

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 29-11-2022		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 10-May-2021 - 9-May-2022	
4. TITLE AND SUBTITLE Final Report: A Next Generation High-Performance Computing Cluster for Research and Teaching at a Primarily Undergraduate Hispanic Serving Institution			5a. CONTRACT NUMBER W911NF-21-1-0266		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 060122		
6. AUTHORS			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES CSU Fullerton Auxiliary Services Corporati 1121 N. State College Fullerton, CA 92831 -3014			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 78204-NC-REP.1		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Sam Behseta
UU	UU	UU	UU		19b. TELEPHONE NUMBER 657-278-8560

RPPR Final Report
as of 02-Dec-2022

Agency Code: 21XD

Proposal Number: 78204NCREP

Agreement Number: W911NF-21-1-0266

INVESTIGATOR(S):

Name: Andrew Scott Petit
Email: apetit@fullerton.edu
Phone Number: 6572785214
Principal: N

Name: Sam Behseta
Email: sbehseta@fullerton.edu
Phone Number: 6572788560
Principal: Y

Name: Willie Peng
Email: wpeng@fullerton.edu
Phone Number: 6572782866
Principal: N

Organization: **CSU Fullerton Auxiliary Services Corporation**

Address: 1121 N. State College, Fullerton, CA 928313014

Country: USA

DUNS Number: 106670755

EIN: 46-5292740

Report Date: 09-Aug-2022

Date Received: 29-Nov-2022

Final Report for Period Beginning 10-May-2021 and Ending 09-May-2022

Title: A Next Generation High-Performance Computing Cluster for Research and Teaching at a Primarily Undergraduate Hispanic Serving Institution

Begin Performance Period: 10-May-2021

End Performance Period: 09-May-2022

Report Term: 0-Other

Submitted By: Sam Behseta

Email: sbehseta@fullerton.edu

Phone: (657) 278-8560

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 0

STEM Participants:

Major Goals: Acquiring a new high performance computing cluster will significantly facilitate the interdisciplinary research and teaching of faculty across the University, particularly in the College of Natural Sciences and Mathematics for the next 10 years.

Accomplishments: Equipment was purchased.

Training Opportunities: We collaborated with a committee of CCAM faculty to develop policies for the Turing HPC cluster.

Results Dissemination: Nothing to Report

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

RPPR Final Report
as of 02-Dec-2022

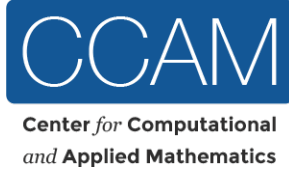
Partners

,

I certify that the information in the report is complete and accurate:

Signature: Sam Behseta

Signature Date: 11/29/22 10:03AM



Final Report

Proposal # **78204RTREP**

Contract # **W911NF2110266**

Project Title: **Next Generation High- Performance Computing Cluster for Research and Teaching at a Primarily Undergraduate Hispanic Serving Institution**

University Affiliation: **California State University, Fullerton**

Principal Investigator: **Dr. Sam Behseta,**
Department of Mathematics

Co-PIs: **Dr. Andrew Petit,**
Department of Chemistry

Mr. Willie Peng,
Information Technology

Technical Consultant: **Mr. Charles Yang,**
Information Technology

A New HPC Cluster Called Turing

An HPC cluster, named Turing, after Alan Turing the father of modern computing, was acquired by the Center for Computational and Applied Mathematics (CCAM) at the California State University, Fullerton (CSUF), through the DoD's contract # W911NF2110266.

In this report, we highlight the set of objectives, approaches, significance, accomplishments, and the timeline associated with this grant. It is our expectation that the full utilization of the system in the upcoming academic year, will result in the creation of a host of scientific publications and other products.

Objectives of the grant

- 1) This instrumentation grant for acquiring a new high performance computing cluster will significantly facilitate the interdisciplinary research and teaching of faculty across the University, particularly in the College of Natural Sciences and Mathematics for the next 10 years.
- 2) It provides much needed computational support for teaching and mentoring a significantly large proportion URM student body (43% Hispanic/Latinx, 58% women, 30% first-generation college students, 61% receive financial aid).
- 3) It will allow outreach to the North Orange County Community College District, and the Anaheim Union High School District (AUHSD) to educate thousands of predominantly Hispanic and underserved students to career paths involving computation and data science.

