University of Arizona Research Computing 2012

Innovative
Multidisciplinary
Collaborative
“HPC systems are key to advancing the computational research being done in various science and engineering departments at the UA. I have routinely used the MPI platforms for the parallel processing of my molecular simulation studies both during my grad and postdoc years and now as a working professional.”

Vivek Kapila, PhD
School of Sustainable Engineered Systems | College of Engineering
Welcome to
The University of Arizona’s new
RESEARCH DATA CENTER
Leslie Tolbert  
Senior Vice President, Research | Graduate Studies & Economic Development

Research excellence at the University of Arizona depends on attracting the top minds in the world to our campus and supporting them with equally top-notch facilities. This is especially true for research computing where our investigators rely on access to state-of-the-art tools to solve society’s most pressing challenges and to create new knowledge and opportunity.

One way to ensure that we have the best research computing resources is to coordinate and support institutional investments for the campus research community. Whether in bioinformatics for advanced healthcare, evaluating environmental sustainability, or engineering new energy sources, our new strategic investments in High Performance Computing and High Throughput Computing (HPC/HTC) and storage will greatly advance the possibilities and impact of research at the UA.

These important new resources are the result of a successful collaboration between the Office of the Chief Information Officer, Office of the Senior Vice President for Research, and University Information Technology Services (UITS) Research Computing, Systems Administration and Data Center Operations. Valuable advice from the HPC Technology Refresh Advisory Committee (HPC-TRAC) enabled a successful partnership between all these groups. I am delighted that the new University of Arizona Research Computing Data Center will house centrally funded new HPC/HTC and research storage systems, and serve as colocation space for college and department research computing systems.
Recognizing the importance of advanced computing in enabling breakthrough discoveries and innovative research, the University of Arizona has completed the installation of a new Research Data Center. Thanks to a productive collaboration across campus, the new Research Data Center houses our next generation of High Performance Computing (HPC), High Throughput Computing (HTC) and high performance, high capacity storage systems and also provides colocation space for college and department research computing systems.

As a research University, it is critical to the mission of the UA that we maintain a competitive position among our peers. Research serves to interconnect the UA campus with our community and university partners. Research today has an increased emphasis on interdisciplinary concepts and on collaborations with industry. Our reliance on high performance computing and the need for sophisticated visualization, simulation and modeling software has far surpassed our expectations. Research computing is strategically important for the UA, is critical to the success of faculty research programs and is an important factor in faculty recruitment and retention. Through high performance computing and technological collaborations, we can increase achievement in research, scholarship and creative expression.
In January 2012, The University of Arizona’s research computing program opened the new Research Data Center (RDC). This new facility is a state-of-the-art computer hub housing five centrally funded research computers. It is the result of collaboration between the University’s Senior Vice President for Research, Leslie Tolbert; Chief Information Officer Michele Norin; and University Information Technology Services (UITS) Research Computing, Systems Administration and Data Center Operations. The new RDC is also home to research data storage systems, and serves as a space for college- and department-owned research computing systems. The RDC is monitored and maintained around the clock, offering a totally secured environment with redundant power and cooling, enabling scholars to focus on their primary research.

- UPS power and chilled water to support 20 water-cooled racks and 15 air-cooled racks
- Increased high density, power efficient disk storage, configured with 350TB (terabytes) of disks, expandable to 2PB (petabytes) in the same footprint
- New systems have approximately twice as many processors as previous UA systems, allowing greater throughput of research computations and more computing capacity
- A concentration of advanced knowledge and expertise of HPC/HTC computing and networks, data perceptualization and visualization personnel
- 10Gbps high speed network access to National LambdaRail (NLR) and Internet2
EQUIPMENT CAPACITY 35 Racks

EQUIPMENT COOLING Water-cooled equipment racks with constant recirculation from central chilled water plant featuring energy efficient thermal storage

ENVIRONMENT COOLING Two 35 ton Stultz Computer Room Air Conditioners (CRAC) maintaining the data center at a constant 72°F

