Poster: Security and Privacy Heterogeneous Environment for Reproducible Experimentation (SPHERE)

Jelena Mirkovic*, David Balenson*, Brian Kocoloski*, David Choffnes†, Daniel Dubois†,
Geoff Lawler*, Chris Tran*, Joseph Barnes*, Yuri Pradkin*, Terry Benzel*,
Srivatsan Ravi*, Ganesh Sankaran*, Alba Regalado*, and Luis Garcia‡
* USC Information Sciences Institute, Email: mirkovic, balenson, bkocolos, glawler, ctran,
jdbarnes, yuri, tbenzel, sravi, sankara, alba@isi.edu
† Northeastern University, Email: choffnes@ccs.neu.edu, d.dubois@northeastern.edu
‡ University of Utah, Email: la.garcia@utah.edu

***Abstract*—To transform cybersecurity and privacy research into a highly integrated, community-wide effort, researchers need a common, rich, representative research infrastructure that meets the needs across all members of the research community, and facilitates reproducible science. To meet researcher needs, USC Information Sciences Institute and Northeastern University have been funded by the NSF mid-scale research infrastructure program to build Security and Privacy Heterogeneous Environment for Reproducible Experimentation (SPHERE). This research infrastructure will offer access to an unprecedented variety of user-configurable hardware, software, and network resources, it will offer six user portals geared toward different populations of users, and it will support reproducible research via a combination of infrastructure services and community engagement activities.**

I. Introduction

Cybersecurity and privacy threats increasingly impact our daily lives, our national infrastructures, and our industry. Recent newsworthy attacks targeted nationally important infrastructure, our government, our nuclear facilities, our researchers, and research facilities. The landscape of what needs to be protected and from what threats is continuously evolving: new technologies are released and the threat actors improve their own capabilities through experience and close collaboration. Meanwhile, defenders often work in isolation, using private data and facilities, and producing defenses that are quickly outpaced by new threats. To transform cybersecurity and privacy research into a highly integrated, community-wide effort, researchers need a common, rich, representative research infrastructure that meets the needs across all members of the research community, and facilitates reproducible science.

To meet researcher needs, USC Information Sciences Institute and Northeastern University have been funded by the NSF mid-scale research infrastructure program to build Security and Privacy Heterogeneous Environment for Reproducible Experimentation (SPHERE). This research infrastructure will offer access to an unprecedented variety of hardware, software, and other resources, all relevant to cybersecurity and privacy research, connected by user-configurable network substrate, and protected by a set of security policies uniquely aligned with cybersecurity and privacy research needs. SPHERE will offer six user portals, closely aligned with needs of different user groups, facilitating widespread adoption. It will provide built-in support for reproducibility, via easy experiment packaging, sharing, and reuse. SPHERE will build a process, a standard, and incentives for community-wide efforts to develop representative experimentation environments for cybersecurity and privacy research, and to continuously contribute high-quality research artifacts. You can learn more about SPHERE by visiting https://sphere-project.net.

II. Community Need

Over the past decade, and especially during the Covid-19 pandemic, both an individual's and society's essential functions (e.g., work, school, entertainment, social, financial, infrastructure, and governance) moved increasingly online. This sharply increased our nation's dependence on correct and reliable functioning of network and computing systems, and has led to increases in the frequency and impact of cybersecurity and privacy (CS&P) attacks. Recent years have seen unprecedented and record-breaking attacks, for example the Solar Winds supply-chain attack [5], which exposed confidential government data, and the Colonial Pipeline attack [8], which shut down our major gas pipeline for several days. Ransomware attacks more than tripled [6], DDoS attacks doubled [2], and data breaches increased by 70% [7]. Simply put, we now live in a world where cybersecurity and privacy are intrinsically intertwined with everything we do, and failures in these domains can have far-reaching monetary and national security impacts, and even jeopardize human lives. **Research progress in cybersecurity and privacy is thus of critical national importance**, to ensure safety of U.S. people, infrastructure and data.

USC Information Sciences Institute ran two workshops in 2022 to learn about community need around cybersecurity and privacy research: the *Cybersecurity Artifacts Workshop* [1] and the *Cybersecurity Experimentation of the Future 2022 Workshop* [4].

CS&P researchers need **common, rich, representative research infrastructure, which meets the needs across all members of the community, and facilitates reproducible science** to move from *piecemeal, opportunistic research* to *pursuing integrated, sophisticated, community-encompassing research*. We also need a well-educated workforce that is knowledgeable about cyber threats, and that has mastery over practical skills to prevent, detect and recover from cyber attacks.