Approach

- 1) Providing unique computational facilities for the main beneficiaries of the cluster on the campus of CSUF, namely faculty and students, as well as creating new projects with CCAM's partners in the North Orange County community colleges, k-12 school districts, and industry.
- 2) Promoting HPC-based projects and coursework via workshops and curriculum and building a network with interested communities across Southern California. This includes defining new interdisciplinary projects, directed by the faculty members of CCAM.

Overcoming Scientific Barriers/Opportunities

- 1) The HPC vendor, Nor-Tech, worked through a number of challenges amidst the post-pandemic's supply chain issues, and the complexities of setting up a HPC cluster, remotely and on-site.

- 2) The CSUF team stayed on track for deploying the HPC cluster to Cal State Fullerton faculty and students. The first group of faculty, with up to 10 students per PI, embarked upon performing their scientific computation on the cluster.
- 3) Once deployed in full capacity, we will provide training resources for faculty and students to maximize the impact of the HPC cluster on campus
- 4) CCAM's Computation Committee, led by Andrew Petit, developed policies to ensure equitable use of HPC resources among stakeholders

Significance

- 1) The new HPC cluster will revolutionize the work of our underrepresented students, scholars, professors, and the entire research community. A significantly large number of research teams will benefit from this. In its early utilization, multiple groups of faculty and students have defined and performed projects via Turing.
- 2) The new HPC cluster will expose a significant number of predominately URM students to computational and data science through its use in teaching and outreach.

Accomplishments

- 1) The PI's negotiated a more capable HPC cluster with the vendor at no additional costs. The acquired cluster's architecture will allow for optimizing parallel computational operations.
- 2) The team at CSUF worked closely with the vendor to finalize the pieces of the HPC, and geared towards a successful installation, initial testing and set up, and its full utilization by the Spring 2023 semester.

Media Coverage of the DoD grant

The PI's have contributed to multiple news articles associated with the acquisition and utilization of Turing.

- [CSUF news coverage](#)
- Orange County Register Article (See Appendix 1)
- [HPC Wire news coverage](#)
- [Webpage and instructional videos](#)

Timeline for the Installation and Utilization of the HPC Cluster at Cal State Fullerton

October 8, 2021 – Hardware acquisition was completed by the vendor Nor-Tech.

October 22, 2021 – Basic software installation was completed by Nor-Tech and CSUF.

November 12, 2021 – Scientific applications installed by Nor-Tech and CSUF.

November 19, 2021 – Initial benchmarking was completed by Nor-Tech and CSUF faculty.

December 10, 2021 – Completion of shipping hardware from Nor-Tech to CSUF.

December 17, 2021 – Completion of hardware installation at CSUF.

January 7, 2022 – Completion of the second round of benchmarking by Nor-Tech, CCAM IT, and CSUF IT.

January 18, 2022 – HPC cluster made available to experienced Cal State Fullerton faculty and student researchers for evaluating performance, and hardware-software synchronization.

January 30, 2022 – Eight nodes were added to the system. These nodes had been acquired through a separate grant from NSF by CCAM member Michael Groves. HPC system was recalibrated to accommodate the expanded compute capacity.

February 25, 2022 – CCAM's Computation Committee, formed through representatives from Turing's stakeholders, began developing policies for the Turing HPC cluster.

March 1, 2022 – Completion of the first round of training resources, including instructional videos for submitting jobs to the HPC's queue for Cal State Fullerton faculty and student researchers.

March 4, 2022 – A Ribbon-Cutting event was held at CSUF with more than 60 guests in attendance. The PI's presented the technical specifications of Turing to the scientific community.

April 2, 2022 – Completion of all scientific software's tests and applications.

May 2, 2022 - A new senior Linux Admin position was advertised by the CSUF's IT, specifically for managing the operational components of Turing. The individual will contribute to the HPC's configuration, optimization, and utilization by the students and faculty.

