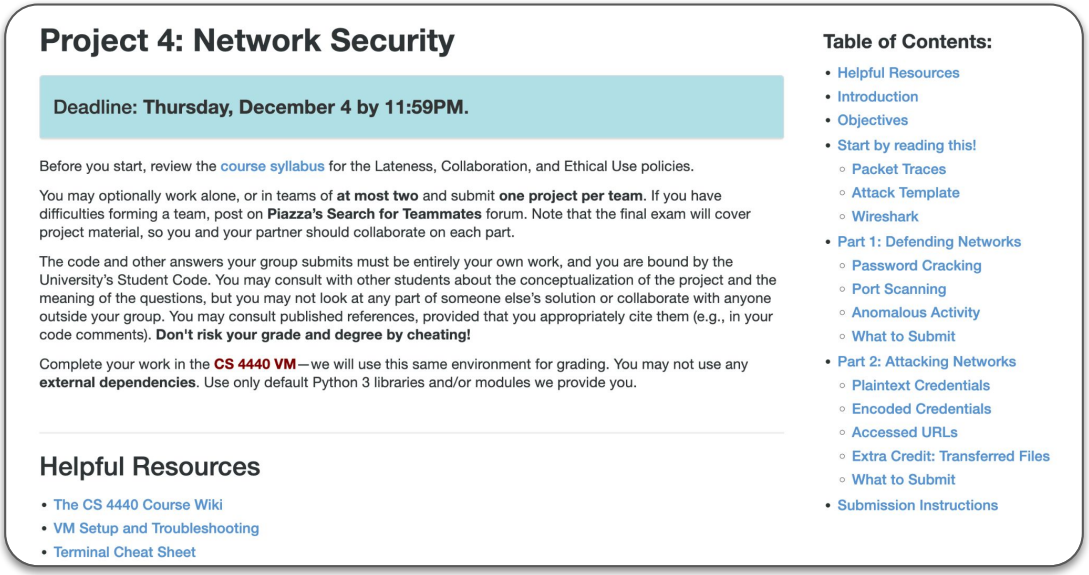
Week 13: Lecture A

Cyber-physical & IoT Security

Tuesday, November 18, 2025

Announcements

- **Project 4: NetSec** released
  - **Deadline:** Thursday, December 4th by 11:59PM



Project 4: Network Security

Deadline: Thursday, December 4 by 11:59PM.

Before you start, review the course syllabus for the Lateness, Collaboration, and Ethical Use policies.

You may optionally work alone, or in teams of **at most two** and submit **one project per team**. If you have difficulties forming a team, post on **Piazza's Search for Teammates** forum. Note that the final exam will cover project material, so you and your partner should collaborate on each part.

The code and other answers your group submits must be entirely your own work, and you are bound by the University's Student Code. You may consult with other students about the conceptualization of the project and the meaning of the questions, but you may not look at any part of someone else's solution or collaborate with anyone outside your group. You may consult published references, provided that you appropriately cite them (e.g., in your code comments). **Don't risk your grade and degree by cheating!**

Complete your work in the **CS 4440 VM**—we will use this same environment for grading. You may not use any **external dependencies**. Use only default Python 3 libraries and/or modules we provide you.

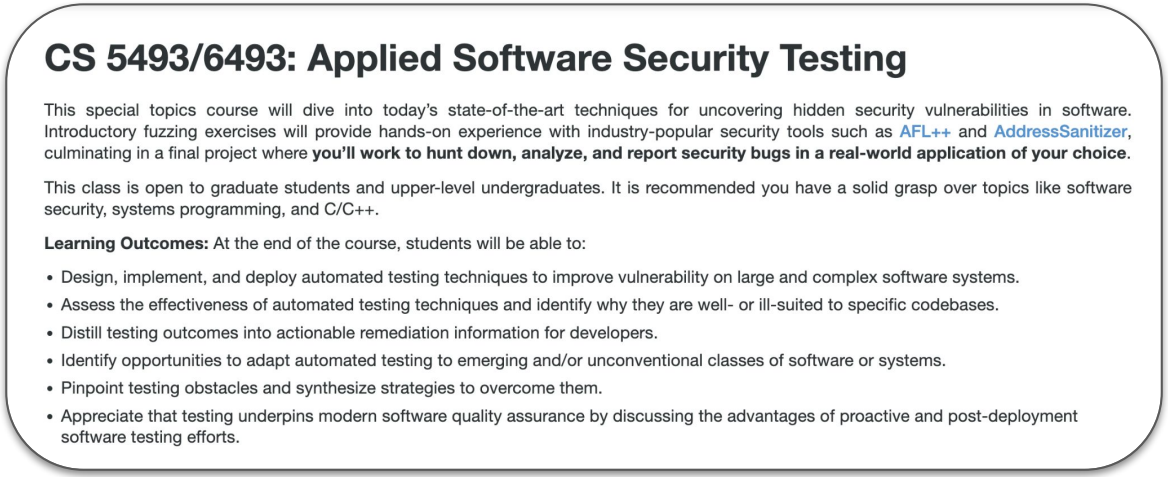
Helpful Resources

- The CS 4440 Course Wiki
- VM Setup and Troubleshooting
- Terminal Cheat Sheet

Table of Contents:

Interested in fuzzing?

- **Spring 2026: CS 5493/6493: Applied Software Security Testing**
  - **Everything you'd ever want to know about fuzzing for finding security bugs!**
  - Course project: team up to fuzz **a real program** (of your choice), and find and report its bugs!
  - http://cs.utah.edu/~snagy/courses/cs5493/



CS 5493/6493: Applied Software Security Testing

This special topics course will dive into today's state-of-the-art techniques for uncovering hidden security vulnerabilities in software. Introductory fuzzing exercises will provide hands-on experience with industry-popular security tools such as AFL++ and AddressSanitizer, culminating in a final project where **you'll work to hunt down, analyze, and report security bugs in a real-world application of your choice**.

This class is open to graduate students and upper-level undergraduates. It is recommended you have a solid grasp over topics like software security, systems programming, and C/C++.

Learning Outcomes: At the end of the course, students will be able to:

- Design, implement, and deploy automated testing techniques to improve vulnerability on large and complex software systems.
- Assess the effectiveness of automated testing techniques and identify why they are well- or ill-suited to specific codebases.
- Distill testing outcomes into actionable remediation information for developers.
- Identify opportunities to adapt automated testing to emerging and/or unconventional classes of software or systems.
- Pinpoint testing obstacles and synthesize strategies to overcome them.
- Appreciate that testing underpins modern software quality assurance by discussing the advantages of proactive and post-deployment software testing efforts.

Announcements

utahsec

See Discord for meeting info!

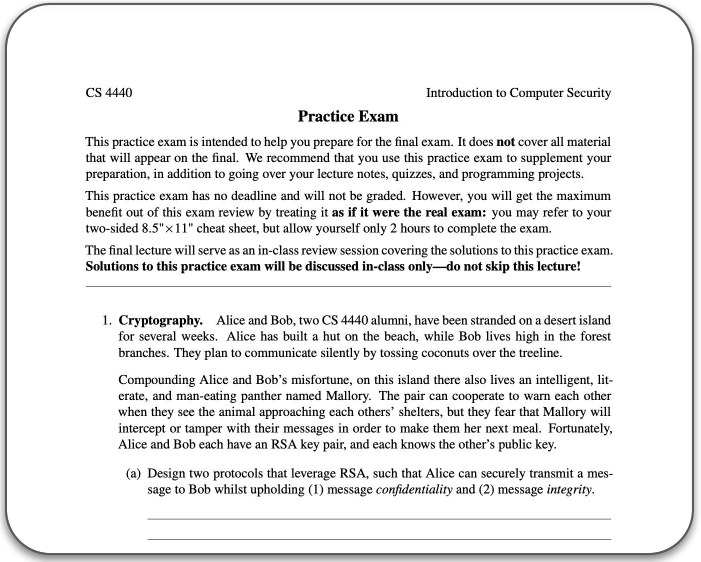
utahsec.cs.utah.edu

Final Exam

- **Save the date: 1–3PM** on **Wednesday, December 10**
  - **CDA accommodations:** schedule exam via CDA Portal
- **High-level details** (more to come):
  - One exam covering all course material
  - Similar to project/quiz/lecture exercises
- **Cheat Sheet**
  - **One 8.5”x11” paper** with handwritten/typed notes on **both** sides
  - **Suggestion:** Don’t just use someone else’s—you’ll learn better making **your own**!
  - **Suggestion:** Don’t just paste lecture slides—you’ll learn better by **writing/typing** it!

Practice Exam

- **Practice Exam** released
  - See **Assignments** page on the CS 4440 website
- **Final lecture** will serve as a **review session**
  - Solutions discussed **in-class only**—don't skip!



CS 4440 Introduction to Computer Security

Practice Exam

This practice exam is intended to help you prepare for the final exam. It does **not** cover all material that will appear on the final. We recommend that you use this practice exam to supplement your preparation, in addition to going over your lecture notes, quizzes, and programming projects.

This practice exam has no deadline and will not be graded. However, you will get the maximum benefit out of this exam review by treating it **as if it were the real exam:** you may refer to your two-sided 8.5"×11" cheat sheet, but allow yourself only 2 hours to complete the exam.

The final lecture will serve as an in-class review session covering the solutions to this practice exam. **Solutions to this practice exam will be discussed in-class only—do not skip this lecture!**

1. **Cryptography.** Alice and Bob, two CS 4440 alumni, have been stranded on a desert island for several weeks. Alice has built a hut on the beach, while Bob lives high in the forest branches. They plan to communicate silently by tossing coconuts over the treeline.

   Compounding Alice and Bob's misfortune, on this island there also lives an intelligent, literate, and man-eating panther named Mallory. The pair can cooperate to warn each other when they see the animal approaching each others' shelters, but they fear that Mallory will intercept or tamper with their messages in order to make them her next meal. Fortunately, Alice and Bob each have an RSA key pair, and each knows the other's public key.

   (a) Design two protocols that leverage RSA, such that Alice can securely transmit a message to Bob whilst upholding (1) message *confidentiality* and (2) message *integrity*.

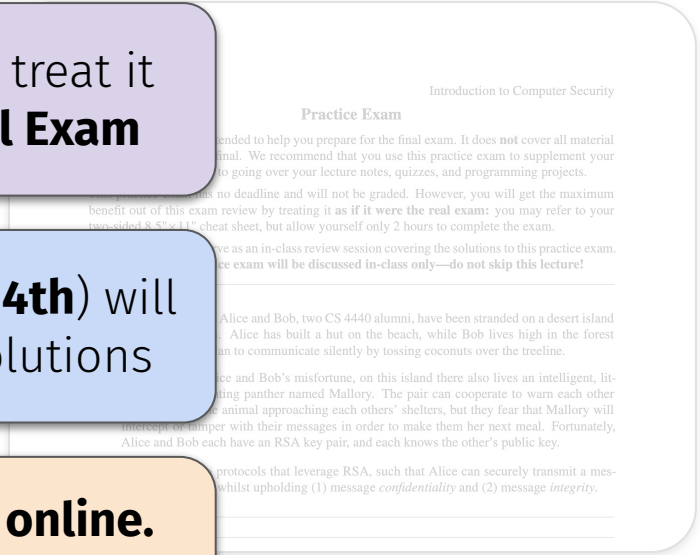
Practice Exam

- **Practice Exam** re
  - See **Assignmen**
- **Final lecture** will **serve as a review session**
  - Solutions discu

To get the most out of this, treat it just **as you would the Final Exam**

Last lecture (**Thursday, Dec. 4th**) will go over the exam review solutions

Solutions won't be posted online.
(Reminder: attendance/participation makes up **5%** of your course grade)



Introduction to Computer Security

Practice Exam

ended to help you prepare for the final exam. It does **not** cover all material
inal. We recommend that you use this practice exam to supplement your
to going over your lecture notes, quizzes, and programming projects.

s no deadline and will not be graded. However, you will get the maximum
benefit out of this exam review by treating it **as if it were the real exam:** you may refer to your
two-sided 8.5"×11" cheat sheet, but allow yourself only 2 hours to complete the exam.

ve as an in-class review session covering the solutions to this practice exam.
ce exam will be discussed in-class only—do not skip this lecture!

Alice and Bob, two CS 4440 alumni, have been stranded on a desert island
Alice has built a hut on the beach, while Bob lives high in the forest
an to communicate silently by tossing coconuts over the treeline.

ce and Bob's misfortune, on this island there also lives an intelligent, lit-
ting panther named Mallory. The pair can cooperate to warn each other
e animal approaching each others' shelters, but they fear that Mallory will
intercept or tamper with their messages in order to make them her next meal. Fortunately,
Alice and Bob each have an RSA key pair, and each knows the other's public key.

protocols that leverage RSA, such that Alice can securely transmit a mes-
whilst upholding (1) message *confidentiality* and (2) message *integrity*.

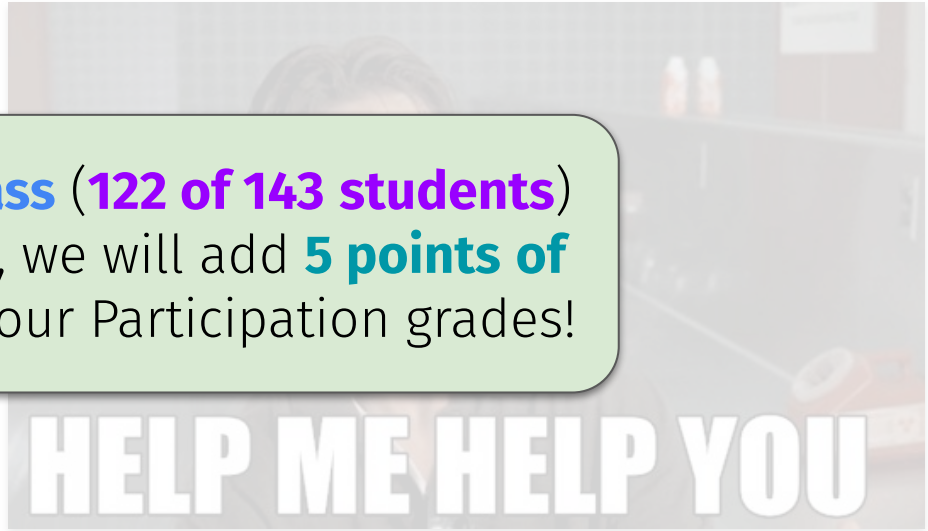
End-of-semester Course Evals

- **I want your feedback!**
  - 3rd time teaching this course 😃
  - **Help me improve the class!**
- Due by **December 15th**
  - https://scf.utah.edu
  - **Please please please!**

End-of-semester Course Evals

- **I want your feedback!**
  - 3rd time teaching this course 😃
  - **Help me improve the class!**
- Due by **Dece**
  - https://sc
  - **Please pl**

If 85% of the class (**122 of 143 students**) submits an eval, we will add **5 points of extra credit** to your Participation grades!



Questions?

Last time on CS 4440...

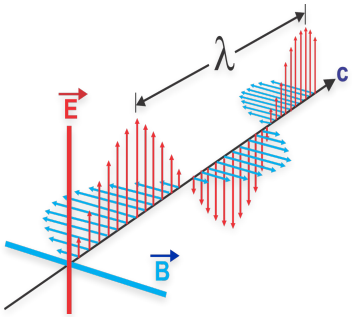
Side Channels
Hardware Security
Hardware Supply Chain Attacks

Side Channels

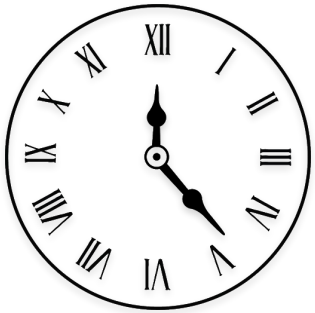
- What are some potential sources of **indirect info** emitted by your computer?
  - **Additional channels** of information beyond what is directly visible/accessible to you

Side Channels

- What are some potential sources of **indirect info** emitted by your computer?
  - **Additional channels** of information beyond what is directly visible/accessible to you



Emitted Radiation

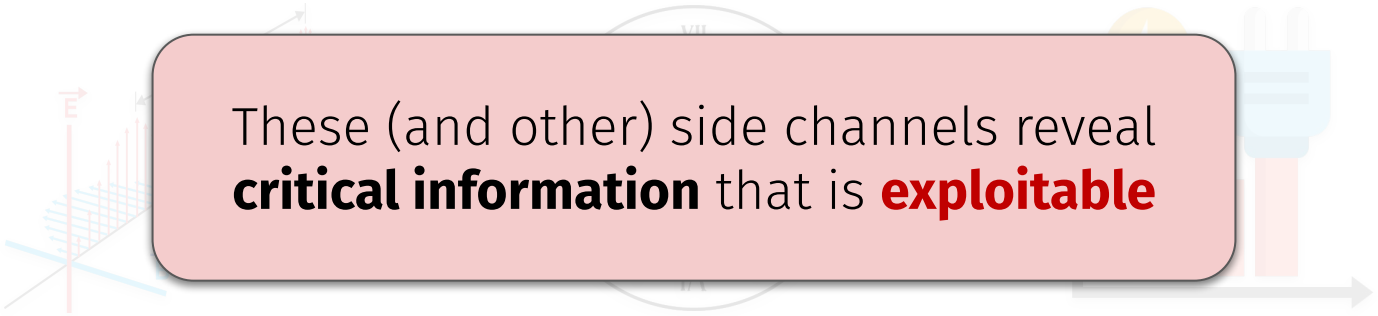


Execution Time

Power Consumption

Side Channels

- What are some potential sources of **indirect info** emitted by your computer?
  - **Additional channels** of information beyond what is directly visible/accessible to you

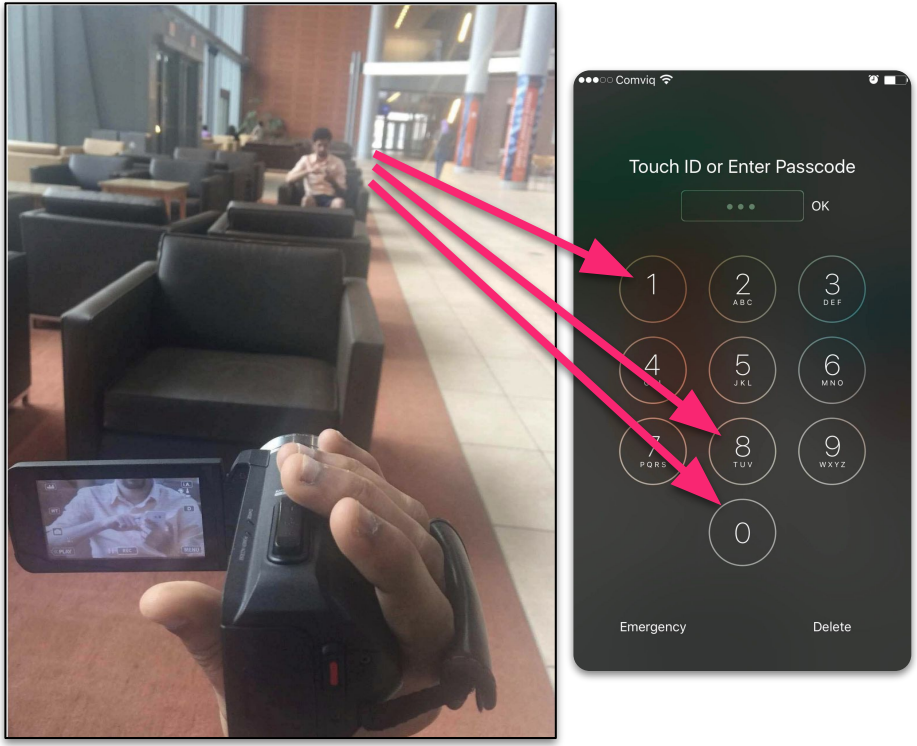


These (and other) side channels reveal **critical information** that is **exploitable**

Emitted Radiation | Execution Time | Power Consumption

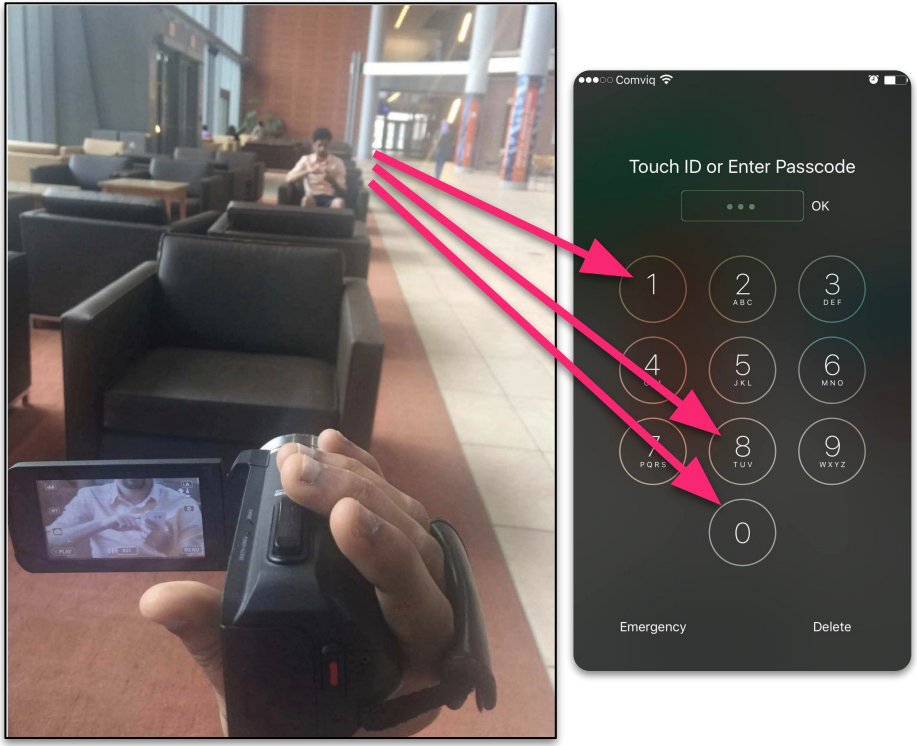
Optical Side Channels

- **Stealing passwords via gestures**
  - ???