POWER SUPPLY A transformerless 825kVA (1130 Amps) Uninterruptable Power Supply (UPS) with a .9 power factor with reduced operating costs and improved efficiency. In addition to a Transient Voltage Surge Suppressor in the switch gear, the UPS is a double-conversion design that completely isolates output power from all input power anomalies and delivers 100-percent conditioned, perfect sine-wave power-regulating both voltage and frequency

POWER BACKUP 1.75MW diesel generator power

POWER DISTRIBUTION Four high-efficiency 500kVA Power Distribution Units (PDUs) with Wavestar Branch Circuit Power Monitoring and overhead PowerWave power bus
The University of Arizona’s Research Data Center is the centerpiece of the University’s Research Computing services. Our extraordinary Research Computing resources provide a world-class facility enabling scholars to tackle complex computational queries in scientific, engineering, social, economic and cultural research. UA Research Computing includes High Performance Computing (HPC) and High Throughput Computing (HTC) systems, dense high capacity data storage, a 3D immersive visualization facility, and consulting services for HPC/HTC, visualization and statistics through UITS Research Computing Support (RCS) services. Combining all of these resources enables breakthrough results that would have been impossible only a decade ago.
MATCHING YOUR NEEDS  Research Computing Support (RCS) works closely with individual researchers and teams to determine each project’s requirements to design custom solutions encompassing hardware, software, automation of tasks and complete operational support. RCS excels in matching research needs with resources.

COMPETITIVE ADVANTAGE  The University of Arizona’s mature, comprehensive HPC/HTC capability creates a competitive advantage for the University when recruiting esteemed faculty and preferred students. With enhanced HPC/HTC cyberinfrastructure and related consulting services, the UA leverages research competitiveness and leadership in computationally intensive research areas.

GOVERNANCE  The governance structure of centrally-funded research computing resources is a collaboration of faculty, staff and University Information Technology Services, defining use policy for resource access and standard procedures. The policies and procedures are monitored and updated by the governance body while the resources are administered, maintained and supported by UITS Client and Infrastructure Services.
Research scholars in fields as diverse as neurobiology, applied physics, aerospace and mechanical engineering, environmental science and psychology depend on state-of-the-art computational systems to analyze, model or simulate massive datasets. HPC/HTC systems achieve enormous computing power by using hundreds of individual processors (cores) to break up huge computing projects into many smaller pieces. The Research Data Center houses three types of computer systems: Shared Memory, Distributed Memory and High Throughput.
WHAT ARE HPC, HTC, SHARED MEMORY AND DISTRIBUTED MEMORY?

- HPC (High Performance Computing) systems enable users to run a single instance of parallel software over many processors.

- HTC (High Throughput Computing) serial systems are more suited to running multiple independent instances of software on multiple processors at the same time.

- Shared Memory refers to a large block of random access memory (RAM) that can be accessed by many individual processors simultaneously.

- Distributed Memory refers to multiple-processor computer systems in which each processor has access to its own dedicated memory.

The University of Arizona differs from most universities by offering a unique “windfall” model that allows researchers and sponsored students to access the HPC/HTC systems during otherwise idle computing time. The windfall model encourages full utilization of all processors in the systems. All centrally funded systems are available to all users with equal priority at no cost. Funded research projects with high computing requirements generally utilize buy-in for memory and nodes, a benefit to projects that require a large amount of resources. Buy-in users have high priority access to their funded capacity to ensure low- or no-queue wait times.
<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Shared Memory</th>
<th>Distributed Memory</th>
<th>High Throughput</th>
<th>Shared Memory</th>
<th>Distributed Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGI</td>
<td>Altix UV</td>
<td>ICE 8400</td>
<td>IBM</td>
<td>Altix 4700</td>
<td>ICE 8200</td>
</tr>
<tr>
<td>RAM</td>
<td>2.7 TB</td>
<td>4.3 TB</td>
<td>3.0 TB</td>
<td>1.2 TB</td>
<td>2.8 TB</td>
</tr>
<tr>
<td>Processor Cores</td>
<td>928</td>
<td>2148</td>
<td>1128</td>
<td>616</td>
<td>1392</td>
</tr>
<tr>
<td>Core Type</td>
<td>Xeon</td>
<td>Xeon</td>
<td>Xeon</td>
<td>Itanium2</td>
<td>Xeon</td>
</tr>
</tbody>
</table>

