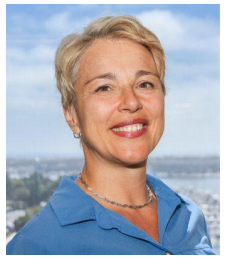
SPHERE - Security and Privacy Heterogeneous Environment for Reproducible Experimentation

University of Southern California Information Sciences Institute,
Northeastern University, University of Utah

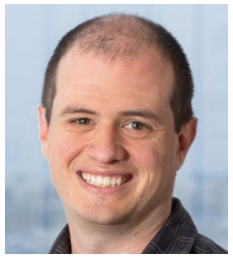
Presented by: Jelena Mirkovic, Principal Investigator <mirkovic@isi.edu> and
David Balenson, Outreach Director <balenson@isi.edu>

SPHERE is based upon work supported by the National Science Foundation under Grant #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.

SPHERE Project Team



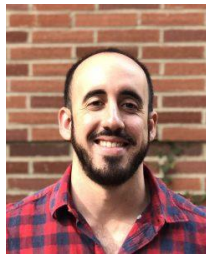
Jelena Mirkovic
USC-ISI
Lead PI



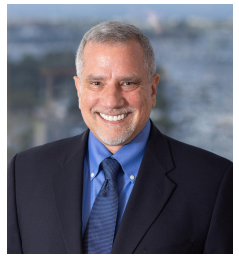
Erik Kline
USC-ISI
Co-PI, Tech Lead

David Choffnes
NEU
Co-PI

Daniel Dubois
NEU
Tech Lead



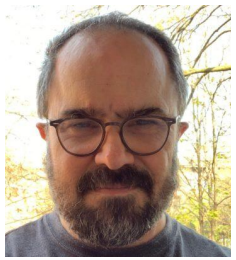
Luis Garcia
Utah
Co-PI, Tech Lead



David Balenson
USC-ISI
Outreach Director

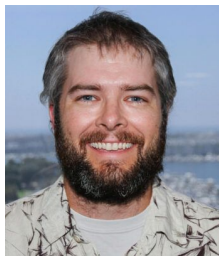
Alba Regalado
USC-ISI
Project Manager

Terry Benzel
USC-ISI



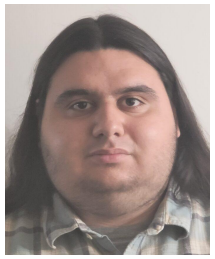
Geoff Lawler
USC-ISI

Joseph Barnes
USC-ISI



Lincoln Thurlow
USC-ISI

Yuri Pradkin
USC-ISI



Kevin Aldana
NEU

Darsh Pandya
NEU

Srivatsan Ravi
USC-ISI

Erika Bobbitt
USC

Government Team

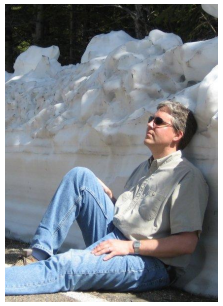
Kevin Thompson
Cognizant PD
NSF/CISE/OAC

Deep Medhi
PD
NSF/CISE/CNS

SPHERE Advisory Board

Sujata Banerjee

Partner Research Manager, Microsoft



Eric Eide

Research Associate Professor, Univ. Utah

Carolyn Ellis

Director, Research Office, Arizona State Univ.

Amir Herzberg

Professor of Computer Science, Univ. Connecticut

Kate Keahey

Senior Fellow, Univ. Chicago; and Computer Scientist, Argonne National Lab

Brian Kocoloski

Performance Architect, AMD

Yoshi Kohno

Professor of Computer Science, Georgetown Univ.

Margaret Martonosi

Professor of Computer Science, Princeton Univ.

