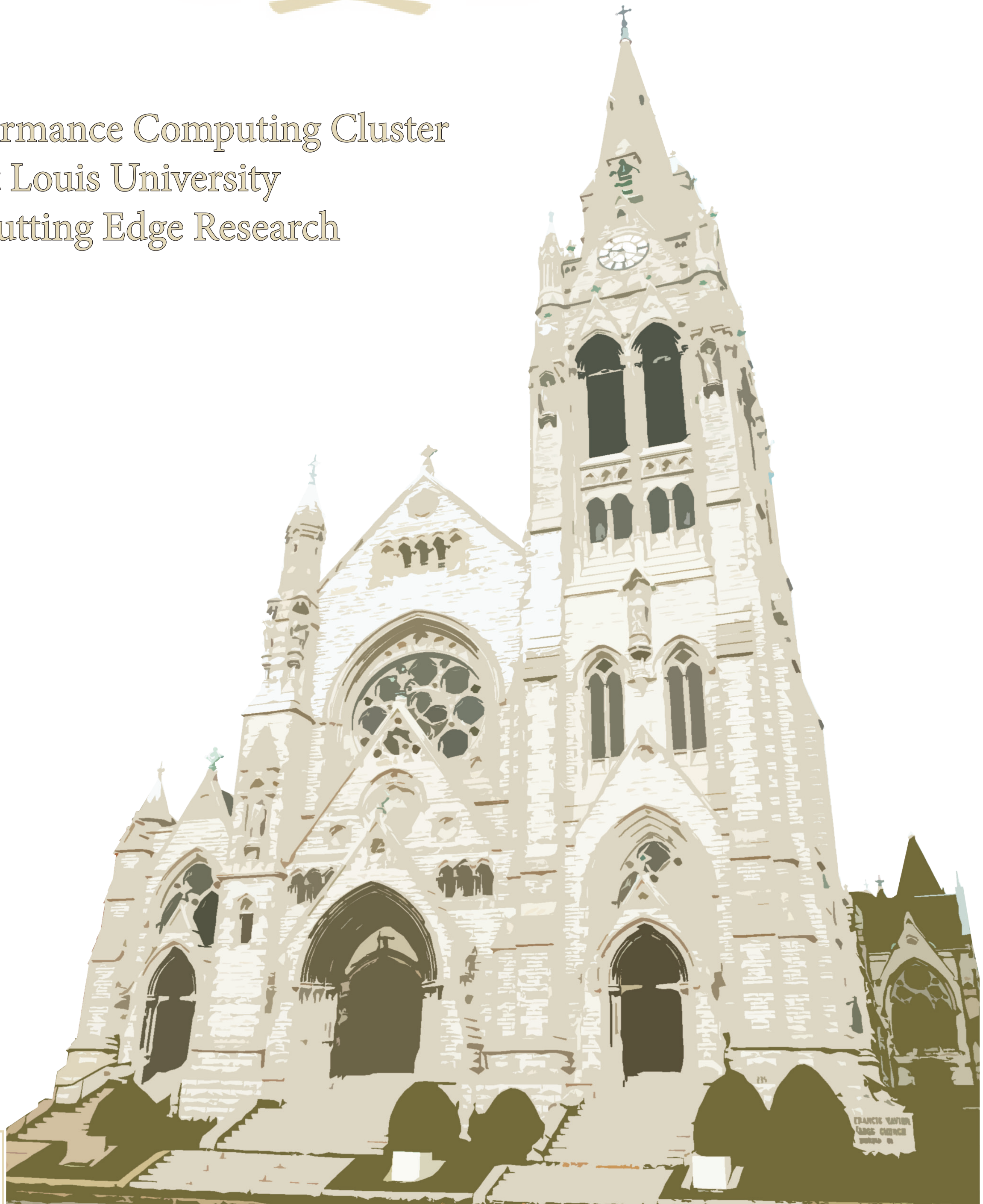
WINNER
SAINT LOUIS UNIVERSITY
RESEARCH CLUSTER GRANT COMPETITION

High-Performance Computing Cluster
Helps Saint Louis University
Conduct Cutting Edge Research