DATA CENTER topology

Information accurate as of February 3, 2012
DATA STORAGE

Today’s research is generating datasets that are increasing exponentially in both complexity and size, making their analysis, archiving and sharing ever more challenging. The support for advanced techniques to transport, store, manipulate, visualize and interpret large datasets is critical to advancing modern science.

The new UA Research Data Center (RDC) offers 350TB (terabytes) of large-scale, redundant HPC/HTC disk storage, expandable to 2PB (petabytes) (2,048 TB). To put 2PB in perspective, the RDC could store the equivalent of 120,000 HD movies, never repeating for more than 26 years, or more than 500 million mp3 songs playing for nearly 4,000 years.

Our experienced Research Computing Support staff can help you identify the special storage requirements your research may need and can provide the appropriate resources. Some solutions will include free services available to the UA research community via central funding. Other research projects may require additional dedicated storage, which is available to buy-in and colocation users.
NETWORKING TECHNOLOGIES

UA’s high-speed networking technologies and services are essential to the performance of the Research Computing systems and their high impact for campus researchers. Using high-capacity 10 gigabit-per-second network connections to National LambdaRail (NLR) and Internet2, scholars’ massive data transfers travel seamlessly between local and remote systems to vastly reduce time-to-solution for users.

Optimized for research, it also provides the ability for multiple research and test nets to coexist on the same infrastructure.

STORAGE | Data Direct Network SFA 1000
300 terabytes expandable to two petabytes
“The ability to run parallelizable calculations on many processors simultaneously means that I can get my results back quickly, allowing me to use my time more efficiently.”

Ben Reynwar, PhD
Postdoctoral Research Associate
Department of Chemistry and Biochemistry
Miles beneath the ocean’s surface is a world as foreign as the deepest reaches of outer space. Unimaginable pressures and frigid temperatures abound. It is a dark, inhospitable, nutrient-poor place that counts as the last great unexplored ecosystem on the planet. This is the environment where University of Arizona’s Bonnie Hurwitz, a PhD candidate in the Sullivan Lab (The Tucson Marine Phage Lab), is discovering new viruses. Viruses are by far the most abundant life forms in the oceans and are the reservoir of most of the genetic diversity in the sea. The estimated 10^30 viruses in the ocean, if stretched end to end, would span farther than the nearest 60 galaxies. Working under the direction of UA faculty advisor Dr. Matt Sullivan, Hurwitz compares 1.8 million protein sequences of her discoveries to 40 million known viral protein sequences. This process takes an enormous amount of computing power through UA’s Research Computing Center.
“Recently I have been working on the viral protein universe and looking at viruses in the ocean. We take samples from the ocean looking at novel communities of viruses that have never been looked at before,” Hurwitz says. Some of her study sites include locations off the coast of Vancouver Island, Monterey Bay, Scripps Pier in San Diego, and the Great Barrier Reef. In the waters off of Vancouver Island, “oxygen is minimal in some areas because of farm runoff, which dumps nutrients into the ocean. That can cause vital plankton to bloom, which in turn causes what is called a ‘dead zone’ or ‘fish kills,’ where oxygen is consumed by these bacteria. So the focus of the study is to examine the viruses and their microbial counterparts in these regions of the ocean with low oxygen levels.”

To help analyze their data, Hurwitz and her team use Research Computing’s High Performance and High Throughput (HPC/HTC) systems. Primarily the HPC/HTC systems are used to check Hurwitz’s data against tens of millions of viruses already listed in public databases. This analysis uses tremendous computing power to help explain the function of the newly discovered viral proteins. With the power of the HPC/HTC systems, they look for protein domains in the sequences. A protein domain is a part of protein sequence and structure that can evolve, function and exist independently of the rest of the protein chain and are considered the most basic building blocks of life. Often Hurwitz’s protein samples have no matches in the databases.

