THIRD INTERNATIONAL WORKSHOP SECURITY IN HIGH PERFORMANCE COMPUTING (S-HPC 2024)

Andrés Márquez on Behalf of the S-HPC Committee

SC24
Atlanta, GA | hpc creates.

ORGANIZATION

Workshop Chairs

Janine C. Bennett, Sandia National Laboratory
Bruno Raffin French Institute for Research in Computer Science and Automation (INRIA); Grenoble Alpes University, France

General Chairs

Nael B Abu-Ghazaleh University of California Riverside
Kevin J. Baker Pacific Northwest National Lab
Yang Guo NIST
Andrés Márquez Pacific Northwest National Lab
Sean Peisert Lawrence Berkeley National Lab

Technical Program Committee

Ryan Adamson (OakRidge National Laboratory)
Amro Awad (University of Oxford)
Purushotham Bangolore (University of Alabama)
Oceane Bel (Pacific Northwest National Laboratory)
Scott Campbell (Energy Science Network)
Clayton Hughes (Sandia National Laboratories)
Dan Kim (University of Queensland)
Erika Alexa Leal (Tulane University)
David McGee (Los Alamos National Laboratory)
Nicholas Multari (Pacific Northwest National Laboratory)
CJ Newburn (NVIDIA)

Organization Chairs

Program Chair: **Joseph B. Manzano (Pacific Northwest National Lab)**
Publication Chair: Yuede Ji (University of North Texas)
Publicity Chair: Sankha Dutta (Brookhaven National Lab)
Web Chair: Haipeng Cai (Washington State University

MOTIVATION
EXPLORING INTERSECTION OF HPC AND SECURITY

- HPC is rapidly permeating multiple areas of public and commercial interest
  - Deficiently secured HPC systems risk…
    - Large capital (hardware & software) and operational investments
    - Endanger large, valuable, vulnerable datasets with privacy/confidentiality expectations

- HPC systems are sufficiently unique to warrant individual security analysis
  - Molding the performance/power Pareto Curve yields unique HPC designs
  - Scale-out effects due to ‘security noise’ can have a large impact on performance/power

PROGRAM

Time	Event	Presenter
09:00 – 09:05	**Welcome**	Andres Marquez (PNNL)
09:05 – 10:00	**Distinguished Speaker:** HPC Cybersecurity in Emerging Computing Research, Development, and Prototyping Challenges	Robinson Pino (U.S. Department of Energy, Advance Scientific Computing Research))
10:00 – 10:30	**Break**	
10:30 – 11:00	**Invited Talk:** Secure HPC: Funding Opportunities from the National Science Foundation	Karen Karavanic (National Science Foundation)
11:00 – 11:30	**Research Paper:** Federated Single Sign-On and Zero Trust Co-design for AI and HPC Digital Research Infrastructures	Sadaf Alam (University of Bristol)
11:30 – 12:00	**Invited Talk:** HPC Security: Why The Time is Now?	Albert Reuther, Andrew Prout (MIT Lincoln Laboratory)
12:00 – 12:30	**Research Paper:** Using Malware Detection Techniques for HPC Application Classification	Thomas Jakobsche (University of Basel)
12:30 – 14:00	**Lunch**	
14:00 – 14:30	**Invited Talk:** The Unintended Effects of Privacy in Decision and Learning Talks	Ferdinando Fioretto (University of Virginia)
14:30 – 15:00	**Research Paper:** Security Testbed for Preempting Attacks against Supercomputing Infrastructure	Phuong Cao (University of Illinois Urbana-Champaign)
15:00 – 15:30	**Break**	
15:30 – 16:00	**Research Paper:** HPC with Enhanced User Separation	Andrew Prout (MIT Lincoln Laboratory)
16:00 – 17:25	**Panel:** HPC Security Guidance, Challenges, and Future Direction	Moderator: Yang Guo (NIST)
17:25 – 17:30	**Closing Remarks**	Kevin J Barker (PNNL)

DISTINGUISHED SPEAKER: DR. ROBINSON PINO

Title: HPC Cybersecurity in Emerging Computing Research, Development, and Prototyping Challenges

Abstract: The DOE Office of Science (SC) mission is to deliver the scientific discoveries and major scientific tools to transform our understanding of nature and advance the energy, economic, and national security of the United States. And the Advanced Scientific Computing Research program goals are delivering world leading computational and networking capabilities to extend the frontiers of science and technology. Currently, our research program has indicated interest in advancing research and development efforts focusing on three emerging areas namely energy efficient computing, analog computing, and neuromorphic computing. Therefore, it will be important to consider cybersecurity and integrity challenges as progress is achieved in these research areas as it involves coordination across potentially highly heterogenous, interoperating, and co-dependent components of future computing systems such as hardware, algorithms, system software, programming models, data management, and applications. This presentation will highlight potential basic research opportunities for emerging computing technologies and cybersecurity challenges for emerging high performance computing applications.