SILICON MECHANICS
RACKMOUNT SERVERS • STORAGE • HPC

University wins highly competitive research grant program with a unique interdisciplinary approach

By Art Mann, Silicon Mechanics

Most researchers would agree that research ideas are abundant and compute resources are always too scarce. At Saint Louis University (SLU), of St. Louis, Missouri, this bane of academic life is being addressed with the assistance of a complete high-performance computing (HPC) cluster, awarded as part of a highly competitive research grant program.

The cluster will help SLU researchers tackle a veritable A to Z of important academic research. Departments sharing the high-performance computing cluster include Aerospace and Mechanical Engineering, Biology, Center for Digital Theology, Chemistry, Sociology (jointly with Earth and Atmospheric Sciences, Center for Sustainability, and School of Public Health), and Political Science, as well as the John Cook School of Business. **Table 1** summarizes the teams that will benefit from the cluster and their research goals.

The compute cluster was developed by Silicon Mechanics, and originally used in 2011 by Boston University to compete in the SC11 Student Cluster Competition. Silicon Mechanics later sponsored a grant program in which educational institutions competed for the cluster. SLU was awarded the grant from a pool of 190 applicants, made up of US and Canadian universities and research institutions.

The cluster includes hardware donated by Kingston Technologies, AMD, NVIDIA, QLogic, Supermicro, and Seagate. The software supplied by Bright Computing uses image-based provisioning, which is important for this type of shared solution, since different research applications have varying configuration requirements to run optimally.

High-performance cluster meets university's computing needs and research goals

Keith Hacke, SLU's interim CIO and Vice President of Information Technology Services, explains the university's approach to its grant application. "We knew that this type of hardware is not used by just one division, and a single department can't keep that much computing busy all the time. Different departments are using it at different times in the year; there are many varied workload schedules for research. So this gives us the ability to add a lot of computational power across multiple departments."

In addition, the cluster will help SLU reach its overarching, long-term goals for research. "One of our university goals is to be a Top-50 research institution, and having these kinds of computer systems is mandatory for a move like that." The new hardware will also be used to help update Saint Louis University's current HPC clusters.

Cluster to handle breadth of research data

The range of research being undertaken with the aid of the high-performance cluster is extremely wide.

For example, the **Aerospace and Mechanical Engineering Department** will use the cluster to study new parachute concepts that may enhance the ability to accurately deliver cargo by airdrop. The team's goals are to study concepts that may offer the potential for more

Table 1 – High-Performance Cluster Research Teams

Department	Research Goals	Principal Reseachers
Aerospace and Mechanical Engineering Department, Parks College of Engineering, Aviation and Technology	Studying parachute concepts to enhance ability to accurately deliver cargo by airdrop.	Mark McQuilling, PhD
Biology	Expanding into the areas of climate change, modeling high elevation plant community ecology in the Andes Mountains.	Gerardo Camilo, PhD Mauricio Diazgranados, PhD (collaborator)
Center for Digital Theology, Department of Theological Studies	Processing large sets of digital images of pre-modern, hand-written, and unpublished manuscripts to support existing research in the field of paleography.	James R. Ginther, PhD Patrick Cuba
Chemistry Department	Studying carbon nanotubes and developing computational models to predict solubility in organic solvents.	Charles C. Kirkpatrick, PhD Michael Lewis, PhD
John Cook School of Business, Department of Operations and Information Technology Management (ITM)	Developing algorithms to solve large-scale stochastic dynamic programs with the goal of improving logistical decision making.	Justin Goodson, PhD
Political Science Department SLU School of Law	Evaluating Institutions, Behavior, and Outcomes in American State Supreme Courts.	Jason Windett, PhD Matthew Hall, PhD Christopher Witko, PhD Michael Wolff
Sociology & Criminal Justice School of Public Health Earth and Atmospheric Sciences Center for Sustainability	Mapping Risk and Resilience in St. Louis (MaRRS).	J.S. Onésimo Sandoval, Phd Kee-Hean Ong, PhD MPH Wasit Wulamu, PhD Sarah L. Coffin, PhD