Optical Side Channels

- **Stealing passwords via gestures**
  - Capture visible **hand movements**
  - Assume attacker **knows (or can easily guess)** the key interface
  - Attacker **maps movements** to pressed keys on the interface



Acoustic Side Channels

- **Stealing passwords via key press noises**
  - ???

Acoustic Side Channels

- **Stealing passwords via key press noises**
  - Build model of key press noises
    - Consider microphone
    - Consider ambient noise
  - Use model to infer entered data
    - Passwords
    - Usernames
    - Phone numbers

Other Side Channels

- How `memcmp()` works under the hood:

```
bool checkPW(char *testPW, char *realPW, int len) {

    for (int i = 0; i < len; i++) {

        if (testPW[i] != realPW[i]) {
            return false;
        }
    }

    return true;
}
```

What is the side channel here?

Other Side Channels

- How `memcmp()` works under the hood:

```
bool checkPW(char *testPW, char *realPW, int len) {

    for (int i = 0; i < len; i++) {

        if (testPW[i] != realPW[i]) {
            return false;
        }
    }

    return true;
}
```

Password Login Attempts:

ABCDEFGH == PASSWORD
- **False** on first iteration

PASSEFGH == PASSWORD
- **True** on iterations **1–4**
- **False** on **fifth** iteration

More code executed
for a **correct** symbol!

Other Side Channels

How can this **side channel** be **exploited**?

Other Side Channels

How can this **side channel** be **exploited**?

Attacker: CRCDEF

Server: False
Server took 2ms to respond

Attacker: CHIDEF

Server: False
Server took 4ms to respond

Other Side Channels

How can this **side channel** be **exploited**?

Attacker: `CHIEFS`

Server: `True`
Server took **7ms** to respond

Through **timing analysis**, attacker can infer the **correctness** of individual **password symbols**!

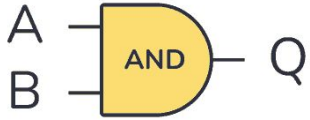
Avoiding Side Channels

- **Solution:**
  - ???

Avoiding Side Channels

- **Solution:**
  - **Constant-time** implementation (e.g., using bitwise **AND**-ing)

```
bool checkPW(char *testPW, char *realPW, int len) {
    bool result = 1; // integer equiv of "true"
    for (int i = 0; i < len; i++) {
        result &= ca[i] == cb[i];
        return result;
    }
}
```



A	B	Q
0	0	0
0	1	0
1	0	0
1	1	1

Avoiding Side Channels

- **Solution:**
  - **Constant-time** implementation (e.g., using bitwise **AND**-ing)

```
bool checkPW(char *testPW, char *realPW, int len) {

    bool result = 1; // integer equiv of "true"

    for (int i = 0; i < len; i++) {

        result &= ca[i] == cb[i];

        return result;
    }
}
```

Guess: PASSEFGH
Bit: 11110000
Result: **False**

Password Login Attempts:

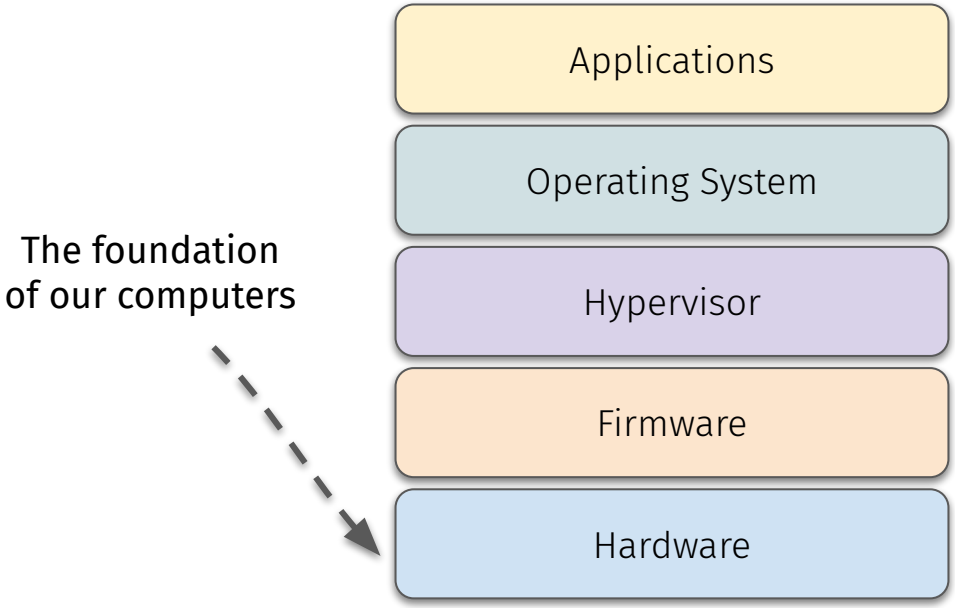
ABCDEFGH == PASSWORD
- **False** on **last** iteration

PASSEFGH == PASSWORD
- **False** on **last** iteration

PASSWORD == PASSWORD
- **True** on **last** iteration

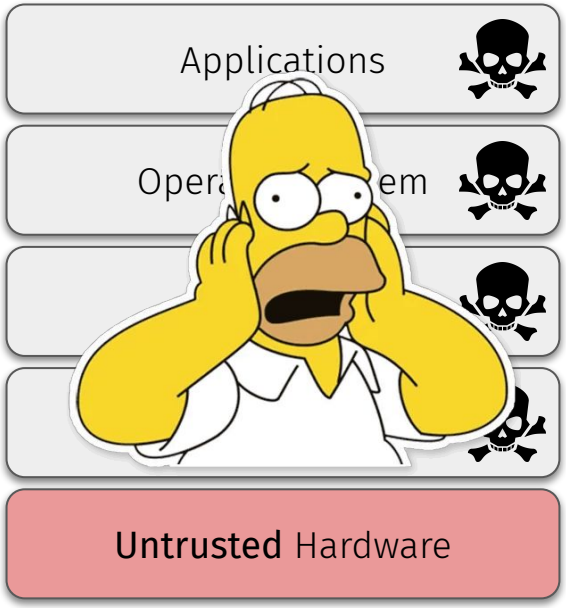
True and **False** run for **identical time**!

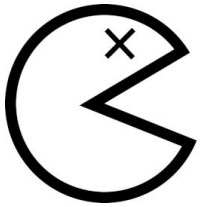
Hardware Threats



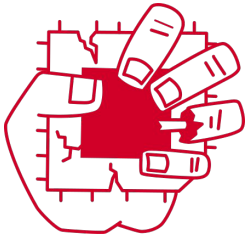
Hardware Threats

The foundation of our computers





Weaknesses weaken the entire system

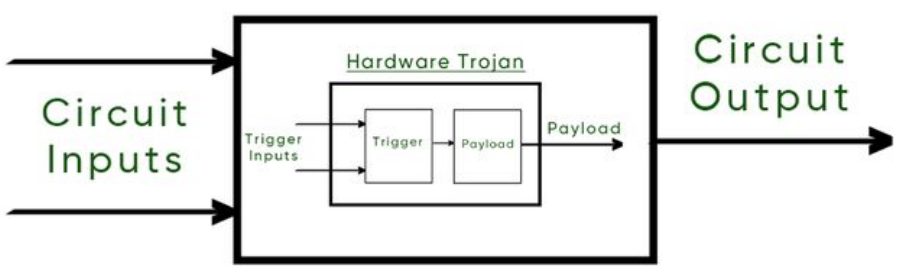


Hardware Threats

- **Hardware Trojans:**
  - ???

Hardware Threats

- **Hardware Trojans:**
  - Attack pre-inserted into chip
  - Will be **exploited** at **run time**
  - **Remotely triggered** by attacker
    - Small
    - Stealthy
    - Controllable



Israeli sky-hack switched off Syrian radars countrywide

Backdoors penetrated without violence

Lewis Page — Thu 22 Nov 2007 / 13:57 UTC

More rumours are starting to leak out regarding the mysterious Israeli air raid against Syria in September. It is now suggested that "computer to computer" techniques and "air-to-ground network penetration" took place.

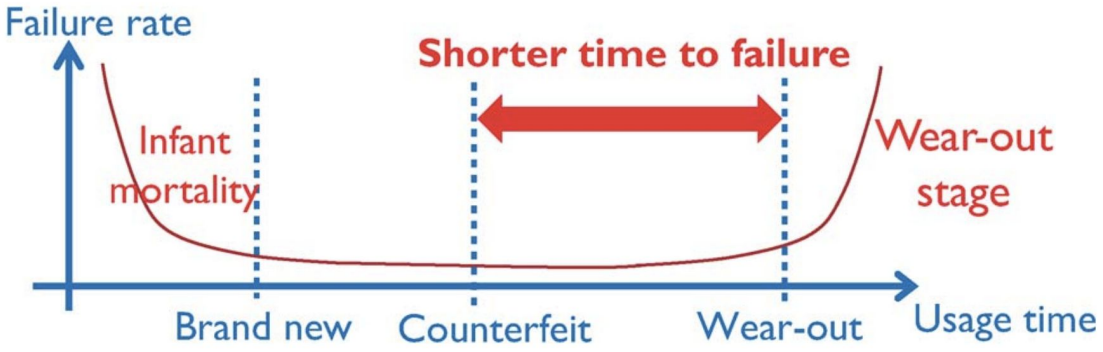
The latest revelations are made by well-connected *Aviation Week* journalists. Electronic-warfare correspondent David Fulghum says that US intelligence and military personnel "provided advice" to the Israelis regarding methods of breaking into the Syrian air-defence network.

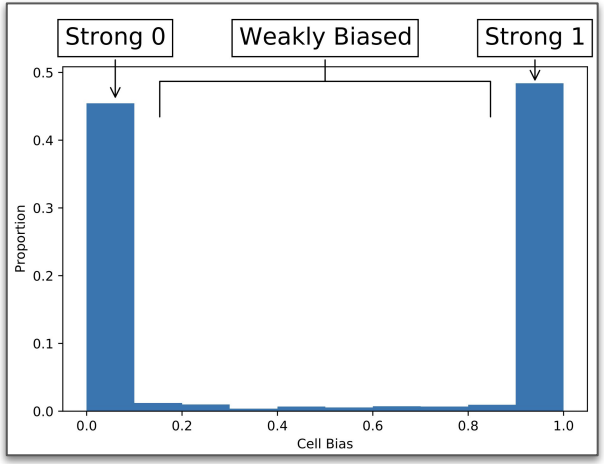
Hardware Threats

- **Counterfeit and recycled chips:**
  - ???

Hardware Threats

- **Counterfeit and recycled chips:**
  - Have a **shorter** lifespan—leads cell bias and/or earlier wear-out
  - Absolutely dangerous for security-critical use cases





Questions?

No Class or Office Hours Next Week

This time on CS 4440...

Cyber-physical Systems &
Internet-of-Things Security

CS 4440
University of Utah

Cyber-physical Systems and IoT Security

A Brief Introduction

https://info.hummingbirdnetworks.com/blog/what-is-iot-and-how-it-works

Luis Garcia

https://iotrustlab.com

My Research: CPS/IoT Security, Safety, & Privacy, XAI for Sensor data, Medical IoT

Course: CS 5464 (Coming back Fall 2026)

Office: MEB 3450

IoTrust Lab

Trustworthy Cyber-physical Systems Lab

U of U
Asst. Prof. (Since July)

USC, ISI
Research Lead (2020-2023)

UCLA
Postdoctoral Fellow (2018-2020)

Rutgers
Ph.D.

U of Miami
M.S. + B.S.

Our society increasingly relies on cyber-physical/IoT autonomy…

From safety-critical infrastructure…

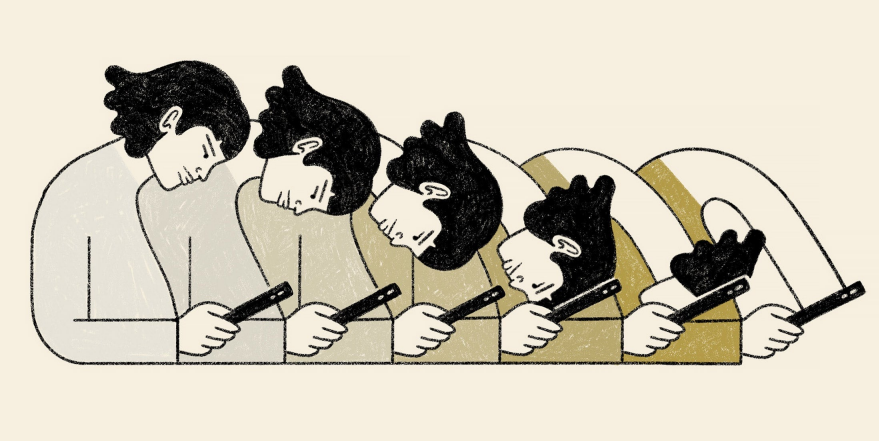
…to making our lives more efficient…

...sometimes at odds with security...

A New Era of Internet Attacks Powered by Everyday Devices

By **DAVID E. SANGER** and **NICOLE PERLROTH** OCT. 22, 2016

...sometimes at odds with privacy



Phones That Can Read Your Mind

Targeted ads may soon show you what you really want before you knew you did.

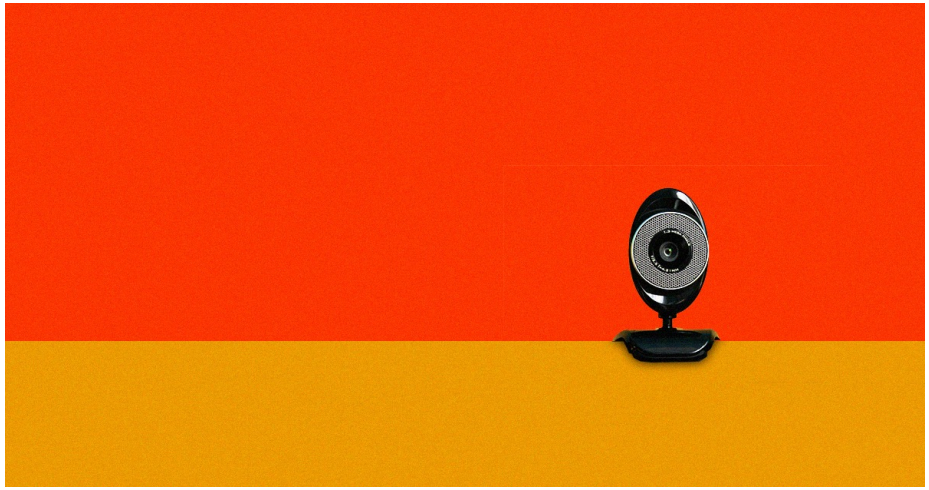
By Andy Kessler Follow

Jan. 26, 2020 3:51 pm ET

Airbnb Hosts Are Spying on Guests With Hidden Cameras

And the platform's botched handling of the issue puts guests in harm's way.

/ Future Society / Airbnb / Gig Economy / Hidden Camera



...sometimes at odds with safety...

Driverless Cars Face Setbacks In San Francisco—Here's What To Know About The City's Problematic Robotaxi Rollout

Mary Whitfill Roeloffs Forbes Staff

I am a Boston-based reporter covering breaking news.

Follow

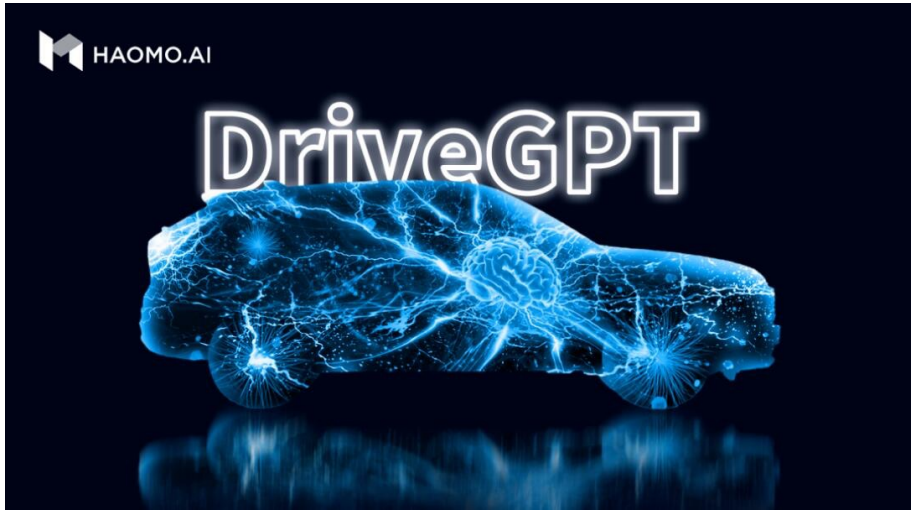
Wing delivery drone crashes into power lines in Australia

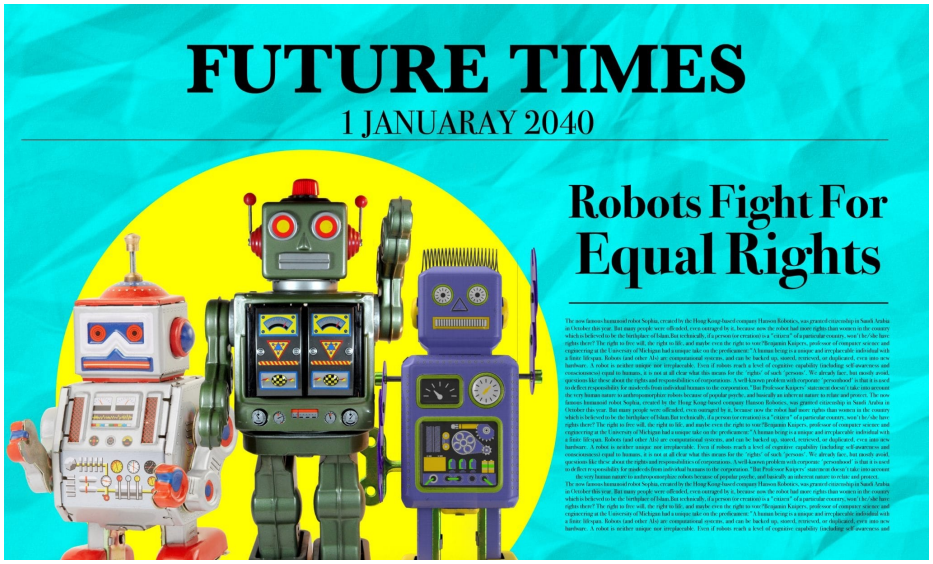
By Brianna Wessling | September 30, 2022

Walmart's latest challenge with its drone delivery system is gun owners shooting packages out of the air

Fortune · Paul Hennessy—Anadolu Agency/Getty Images

Innovation only leads to more questions...



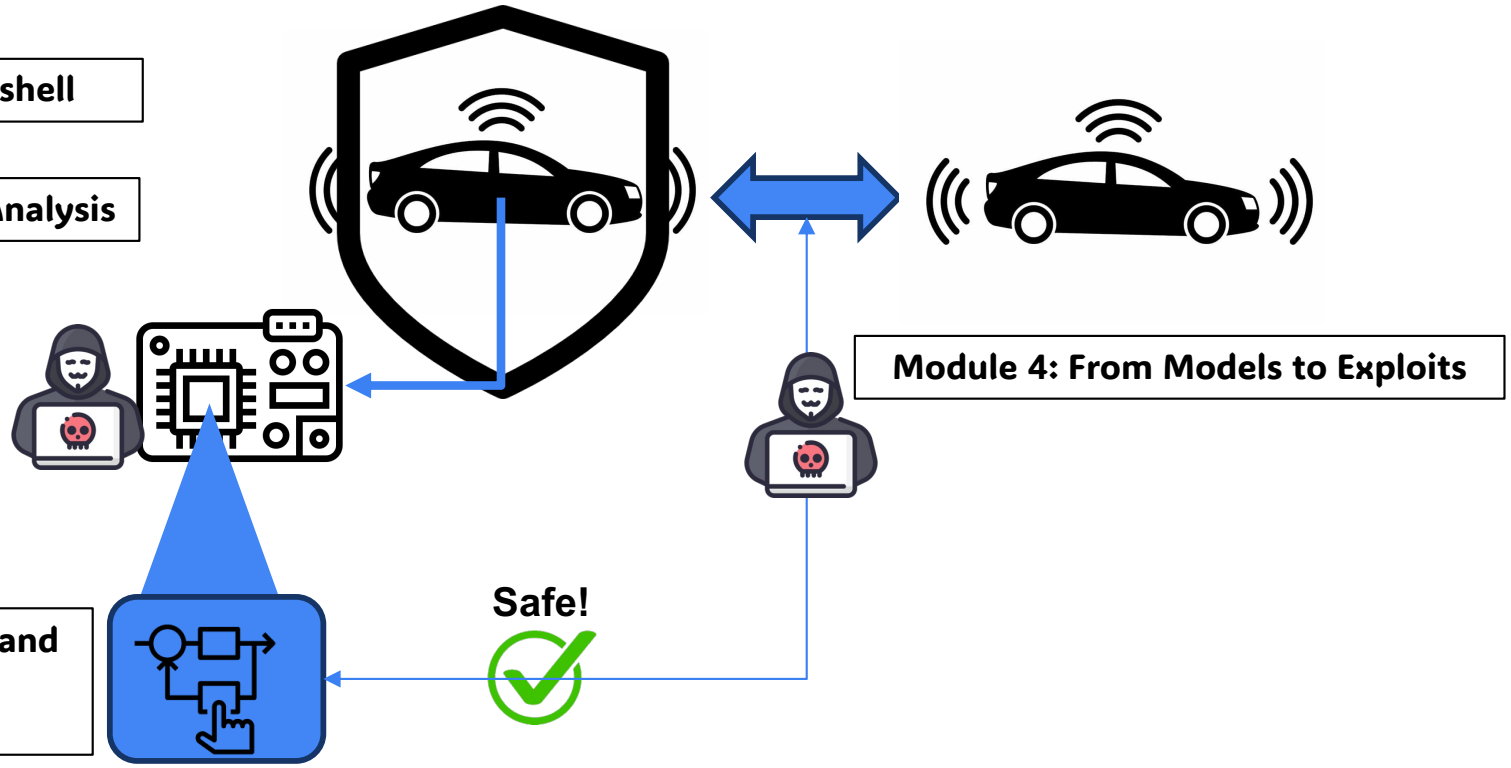


Black-box models == black-box guarantees?

