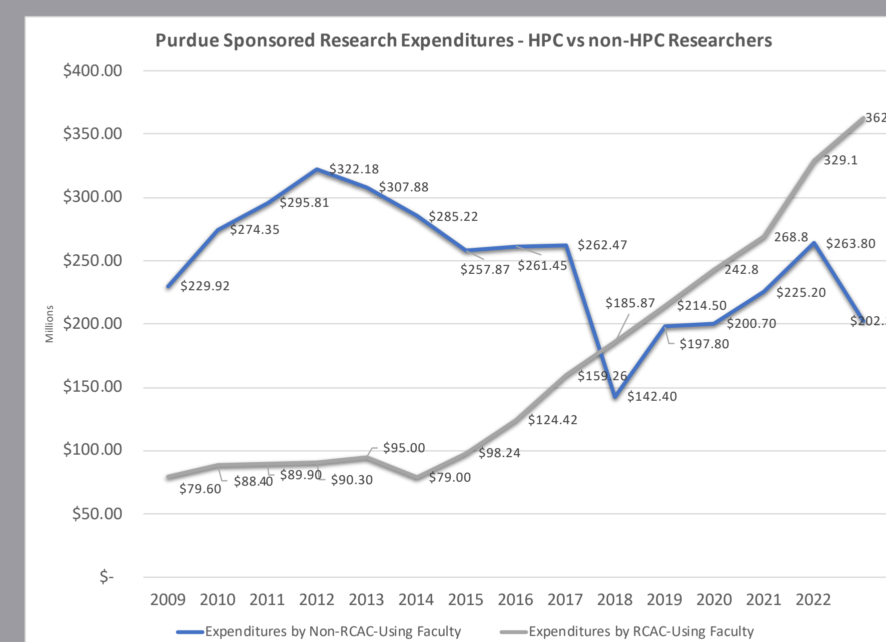
Arman has a doctorate in mechanical engineering from the University of Wisconsin-Madison and a bachelor's degree from the University of Tehran.

At RCAC, Arman leads the scientific applications group, which is accountable for researcher-facing functions of scientific computing, and responsible for the design and delivery of services to support researchers' use of computing algorithms, infrastructure, and data. He can be reached at apazouki@purdue.edu.

Awards: *RCAC-using faculty partners accounted for 60% of awards in FY23, for a total of $395M. (From 61%/$350.8M in FY22) This is a $45M increase in new awards to RCAC PIs.*

Expenditures: *RCAC-using faculty accounted for 64% of sponsor expenditures in FY23, a total of $362.7M. (Up from 56%/$329.1M in FY22) This is a $60M increase in expenditures by RCAC PIs.*

F&A: *RCAC-using faculty paid $77.5M of F&A to Purdue's general fund in FY23 – 88% of all F&A recovered. (Up from $69.9M/79% in FY22)*



RCAC DIRECTORS AWARDED NSF FUNDING FOR PEARC23 STUDENT PROGRAM

RCAC directors Laura Theademan and Betsy Hillery were the co-chairs of the PEARC23 student program and received NSF funding (award #2328478) "Conference: Practice and Experience in Advanced Research Computing (PEARC23)" to provide registration and lodging costs for 22 of the 57 students accepted into the program.

Students participating in the PEARC student program had the opportunity to contribute to the technical program by submitting a student paper and/or poster, allowing them to share their research efforts and gain feedback, insights and inspiration from like-minded experts at the conference. The program also aided the goal of improving diversity in the STEM workforce, with half the funded participants being women and more than half being minorities.

PURDUE SPONSORS INDYSCC COMPETITION AT '22 SUPERCOMPUTING CONFERENCE

Purdue's Rosen Center for Advanced Computing (RCAC) recently donated hardware infrastructure, including an entire cluster of nodes of the recently retired Rice cluster, and provided staff support to a supercomputing competition for students just getting started in the field of high performance computing (HPC).

IndySCC is a remote, cloud-based competition that's part of the larger Student Cluster Competition in which Purdue fielded a joint team with Indiana University this year. Now in its second year, the IndySCC seeks to minimize entry barriers for students interested in HPC, with the goals of education and inclusion for beginner-level teams. This year's event took place from Nov. 4 to 6, where the teams competed remotely utilizing provided hardware as part of an educational experience backed by HPC industry professionals.