drag for the same weight as existing parachutes, so they are more efficient at deceleration. They are also developing new parachute inflation models that can more accurately predict peak loading in the parachute (so they know it won't rip apart), parachute inflation time (so they know how long it takes), and trajectory (so they know where it will end up).

Funded largely by the US Army Natick Soldier Research, Development, and Engineering Center, the research is led by Dr. Mark McQuilling, who will be the primary HPC user, with collaboration from parachute inflation expert Dr. Jean Potvin, of SLU's Physics Department.

"One of the concepts being studied is an annular parachute, because annular geometries hold the potential for the highest achievable drag for non-gliding, hemispherical parachutes," said Dr. McQuilling. "Experiments show the possibility of instability, so computations are further studying the aerodynamics to better understand flow physics around annular geometry at various descent speeds."

Mark McQuilling

Other cluster computational fluid dynamics (CFD) research includes studying the aerodynamics of various payload geometries, where a better understanding of the flow around modern airdrop payloads will also help improve prediction of the airdropped system's trajectory.

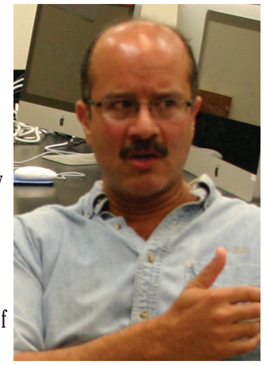
The sizes of geometries and flow speeds of interest require a high number of mesh elements to properly resolve the flow physics involved, and this requires a significant amount of computation time; thus, a compute cluster speeds up the simulation times and means the team can perform simulations not otherwise possible without the cluster.

The main parallel computing application to be used is SC/Tetra, a CFD application, along with the NASA Langley Stability and Transition Analysis Code (LASTRAC), a code used on a Unix/Linux system to analyze velocity profiles near a surface for instability characteristics. Previously the researchers had been using an 80-processor cluster funded by the US Army NSRDEC group, and this new cluster will allow them to expand their simulation capability to more airdrop system geometries.

"With increased simulation capacity, we'll also be able to simulate flow physics around gas turbine engine components and the hydrodynamics of marine mammals. Both situations involve simulations requiring a high number of mesh elements to properly simulate the physics involved," said Dr. McQuilling. Researchers from the Air Force Research Laboratory and the University of British Columbia are also involved with these projects.

Mike Lewis

Furthering research on climate change, the **Biology Department** will be using the high-performance cluster to model high-elevation plant community ecology in the Andes Mountains. Explains Dr. Gerardo Camilo, "High-performance computing will ultimately be used to validate the many projections made by the Intergovernmental Panel on Climate Change (IPCC) on what is likely to happen to this important area as a function of specific scenarios, or "story lines.""



Gerardo Camilo

SLU researchers conducted exhaustive research on plants in the Andes, gathered by viewing and documenting collections in South American universities, as well as trekking through the mountainous region.

PhD candidate Mauricio Diazgranados took more than 20,000 high resolution photos over one recent grueling 9-month trip through mountains from northern Venezuela to Columbia. The trip was made as part of his dissertation research on the study of the phylogenetic and biogeographic relationships of a group of plants in the páramos of South America that comprises nearly 145 species, most of them endangered by global warming and land use change.

Mauricio Diazgranados

Using all of the data collected, combined with long-term data from world climate observatories, researchers developed a climatic model showing plant distribution ranges (tolerances and optimum temperatures) for their area of interest. This was then spliced onto the actual topography, allowing development of real-world models of the local situation.