Inder Monga

Executive Director of ESNet and Division Director of Scientific Networking

Peter Peterson

Associate Professor of Computer Science, Univ. Minnesota Duluth

Jonathan Petit

Head of AI Security, Safety and Privacy Research & Standardization, Qualcomm

Shokoufeh Mirzaei

Chair & Professor of Engineering, Cal Poly Pomona

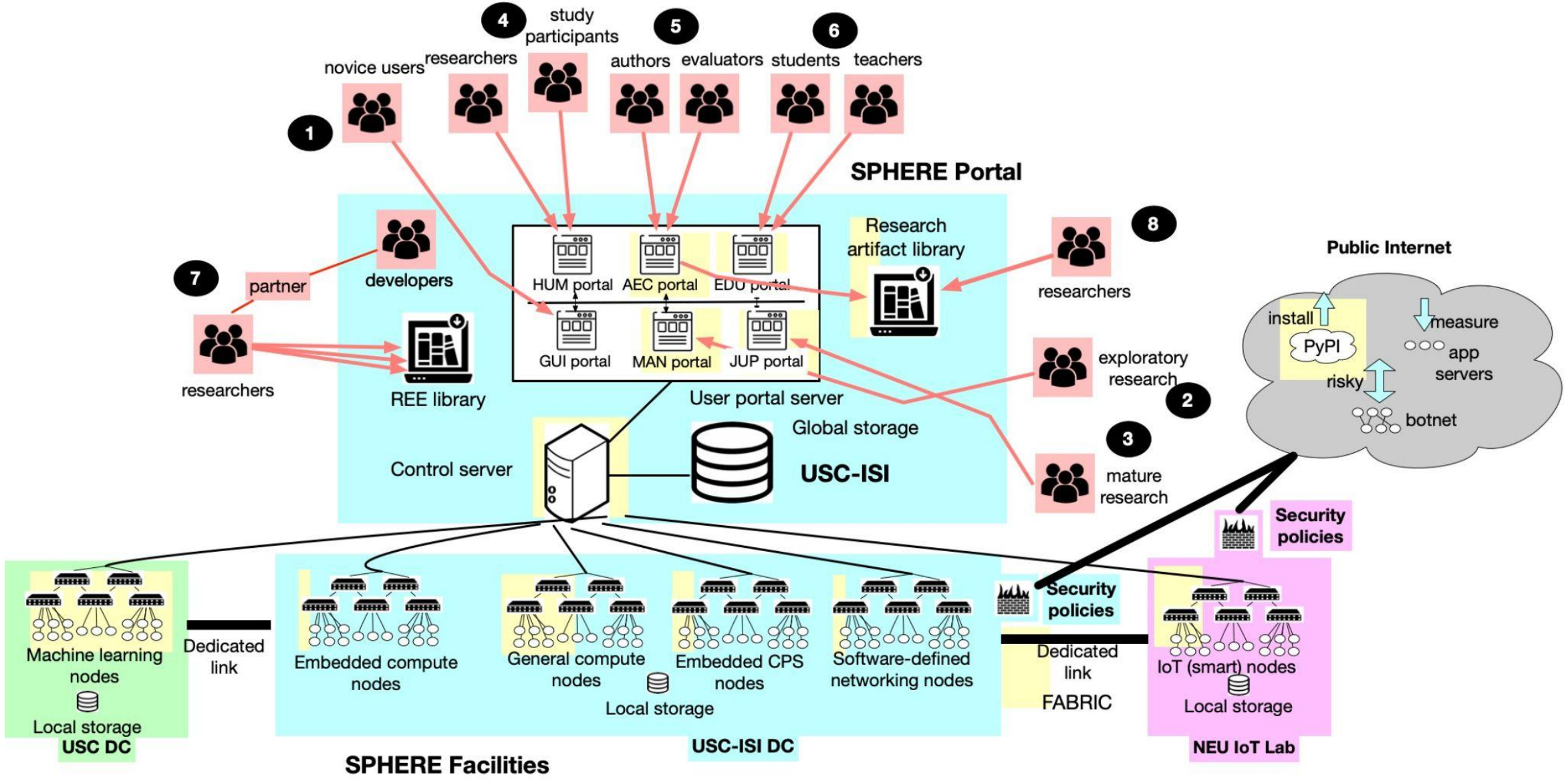
Patrick Traynor

Professor and Chair in Engineering, Univ. Florida

Motivation and Need

- **Motivation:** Cyber threats affect every aspect of our daily lives, critical infrastructure, science, and government. Research solutions are simplistic, piecemeal, and opportunistic, and slow to reach the market
- **Community need:** Common, rich, representative research infrastructure, which meets the needs across all members of the community and facilitates reproducible and open science → vertical progress, integrated research, and more sophisticated solutions
- **Proposed:** SPHERE research infrastructure
  - Heterogeneous resources to meet 89% of research needs in the community
  - Multiple user portals to meet the unique needs of different classes of users
  - Processes/incentives for the community to create representative experimentation environments (REEs) on SPHERE
  - Integrated reproducibility support and processes/incentives for stakeholders to share/reuse research artifacts