Topics covered in the first half of my course*

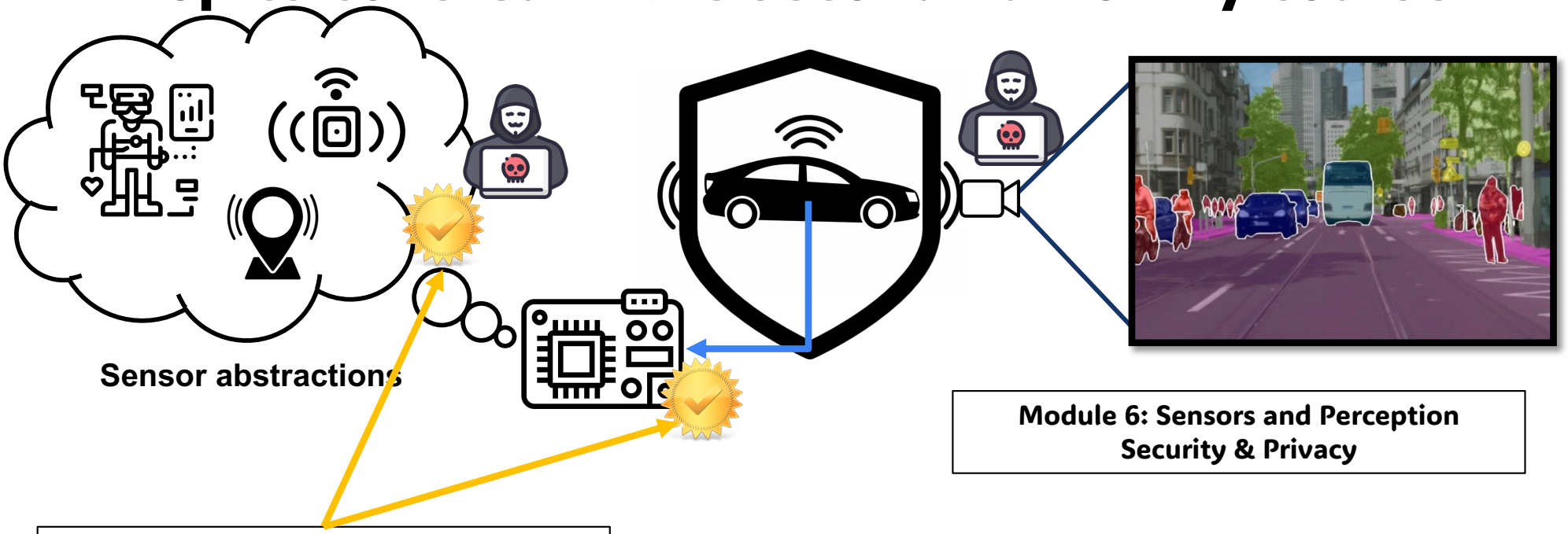
Module 1: Security in a nutshell

Module 2: CPS/IoT Program Analysis



Module 3: Formal Modeling and Verification of CPS

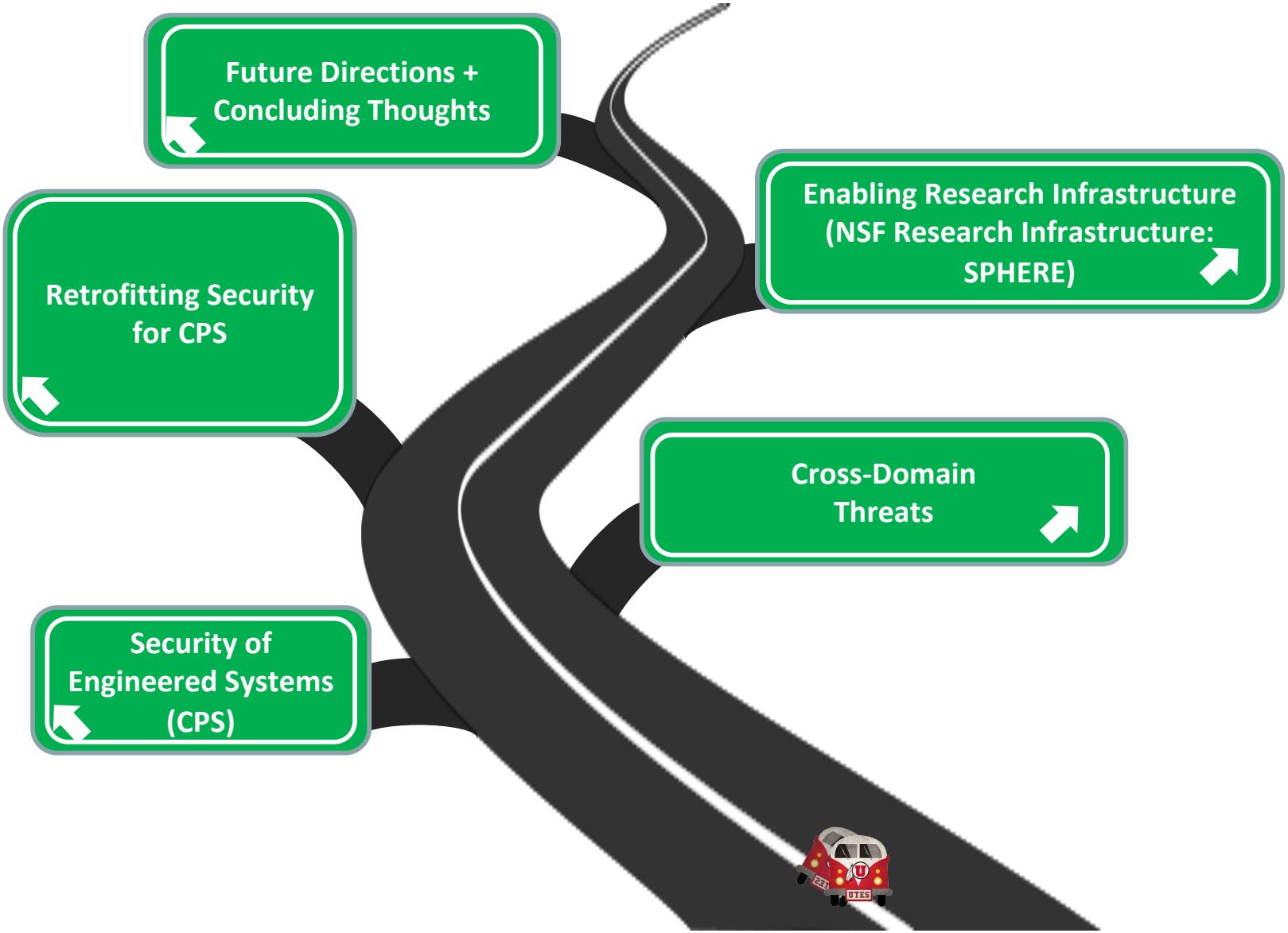
Topics covered in the second half of my course*



Module 6: Sensors and Perception Security & Privacy

Module 7: Establishing Trust for CPS

- Trusted computing/Hardware Support
- Remote Attestation
- Explainable and Verifiable Inferences


Future Directions + Concluding Thoughts
Enabling Research Infrastructure (NSF Research Infrastructure: SPHERE)
Retrofitting Security for CPS
Cross-Domain Threats
Security of Engineered Systems (CPS)

Cyber-Physical Industrial Control Systems (ICS)

Wate

INSIDE THE CUNNING

THE WORLD'S FIRST

UNPRECEDEN

ON

Attacker Objective:

Maximize Physical Impact through Stealthiness

ANDY GREENBERG

A Hac

The attacker u

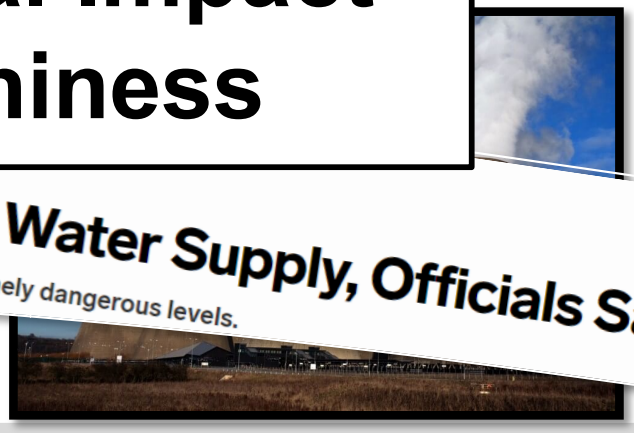
sodium hydroxide levels in the Oldsmar, Florida, water supply to extremely dangerous levels.

Hack attack

steel works

22 December 2014 | Technology

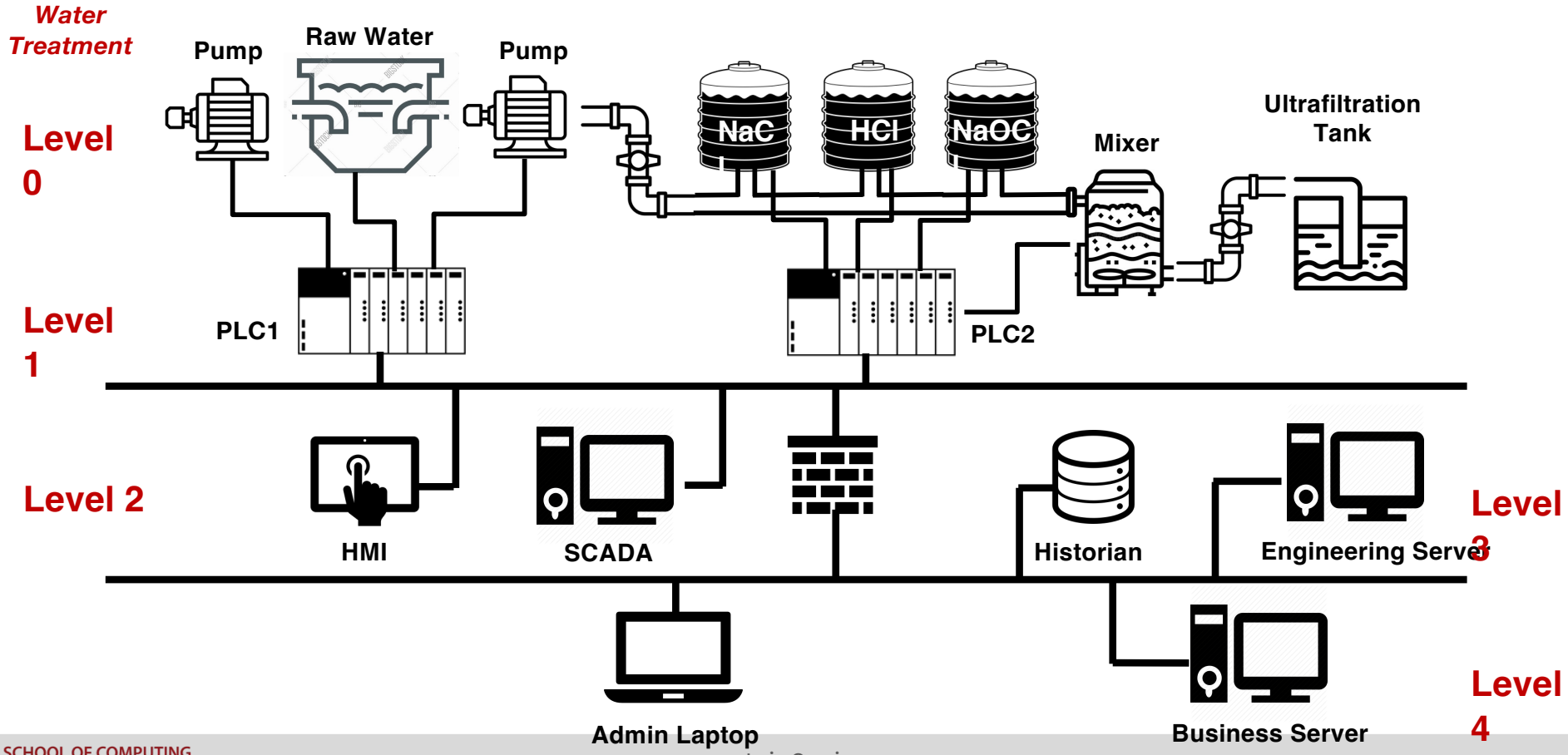
a Florida City's Water Supply, Officials Say