May 10 - Early production of Turing was announced to the community. Faculty researchers were initially allowed to request accounts on Turing for up to two students. This has gradually increased and now each faculty is allowed to have up to 10 students on Turing.

Aug 12, 2022 – Streamlining students' onboarding process was completed using the application *ServiceNow* ticketing system. News users can now apply for an account on Turing via CCAM's online ticketing process.

Aug 23, 2022 – In collaboration with Nor-Tech, InfiniBand connectivity was optimized.

September 16, 2022 – Streamlining external campus user provision process was completed using the environment *Adobe Acrobat Sign*.

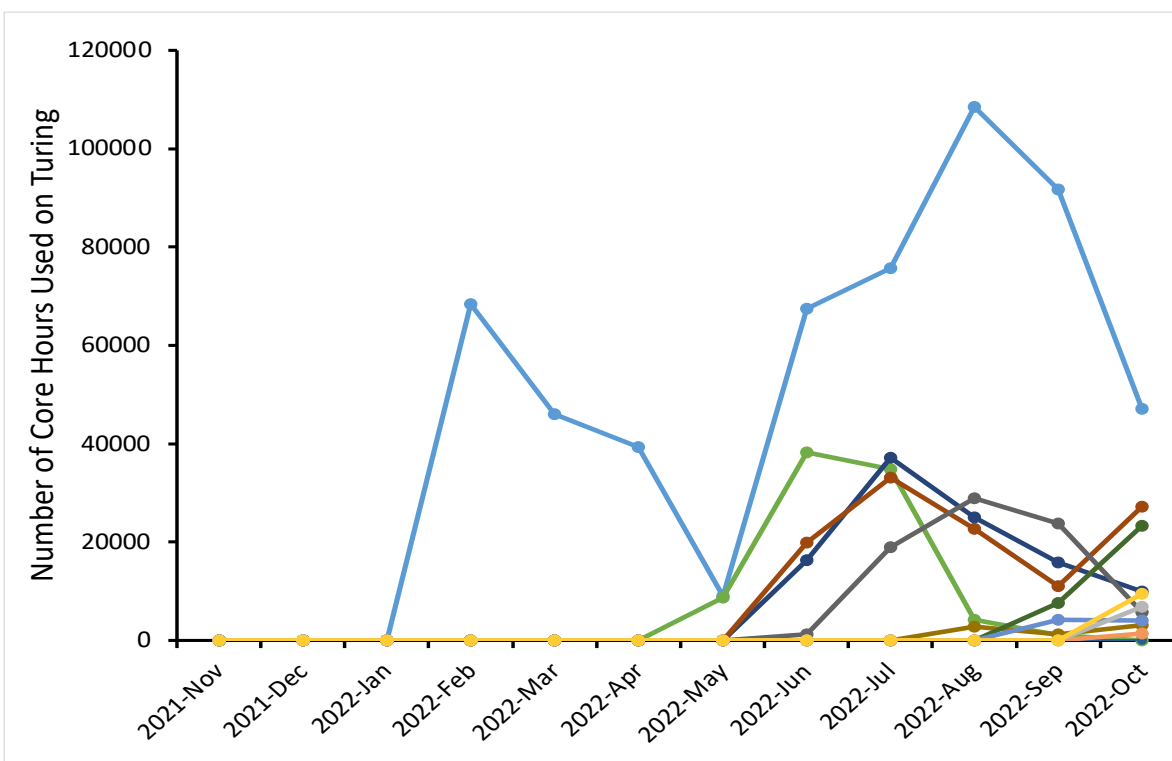
September 30, 2022 – An extensive workshop with engineering team at Nvidia was administered to acquire knowledge about the specifications and utilization of the new generation of the GPUs installed on Turing.

Nov 9, 2022- CSUF's CCAM and IT hiring committee finalized their selection of the Linux Admin position for Turing.

Nov 14, 2022 – CSUF IT Created and posted a student assistant position to help with the operations of Turing.

HPC Initial Utilization by Users

The chart below specifies the usage of the Turing in its experimental and preparatory phases. More users began performing jobs on the HPC, as more computational resources became available to them during the summer and fall months of 2022. A full utilization of the system is expected during the Spring of 2023 semester. In the figure below, the x-axis represents time (in month), the y-axis gives the total core hours used each month, and each color represents a different user.