High-Level Architecture

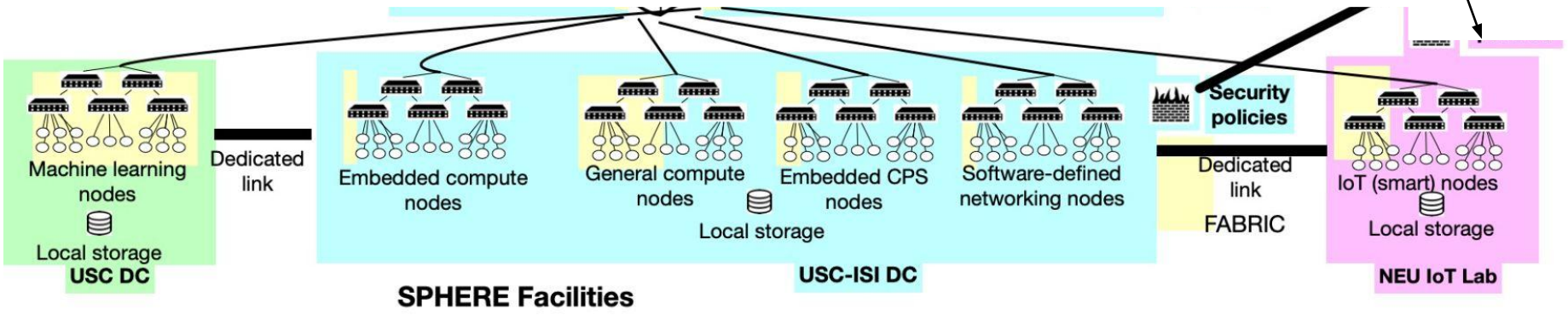


Multiple Types of Resources

- **Multiple** types of resources, needed for **emerging cybersecurity and privacy research**:
  - General compute nodes with trusted computing technology – research on network, cloud computing and system threats
  - Embedded compute nodes (e.g., in phones, tablets, etc.) – research on distributed threats, threats on distributed computing, attacks on specific CPU architectures
  - Cyber-physical nodes (PLCs) – research on threats on industrial systems and critical inf.
  - GPU nodes – incorporate machine learning into solutions
  - Programmable nodes (FPGAs) and switches – facilitate transition to market
  - IoT nodes (smart home nodes and personal devices) – research on threats on IoT

GPU – graphical processing unit
CPU – central processing unit
PLC – programmable logic controller
FPGA – field-programmable gate array
IoT – Internet of Things