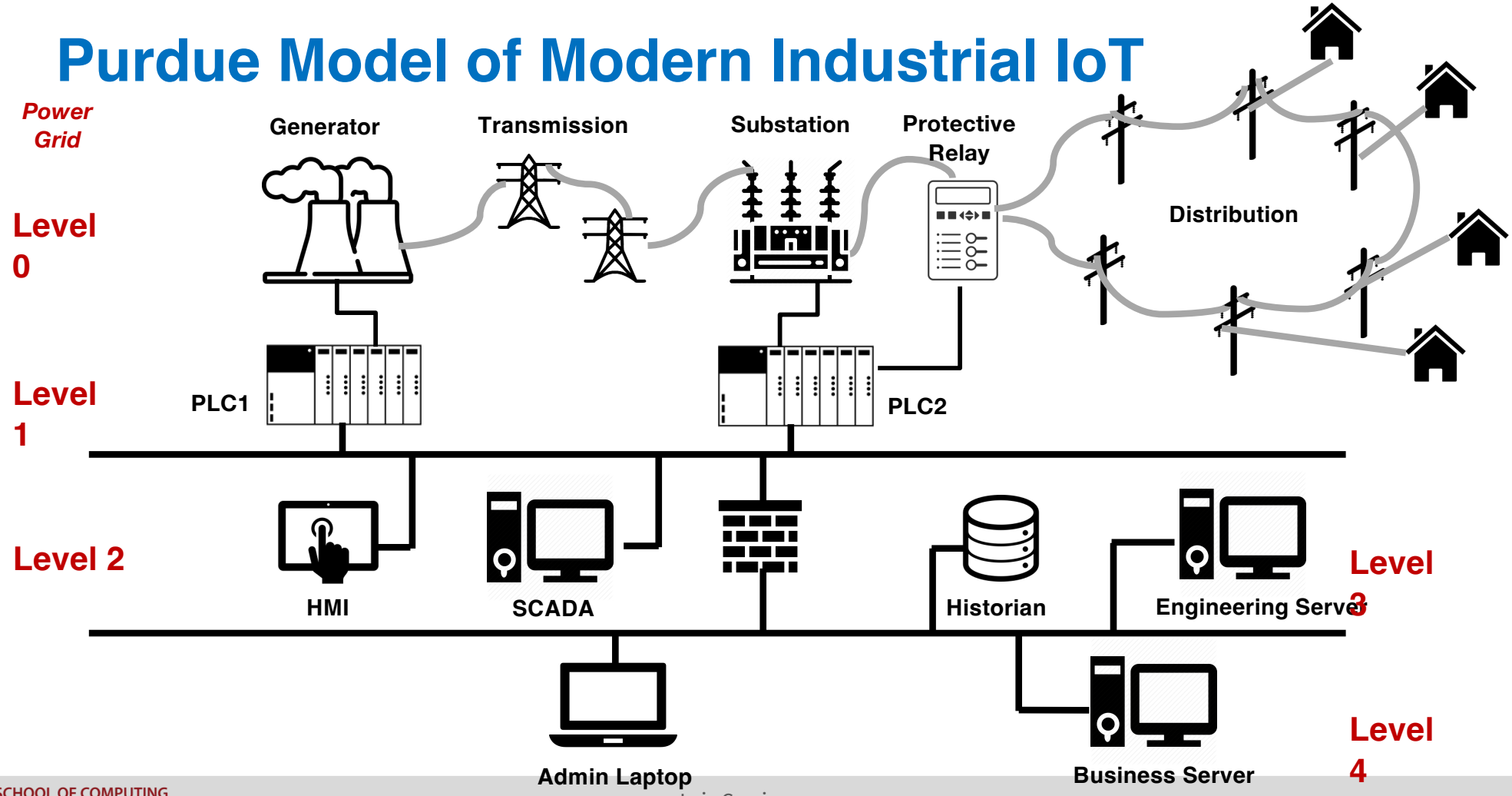
Electric Power Grid

Nuclear Reactor

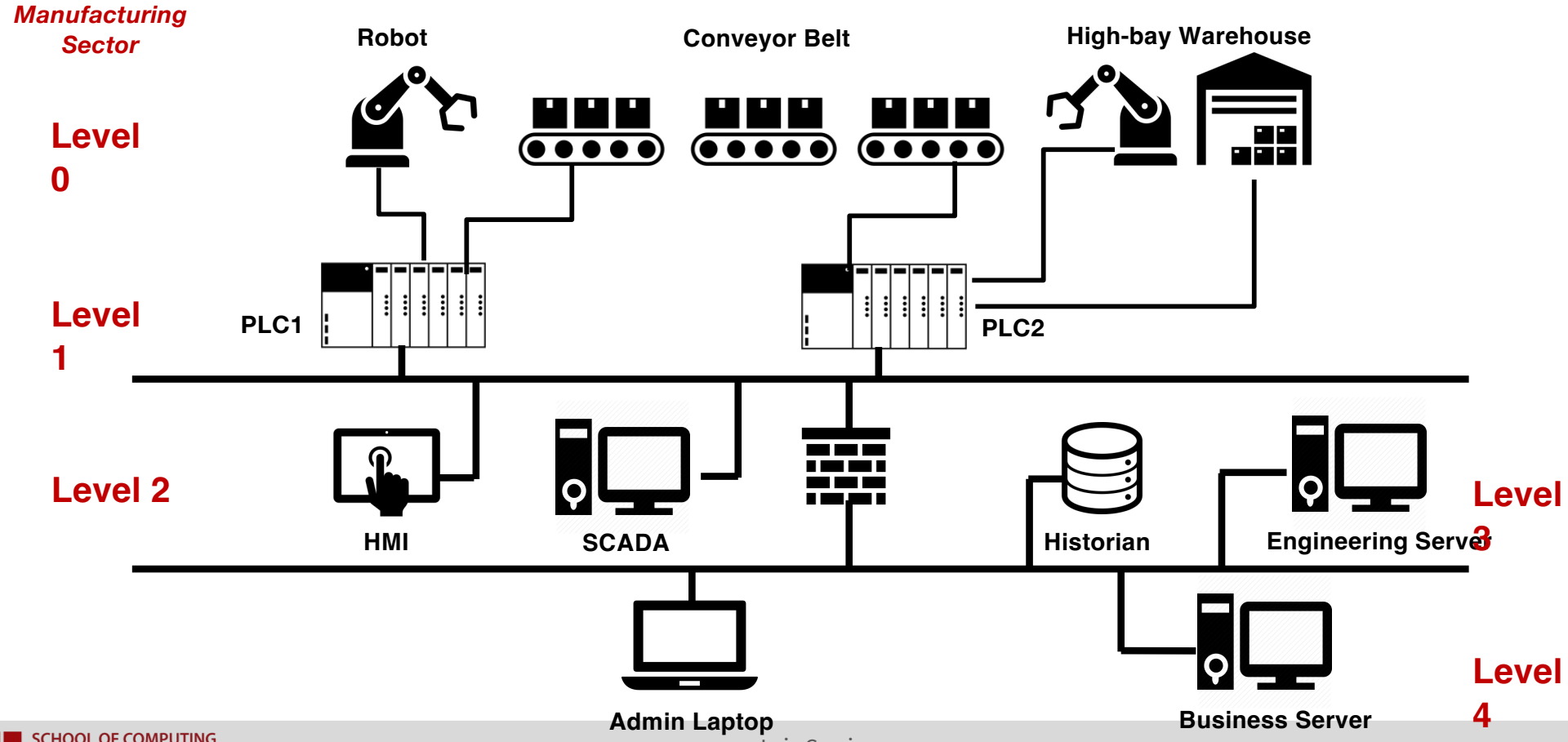
Purdue Model of Modern Industrial IoT



Purdue Model of Modern Industrial IoT

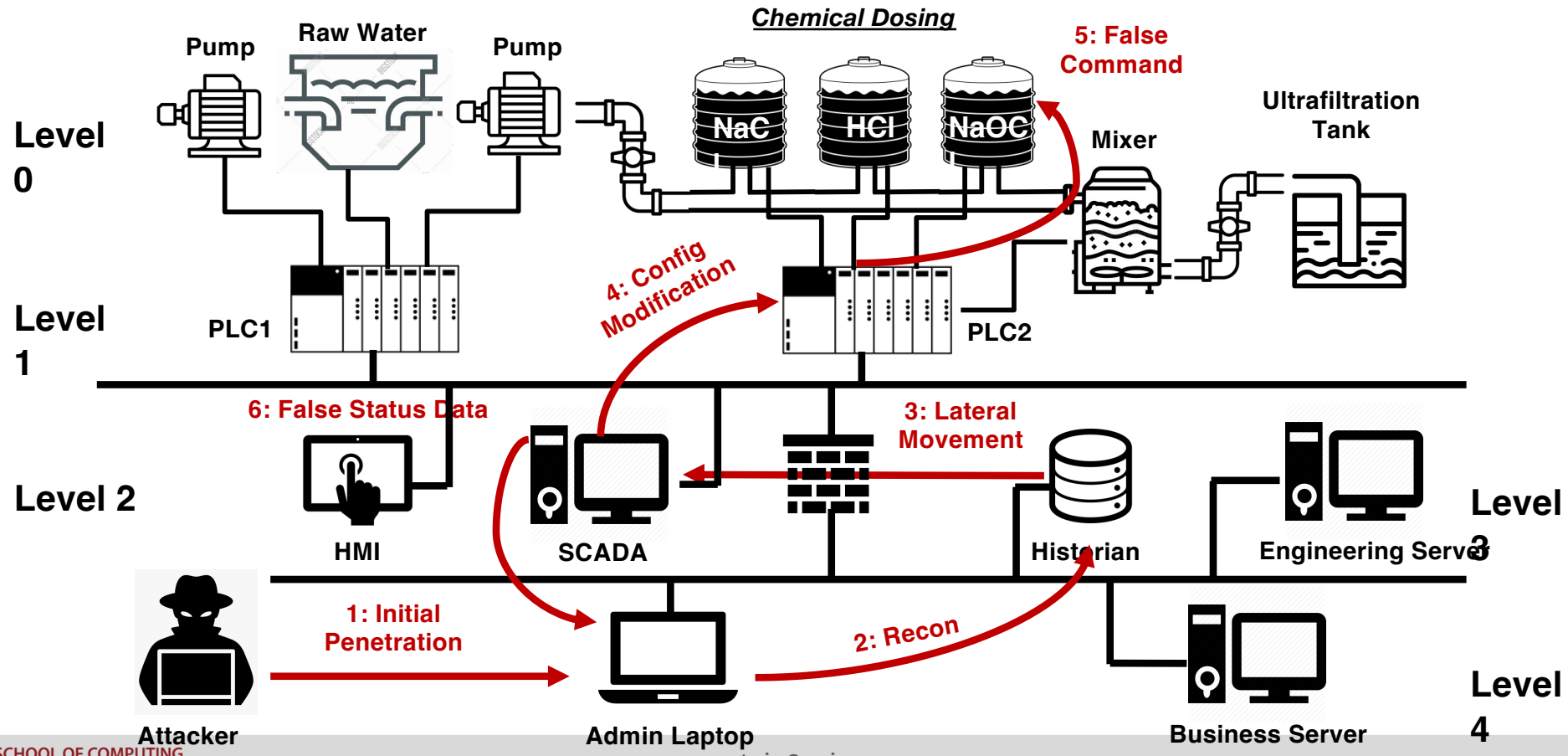


Purdue Model of Modern Industrial IoT



IIoT Attacks

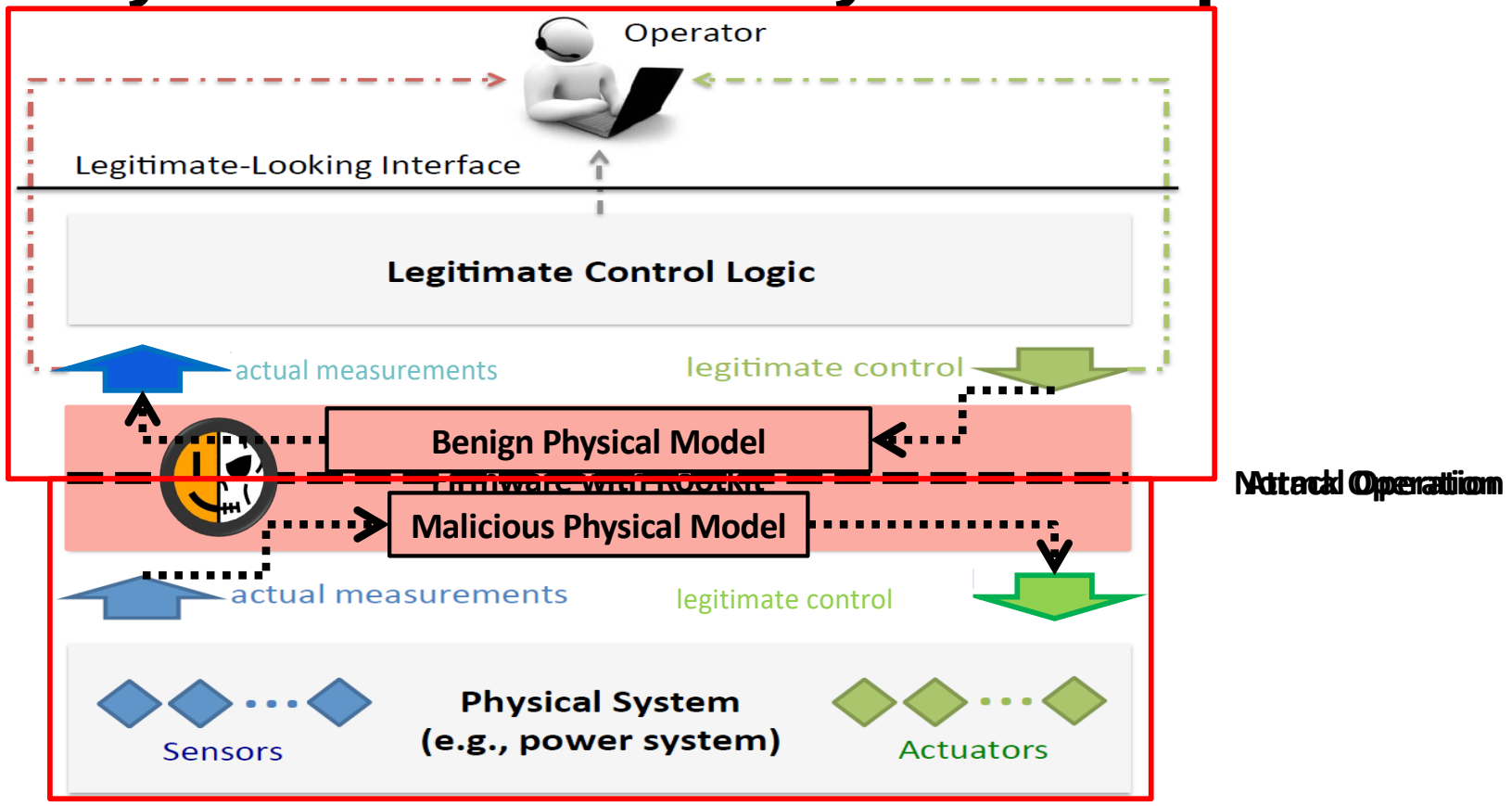
Oldsmar, Florida Water Treatment Plant Attack+



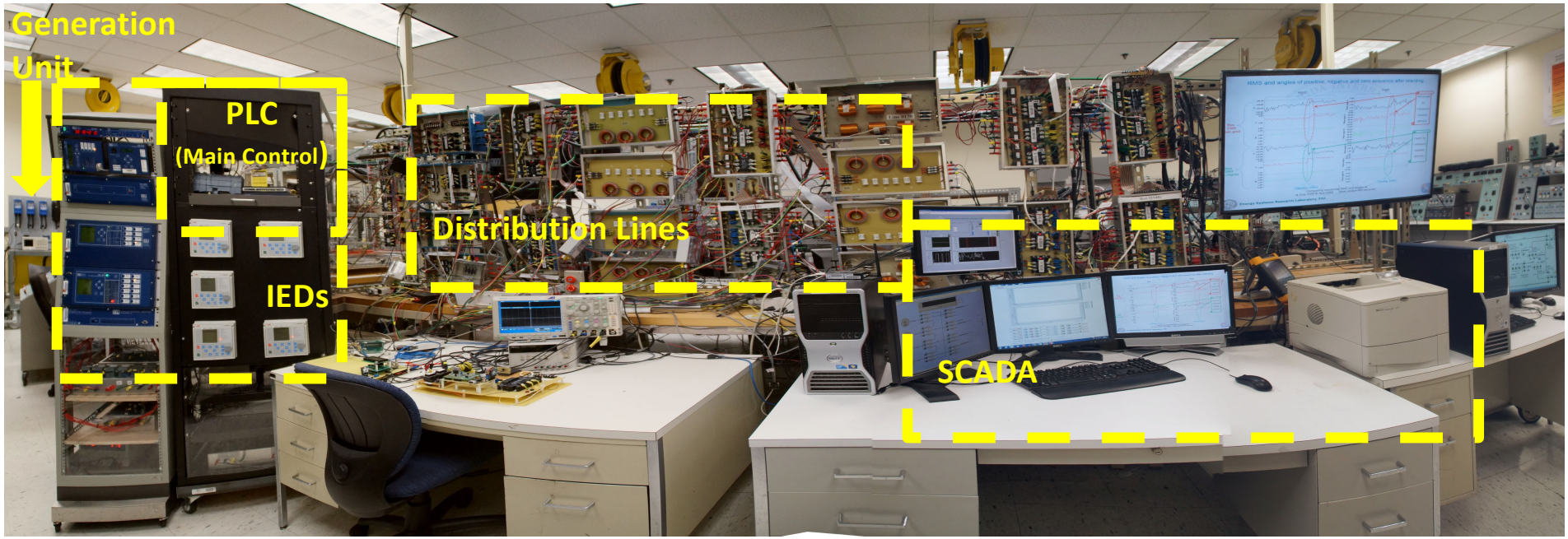
More Advanced Attack: Harvey: Model-Aware Rootkit (NDSS 2017)

- A malware that takes into account the physical topology of the ICS
- Model
  - Uses physical models to optimize control commands for an adversarial objective function
- PLC infection: compromising the PLC's firmware
  - Utilize the firmware update mechanism to replace firmware over the network
  - Local firmware modifications, e.g., SD card or JTAG implantation
  - Run-time attacks, e.g., network exploits or remote code execution vulnerabilities (FrostyURL)

Physics-Awareness: 2-Way Data Manipulation

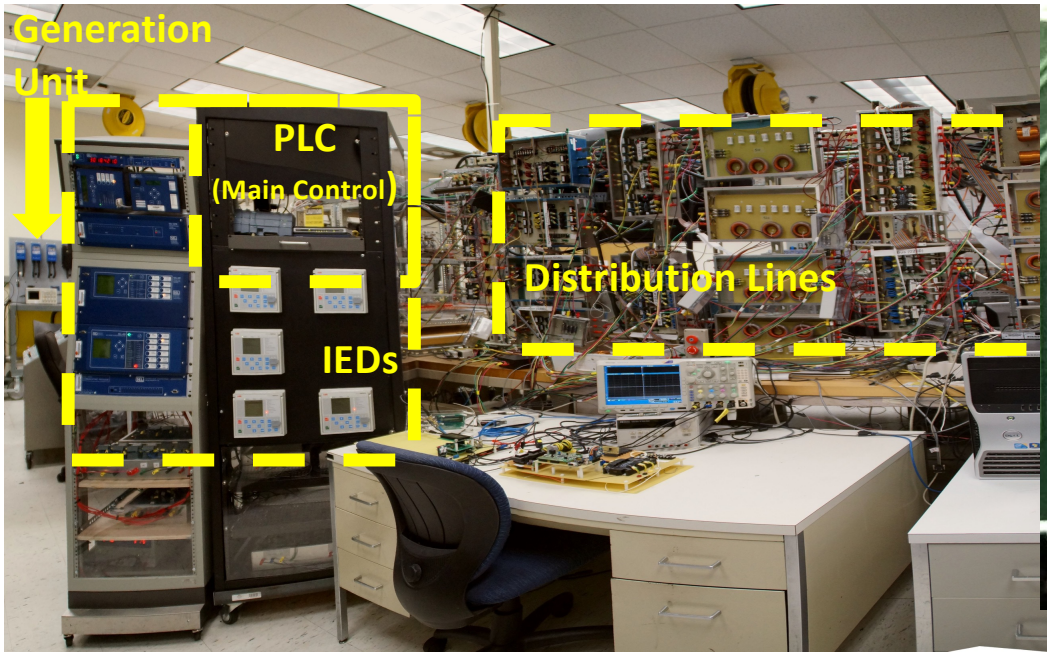


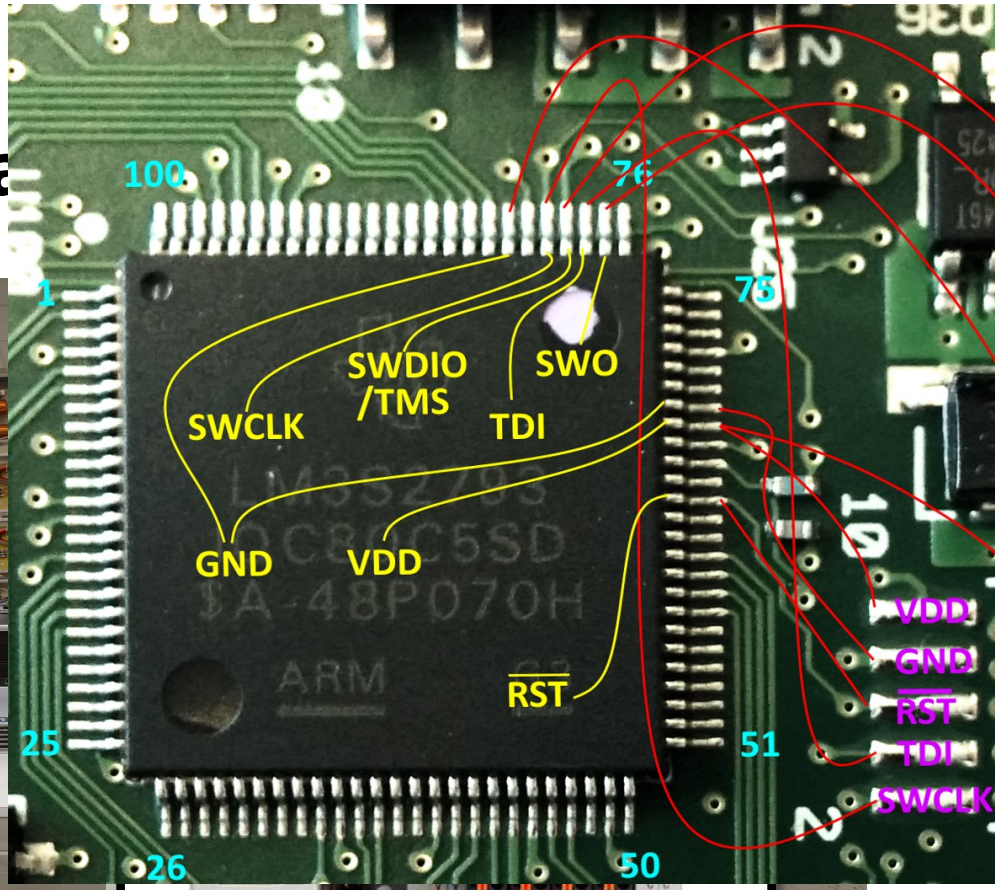
Evaluating Physics-Aware Malware on a Power System



FIU Power System Testbed

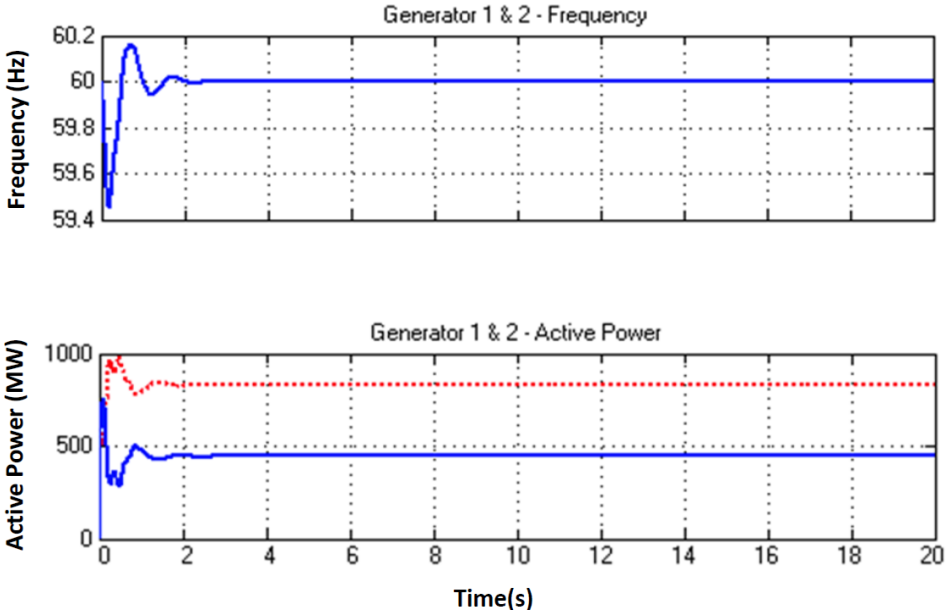
Evaluating Physics-Aware Ma

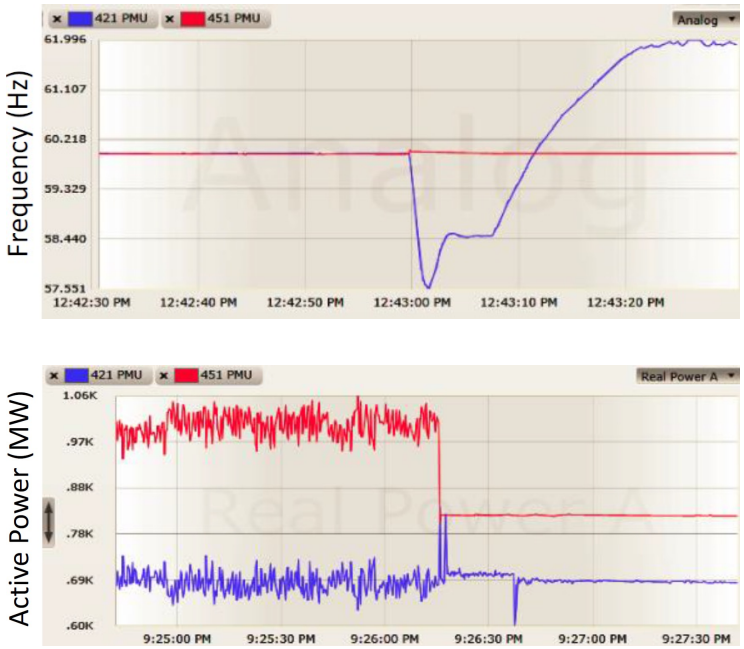




FIU Power System Testbed

Evaluating Physics-Aware Malware on a Power System





HMI Measurements

looks like stable operation…

Actual System Measurements

…in reality, it's unstable!

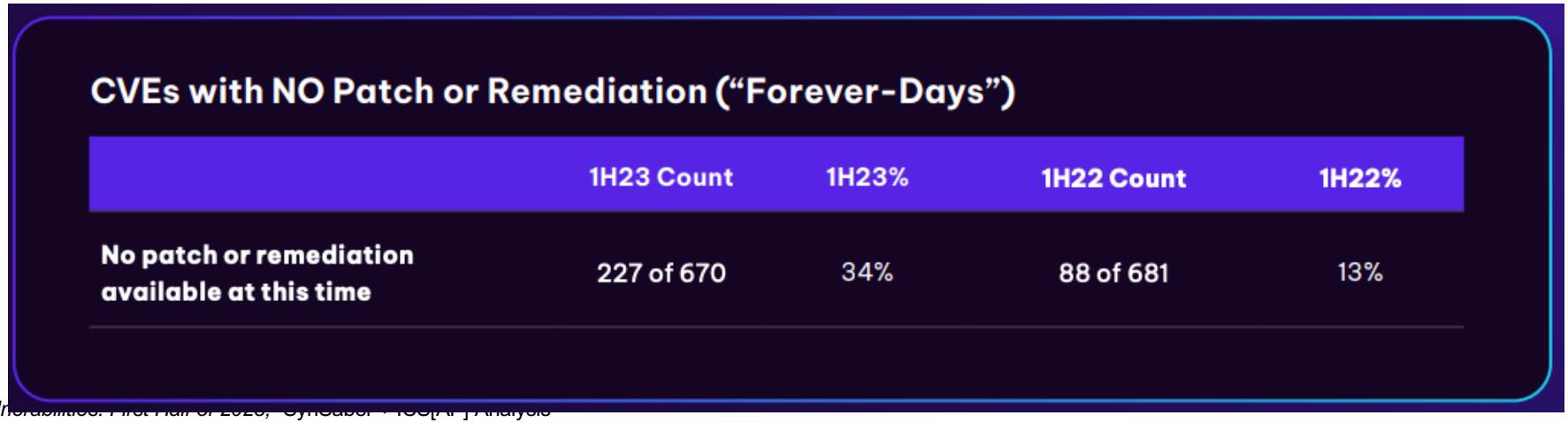
The Trouble with SCADA

- **Designed over 20 years ago**
  - Isolated
  - No security in mind
  - Access was local (not remote; no internet)
- **Today**
  - Connected to too many things
  - Remote control access
- **General Issues**
  - Lots of (vulnerable) proprietary protocols
  - Very hard to patch (reluctant vendors)

”Legal Informatics, Privacy, and Cyber Crime,” Sandro Etalle, ‘18

ICS Vulnerability Trends

- CISA ICS Advisory Report 2023
- **“Forever-Day” Vulnerabilities:**
  - Architectural and interoperability impacts
  - No simple way to “patch” a protocol vulnerability
  - Orgs have to deal with these CVEs for a long time

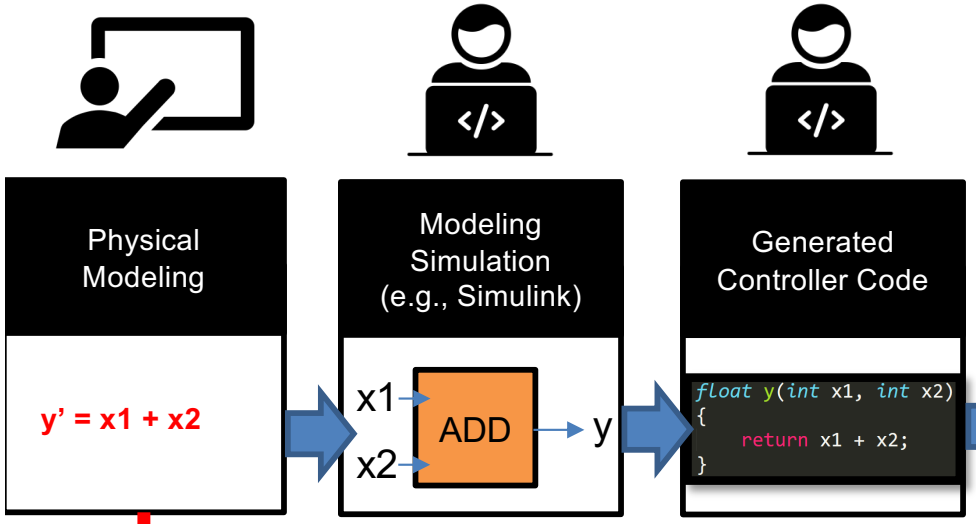


CVEs with NO Patch or Remediation (“Forever-Days”)