“By looking for the domains, even though we do not know what the protein does, we generally know what the domain is involved in, so we can generate some hypotheses to what our protein is doing,” says Hurwitz. Although much is known about the specific viruses that impact commercial fisheries and at-risk mammal populations, little is known about their natural reservoirs. However, environmental genomic approaches like the work of Hurwitz and her team are providing insights into the enormous genetic diversity of viruses in the sea, and hold promise for revealing the sources of these pathogens.
Greg Byrne, University of Arizona Director of Athletics (right), and Scott Shake, Senior Associate Athletics Director for Development (left), explore a three-dimensional, virtual reality pre-construction model of Arizona Stadium's new north end-zone facility.
To support the world-class research being performed on our high performance computing systems, Research Computing Support provides advanced visualization resources and consulting services. Scholars often require innovative ways to display and share complex research results and simulations in high resolution, visual forms from enormous datasets.

**IMMERSIVE ENVIRONMENTS**

Imagine walking through a human heart, floating through a complex molecule or visiting an ancient Roman spa. AZ-LIVE (Arizona Lab for Immersive Visualization Environments) is a physical space where users are immersed in three-dimensional (3D), computer generated worlds. The system combines 3D computer graphics, stereoscopic projection technology, acoustical tracking devices and four-channel audio to allow the users to walk around in a wide variety of virtual worlds.

**RESEARCH APPLICATIONS**

Breaking education and research barriers with 3D visualization technology enhances existing educational and research programs. A 3D visual computing environment reconstructs and simulates scientific, engineering, architectural and imaginary artistic spaces and allows researchers and scholars to interact and visualize their complex data in an interactive 3D environment.

**RESULTS—2D AND 3D IMAGES IN THE FOLLOWING FORMATS:**

- 3D Stereo Immersive Environment
- 2D DVD
- 3D Blu-ray DVD
- RealD© stereoscopic 3D video technologies
- High resolution still images suitable for slides or printing
Spiders—the word alone can trigger automatic anxiety in some people. Arachnophobia is one of the most common phobias with reactions ranging from mild (spider fearful) to severe (phobia).

Sarah Burger, a graduate student with the College of Psychology, working under the direction of her advisor, W. Jake Jacobs, PhD, is using the Arizona Laboratory for Immersive Visualization Environments (AZ-LIVE) to measure the reactions of study participants when confronted with their fear in a virtual setting.
After a tour of the AZ-LIVE facility Burger realized the potential of visualization for testing and advancing her theories in both a live laboratory setting and in a virtual setting. She and her team collaborated with Marvin Landis, a scientific visualization specialist and consultant, to create a 3D environment complete with a lifelike animated tarantula. In the 3D setting, study participants are seated at the end of a virtual long table wearing special 3D glasses. At the opposite end sits the spider, which the participants guide closer toward themselves via a joystick. Researchers record the participants’ reactions as the spider approaches. Once the comfort level of the participant is no longer tolerable, they have the option to hit a panic button, which stops the spider from moving closer and suspends the simulation. The 3D visualization services at AZ-LIVE provide a virtual way for the participant to interact with the spider, and the simulation mirrors a real-life version of the experiment with an actual live spider.

Using the virtual world of AZ-LIVE for her research has opened more doors for Burger’s research. The next step is to focus on particular anatomical areas of the virtual spider to help learn which parts of the spider (e.g., mandibles, eyes, legs) trigger the most fear and phobia. She also wants to use the virtual spider to desensitize participants prior to interacting with a real spider. “If we have a better sense of what the experience is like in the virtual setting relative to a real setting and whether it translates to a real setting, we can see if that gives them any jump towards being ready to deal with a live tarantula,” Burger said.