Flexible security policies, support select, safe communication with the Internet



NEU IoT Lab

SPHERE RESEARCH INFRASTRUCTURE

Remote access to a set of 500 consumer IoT devices

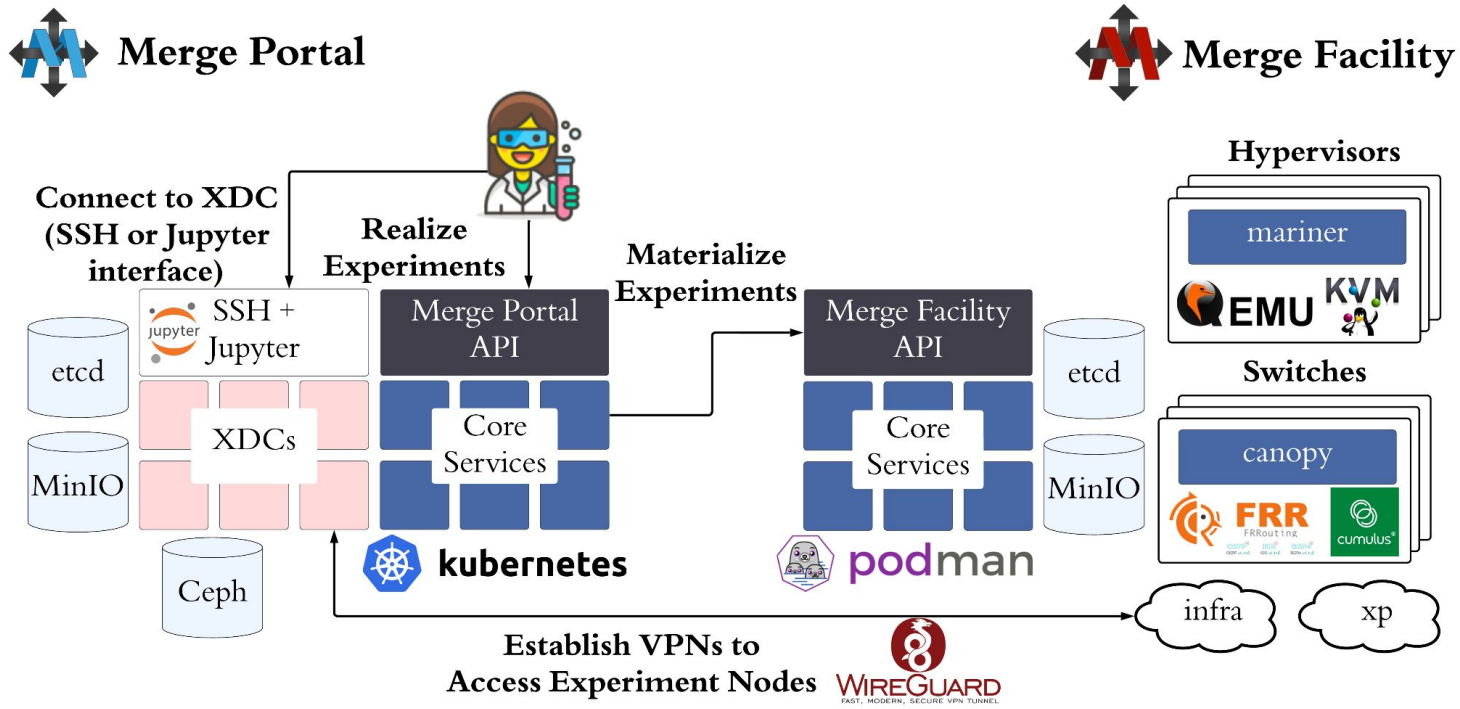
- Device reservation (individual devices or groups of devices)
- Network isolation and management:
  - Remote network access to the devices (VPN / VLANs)
- Remote sensor access
  - Cameras, microphones
- Remote actuator access
  - Speakers, button pushers, robotic arms, power control, infrared control
- Companion apps on physical/virtualized phones
- Logging and access to network traffic and sensor/actuator data