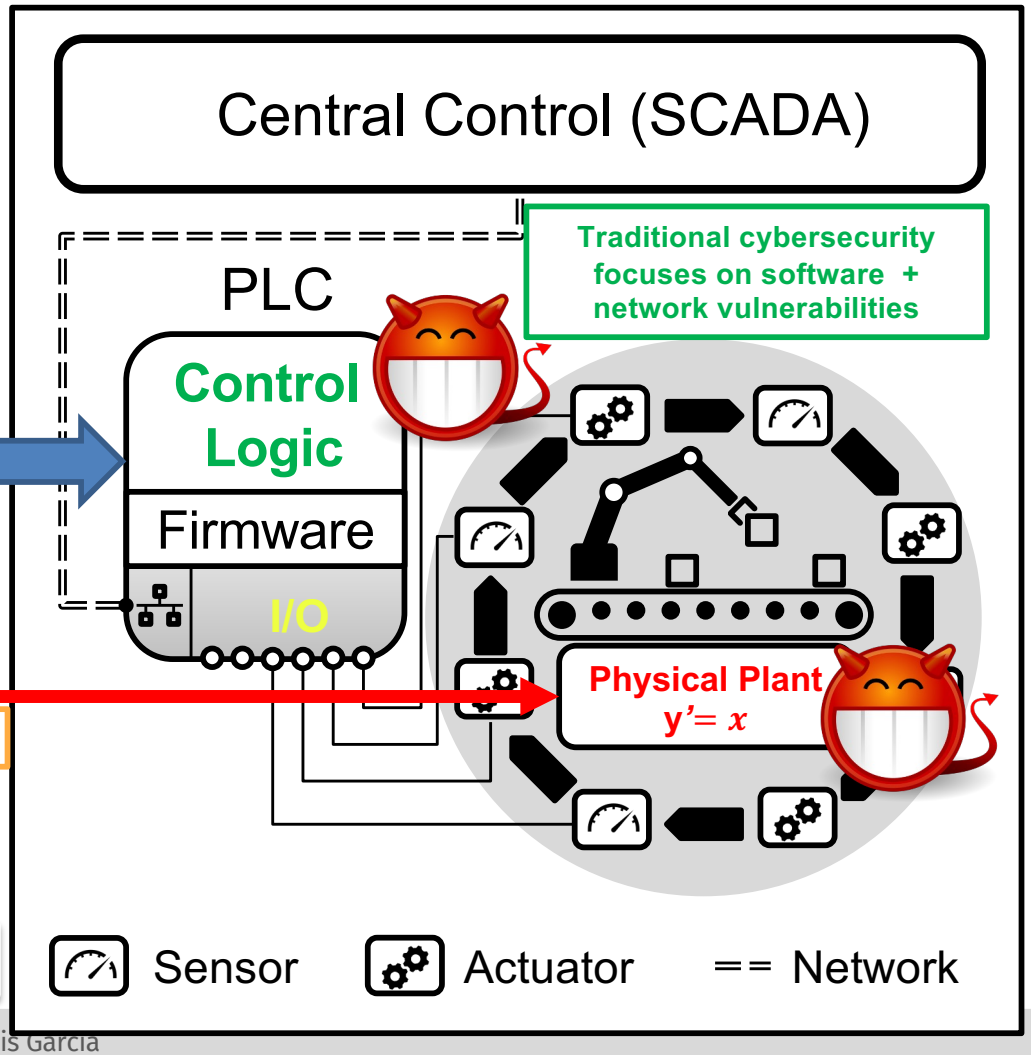
	1H23 Count	1H23%	1H22 Count	1H22%
No patch or remediation available at this time	227 of 670	34%	88 of 681	13%

“ICS Vulnerabilities: First Half of 2023,” CyHoBer + ICS[AI] Analysis

Control Systems Engineering Workflow

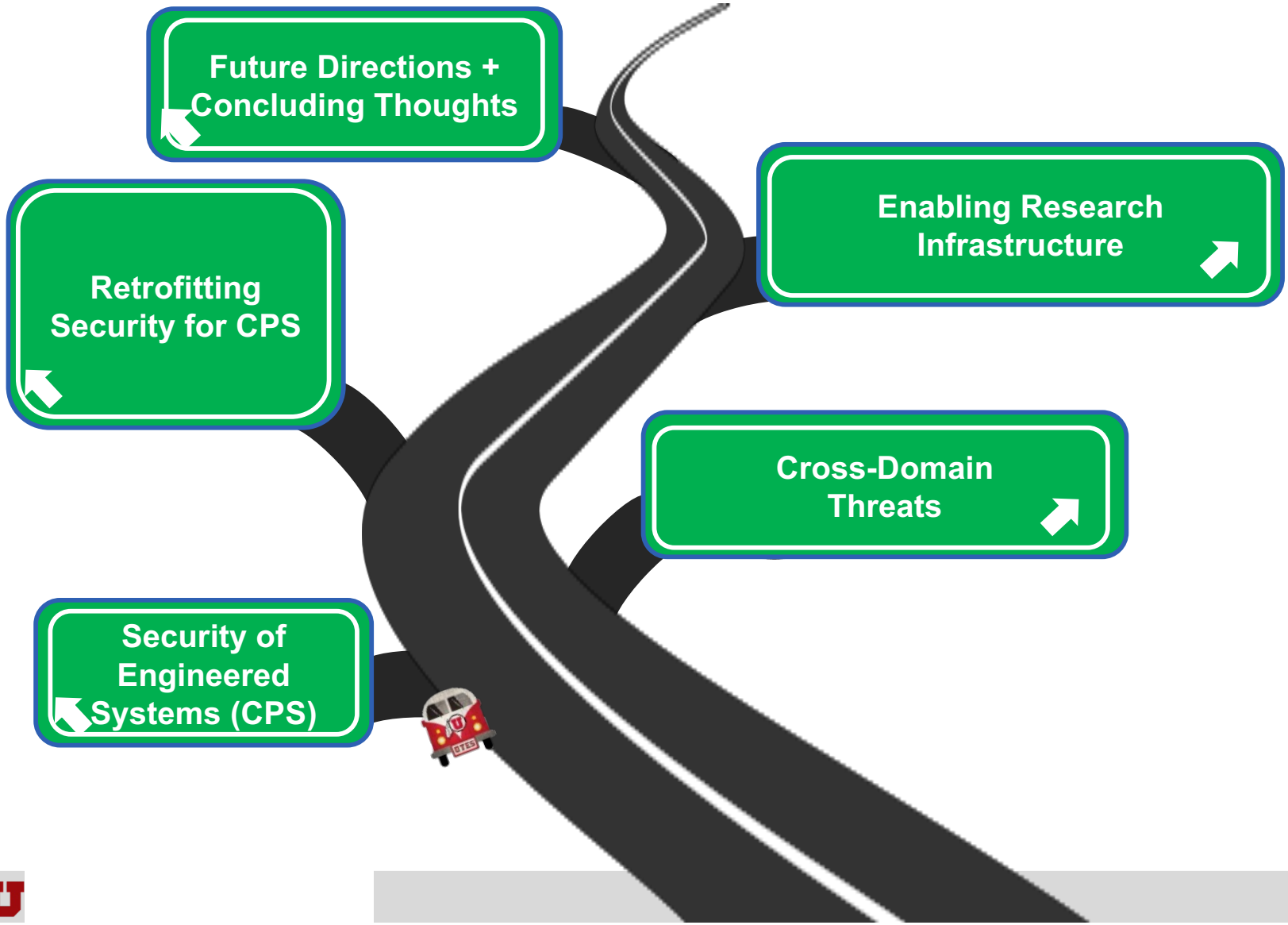


Industrial Control System



Opportunity: Leverage semantics to *retrofit* security!

CPS Security == "Security of Engineered Systems"

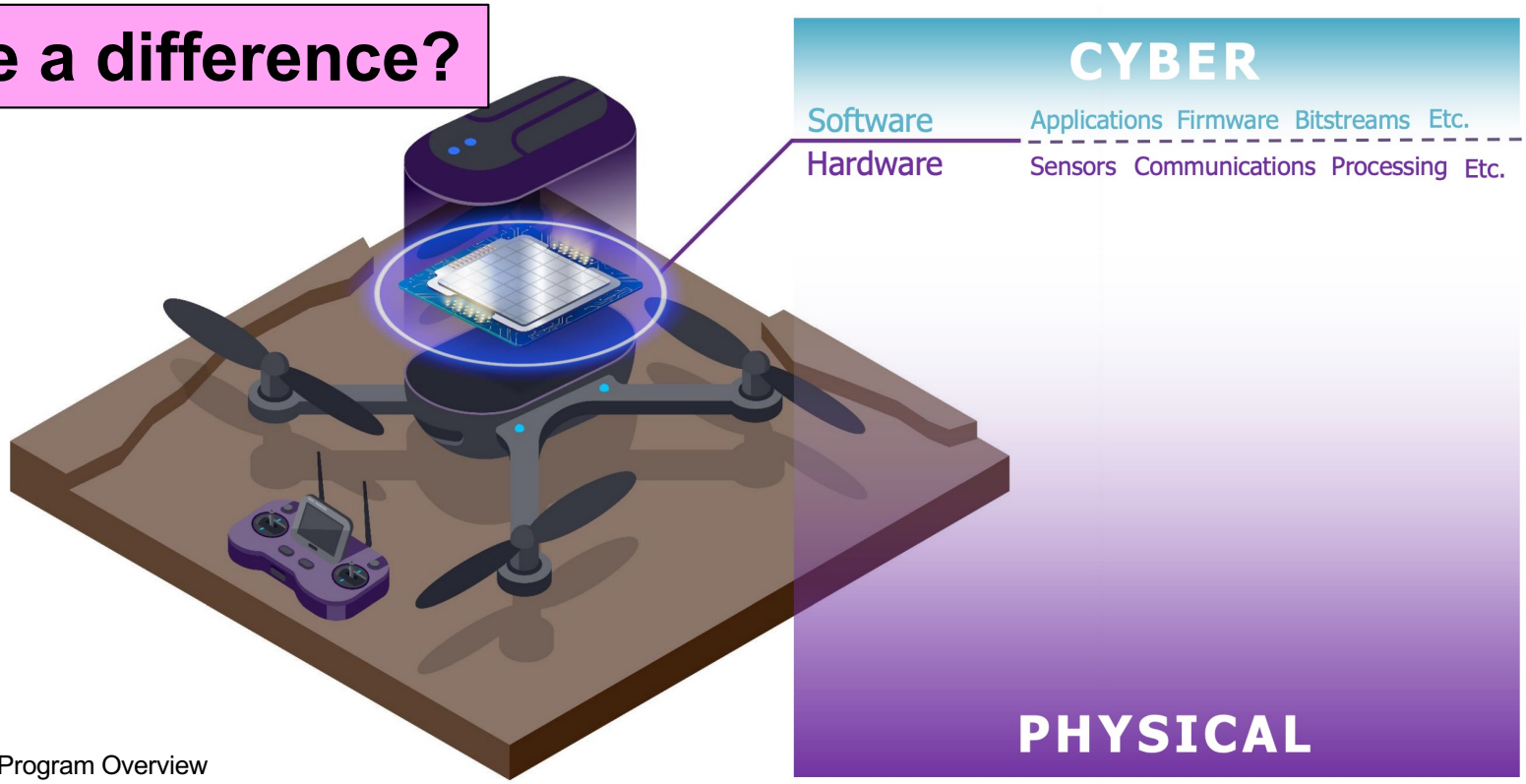

Future Directions + Concluding Thoughts
Enabling Research Infrastructure
Retrofitting Security for CPS
Cross-Domain Threats
Security of Engineered Systems (CPS)

Cross-domain (Cyber-physical) Threats: Information Technology (IT) vs Operational Technology (OT) Convergence?

- Security or Privacy Failures that may ***intentionally lead to physical threats***
- For instance, a cyber threat in a smart home may lead to
  - *physical harm to things, environments, or occupants*
  - *Financial harm to owners of environments*
  - *Reputational or financial harm to owner or occupants through exposure of personal information*

Cyber Threats vs. Cyber-physical Threats

Is there a difference?



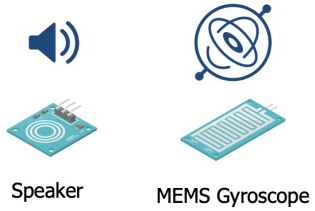
L. Yan, DAPRA FIRE Program Overview

Cyber-physical Threats

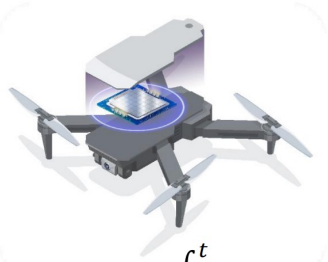
DARPA FIRE Program:

"Vulnerabilities that arise from the composition of hardware, software, and physical sub-systems where each component may not be vulnerable in-and-of itself."

0. Wait until quadcopter is in flight



1. Use speaker to inject false readings in Z-axis



$$u(t) = K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{de(t)}{dt}$$

2. Errant readings lead to wild swings in error $e(t)$ causing wild swings in output $u(t)$

3. Crash

L. Yan, DAPRA FIRE Program Overview

Recent ICS Vulnerability Trends: Common Vulnerabilities Across *Physical Domains*

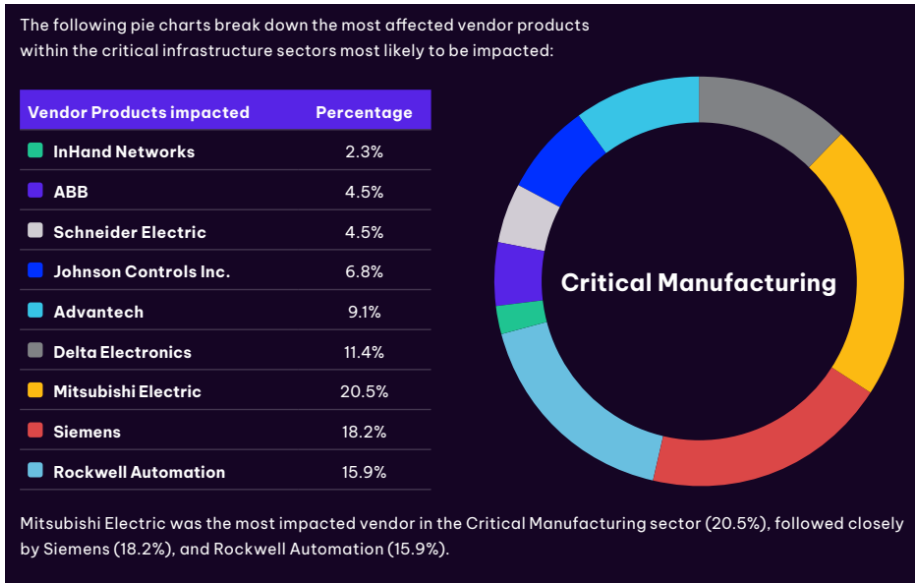
- CISA ICS Advisory Report 2023: **670** new ICS-related vulnerabilities in **first half** of 2023

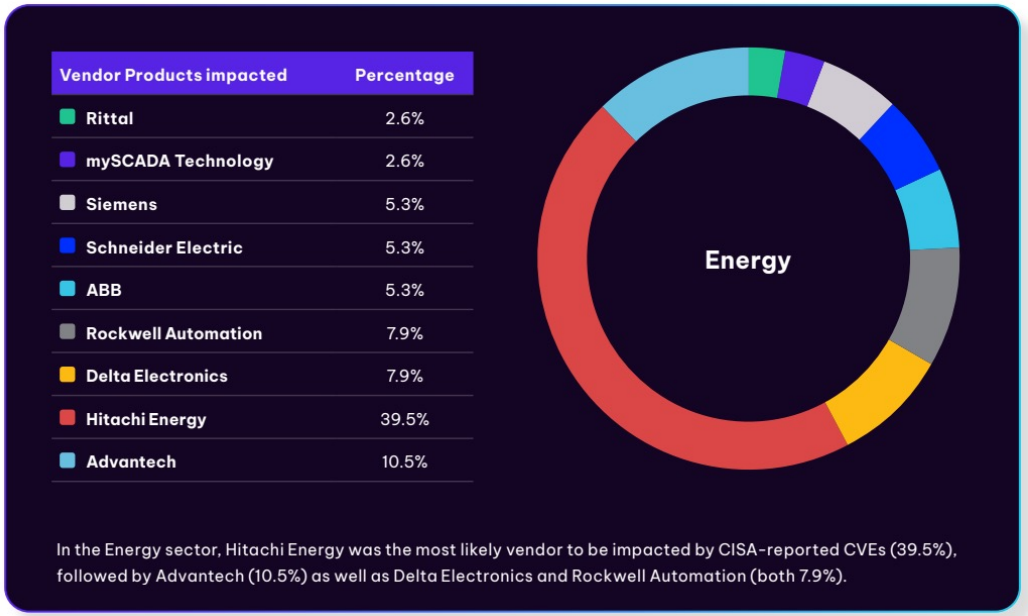
Critical Infrastructure Sector Most Likely to be Impacted by the CISA ICS Advisory	1H23 CISA ICS Advisory Count (note that some advisories impact multiple sectors)
Chemical	7
Commercial Facilities	11
Communications	10
Critical Manufacturing	69
Dams	2
Energy	45
Food & Agriculture	6
Government Facilities	4
Healthcare & Public Health	7
Information Technology	3
Multiple Critical Sectors*	66
Transportation Systems	8
Water & Wastewater Systems	16

"ICS CVE Research Report: First Half of 2023," SynSaber + ICS[AP] Analysis

Recent ICS Vulnerability Trends: Common Vulnerable Vendors Across *Physical* Domains

- CISA ICS Advisory Report 2023: **670** new ICS-related vulnerabilities in **first half** of 2023

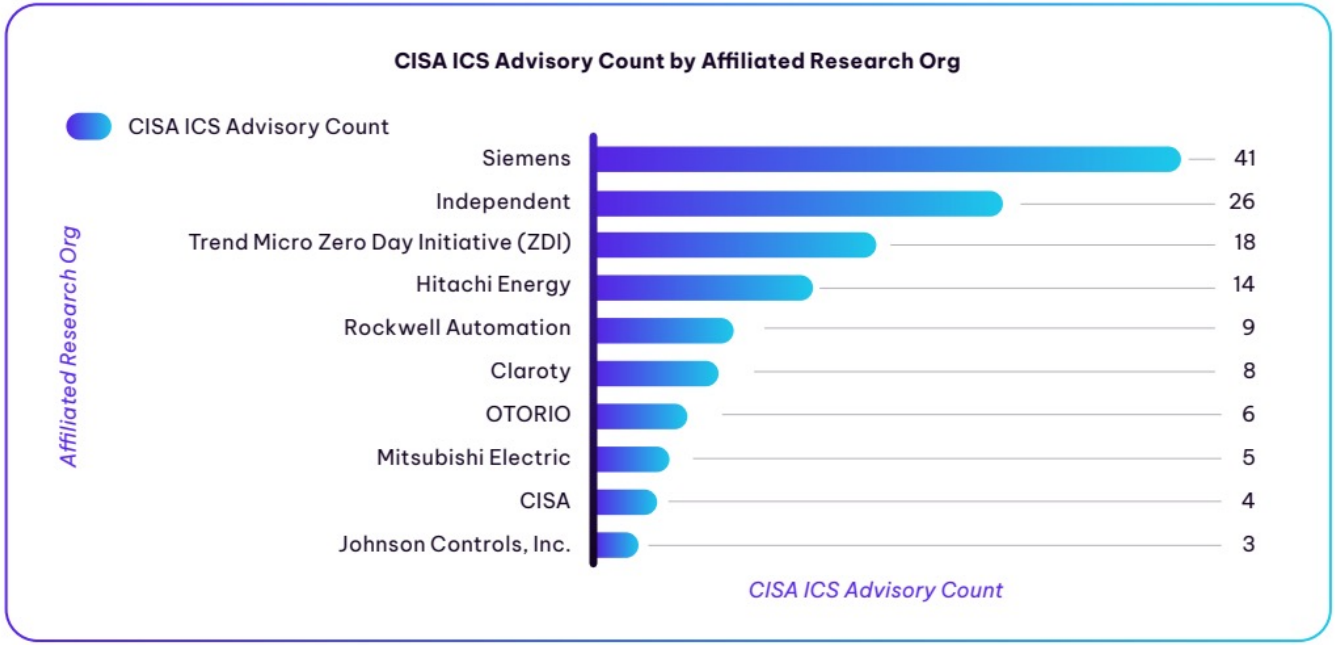




"ICS CVE Research Report: First Half of 2023," SynSaber + ICS[AP] Analysis

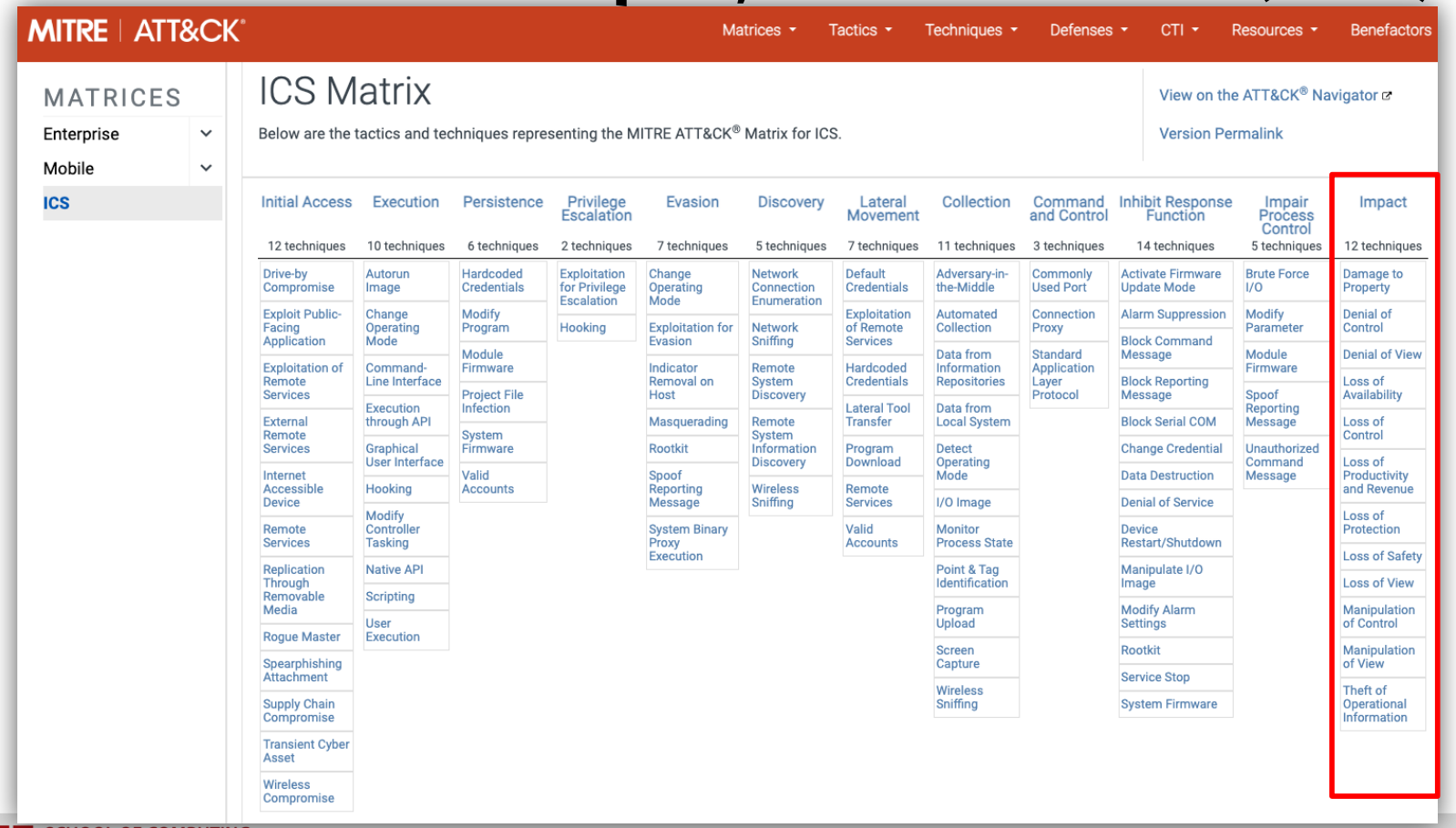
Recent ICS Vulnerability Trends: Common Vulnerable Vendors Across *Physical* Domains

- CISA ICS Advisory Report 2023



"ICS Vulnerabilities: First Half of 2023," SynSaber + ICS[AP] Analysis

More General Taxonomy of ICS Attack Tactics, Techniques, and Procedures (TTPs)



MITRE | ATT&CK®

Matrices ▾ Tactics ▾ Techniques ▾ Defenses ▾ CTI ▾ Resources ▾ Benefactors

MATRICES

Enterprise ⌄

Mobile ⌄

ICS

ICS Matrix

Below are the tactics and techniques representing the MITRE ATT&CK® Matrix for ICS.