Scientific Visualization consultant Marvin Landis is intrigued by the opportunity to expand the use of 3D visualization beyond the classic rendering of statistics and models for the hard sciences and into other arenas.
“Most of the users like me only know our own algorithms and software and do not know setup, environments and libraries of the Research Computer systems. The help I receive from RC’s resident consultants and experts is extremely useful and helps make my work more efficient.”

Jialun Li, PhD

Research Computing Support services provides expert consultation to scholars in the areas of HPC/HTC use, visualization, and statistical research and analysis. Consulting services are provided to help researchers utilize research computing resources more successfully.

HPC/HTC CONSULTING

Expert guidance means that hardware will be used to its fullest potential. HPC/HTC computational staff is available to assist users with HPC/HTC computer services. This assistance may take the form of guidance or direction when determining the best methods of evaluating and recommending high-end computing technology and software. This occurs usually after an initial assessment that solves issues using HPC/HTC services (consulting) or more general sustained assistance to a project HPC/HTC services (collaboration).

Research Computing Support consulting expertise includes:

- Multicore/multiprocessor programming
- Individual code clinics
- Scaling issues and analysis of parallel codes
- Optimization of serial codes
- Use of efficient serial and/or parallel algorithms
- Porting parallel code on different platforms
- Debugging
VISUALIZATION CONSULTING

Visualization Consulting is a service that allows researchers to create graphical representations of data and simulations. These computer-generated images and animations allow researchers to visually analyze their data and explore infinite results while changing parameters. Services include consulting, training, demonstrations and equipment for high-resolution visualization output and explorations.

STATISTICAL CONSULTING

In research, statistics is a central consideration, from planning to interpretation of results. Research Computing Support services expert statisticians provide assistance with statistical analysis and software to researchers in any discipline. They consult on experimental design, data analysis, interpretation and plotting of results, software selection, programming, biostatistics, design and analysis of questionnaires, and creation of online surveys. Researchers use Statistical Consulting services when preparing a grant proposal, thesis or dissertation; designing an experiment or survey; selecting software; writing result sections of papers; or responding to reviews.
PARTNERSHIPS

BUY-IN PARTNERS

The University of Arizona’s Research Computing environment is a mix of HPC/HTC resources including shared memory, distributed memory and high throughput computer systems, and a high capacity data storage system. These systems are designed to easily grow with the campus needs. The base systems are purchased and supported with central funding.

Buy-in permits research groups to add resources such as processors, memory, storage, etc. to the base HPC/HTC systems. Buy-in research groups have highest priority on the resources they add to the system.

The concept of the base system plus expansion model allows research groups to take advantage of central machine room space, maintenance, administration and redundant power and cooling of the computing systems while maintaining highest priority access for the resources they purchase. An additional benefit to buy-in research group members is that expansion resources, when integrated into the base systems, can be used in conjunction with base resources. This allows computational projects that are well beyond the capacity of the group’s sole resources if the new equipment were configured as a stand-alone independent system.
The University research computing community benefits from expansions to the base systems by having additional computing resources available. During time when the expansion resources are not fully utilized by the Buy-in groups, the expanded capacity is made available to all users. Not only is this an opportunity for students and other researchers who will benefit from acquiring data through the HPC, but also keeps the HPC/HTC systems running 100% of the time. This joint operation of base and Buy-in expansion resources is designed to maintain economical and full use of all HPC/HTC resources.

HPC Consulting will assist research groups in configuring Buy-in resources and options to meet the specific requirements of the group. Consultants will request detailed cost quotes from the appropriate vendors for researchers. Buy-in resource configurations will be approved and funded by the Buy-in group, and ordered, installed, administered and maintained by University Information Technology Services staff.

COLOCATION

Colocation space in the Research Data Center (RDC) is available for research computing systems purchased and administered by colleges, departments and research projects. The policies and procedures for requesting and allocating RDC space, power, cooling, access and networking are determined by Research IT governance group(s) designated by the Research Computing Steering Committee.