III. SPHERE

We are building innovative, transformative research infrastructure (RI) for CS&P experimentation: SPHERE – Security and Privacy Heterogeneous Environment for Reproducible Experimentation. In this section we describe the architecture, services, and community-building activities we plan to undertake to transform CS&P research from piecemeal and opportunistic to highly integrated, community-wide effort that is sophisticated and reproducible.

The SPHERE research infrastructure will offer rich, abundant, and diverse hardware resources, which would meet the experimental needs of 90% of researchers today [3]. The devices we plan to purchase and integrate with SPHERE as *experimental nodes*, and the research that benefits from these are as follows: (1) **General compute nodes:** 48 from DeterLab, 144 new nodes, with Intel TDX, ARM CCA/TrustZone, and AMD SEV; **Research supported:** application, system and network security, measurement, human user studies, large-scale experiments, education, trustworthy computing; (2) **Machine learning nodes:** 10 GPU-equipped servers; **Research supported:** security with machine-learning in the loop; (3) **Cyber-physical nodes:** 15 Rockwell Automation ControlLogix PLCs, I/O modules; **Research supported:** critical infrastructure security; (4) **Embedded compute nodes:** 600 from DCOMP, 312 new (Intel Atom, Intel Xeon D, ARM Cortex-A57, and NVIDIA Jetson NX Volta GPUs); **Research supported:** edge computing security, blockchain security, private computing, trustworthy edge computing, federated learning; (5) **IoT nodes:** 500 IoT nodes (a variety of smart home, smart speaker, camera, doorbell, TV, appliance, medical, office, wearable, and miscellaneous devices); **Research supported:** IoT security, user privacy; and (6) **Programmable nodes:** 8 Tofino switches, 16 Xilinx Virtex-7 NetFPGA development boards (smartNICs); **Research supported:** dynamic (programmable) network security, SDN security. SPHERE will support most popular and relevant devices for CS&P research today. If CS&P research trends change in the future, new devices can be easily added by adding new installation and control scripts.

Many CS&P researchers study phenomena that interact closely with network topology, protocols and actors – SPHERE will meet the field's unique needs by offering a dedicated, user-configurable network substrate. CS&P experiments further may include generation of harmful traffic, taking live measurements from the real Internet, running human user studies, and even interacting with malicious Internet actors. To support these different research needs, and protect the Internet, SPHERE will provide safe network security policies.

A. Services and Community Building

All SPHERE nodes will be accessible via a single user interface. To meet the needs of various classes of users, SPHERE will provide six user portals: MAN (manual) - for exploratory research, JUP (Jupyter) – for mature research, GUI – for novice users, EDU – for use in education, AEC – for artifact evaluation committees, and HUM – for human user studies. Users will be able to access all portals from the user interface, and obtain a consistent view of their experiments, while being able to switch between portals as their needs evolve.

SPHERE hopes to serve not just as environment for experimentation but also for sharing and reuse of high-quality research artifacts, to promote integrated research in cybersecurity and privacy. SPHERE will facilitate reproducible science by building a streamlined process, standards, and incentives for the community to develop representative (realistic) experimentation environments and offering built-in infrastructure supports and community engagement process for artifact sharing and reuse. The SPHERE team will first engage with research and education communities in CS&P to learn about their experimentation needs and about needs around artifact sharing and reuse. SPHERE will further build infrastructure services that include extensive logging of user actions and support for various approaches to capture experiment topology, setup and workflow. In addition to these technological advances, SPHERE team will engage with artifact evaluation committees at conferences and journals to support artifact evaluation and hosting on SPHERE. Additionally, SPHERE will issue an open call for mature research artifacts to be deployed on SPHERE as representative experimentation environments.

This poster describes SPHERE[1] , a new research infrastructure for cybersecurity and privacy that will be built in the next four years by USC-ISI and Northeastern University. It is our hope that SPHERE will transform and propel CS&P research to new advances, by providing a common experimentation platform for the research community.

REFERENCES

[1] D. Balenson, J. Mirkovic, E. Eide, L. Tinnel, T. Benzel, D. Emmerich, and D. Johnson, "Cybersecurity artifacts workshop – report," https://bit.ly/CyberArtifactsWkshp2022, 2022.

[2] Government Technology, "Hacktivism and DDOS Attacks Rise Dramatically in 2022," https://www.govtech.com/blogs/lohrmann-on-cybersecurity/hacktivism-and-ddos-attacks-rise-dramatically-in-2022.

[3] J. Mirkovic, "Survey of Experimentation Approaches in Cybersecurity and Privacy Papers," https://bit.ly/CyberPapersSurvey2022, 2022.

[4] J. Mirkovic, D. Balenson, S. Ravi, L. Garcia, and T. Benzel, "Cybersecurity Experimentation Workshop – 2022 – Report," https://bit.ly/CyberExperWkshp2022, 2022.