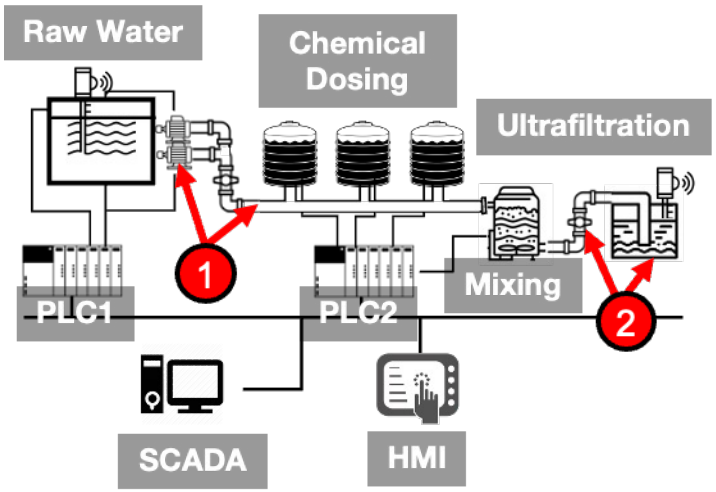
View on the ATT&CK® Navigator

Version Permalink

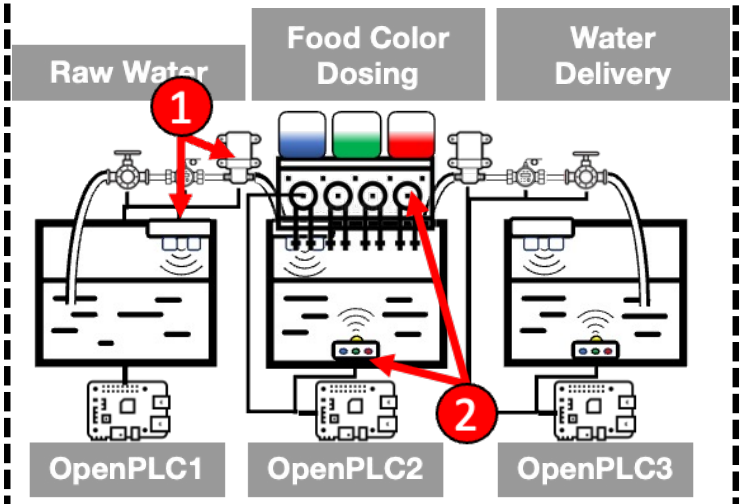
Initial Access	Execution	Persistence	Privilege Escalation	Evasion	Discovery	Lateral Movement	Collection	Command and Control	Inhibit Response Function	Impair Process Control	Impact
12 techniques	10 techniques	6 techniques	2 techniques	7 techniques	5 techniques	7 techniques	11 techniques	3 techniques	14 techniques	5 techniques	12 techniques
Drive-by Compromise	Autorun Image	Hardcoded Credentials	Exploitation for Privilege Escalation	Change Operating Mode	Network Connection Enumeration	Default Credentials	Adversary-in-the-Middle	Commonly Used Port	Activate Firmware Update Mode	Brute Force I/O	Damage to Property
Exploit Public-Facing Application	Change Operating Mode	Modify Program	Hooking	Exploitation for Evasion	Network Sniffing	Exploitation of Remote Services	Automated Collection	Connection Proxy	Alarm Suppression	Modify Parameter	Denial of Control
Exploitation of Remote Services	Command-Line Interface	Module Firmware		Indicator Removal on Host	Remote System Discovery	Hardcoded Credentials	Data from Information Repositories	Standard Application Layer Protocol	Block Command Message	Module Firmware	Denial of View
External Remote Services	Execution through API	Project File Infection		Masquerading	Remote System Information Discovery	Lateral Tool Transfer	Data from Local System		Block Reporting Message	Spoof Reporting Message	Loss of Availability
Internet Accessible Device	Graphical User Interface	System Firmware		Rootkit	Wireless Sniffing	Program Download	Detect Operating Mode		Block Serial COM	Unauthorized Command Message	Loss of Control
Remote Services	Hooking	Valid Accounts		Spoof Reporting Message		Remote Services	I/O Image		Change Credential		Loss of Productivity and Revenue
Replication Through Removable Media	Modify Controller Tasking			System Binary Proxy Execution		Valid Accounts	Monitor Process State		Data Destruction		Loss of Protection
Rogue Master	Native API						Point & Tag Identification		Denial of Service		Loss of Safety
Spearphishing Attachment	Scripting						Program Upload		Device Restart/Shutdown		Loss of View
Supply Chain Compromise	User Execution						Screen Capture		Manipulate I/O Image		Manipulation of Control
Transient Cyber Asset							Wireless Sniffing		Modify Alarm Settings		Manipulation of View
Wireless Compromise									Rootkit		Theft of Operational Information
									Service Stop		
									System Firmware		

What about common cross-physical semantics?

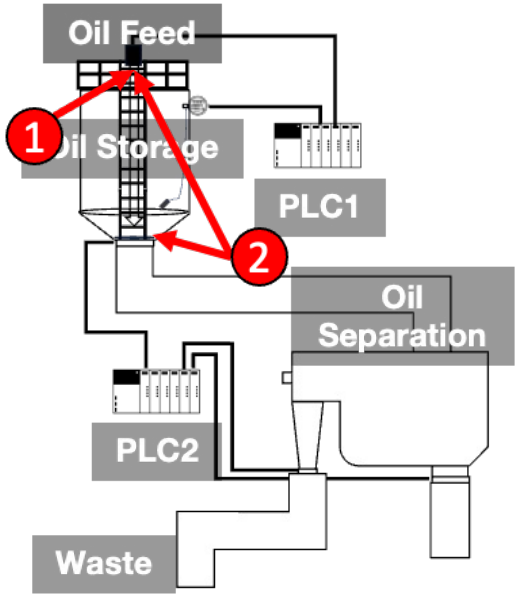
Recent Work: Inferring Cross-Domain *Physical* Threats (RICSS '24)



1. Attack: P1 + P2 ON
   Impact: Pipe Burst
2. Attack: MV2 Open
   Impact: Tank 2 Overflow
3. ...



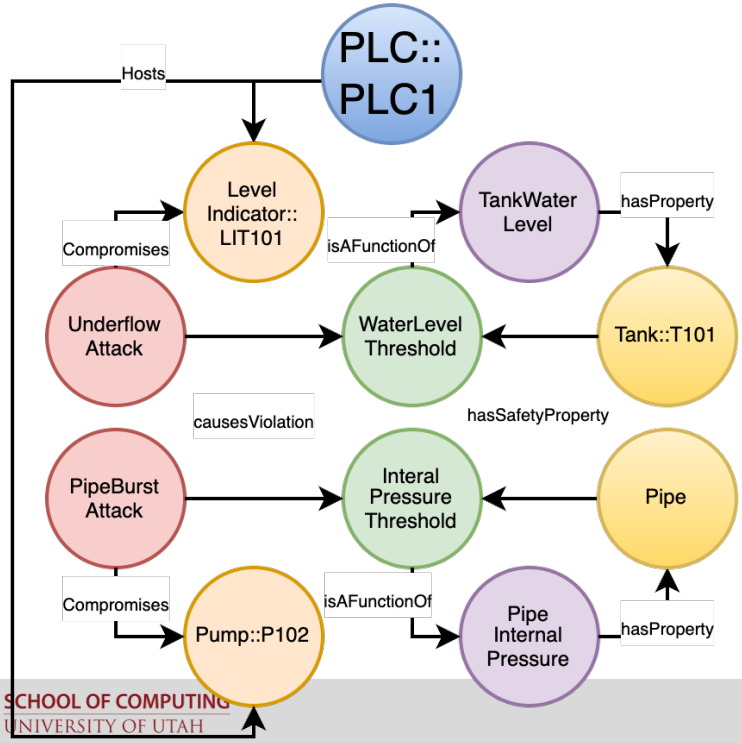
1. Attack: P1 ON
   Impact: Underflow
2. Attack: Doser Pump 3 ON
   Impact: Poisoned Treatment
3. ...

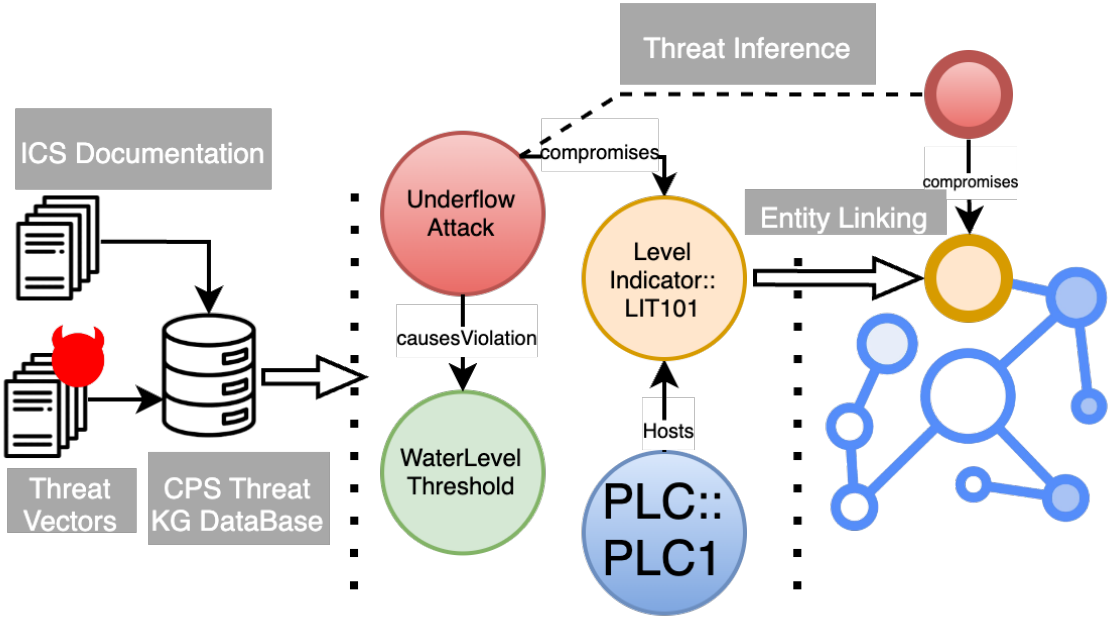


1. Attack: Feed Pump (FP)
   Impact: Overflow Tank1
2. Attack: FP + Valve
   Impact: Overflow Tank2
3. ...

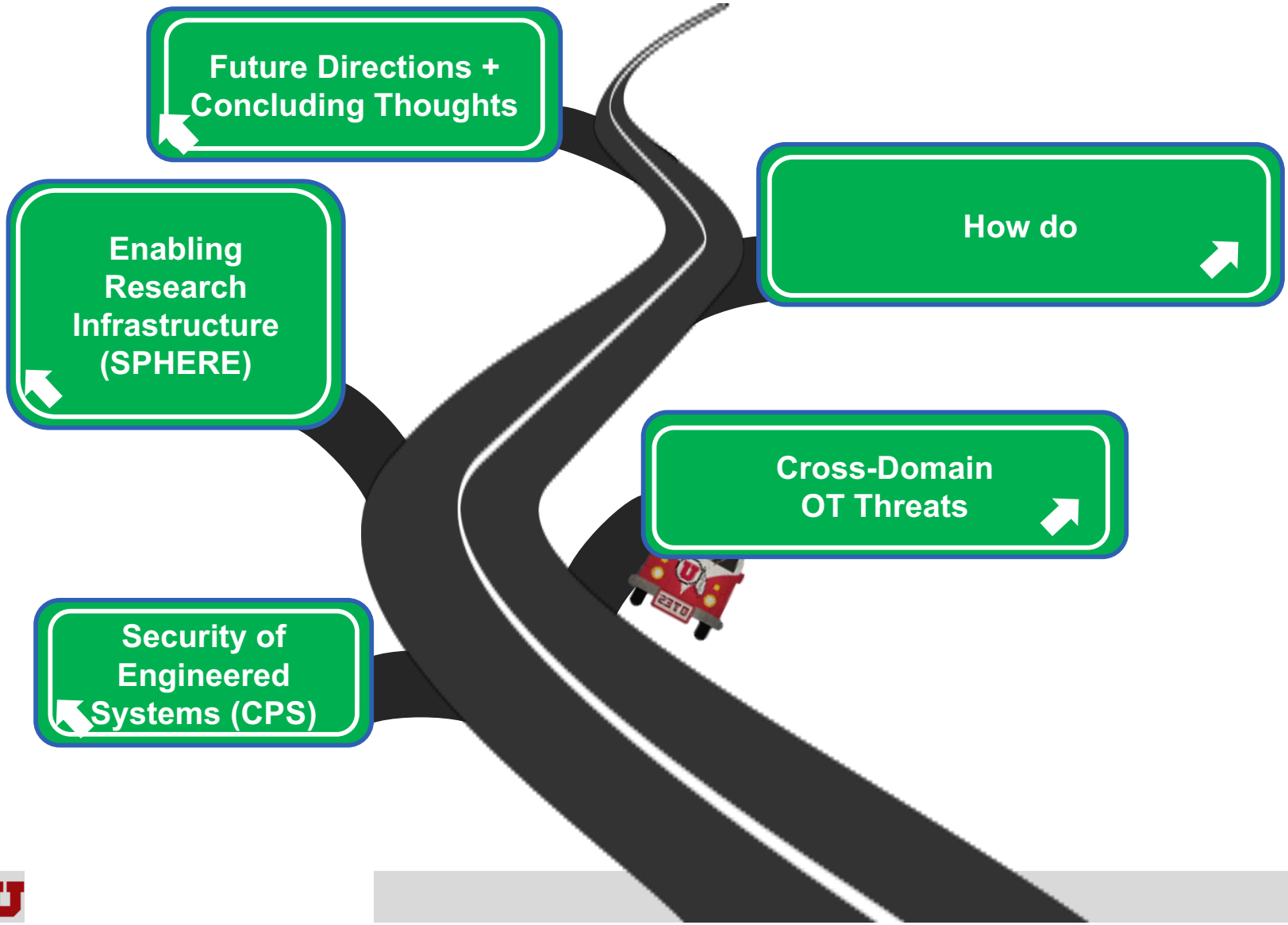
Recent Work: Inferring Cross-Domain *Physical* Threats (RICSS '24)

Create knowledge base of known attacks and their physical impact grounded in formal ontology for representing sensors/actuators



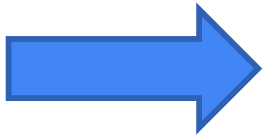


Leverage Knowledge Base to Infer Threats *Across* Domains


Future Directions + Concluding Thoughts
Enabling Research Infrastructure (SPHERE)
How do
Cross-Domain OT Threats
Security of Engineered Systems (CPS)

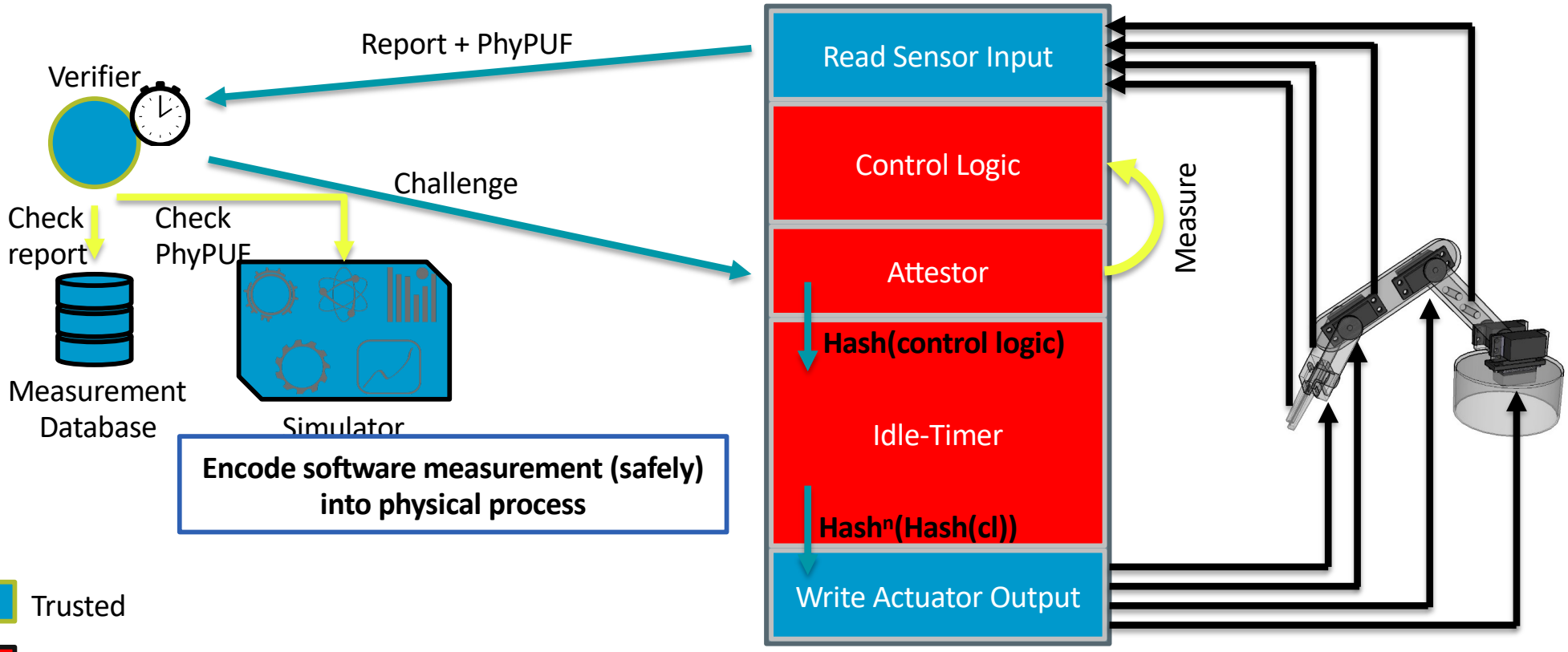
Retrofitting Security: Physics for the Sake of Security

- **Problem:** Legacy devices (such as PLCs) don't support a hardware trusted computing base (TCB) for remote attestation

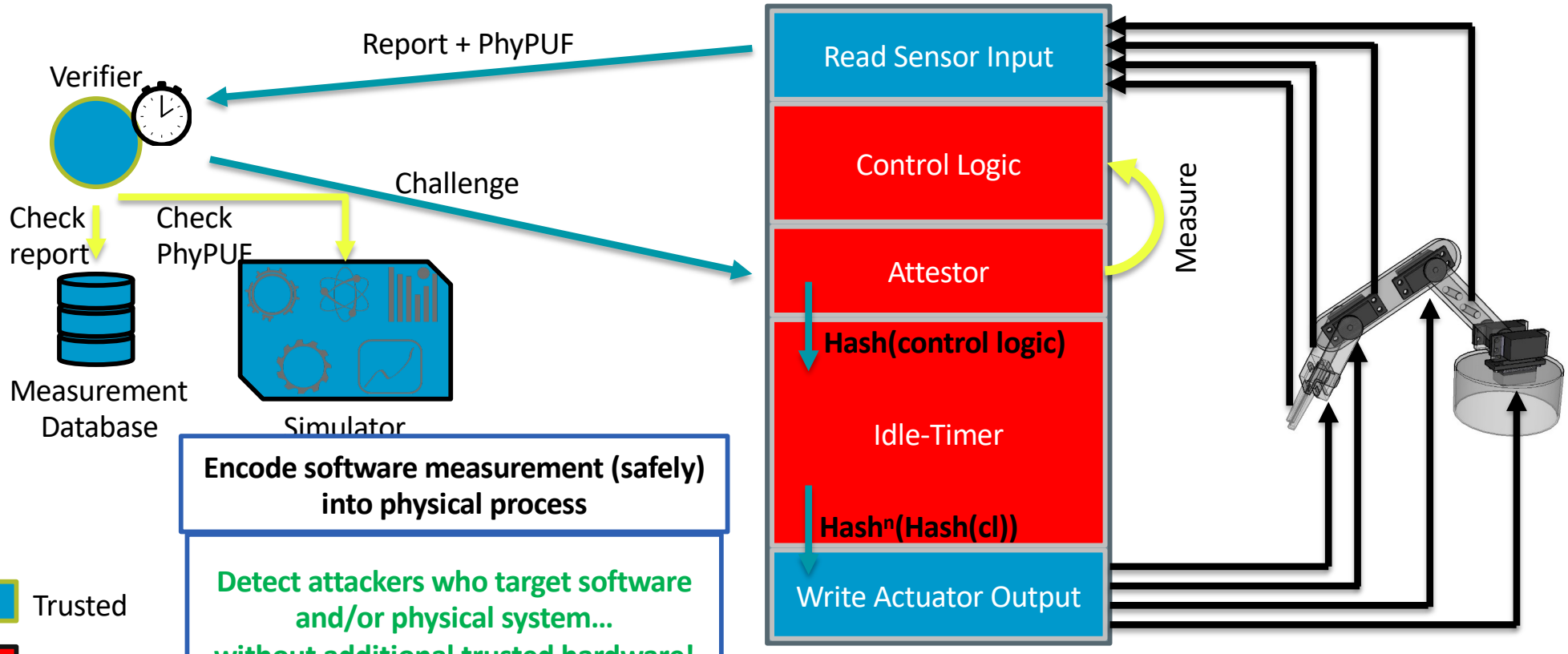


Patt: Physics-based Attestation of Control Systems (RAID '19)