Now comes the difficult task of comparing the fates of every species in the local situational model among the many possible "story lines" suggested by global situation models developed around the world. The HPC cluster will be used to compare 100 different IPCC scenarios, including alternatives looking at the years 2020, 2040, and 2060 – about 2.7 terabytes of raw data! Future work will focus on additional global situational models based on IPCC's upcoming expected projections of a 4-degree minimum rise in world temperatures.

The **Chemistry Department** is using the HPC cluster for basic research that may advance cancer therapeutics using small interfering RNA (SiRNA) to silence genes by binding to messenger RNA and inhibiting its ability to produce certain proteins. The cluster will enable researchers to get their data more quickly than before and address more research questions. Other investigations include predicting the solubility of carbon nanotubes in organic solvents.

According to principal researchers Dr. Michael Lewis and Dr. Charles C. Kirkpatrick, one of the significant challenges with carbon nanotube applications is their prosperity to self-aggregate. "To improve their solubility we need to develop solvents that are more attracted to the nanotubes than they are to each other," said Dr. Lewis. Such solvents would be useful to researchers working on applications for carbon nanotubes.

The computational algorithm the team is using to find these solvents is much more efficient when implemented in a parallel environment.


Chuck Kirkpatrick

"We have realized efficiencies of a factor of two or greater in this way," said Dr. Kirkpatrick. "For small systems this may save only minutes in a calculation, but for large systems the difference is completing a calculation in one week rather than in two."

The team had been using SLU's existing high-performance computational cluster, and prior to that they used the supercomputer clusters available through NCSA (National Center for Supercomputing Applications). These systems offer traditional FORTRAN compilers for program design and development. With the new high-performance computing cluster, they can compile the code and run on the GPUs, which should enable them to further develop and exploit the parallel algorithms.

Researchers from the **John Cook School of Business** Department of Operations and Information Technology Management (ITM) are using the cluster to help develop algorithms to solve large scale decision problems, often referred to as stochastic dynamic programs, with the goal of improving logistical decision making.

As principal researcher Dr. Justin Goodson explains, a sequential decision problem requires a decision maker to make a series of decisions over time, often while facing uncertain future events. "Think of rebalancing an investment portfolio each month given uncertain asset returns," says Dr. Goodson.

"This research focuses on solving sequential decision problems in transportation and logistics. Specifically, I develop procedures to dynamically adjust vehicle delivery routes when customer demands are uncertain. As customer demands become known, I use that information to re-plan delivery routes in the most efficient manner."

Research outcomes typically include development of methods that perform better than some industry standard. Recently, Dr. Goodson's dynamic methods have been shown to improve on static fixed-route policies, a commonly employed routing methodology. The work is of great interest to third-party logistics providers, less-than-truckload (LTL) trucking companies, and vendor-managed distribution systems.

These advances in the methodology for solving stochastic dynamic programs are also applicable to companies facing sequential decision problems, like those that may arise in the management of supply and distribution networks, health care delivery, energy, and financial portfolios.

The volume of computation required to evaluate the effectiveness of an algorithm is very large. For example, recently completed calculations required more than six CPU years to evaluate whether the procedure delivered acceptable results. "Such a large amount of computing time just isn't practical on a desktop computer – this kind of research requires a cluster."


Justin Goodson

Adds Dr. Goodson, "Prior to the arrival of the Silicon Mechanics cluster, I was using SLU's HPC cluster. The main advantage of the new cluster from Silicon Mechanics is an increase in the number of cores available for computational experiments. This additional capacity will accelerate the pace of my research."

Dr. Goodson will be writing his own computer code, using the C++ programming language. He is collaborating with Jeffrey W. Ohlmann and Barrett W. Thomas at the University of Iowa's Tippie College of Business.

The HPC cluster is not just a tool for the hard sciences; it is also being used in the social sciences at SLU. Analyzing large data sets is difficult with traditional resources. By enabling large-scale parallel processing, the high-performance cluster will make data analysis quicker and more efficient. Perhaps as important, SLU researchers expect to improve the quality of their social science research by improving research methodology, as social sciences have typically not had much direct access to HPC clusters.