Development and Administration of Policies for the Turing HPC Cluster

During Spring 2022, we collaborated with a committee of CCAM faculty to develop policies for the Turing HPC cluster. In developing these policies, we strived to promote efficient use of the HPC resources, equity among users, and future growth. Specific examples of notable policies include:

- 1) The SLURM scheduler, along with constraints on the maximum number of cores each user can simultaneously use, ensures a fair-share allocation of computational resources. Similarly, we developed a policy for fairly allocating the data storage space.
- 2) One compute node is reserved for a QOS which has a maximum wall time of 1 hour. This allows users with lower computational requirements, such as new users and users who

are debugging, to avoid having to wait in the queue for hours or days before their relatively small job runs. Based on feedback from CCAM faculty, this was a major source of frustration with our prior HPC cluster, Kepler.

- 3) We have established a shared-cost condo model which encourages faculty to acquire compute nodes for the Turing HPC cluster using external grants. In return, faculty receive priority access to their HPC equipment as well as benefit from the existing HPC infrastructure. Michael Groves has already purchased 8 CPU compute nodes which are incorporated into the Turing HPC cluster.
- 4) CCAM is developing a strategy to charge faculty with external grants a modest usage fee to support maintenance of the Turing HPC cluster. Based on our discussions in spring 2022, there is significant support among faculty users for this.

Examples of Research Being Performed on the Turing HPC Cluster

Co-PI Andrew Petit currently has 2 M.S. students and 5 undergraduate students regularly utilizing the Turing HPC cluster. These students are engaged in a variety of research projects that broadly involve computationally modeling light-driven chemistry (photochemistry). The quantum chemical calculations involved, which employ the Q-Chem and Orca packages, require parallelization across many CPU cores as well as significant memory.

Specific projects include:

- 1) Students Aerin Bridgers (M.S.) and Ken Jones (undergraduate) are working to understand pathways for energy exchange and chemical reaction when electronically excited NO collides with molecular partners that are relevant in combustion and atmospheric chemistry. The specific molecular partners of current interest are CH₄, CO₂, and C₂H₂. This work has the potential to provide new insights into the chemical physics of reactive, open-shell molecules like NO as well as atmospheric photochemistry.
- 2) Students Geovanny Gallardo (M.S.), Allyson Millan (undergraduate), and Emily Birlea (undergraduate) are working to understand the underlying mechanisms for a series of blue light-driven reactions between alcohols and diazo compounds. Such reactions have been demonstrated to have enormous potential for the synthesis of a wide variety of pharmaceutically and industrially important compounds. Part of this work involves reconciling controversy in the literature about how the diazo compounds fall apart after absorbing light, producing N₂ and reactive carbene intermediates. Other aspects involve understanding how changes to the molecular structure of the reactants affect the reaction outcome.
- 3) Students Mark Cruz and A.J. Kam (undergraduates) are working to develop fundamental structure-function relationships in photobases, molecules that become many orders of magnitude more reactive after absorbing light. Current work involves mapping out alternative light-driven pathways that compete with the desired photochemical reactions as well as investigating interactions between photobases and other Lewis acids such as metal ions and boron-containing molecules. Ultimately, this work has the potential to lead to the development of novel photocatalysts.

Dr. Michael Groves from the Chemistry Department currently has a team of 8 students working on Turing, including 3 M.S. students and 5 undergraduate students. Their work involves research is broadly related to surface chemistry and heterogenous catalysis. The required calculations, which utilize the GPAW software, benefit from parallelization across many CPU cores as well as GPU cores. Specific projects include:

- 1) Determining the role of oxygen defects in graphene on its ability to catalyze the formation of hydrogen peroxide. In doing so, the Groves group must sample a wide variety of relative placements of the oxygen defects.
- 2) Investigating the structure, electronic properties, and chemistry of two-dimensional boron-containing materials. Specific materials of interest involve borophene and boron nitride.
- 3) Identifying the mechanism for the oxidation of biofuels in fuel cells that utilize a palladium catalyst or a palladium alloy catalyst. The Groves group is especially interested in how changes to the molecular structure of the biofuel affect the distribution of oxidation products.
- 4) Determining how diamond-based electronics are affected by modifications to the surface of the diamond as well as the interface between the diamond and other insulators. Diamond is a wide-bandgap material with potential application in high-frequency and high-power electronics. Note that this work is in collaboration with the Army Research Laboratory.