Colocation is provided without annual fees. Initial installation costs associated with connecting colocation equipment to the power, network, and cooling systems are the responsibility of the research group.
You say tomato, I say tomahto. Research on this pulpy edible fruit, also known as *Solanum lycopersicum*, is the topic of research for Mark Kroggel, research specialist for the School of Plant Sciences. Kroggel has been utilizing Research Computing Support (RCS) Statistical Consulting services as a resource to create predictive models about tomato plant growth. In this research, Kroggel’s primary role is to produce tomatoes and analyze the data as a member of Dr. Chieri Kubota’s research team, who works in collaboration with Dr. Cynthia Thomson in the Department of Nutrition.
The research team is investigating the science of tomato growth. “We’re always interested in how the environment is affecting the plants,” Kroggel said. That is where the statistical consulting staff became a valuable resource. Kroggel wanted predictive models of how greenhouse-grown tomato plants would respond in different climates, from our Sonoran desert to more humid and temperate climates. For that, he needed data processed through statistical software. “There’s a lot of analysis that I do know how to do, but at some point I don’t.” Kroggel said. For some of the more intricate statistical data, he turned to Dr. Mohammad Torabi, a statistical consultant on the RCS staff.

Kroggel works closely with Torabi to acquire, process, archive, analyze, model, visualize, simulate and disseminate complex data to close the data-to-knowledge gap across multiple time and length scales. “We’ll work together, and make sure that he understands the data, and that I understand what he’s doing with the data, and then he can just generate predictive models and hand it to me,” Kroggel recalls.

Sound simple? Torabi states that the process is a little more involved. When Kroggel first approached him for this project, Torabi helped design the research. Torabi is actively involved in helping researchers design projects to be more efficient, working one-on-one with the researcher. Even if Kroggel does much of the work, the statistical analysis model built by Torabi is what produces the final results.

This isn’t the first time that Kroggel has worked with Torabi. Kroggel and graduate students from Kubota’s team in Plant Sciences have worked with Torabi for five years. The consulting resource has allowed Kroggel and others to analyze their data in a way that is not necessarily part of their expertise.
University of Arizona Colleges and Departments using RESEARCH COMPUTING SERVICES

**COLLEGE OF AGRICULTURE AND LIFE SCIENCES**
- Agricultural and Biosystems Engineering
- Environmental Research Lab
- School of Natural Resources and the Environment
- Soil, Water and Environmental Science
- Veterinary Sciences and Microbiology

**ELLER COLLEGE OF MANAGEMENT**
- Economics
- Finance
- Management Information Systems

**COLLEGE OF ENGINEERING**
- Aerospace and Mechanical Engineering
- Biomedical Engineering
- Chemical and Environmental Engineering
- Civil Engineering and Engineering Mechanics
- Electrical and Computer Engineering
- Materials Science and Engineering
- Systems and Industrial Engineering

**COLLEGE OF MEDICINE | PHOENIX**
- NanoBioscience and Medicine

**COLLEGE OF MEDICINE | TUCSON**
- Arizona Respiratory Center
- Biochemistry and Molecular Biophysics
- Cellular and Molecular Medicine
- Immunobiology
- Medicine
- Physiology
- Radiation Oncology
- Radiology

**COLLEGE OF OPTICAL SCIENCES**
- Optical Sciences

**COLLEGE OF PHARMACY**
- Pharmacology and Toxicology

**COLLEGE OF PUBLIC HEALTH**
- Epidemiology and Biostatistics

**RESEARCH DIVISION**
- Arizona Research Labs
- BIO5 Institute
COLLEGE OF SCIENCE
Astronomy
Atmospheric Sciences
Chemistry and Biochemistry
Computer Science
Ecology and Evolutionary Biology
Geosciences
Hydrology and Water Resources
Institute Atmospheric Physics
Lunar and Planetary Laboratory
Mathematics
Molecular and Cellular Biology
Neuroscience
Physics
Psychology
School of Information: Science, Technology and Arts
Steward Observatory

COLLEGE OF SOCIAL AND BEHAVIORAL SCIENCES
Linguistics
Sociology

Research Computing at the University of Arizona

Contact us:
Visit rc.arizona.edu
Email rc@list.arizona.edu