MERGE Control Software

Microservice Architecture for Modularity and Resilience

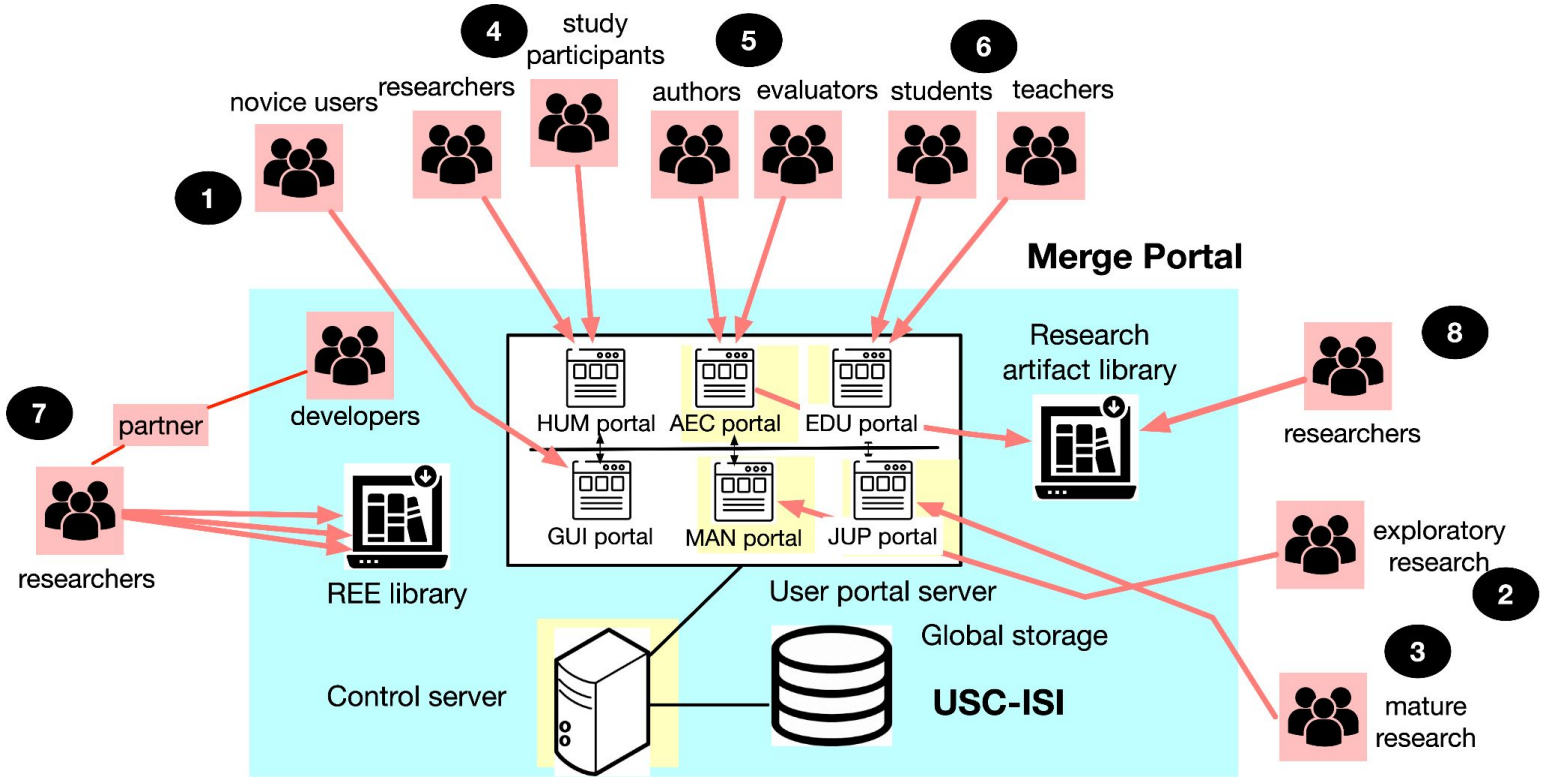
The Merge Portal and Facilities codebases use microservice architectures to flexibly integrate homegrown and third-party services to implement the Merge APIs

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

User Portals



- Multiple user portals, supporting different types of users and use modalities
  - Manual, scripted, and GUI-only use support exploratory, mature, and novice research
  - Dedicated support for AECs, education, Internet measurement, and human user studies

GUI – graphical user interface
AEC – artifact evaluation committee

REEs and Research Artifacts

Representative Experimentation Environments (REEs)

- Make research more relevant, vertical and sophisticated
- "Standard" for experimentation in each CS&P area
- Integrated by their authors into SPHERE (funded)

Research Artifacts

- Make research more vertical and reproducible
- Acquired via partnership with artifact evaluation committees (AECs)
- Integrated by their authors into SPHERE as part of artifact evaluation for a conference

CS&P – cybersecurity and privacy
AEC – artifact evaluation committee

History

- **Over the past 20 years:** USC-ISI designed, built and operated DETERLab
  - 389 research project teams from 278 institutions, and involving 1,042 researchers from 205 locations and 46 countries
  - 230 classes from 147 institutions and helped educate more than 20,000 students
- **2019:** Merge software for testbed control and management
  - Built w/ modern open-source tools for large-scale, high-fidelity, robust experimentation
  - Merge has run several of our testbeds for the past four years – DCOMP, Searchlight, RedStar, and modernized DeterLab
- **Modernized software and hardware:** via NSF CCRI grant 2019-2022 and ARO DURIP grant 2019-2021
  - 48 new nodes, 6 new switches, Merge software, user transition
- Modernized DETERLab became the first seed to grow SPHERE as part of its general compute enclave