One fascinating social science project that will make use of the HPC cluster is jointly sponsored by the **Sociology & Criminal Justice Department**, in cooperation with **Earth and Atmospheric Sciences, Center for Sustainability**, and the **School of Public Health (Department of Environmental and Occupational Health)**. The project, known as Mapping Risk and Resilience in St. Louis (MaRRS), will look into the spatial aspects of social and economic, environmental and ecological phenomena. MaRRS is designed to give local government and community leaders scientific research, to work with them to identify strategies to rebuild the social, economic, and environmental systems.


Kee-Hean Ong

"This project is an exciting collaboration that is going to develop a model to help the City of St. Louis come up with a decision matrix on where to focus resources and how to use the laboratory to incorporate surveillance of diseases in the future," said Dr. Kee-Hean Ong, of the School of Public Health's Environmental & Occupational Health Department.

Sociology Professor J.S. Onésimo Sandoval puts the project into perspective by explaining that the depopula-

tion of St. Louis (from about 900,000 residents in 1950 to about 320,000 in 2010), has contributed to rapid changes in land use, shifting local and regional economies, and environmental degradation. The socio-environmental changes have strained the capacities of the local government to study and understand the complex relationship between the social, economic, and environmental systems.

Ness Sandoval

He says, "The challenges that face the city of St. Louis and the surrounding region require a new kind of scientific synthesis research." The goals of MaRRS, developed to meet this challenge, include advancing the frontiers of the scientific understanding of the complexity of social, economic, and environmental systems in St. Louis; working closely with policy makers and community leaders to anticipate and manage emerging social environmental challenges; and harness existing and emerging datasets and advanced computational tools to envision and understand possible future visions for the city of St. Louis. MaRRS will also help to build socio-environmental synthesis capacity among community leaders, students, and researchers; and create a synthesis that will be responsive and adapt to changing social, environmental, and economic systems.

"MaRRS will give us the opportunity to identify neighborhoods at risk and to envision possible futures and understand the actions required to access or avoid those futures," says Dr. Sandoval. "It will also give us the opportunity to identify neighborhoods that are able to produce and accumulate human, economic, symbolic, cultural, and physical capital, which fosters resilience."

Another example of the novel ways in which high-performance computing can be used comes from the **Center for Digital Theology**, which is using the HPC cluster to help process large sets of digital images of pre-modern, hand-written, and unpublished manuscripts to support existing research in the field of paleography.

The researchers are developing tools to help scholars analyze medieval European manuscripts, usually those with religious or theological meaning, written from about 500 AD to 1500 AD. About a year ago, they developed a tool called T-PEN (transcription for paleographical and editorial notation), a web-based tool for working with images of manuscripts, which allows transcriptions to be created, manipulated, and viewed in many ways. Users attach transcription data to the actual lines digital images of the original manuscript. This allows scholars to be able to combine the benefits of digitization with the expert scholarly eye needed to review and compare aspects of the manuscript they are interested in, like handwriting, script, or page layout.

Since scholars are looking at thousands of manuscripts, each of which can be as long as 200 to 600 pages, data sets grow exponentially, and the HPC cluster will let the CDT expand the tool in several possible ways. Scholars would be able to review the information in real time. A broader application would be to provide pre-processing for manuscript repositories now seeking to digitize their collections. This would allow them to provide immediate access to their manuscripts without endangering conservation by handling of the documents.

Patrick Cuba

Last, but definitely not least, is the **Political Science Department**, which is using the cluster in research that is evaluating institutions, behavior, and outcomes in American State Supreme Courts. Dr. Jason Windett is the principal researcher, along with Dr. Matthew Hall, and Dr. Christopher Witko, and Michael Wolff, of the SLU School of Law.