Patt: Cyber-Physical Remote Attestation for PLCs

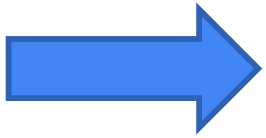


Patt: Cyber-Physical Remote Attestation for PLCs



More Defenses: Physics for the Sake of Security

- **Problem:** Legacy devices (such as PLCs) don't support a hardware trusted computing base (TCB) for remote attestation



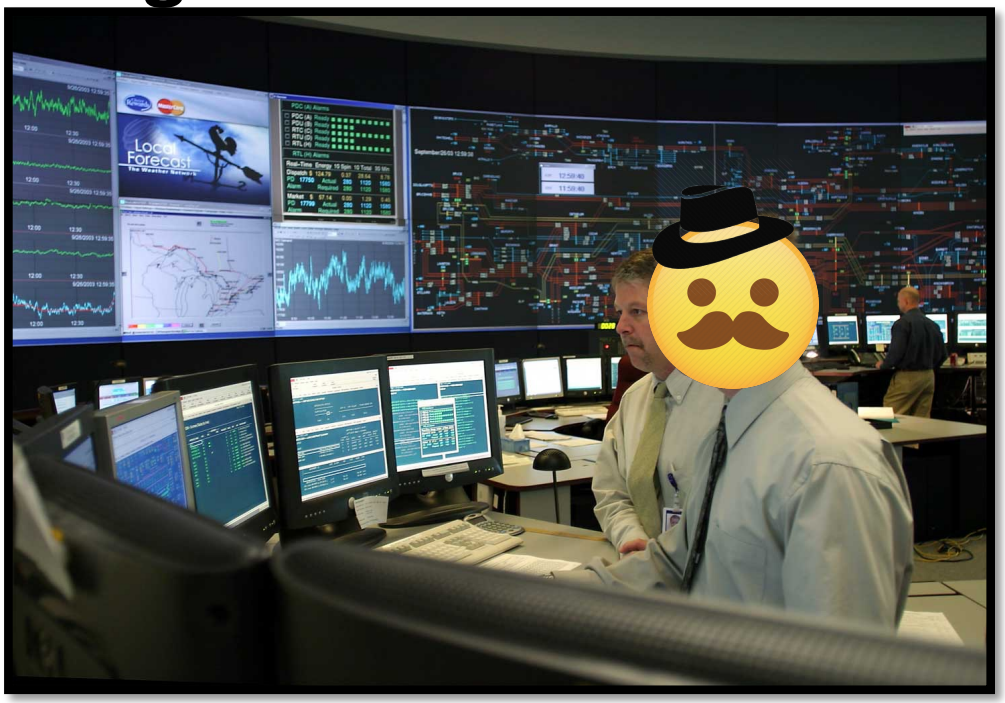
Patt: Physics-based Attestation of Control Systems (RAID '19)

- **Problem:** How can we monitor physical semantics in more general multi-process settings?

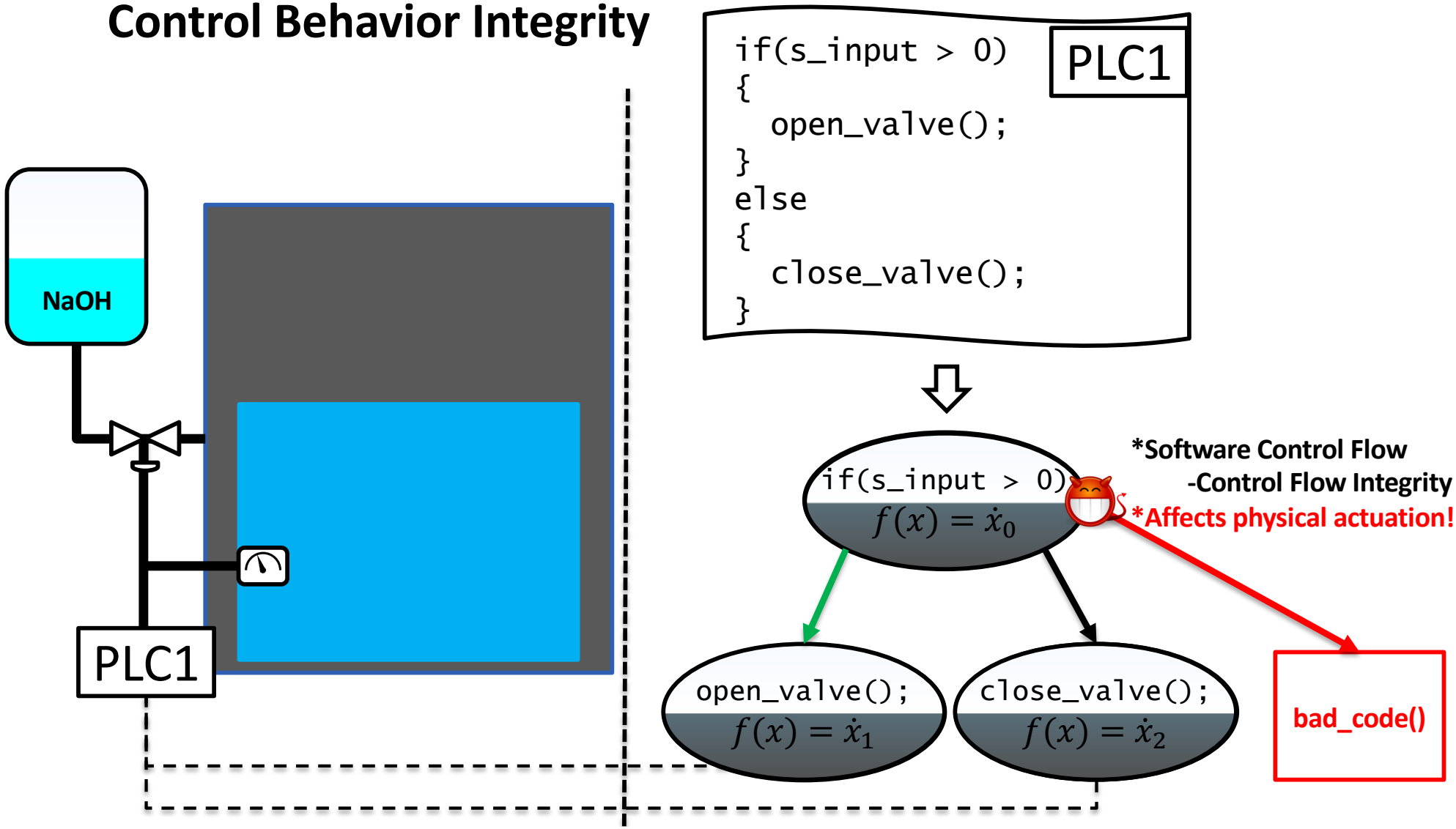
Cyber-physical Control Behavior Integrity (ICCPS '20)

Scadman: Control Behavior Intrusion Detection using Cyber-physical Digital Twins

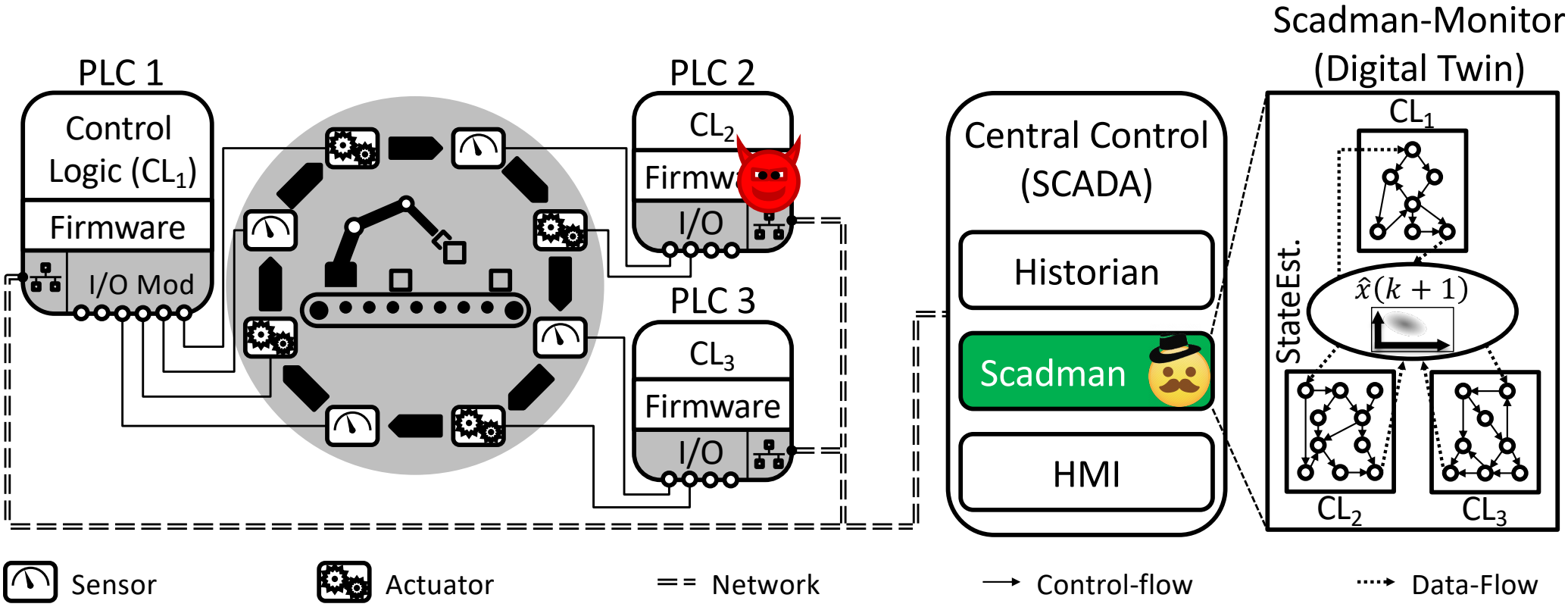
- An intrusion detection solution for distributed ICS
- **Hybrid Digital Twin Model**
  - Uses physical state estimation for IDS
  - Updates physical state estimation based on software control flow

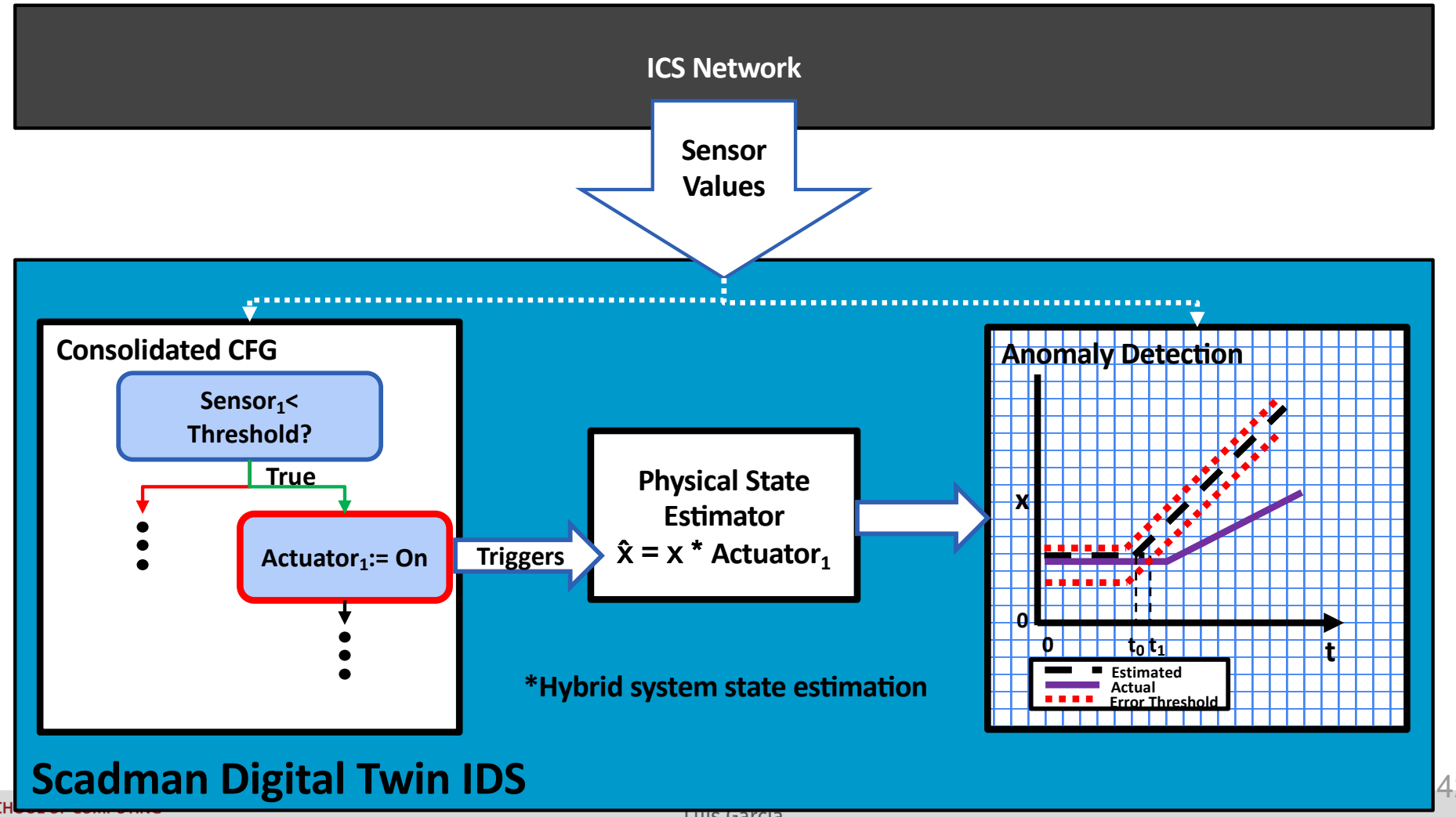


Control Behavior Integrity



Scadman Overview

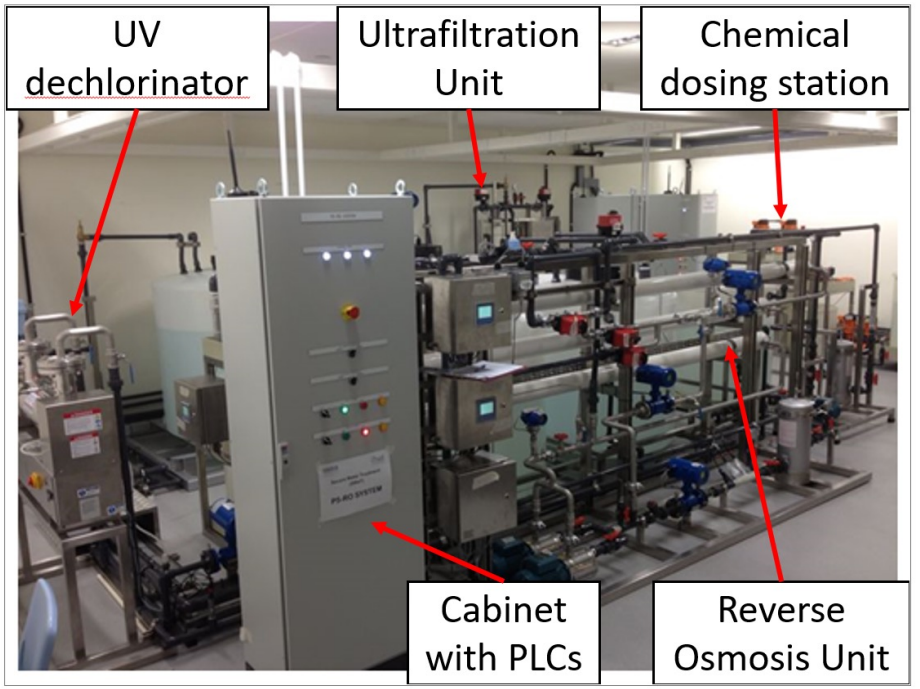


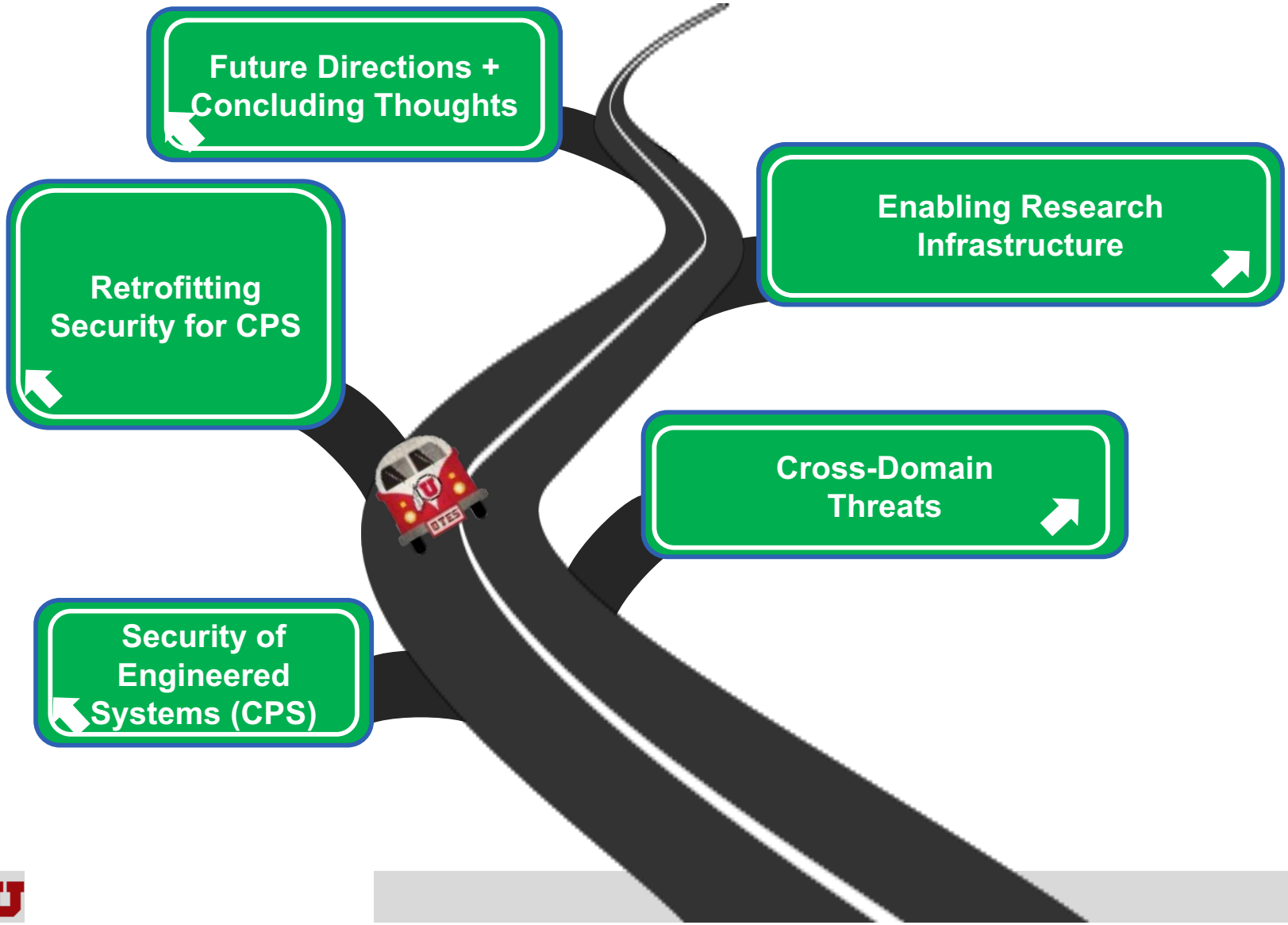

ICS Network
Sensor Values
Consolidated CFG
Sensor$_1$< Threshold?
True
Actuator$_1$:= On
Triggers
Physical State Estimator
$\hat{x} = x * Actuator_1$
*Hybrid system state estimation
Anomaly Detection
x
0
0
t_0 t_1
t
Estimated
Actual
Error Threshold
Scadman Digital Twin IDS

Evaluation: Water Treatment Testbed

- Evaluated against known set of ICS attacks from
  - 7 days worth of data
  - Multi-point attacks included
- Detected all attacks
  - Also detected faulty sensor data
  - Zero false positives
- No overhead on ICS operation
  - Scadman utilizes historian data