[5] NPR, "A 'Worst Nightmare' Cyberattack: The Untold Story Of The SolarWinds Hack," https://www.npr.org/2021/04/16/985439655/a-worst-nightmare-cyberattack-the-untold-story-of-the-solarwinds-hack.

[6] Statista, "Annual number of ransomware attacks worldwide from 2016 to first half 2022," https://www.statista.com/statistics/494947/ransomware-attacks-per-year-worldwide/.

[7] Sumeet Wadhwani, Spiceworks, "Data Breaches Soared by 70% In Q3 2022 in an Otherwise Dull Year," https://www.spiceworks.com/it-security/data-security/news/data-breach-report/, 2022.

[8] TechTarget.com, "Colonial Pipeline hack explained: Everything you need to know," https://www.techtarget.com/whatis/feature/Colonial-Pipeline-hack-explained-Everything-you-need-to-know.

[1] SPHERE is based upon work supported by the National Science Foundation under award number 2330066. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Security and Privacy Heterogeneous Environment for Reproducible Experimentation

Jelena Mirkovic and Brian Kocoloski (USC-ISI), David Choffnes and Daniel Dubois (Northeastern University), Geoff Lawler, Christopher Tran,
Joe Barnes, Yuri Pradkin, Terry Benzel, David Balenson, Srivatsan Ravi, Ganesh Sankaran, and Alba Regalado (USC-ISI), Luis Garcia (University of Utah)

Societal Need

- Our nation depends on correct and reliable functioning of network and computing systems
- Frequency and impact of cybersecurity and privacy attacks are constantly increasing:
    - Solar Winds supply-chain attack, which exposed confidential government data
    - Colonial Pipeline attack, which shut down our major gas pipeline for several days.
    - Ransomware attacks more than tripled
    - DDoS attacks doubled
    - Data breaches increased by 70%
- **Research progress in cybersecurity and privacy is of critical national importance, to ensure safety of U.S. people, infrastructure and data.**

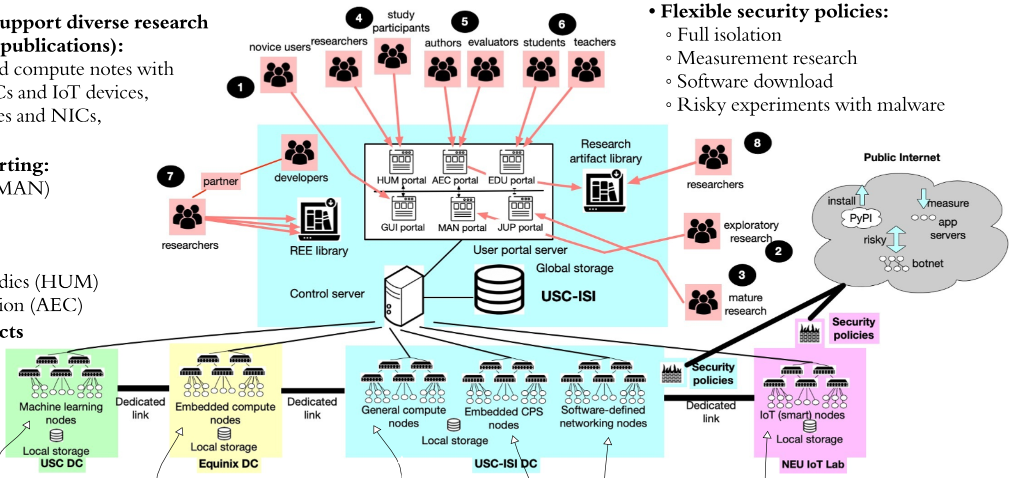
Research Need

The cybersecurity and privacy research community needs a common, rich, representative research infrastructure, which meets the needs across all members of the community, and facilitates reproducible science

- **Common, rich infrastructure**:
    - Security and privacy issues affect different technologies differently (e.g., different CPU architectures)
    - Some emerging technology can create new vulnerabilities (e.g., IoT)
    - New technologies can be used for defense (e.g., trusted hardware, SDN)
    - Infrastructure must have diverse hardware to meet wide research needs
- **Meet needs across all members of the community**:
    - Experienced and novice users, researchers and students
- **Facilitate reproducible science**:
    - Help researchers create, share, and reuse research artifacts