Research Products Obtained Using the Turing HPC Cluster

Despite the fact that users only began utilizing the Turing HPC cluster in May 2022, results obtained with this cluster have already appeared in publications and presentations (CSUF students underlined):

Publications by Petit group including data generated on Turing:

- 1) Guardado, J.S.; Urquilla, J. A.; Kidwell, N.M.; ***Petit, A.S.** Reactive Quenching of NO ($A^2\Sigma^+$) with H₂O Leads to HONO: A Theoretical Analysis of the Reactive and Nonreactive Electronic Quenching Mechanisms, *Phys. Chem. Chem. Phys.*, **2022**, *24*, 26717-26730.

Presentations by Co-PI Petit including data generated on Turing:

- 1) *Unraveling the Mechanism of the Electronic Quenching of NO ($A^2\Sigma^+$) with Molecular Partners* (poster) 2022 Molecular Interactions and Dynamics Conference, Stonehille, MA
- 2) *Unraveling the Mechanism of the Electronic Quenching of NO ($A^2\Sigma^+$) with Molecular Partners* (talk) 2022 Spectroscopy and Dynamics on Multiple Potential Energy Surfaces, Telluride, CO
- 3) *Computational Development of Structure-Function Relationships in Photobases* (talk), 2022 Fall National American Chemical Society Meeting, Chicago, IL

Presentations by students in Petit group including data generated on Turing:

- 1) *Extending Structure-Function Relationships in Photobases to Curved N-Heterocycles* (talk), A.J. Kam, 2022 Fall National American Chemical Society Meeting, Chicago, IL
- 2) *Theoretical Study of the Electronic Quenching of NO ($A^2\Sigma^+$) with C_2H_2* (poster), Ken Jones, 2022 Fall National American Chemical Society Meeting, Chicago, IL
- 3) *Theoretical Study of the Electronic Quenching of NO ($A^2\Sigma^+$) with C_2H_2* (poster), Ken Jones, 2022 Informal Gathering on Atmospheric Science and Photochemistry, Irvine, CA

Plans to Expand User-Base for Turing HPC Cluster

CCAM is developing strategies to encourage a more expansive user-base for the Turing HPC cluster. Specific examples of our plans include:

- 1) Led by Co-PI Willie Peng, we are in the process of hiring a student assistant with IT to help with routine, day-to-day tasks associated with managing the cluster. The student assistant will also help to train new users of the Turing HPC cluster.
- 2) In Spring 2023, CCAM will host a series of 3 workshops that will introduce interested students and faculty to a) the foundations of neural networks and machine learning, b) accessing and using Linux HPC clusters, and c) using Python scripts to perform machine learning.
- 3) In Spring 2023, CCAM will continue to develop videos and training materials to help new users (both faculty and students) to access the Turing HPC, run calculations using it, and download the resulting data.
- 4) CCAM is actively seeking out external and internal programming funding to support undergraduate and M.S. student researchers who are engaged in computational research that makes use of the Turing HPC cluster.
- 5) Led by PI Sam Behseta, we will develop outreach workshops in spring 2023, with a plan to first run them with students from the North Orange County Community College District in summer 2023 or fall 2023.

Appendix 1: Orange County Register's News Coverage

ARMY AWARD

Supercomputer will broaden research, teaching

CSUF News Media Services

Cal State Fullerton researchers are getting a powerful computing tool to solve a wide range of large computational problems in science and mathematics – from understanding the spread of COVID-19 to preparing for future pandemics to developing advanced math models of crime and cybersecurity.

A \$583,900 grant from the U.S. Army is funding the acquisition of a new high-performance computing cluster to enable science and mathematics faculty and students to engage in leading-edge research activities.

The supercomputer, which will be part of the Center for Computational and

Applied Mathematics in the College of Natural Sciences and Mathematics, gives researchers the ability to process a large amount of data in a fraction of the time.