Bio: Dr. Robinson E. Pino is a Program Manager for the Advanced Scientific Computing Research (ASCR) program office in the U.S. Department of Energy's (DOE) Office of Science and prior Senior Advisor to the CHIPS Program Office in the U.S. Department of Commerce (DOC), National Institute of Standards and Technology (NIST). In his portfolio, Dr. Pino focuses on revolutionary basic research and development efforts for high performance computing (HPC), edge computing, neuromorphic computing, machine learning, artificial intelligence, photonics, microelectronics, advanced wireless, and applications that will enable our continued leadership through exascale and beyond computing and energy efficient technologies. Dr. Pino has expertise in technology research and development, program management, government, industry, and academia. He previously worked as Director of Cyber Research at ICF International advancing the state of the art in cybersecurity by applying autonomous concepts from computational intelligence and neuromorphic computing for the U.S. Department of Defense (DoD) Army Research Laboratory (ARL) and various collaborators in National Laboratories, industry and academia. In addition, Dr. Pino was a Senior Electronics Engineer at the U.S. Air Force Research Laboratory (AFRL) where he was a program manager and principal scientist for the computational intelligence and neuromorphic computing research efforts. He also worked at IBM as an Advisory Scientist/Engineer Development enabling advanced CMOS technologies and as a Business Analyst within IBM's photomask business unit where he was responsible for capacity planning, and manufacturing and development spending. Dr. Pino served as an adjunct professor at the University of Vermont where he taught electrical engineering courses. Dr. Pino has a Ph.D. and M.Sc. degrees in Electrical Engineering with honors from Rensselaer Polytechnic Institute and a B.E. in Electrical Engineering with honors, summa cum laude, from the City University of New York, City College. He is the recipient of numerous awards and professional distinctions; has published over 55 technical papers/reports, including four books; and holds nine patents.

INVITED TALK: DR. KAREN KARAVANIC

Title: Secure HPC: Funding Opportunities from the National Science Foundation

Abstract: In this talk we will provide a brief overview of the National Science Foundation followed by a more in-depth discussion of selected programs of interest to researchers at the intersections of security, privacy, performance, and high-end computing.

Bio: Karen Karavanic is currently serving as a Program Director at the National Science Foundation where she is interim co-lead of NSF's Secure and Trustworthy Cyberspace Program. She is a Professor of Computer Science at Portland State University in Oregon where she conducts research in performance and security of large-scale systems, and teaches classes in systems, accelerated computing, security, and performance. Dr. Karavanic holds a BA in Computer Science from New York University and an MS and Ph.D. in Computer Science from the University of Wisconsin - Madison.

INVITED TALK: DRS. ALBERT REUTHER, ANDREW PROUT

Title: HPC Security: Why the Time is Now?

Abstract: Supercomputing has been a discipline for at least four decades, but why has HPC security become such a hot topic the past several years? Just seven years ago the first HPC security papers were accepted and presented at SC17, yet efforts at NIST and elsewhere gained little traction. In this talk we explore some of the reasons why security of HPC systems has received so much more attention recently. We will discuss the expansion of scientific computing into new disciplines, changes in enterprise cybersecurity policy driving scientific computing away from general purpose devices to dedicated research computing assets, and the expansion of big data beyond what can be supported by single researcher workstations. We will also discuss how this expansion has increased the variety of codebases, languages, computational frameworks, and parallel computing models bringing new security challenges with them to the HPC space. We will explore why HPC is increasingly becoming a target due to its attractiveness to cybercriminals for cryptocurrency mining, the fact HPC centers host an increasing volume of non-public data, and how insider threat concerns are changing with the expanded userbase. Finally, we will look at why HPC security is different than enterprise security, discussing why existing security research and common practices are not automatically usable for HPC operators, and how the feedback and incentive loop for vendors is broken.

Bio: Albert Reuther is a Senior Technical Staff Member of the MIT Lincoln Laboratory Supercomputing Center (LLSC). In this role, he oversees the Computational Science and Engineering team of the LLSC, which works with users and research teams to most effectively use LLSC systems, software frameworks, and tools. He is also part of the leadership team overseeing the operations of the LLSC. His current areas of research involve interactive high-performance computing, machine learning, novel computer architectures, and graph analytics. He earned a dual BS degree in Computer and Electrical Engineering (1994), an MS in Electrical Engineering (1996), and a Ph.D. in Electrical and Computer Engineering (2000), all from Purdue University. He subsequently earned an MBA (2001) from the College des Ingrenieurs in Paris, France and Stuttgart, Germany.