SPHERE Research Infrastructure

- **Diverse hardware to support diverse research needs (85% of today's publications):**
    - General and embedded compute notes with trusted hardware, PLCs and IoT devices, programmable switches and NICs, GPU-equipped nodes
- **Six user portals supporting:**
    - Exploratory research (MAN)
    - Novice users (GUI)
    - Mature research (JUP)
    - Use in classes (EDU)
    - Use in human user studies (HUM)
    - Use for artifact evaluation (AEC)
- **Libraries of artifacts**
    - REEs and other artifacts
    - Easy reuse on SPHERE



- **Flexible security policies:**
    - Full isolation
    - Measurement research
    - Software download
    - Risky experiments with malware

10 GPU-equipped servers
Research supported: security with machine-learning in the loop

600 from DCOMP, 312 new (Intel Atom, Intel Xeon D, ARM Cortex-A57, and NVIDIA Jetson NX Volta GPUs)
Research supported: edge computing security, blockchain security, private computing, trustworthy edge computing, federated learning

48 from DeterLab, 144 new nodes, with Intel TDX, ARM CCA/TrustZone, and AMD SEV
Research supported: application, system and network security, measurement, human user studies, largescale experiments, education, trustworthy computing

8 Tofino switches, 16 Xilinx Virtex-7 NetFPGA development boards (smartNICs)
Research supported: dynamic (programmable) network security, SDN security

500 IoT nodes (a variety of smart home, smart speaker, camera, doorbell, TV, appliance, medical, office, wearable, and miscellaneous devices)
Research supported: IoT security, user privacy

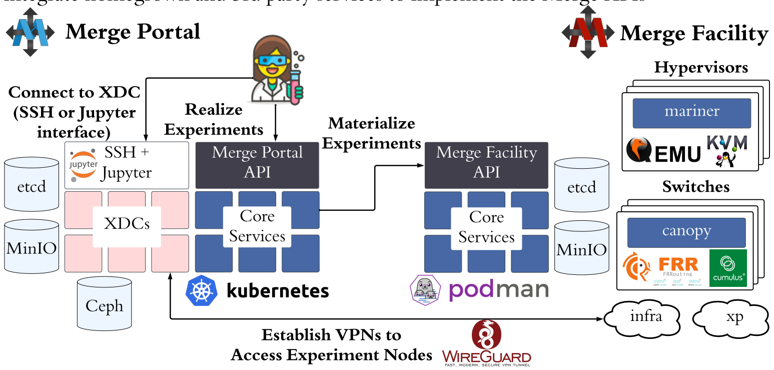
5 Rockwell Automation ControlLogix PLCs, I/O modules
Research supported: critical infrastructure security

- **Reproducibility support by research infrastructure**
    - User action logging to alleviate cognitive load
    - Help package artifacts on SPHERE (including workflows)
    - Automatically verify completeness of an artifact and: stability, consistency of results and portability
- **Dedicated team of researchers, developers and managers**
    - Operated the only public cybersecurity testbed – DeterLab (20 years)
    - Built and operated the largest IoT testbed – Mon(IoT)r Lab
    - Developed and shared Merge and IoT testbed software

Merge SW for Research Infrastructure

Microservice Architectures for Modularity and Resilience

The Merge portal and facility codebases use microservice architectures to flexibly integrate homegrown and 3rd party services to implement the Merge APIs



Merge supports multiple facilities, which may be managed by different teams and contain different hardware and software.

Any compute/network infrastructure implementing the *Merge Facility API* can be commissioned as a Merge testbed facility

Transforming Research Community

- **Need-discovery workshops and surveys**
    - Presentations and BoFs at major conferences
    - Direct engagement with researchers via surveys and interviews
    - Discover needs of all community members and adjust SPHERE development to meet them
- **Help develop standards for artifacts**
    - Engage wide research community in discussion arout artifacts
    - Help produce specifications around proper and complete artifact documentation
- **Representative experimentation environments (REEs)**
    - Used by multiple researchers for a given experimentation task, become a standard for evaluation in a sub-field of cybersecurity and privacy
    - Contributed by research community – researchers receive supplemental funding to deploy their high-quality artifacts as REEs on SPHERE
- **Streamlining artifact evaluation**
    - Work with artifact evaluation committees (AECs) to have artifacts evaluated on SPHERE
    - Artifact authors can submit their artifacts by deploying them on SPHERE
    - AECs evaluate on SPHERE, make recommendations for improvement
    - Artifacts remain hosted on SPHERE
- **Broadening participation in computing**
    - Host 20 minority students per year, involve them in SPHERE development
    - Provide research infrastructure to underresourced institutions
    - Improve cybersecurity education via EDU portal, hosting of education materials

Visit us at https://sphere-project.net

USC Viterbi School of Engineering | USC Information Sciences Institute | Northeastern University Khoury College of Computer Sciences

SPHERE is based upon work supported by the National Science Foundation under grant number 2330066. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.