"This is groundbreaking because of the advanced technology it brings to campus, which will allow for the continued growth of leading computational research and teaching," said Sam Behseta, professor of mathematics and director of the Center for Computational and Applied Mathematics (CCAM).

"We were able to demonstrate to the Army that as a Hispanic-Serving Institution, we will efficiently utilize the cluster for furthering interdisciplinary research across STEM areas and better equip our students with the essential skills to join the workforce."

The HPC cluster is essentially a machine with multiple units, called servers or nodes, which can process large and complex computing tasks in a parallel fashion, Behseta explained.

The new machine is not the university's first high-performance computing cluster. In 2015, CCAM acquired its first cluster, known as Kepler, and the college's gravitational-wave scientists also use a HPC cluster for their research.

"This new HPC cluster is unique and very powerful and will significantly expand our cluster resources at CCAM and in the college," Behseta said.

The new HPC cluster will be housed in the university's Data Center managed by the Division of Information Technology.

Advanced solutions

The new supercomputer, named "Turing," after English mathematician Alan Turing for his contributions to computing and theoretical computer science, is expected to be operational by January 2022.

Behseta is directing the grant project in collaboration with Andrew Petit, assistant professor of chemistry and biochemistry; Emerio Martinez, information technology specialist and consultant for the College of Natural Sciences and Mathematics who contributed significantly to the design of the cluster; and Willie Peng, assistant vice president for information technology/infrastructure services in Information Technology. Peng, as director of the Data Center, provided the infrastructure needed for the operation of such a complex machine.

Petit noted that the new computer cluster will be a catalyst for faculty members to achieve their research agendas.

"My research group, for example, uses HPC to gain new insights into how the chemical properties of molecules are fundamentally transformed when they absorb light and become electronically excited," Petit said. "This informs the

development of new catalysts and dyes. The necessary calculations require a large amount of computing power, and as such, my research program will greatly benefit from the new HPC cluster."

Workforce ready

Chemistry graduate student José Luis Guardado '16 (B.S. biochemistry), who works under the mentorship of Petit, is excited about using such a powerful computing tool to solve some of the missing gaps needed to better understand his research on photochemical pathways.

"With this new HPC cluster, I will continue progressing in my research projects and making reliable predictions. This cluster will allow us to efficiently run our calculations and obtain results in a timely manner," Guardado said. "This experience will also help students to compete for jobs in industries that innovate our society by tapping into the symbiotic human/computer relationship."

Computational modeling, data analysis and machine learning are increasingly important aspects of modern science, both in academia and in industry. To keep up with the pace of scientific inquiry, Behseta and Petit emphasized that it is imperative that faculty and students have access to such computational tools that can respond to the research and workforce demands.

"The new HPC cluster not only meets the current demands for computational resources in the college, but facilitates future expansion," Petit said.

The grant project also will provide students with hands-on training and prepare them for careers in STEM (science, technology, engineering and mathematics) – and the growing demand for high-performance computing and big data professionals.

"We need HPC clusters to advance research and to expose students to the benefits and challenges of modern computing," Behseta said.

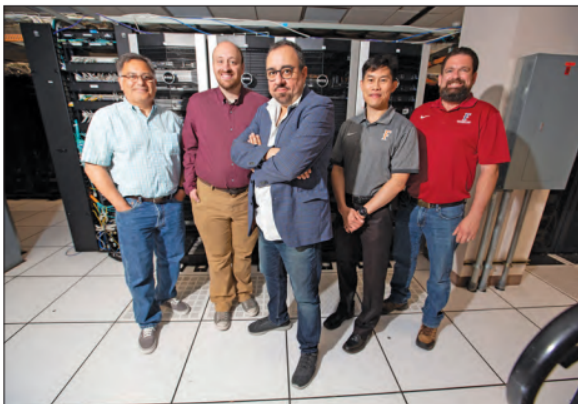


PHOTO COURTESY CSUF NEWS MEDIA SERVICES

From left, Emerio Martinez, Andrew Petit and Sam Behseta, along with Willie Peng and John King of CSUF's Data Center, are collaborating on the new high-performance computing cluster grant project.