Prior Government Programs

USC-ISI has leveraged DETERLab or other Merge-based testbeds to provide T&E support (methodologies, tools, scenarios development, testbeds) for numerous DARPA programs:

- **Safer Warfighter Communications (SAFER):** DETERLab served as the testbed and experimentation platform for the development and evaluation of technologies advancing the state-of-the-art in non-blockable and anonymous communication
- **Edge-Directed Cyber Technologies for Reliable Mission Communication (Edge-CT):** DETERLab served as the testbed and experimental platform build realistic models of networks and workloads to help develop and evaluate prototypes
- **Dispersed Computing (DCOMP):** Merge-based DCOMP testbed was used for program evaluation with simultaneous experiments with hundreds of nodes interconnected through complex network topologies
- **Extreme DDoS Defense (XD3):** DETERLab served as the test platform to evaluate performer technologies
- **Searchlight:** Merge-based Lighthouse testbed supported extensive virtualization built around QEMU/KVM and Sandia's Minimega virtual machine automation platform to evaluate performer technologies

Unique Research Capabilities

SPHERE RESEARCH INFRASTRUCTURE

- **Relevance:** Experiments with emerging technology and specialized hardware, not currently available to many researchers, support 90%
- **Realism:** Experiments that combine different hardware devices to create realistic scenarios
  - e.g., IoT nodes with GPU nodes and programmable switches to filter attacks
- **Reproducibility:** Experiments on common RI, with extensive support for artifact sharing and reuse, facilitate vertical development
- **BPC:** Different experimentation portals cater to users with different abilities and interests, lowering barrier to entry
- **Impact:** Faster pace of innovation in CS&P and faster technology transition to practice

IoT – Internet of Things
GPU – graphical processing unit
RI – research infrastructure

BPC – Broadening Participation in Computing
CS&P – cybersecurity and privacy

Team Background

- **DeterLab:** the only public cybersecurity testbed for **18 years**
- Additional testbeds for formal eval. of DARPA programs
- **Merge:** mature testbed management software, running all three testbeds
- **Mon(IoT)r:** largest private IoT testbed and datasets
- Prior NSF funding: **SEARCCH** (reprod.), **DEW** (reprod., usab.), **DeterLab modernization** (RI)
- Many publications on experimentation, reproducibility, IoT privacy
- Founded **CSET workshop**, led NSF-funded **CEF study**, organized **CEF 2022** and **Cybersecurity Artifacts 2022** workshops, pioneered use of testbeds in education

389 research groups
1K researchers
237 classes
20K students

ported to 4 partner institutions

560 downloads

IoT - Internet of Things
SEARCCH – Sharing Expertise and Artifacts for Reuse through a Cybersecurity Community Hub

DEW – Distributed Experiment Workflows
CSET – Cyber Security Experimentation and Test, running for 17 years
CEF – Cybersecurity Experimentation of the Future

Research Value

- Transform CS&P research from piecemeal, opportunistic to integrated; and from reactive to proactive
- Enable reproducible experimentation that is easily and remotely accessible to all researchers
  - Benefits broad researcher populations (evidence from DeterLab)
- Undergraduate and graduate students recruited for paid internships
- Work with AECs to transform the research process and host artifacts
- REEs and artifacts will lead to increase in publications and data products

CS&P – cybersecurity and privacy
AEC – artifact evaluation committee
REEs – representative experimentation environments

Societal Benefits

- Faster pace of innovation in CS&P and more mature solutions on the market
- Protect scientific infrastructure and society from various threats: ransomware, data theft, data corruption, supply chain attacks, denial of service, etc.
- Produce larger, more diverse, better educated and prepared CS&P workforce
- Help integrate CS&P solutions into new and emerging technologies before they get widely deployed