Purdue was the only university to donate infrastructure to the IndySCC.

Several members of RCAC also contributed to the event. Amiya Maji, senior computational scientist at RCAC, was in-charge and judged the HPL (high performance LINPACK) benchmark. Lead research solutions engineer Patrick Finnegan was responsible for managing the entire infrastructure and plugging it into the Chameleon cloud test-bed the students used to access the Rice nodes. Lev Gorenstein, senior computational scientist, was secondary judge of a molecular dynamics application the students used, and also co-judged the poster session of the event.

"We had given students access to up to 300 nodes to run the HPL benchmarks which is pretty unique for any competition," says Maji. "I have not heard of any other cluster challenge competition that could give such large-scale resources to the competitors."

The teams participated in a final challenge that was split into two parts to test their newly acquired knowledge. Each of the 10 teams had 24 hours and a maximum of 300 nodes during the first phase to set up a cluster and run HPL across the nodes to achieve the highest HPL score. The teams then battled for two days, using their knowledge of scientific software applications, including a mystery application that they were only made aware of at the beginning of the competition. Teams were graded in a zoom session on their application expertise as well as their completion and accuracy rates.

"We are extremely grateful for Purdue's contributions to this year's competition," says Dan Dietz, the chair of student educational competitions for the Supercomputing conference. Dietz is an HPC engineer in the user assistance and outreach group at Oak Ridge National Laboratory, and a former RCAC staff member who reached out to RCAC about participating in the competition.

"The scale at which the teams were able to compete gave them a unique experience.", says Dietz.

RCAC CHIEF SCIENTIST CAROL SONG IS CO-PI IN $3M NSF FUTURE MANUFACTURING RESEARCH GRANT

RCAC is a partner in a $3M NSF Future Manufacturing Research Grant awarded to Shweta Singh, professor of agricultural and biological engineering. The award is titled "Cyber enabled transformation to circular supply chains (CSC) for sustainable pharmaceutical manufacturing networks." The overall goal of this project is to enable the design of zero-waste future pharmaceutical manufacturing networks at a macroscale using an integrative, multidisciplinary approach focused on novel process chemistry and separation methods, macroscale modeling of manufacturing networks to integrate new processes at scale, and pricing optimization for collaborative decision-making through dedicated cyberinfrastructure development.

Shweta Singh, associate professor of agricultural and biological engineering at Purdue University, will adapt a circular economy design computational tool that she co-invented for application to the pharmaceutical industry. (Purdue University photo/Tom Campbell)

RCAC chief scientist Carol Song is a Co-PI of this award. She leads the development of the cyberinfrastructure that aids in understanding the feedback loops between plant-scale automated manufacturing and macroscale manufacturing networks in making optimal decisions for adoption of recycling processes at scale and redesign of material flows towards CSCs. The work will build upon RCAC's past collaboration with Singh that developed PIOTHub, a collaborative cloud tool deployed on MyGeoHub (mygeohub.org) for generation of physical input-output tables (PIOT) using mechanistic engineering models. The automation and sharing features provided by PIOT-Hub will help to significantly reduce the time required to develop PIOT and improve the reproducibility/continuity of PIOT generation, thus allowing the study of the changing nature of material flows in regional economy.

PURDUE UNIVERSITY

Rosen Center for Advanced Computing

597.8 Million
CPU hours provided

1.1 Million
GPU hours provided

228
Community cluster PIs

2688
Community cluster users

60
Community cluster departments

43,690
Community cluster OOD sessions

27%
% of doctoral graduates using community clusters

28
RCAC Outreach events

55
Students supported by PRACTICE internships

64%
% of Purdue research expenditures to RCAC users

31
RCAC Training sessions

Sponsored research expenditures within RCAC
3.7 Million

RCAC provides access to leading-edge computational and data storage systems as well as expertise in a broad range of high-performance computing activities.

To learn more please contact us here **rcac-help@purdue.edu**

PURDUE UNIVERSITY | Rosen Center for Advanced Computing

101 Foundry Dr, suite 3314
West Lafayette, IN 47906-1394
rcac.purdue.edu