The research team is interested in how the judges' stated ideological position and method of appointment affects their decisions. The project received a substantial grant from SLU's President's Research Fund (PRF), which supports promising research projects that have strong potential to attract external grant funding.

Dr. Windett explains that the team recently successfully wrote computer code that allows it to automate the collection of data on state Supreme Court decisions around the country, including information on the voting behavior of state court judges over every single case they heard over every single year data was collected. They then needed to analyze the mountain of data.

Putting the size of the computing task into perspective, Dr. Windett explains that a similar computer analysis for the US Supreme Court, which hears about 65 to 90 cases per year, takes roughly 13 hours to run. By contrast, state courts hear thousands of cases per year and have more judges. During the first iteration of their analysis, before gaining access to the high-performance cluster, they ran 250 simulations on the state with the smallest docket – and that took 3 days on a personal computer.

James Ginther

The team was one of the first to exploit the new HPC cluster and has been using it extensively since day one, reporting that "what would have taken days can now be done in hours." The work involves the use of "R," a language and environment for statistical computing and graphics. The model looks at the judges' ideological position by stated ideology and how they were appointed or elected (partisan or non-partisan elections), and then analyzes all the votes taken over the studied time period. It then runs about 275,000 iterations to estimate groupings of voting behavior, arriving at a latent ideology score. The information has been run for 46

Jason Windett

out of 50 states.

According to Dr. Windett, the information will be used in three main areas of inquiry. First, it will further research on latent ideological preferences in America, which is not yet available for state Supreme Courts. Second, it will be used by those studying judicial reform, who will be able to evaluate if voting patterns differ based on how judges are selected, and what factors may influence polarization of courts. Lastly, judicial politics scholars will be able to review the cases heard and how they change over time, what opinions look like, and the impact of turnover and diversity on courts.

"I am convinced this is really going to be a large scale project with a wide applicability," said Dr. Windett. "After a brief mention in a journal of the preliminary results on ideology score data in 36 states, I sent the information to a LISTSERV and I've already received many inquiries from scholars looking to share data and jointly publish research."

According to Art Mann, education, research, and government account manager at Silicon Mechanics, who championed the grant program, "We had a variety of proposals, but Saint Louis University's stood apart, benefitting an incredibly wide range of academic pursuits."

The success of this first campaign, for both Saint Louis University and Silicon Mechanics, has inspired the start of a second annual research grant. Says Mann, "Seeing how positively this grant has affected the university, Silicon Mechanics will be awarding another HPC cluster grant in 2012. Research driven by today's educational institutions can change the world, and we want to be part of that change."

Art Mann is a software sales and business development professional with more than 25 years of experience serving the educational and research market. He has served as Silicon Mechanics' education/research /government vertical group manager for the past eight years, focusing on strategic planning and developing and growing new and installed base opportunities. Mr. Mann's long career supporting educational and research needs gives him a special expertise, and he recently initiated a series of technology-focused seminars for research faculty and their information technology support.

Art Mann
Education/Research/Government Vertical Group Manager at Silicon Mechanics

Prior to Silicon Mechanics, Mr. Mann founded Mann Technology Advisors, an information systems solutions company specializing in the higher education market, which was successfully sold to Denali Advanced Integration in 2003. He previously worked for Compaq Computer Corporation/Digital Equipment Corporation (now HP) for 17 years.

Mr. Mann served as a lecturer in the Math-Sciences Division and assistant dean of the Graduate School at Babson College, and as a lecturer at the University of Washington's Graduate School of Business. He has an MBA from the University of Massachusetts at Amherst, and a BS in Management from State University of New York at Buffalo. Among other accomplishments, he is also producer and creative consultant for Razzle Dazzle Kids, an educational television/video series for children.

Research Cluster Grant made possible by:

SILICON MECHANICS

RACKMOUNT SERVERS • STORAGE • HPC

and our partners:

QLOGIC AMD Kingston TECHNOLOGY NVIDIA

Seagate SUPERMICRO Bright Computing