Andrew Prout is a principal HPC systems engineer in the MIT Lincoln Laboratory Supercomputing Center. He developed the dynamic database management system, the dynamic virtual machine system, the dynamic web application portal, and user-based firewall technologies used to provide advanced supercomputing capabilities to Lincoln Laboratory staff. He has contributed to many open-source projects and is experienced with low-level systems and kernel programming. He holds a Certified Information Systems Security Professional certification and a BS degree in information technology security from Western Governors University.

INVITED TALK: DR FERDINANDO FIORETTO

Title: The Unintended Effects of Privacy in Decision and Learning Talks

Abstract: Differential Privacy has become the go-to approach for protecting sensitive information in data releases and learning tasks that are used for critical decision processes. For example, census data is used to allocate funds and distribute benefits, while several corporations use machine learning systems for financial predictions, hiring decisions, and more. While differential privacy provides strong guarantees, we will show that it may also induce biases and fairness issues in downstream decision processes. In this talk, we delve into the intersection of privacy, fairness, and decision processes, with a focus on understanding and addressing these fairness issues. We first provide an overview of Differential Privacy and its applications in data release and learning tasks. Next, we examine the societal impacts of privacy through a fairness lens and present a framework to illustrate what aspects of the private algorithms and/or data may be responsible for exacerbating unfairness. Finally, we propose a path to partially mitigate the observed fairness issues and discus challenges that require further exploration.

Bio: Ferdinando (Nando) Fioretto is an assistant professor of Computer Science at the University of Virginia. His research focuses on addressing foundational challenges to advance artificial intelligence, privacy, fairness, and the intersection between machine learning and optimization. His group focuses on two key questions: (1) How to endow discriminative and generative ML models the ability to comply with constraints, uphold physical principles, and adhere to safety standards, and (2) How to ensure that ML models and decision-making systems adhere to safety, privacy, and fairness principles. While the focus of his research is foundational, Nando's research is motivated by the application of ML in science and engineering, with applications to power systems, material science, policy optimization, and beyond. His work has been recognized with the 2022 Caspar Bowden PET award, the IJCAI-22 Early Career spotlight, the 2017 AI*AI Best AI dissertation award, and several best paper awards. Nando is also a recipient of the NSF CAREER award, the Google Research Scholar Award, the Amazon Research Award, the ISSNAF Mario Gerla Young Investigator Award, and the ACP Early Career Researcher Award in Constraint Programming. He is a board member of the Artificial Intelligence Journal (AIJ) and has been a member of the organizing committee of several workshops, tutorials, and events with focus on privacy, fairness, and optimization at premier AI and ML venues. He holds a dual PhD degree in Computer Science from the University of Udine and the New Mexico State University. Before joining the University of Virginia, Nando was an assistant professor at Syracuse University, a postdoctoral research associate at the Georgia Institute of Technology and a research fellow at the University of Michigan.

THANKS TO …

- Center for Advanced Technology Evaluation **(CENATE)** from Pacific Northwest National Laboratory for their support in this effort.
- Funded through the Department of Energy's Advance Advanced Scientific Computing Research
- Dedicated to better understand the applicability of novel architecture concepts to scientific discovery
- Provides a proving ground that is advancing scientific computing, artificial intelligence, machine learning, and cybersecurity through assessment and evaluation of advanced technologies and concepts
- Established in launched in 2015 with collaborations with University of Florida, University of California Riverside, Worchester Polytechnic Institute, North Carolina State University, among others.

- Memory has always been a second-class citizens
  - Concepts in memory research and tools (PIM / dataflow designs, memory profilers, memory centric runtimes, etc) have been "in the works" for decades
- Novel architectures has exposed the need orchestrate memory more carefully due to the challenges of heterogeneity: i.e., the limitations of current memory hierarchies
- A novel view is being taken across several research domain, especially exemplify in the PNNL's **AMAIS** project
  - It aims to provide end to end solution to memory challenges using a comprehensive approach to software, codesign hardware, and accelerators
- Currently collaborators with major industry partners with full support of the DOE ASCR office of science

Thanks also to the NIST **working group in High Performance Computing Security** for all their help and support in organizing this event: https://csrc.nist.gov/projects/high-performance-computing-security