CS&P – cybersecurity and privacy

Community Outreach

- Needs discovery to inform our project activities and awareness to encourage beta use
- Presentations, posters, and other activities at major conferences
  - Major cybersecurity conferences: NDSS, S&P, USENIX Security, CCS, ACSAC
  - Other conferences: IoT, CPS, SC, etc.
  - NSF events: RIW, Cybersecurity Summit, MERIF, SaTC PI meeting
    Other communities: Tapia, Grace Hopper, SACNAS NDiSTEM
- Engage researchers via surveys and interviews
  - Target diverse research communities, such as networking, CPS, IoT, SDN, etc.
- Adjust SPHERE development to meet community needs

Google form at https://bit.ly/SPHERE-Needs-Survey

NDSS – Network and Distributed System Security
S&P – IEEE Symposium on Security & Privacy
CCS – ACM Conference on Computer & Communication Security
ACSAC – Annual Computer Security Applications Conference

RIW – Research Infrastructure Workshop
SaTC – Secure and Trustworthy Cyberspace
SACNAS – Advancing Chicanos/Hispanics & Native Americans in Science
NDiSTEM – National Diversity in STEM

Open to Beta Users

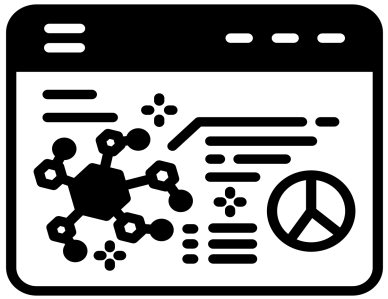
- These users help us grow and improve before we open to larger audience
- Currently around 130 researchers and 200-1000 students per semester
- Users get access to cool new hardware and features
  - Log in remotely via browser, create custom topologies of general purpose VMs (control VM resources, network topology, bandwidth and delay)
  - Access nodes via SSH w/ sudo privileges
  - Experiment directly on nodes or via Jupyter notebooks
  - Able to reach into the Internet, can also support incoming connections
  - Chat-based user support

Component	Dev Started	Available for Use
SPHERE Infrastructure	Oct 2023	Mar 2024
General purpose nodes	May 2024	Oct 2025
GPU nodes	Nov 2024	Apr 2025
CPS nodes	Nov 2024	Oct 2025
Embedded compute nodes	May 2025	Jan 2026
IoT nodes	Oct 2023	Aug 2025
Programmable nodes	Sep 2025	Mar 2026

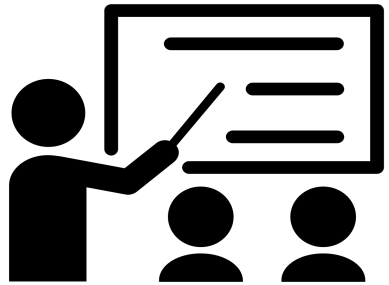
How You Can Help!

Promote and leverage SPHERE at your organizations!

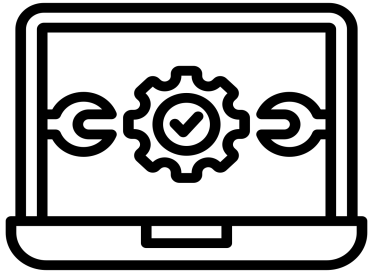
Researchers can use SPHERE to conduct new, innovative research



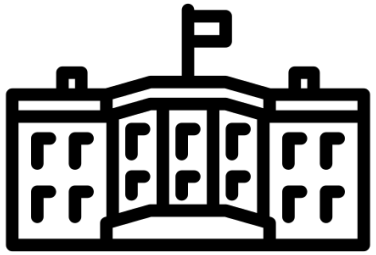
Faculty and students can use SPHERE for educational purposes



IT staff can use SPHERE to test and evaluate new solutions and technologies



Govt PMs can use SPHERE (or other Merge testbeds) to support their programs



Thank you!

If you use SPHERE in your research and a publication, we would appreciate a citation. Currently, the best reference for SPHERE is:

Jelena Mirkovic, David Balenson, Brian Kocoloski. Enabling Reproducibility through the SPHERE Research Infrastructure. USENIX ;login: Online, USENIX Association. December 16, 2024. https://www.usenix.org/publications/loginonline/enabling-reproducibility-through-sphere-researchinfrastructure

https://sphere-project.net

contact@sphere-project.net