We can effectively monitor software and physical states across an ICS

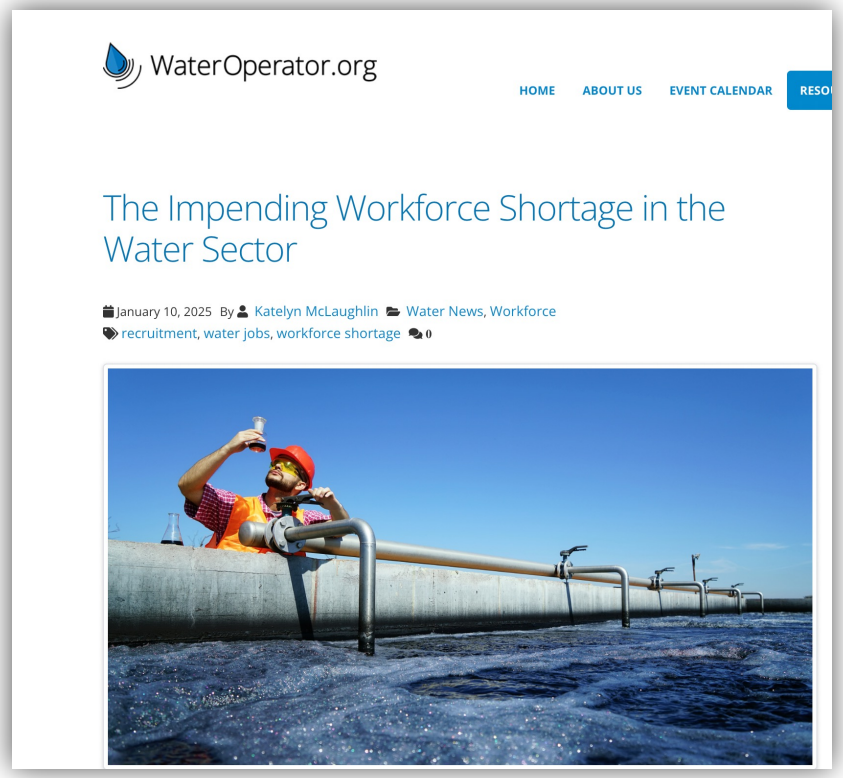



Future Directions + Concluding Thoughts
Enabling Research Infrastructure
Retrofitting Security for CPS
Cross-Domain Threats
Security of Engineered Systems (CPS)

(Some) Critical Challenges in Modern Critical Infrastructure

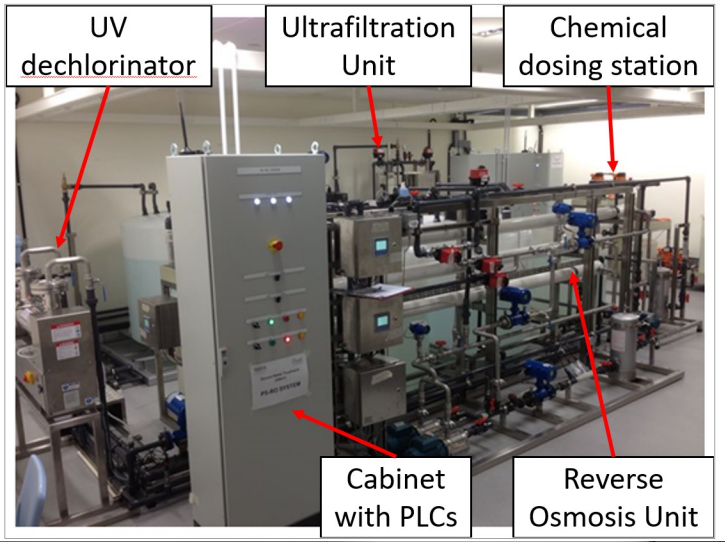
- Rising Cyber/Physical Threats
- Severe Operator Shortages
- Difficult to test realistic attack/failure scenarios on *live* systems
  - (also a bottleneck for research)

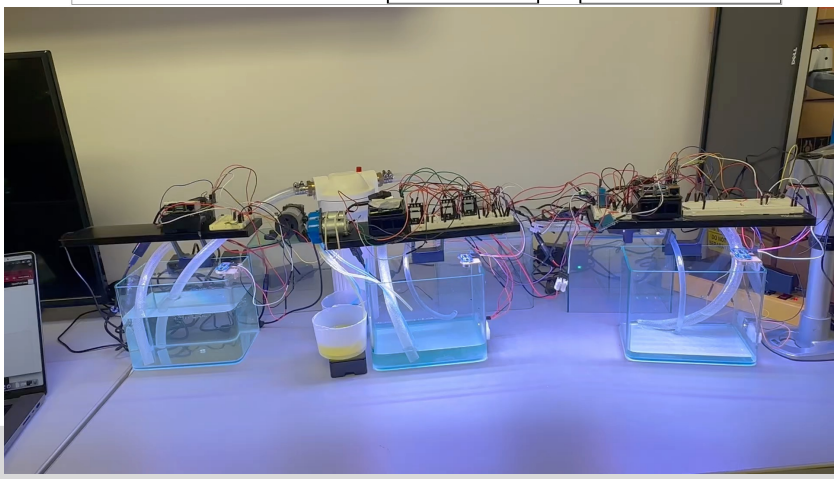
Opportunity: What role should intelligent agents have in CPS decision-making and develop standards, datasets, and tools that guide their use in high-stakes domains.



How do we enable cybersecurity experimentation research for ICS?

- Several testbeds available
  - Water treatment, power grid, smart manufacturing…
- Limitations of ICS testbeds:
  - Requires domain experts to actively configure and operate experiments
  - Expensive!
  - Inflexible architectures
  - Testbed maintainers are often very closed/cautious
  - **Often lack the "human" factor!**



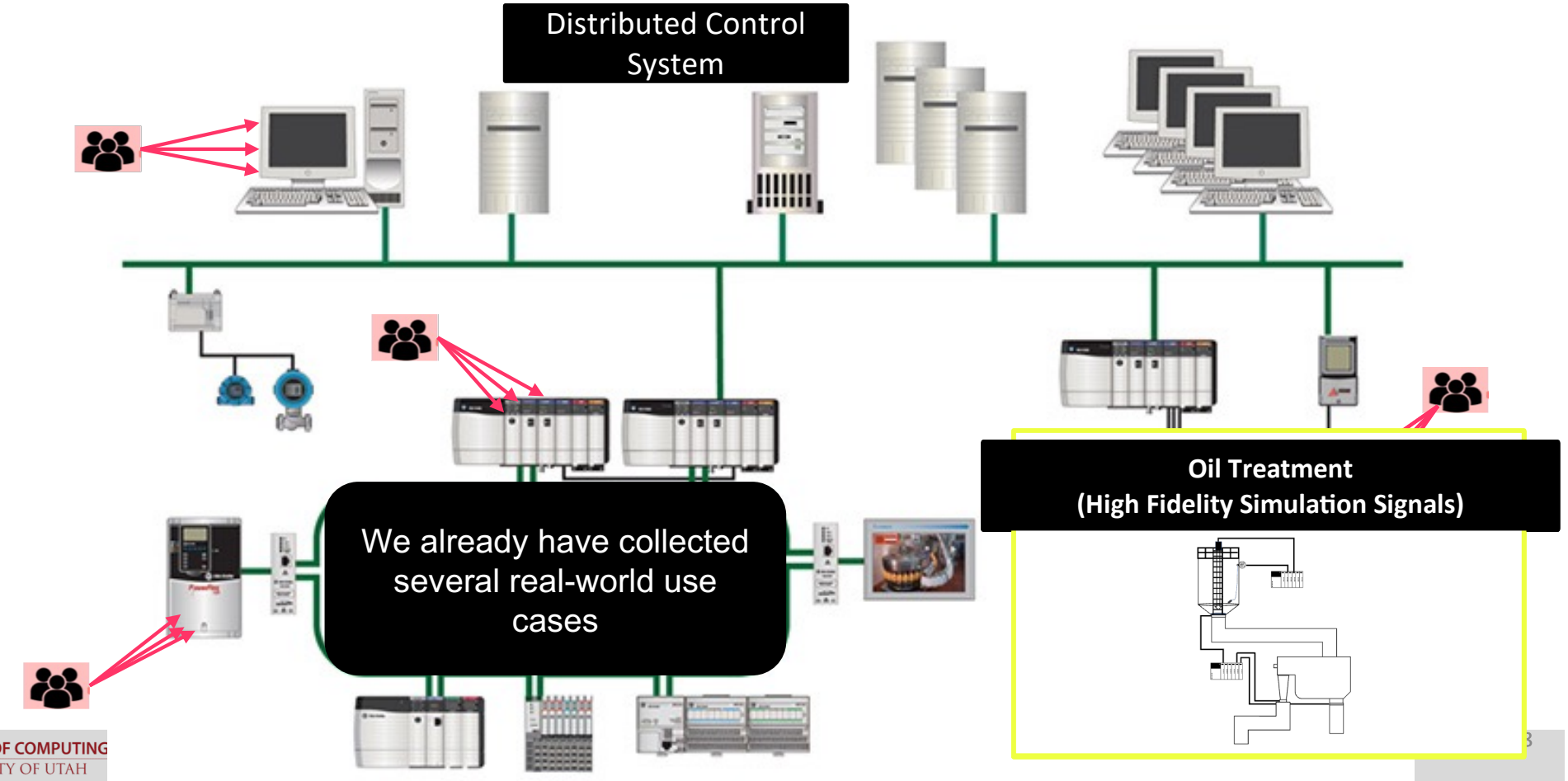


SPHERE CPS Enclave: Human–AI Teaming for Securing Critical Infrastructure

Luis Garcia

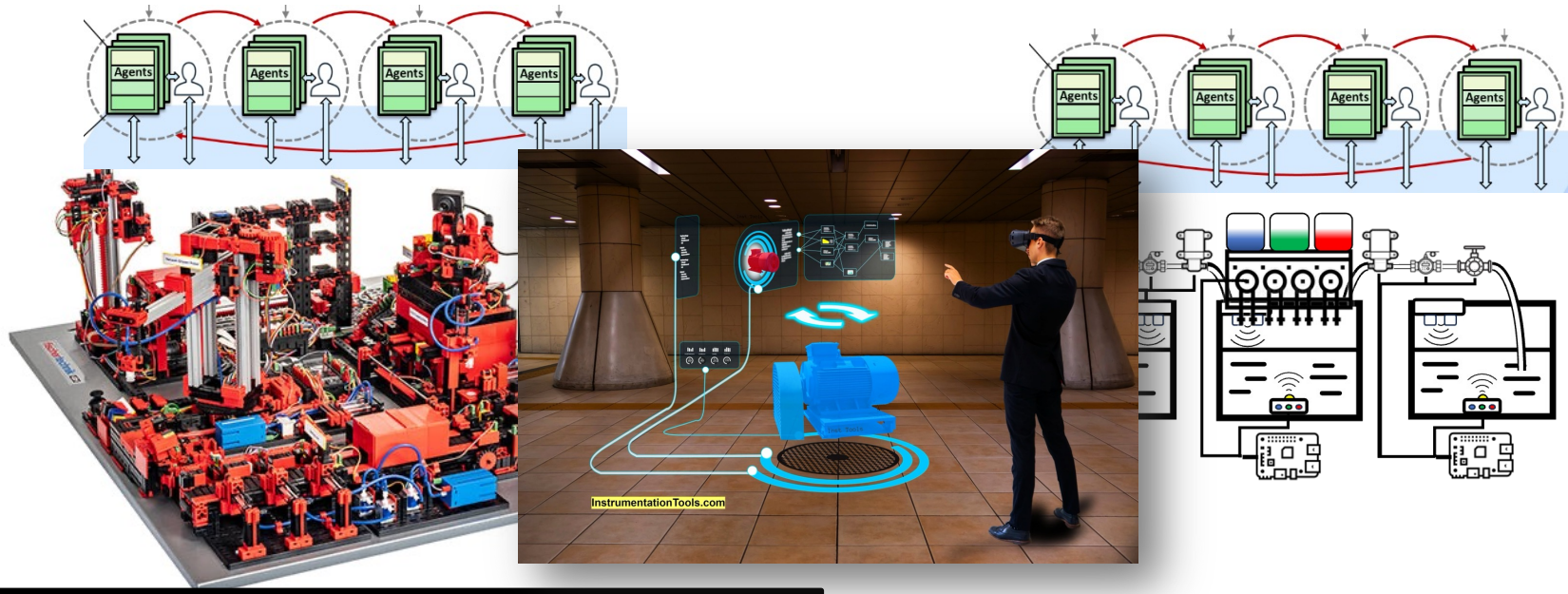
SPHERE is based upon work supported by the National Science Foundation under Grant #2330066. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Vision for CPS Nodes: Remotely Accessible ICS Cybersecurity Experimentation



Near-Future/Ongoing Plans

- **BYOT**: Provide abstractions to add a high-fidelity ICS infrastructure to control existing testbed equipment
- **Agentic infrastructure** for cyber intelligence and operation
- **XR Abstractions** for human feedback/training/interaction on live systems

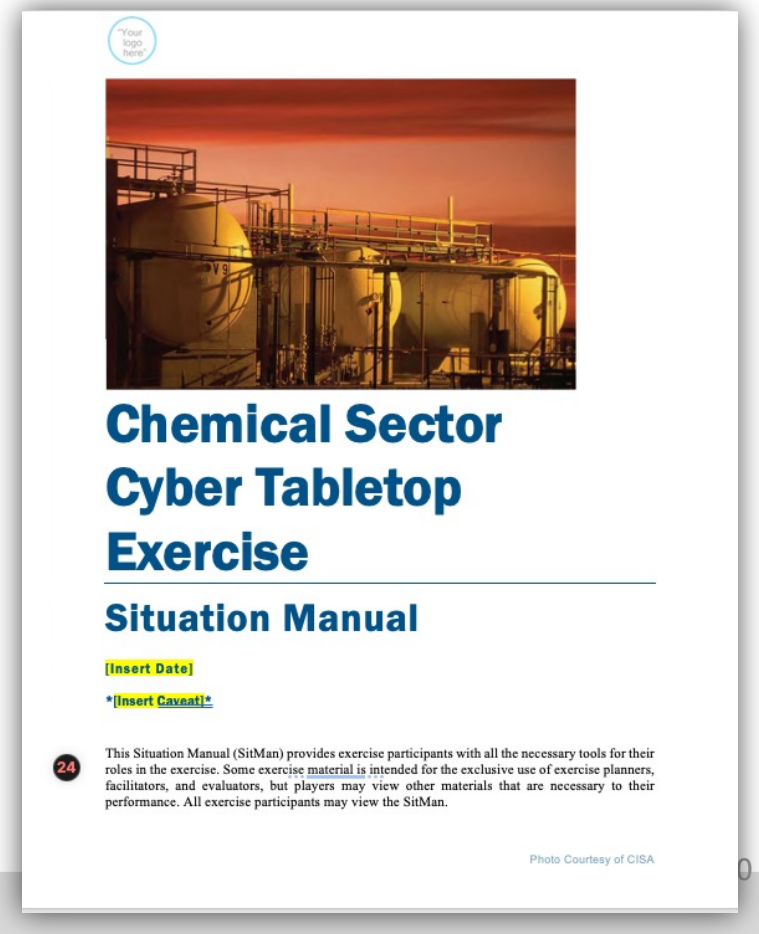


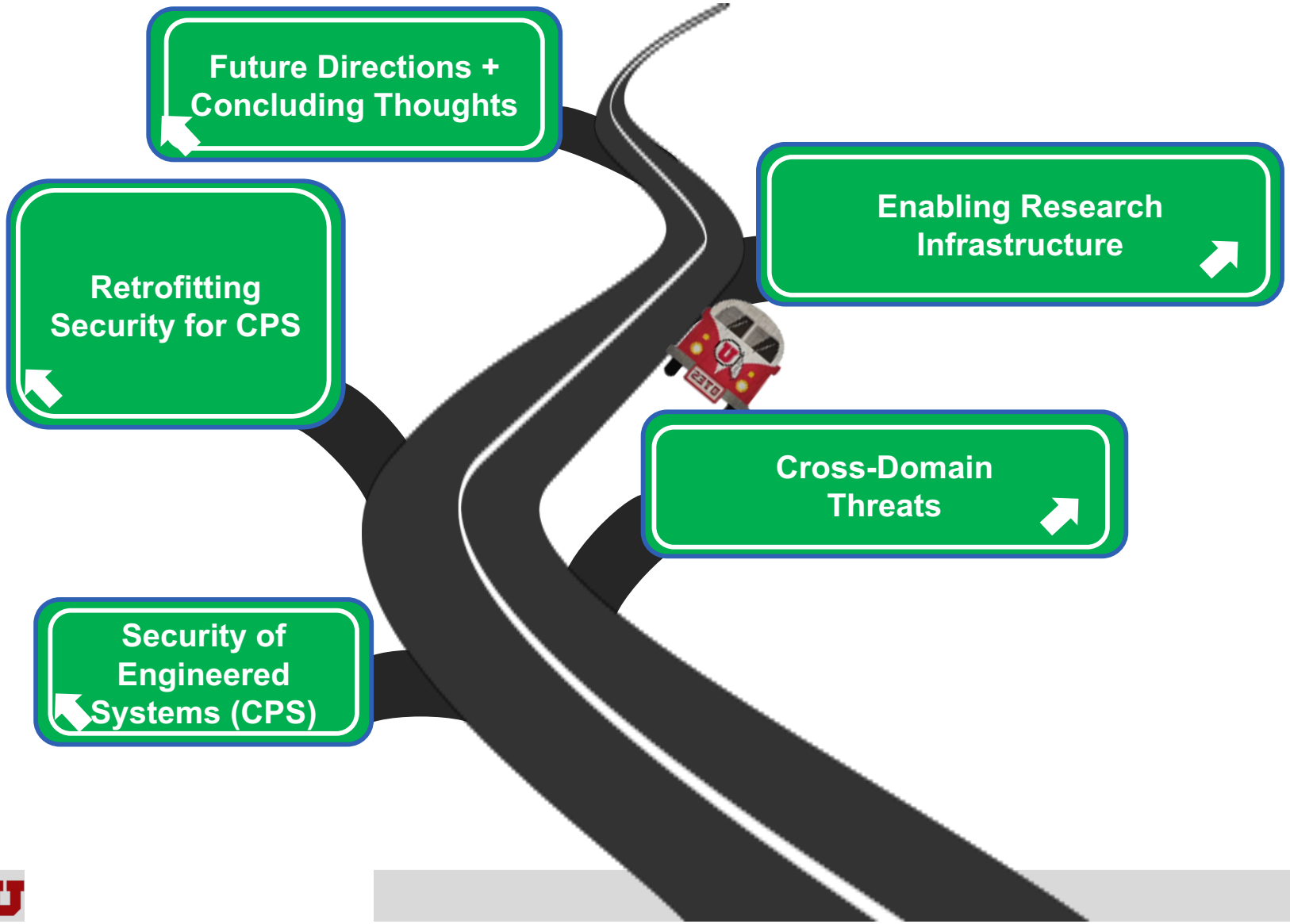
Smart Manufacturing Testbed (U of U- Mu Zhang)

miniSWaT (U of U- Luis Garcia)

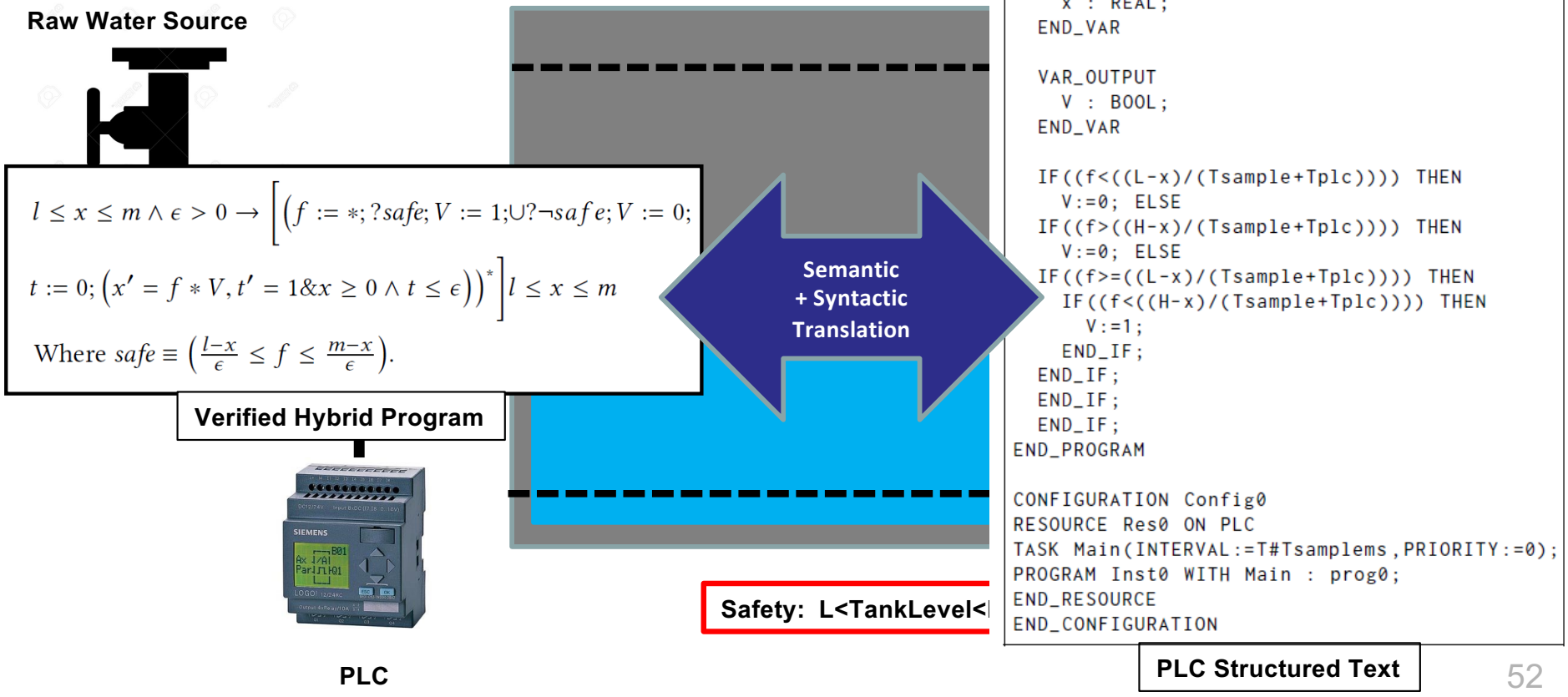
How else can “agents” help?

- Cyber-intelligent agents: this is the more traditional agent being considered in these contexts (and the original idea of our NSF ACTION collab) where **agents can help with situational awareness and communicate with operators**.
- Simulating live attackers
- Modeling human factors:
  - Simulating operator interaction for testing
  - Augmenting (or filling in for) real operators



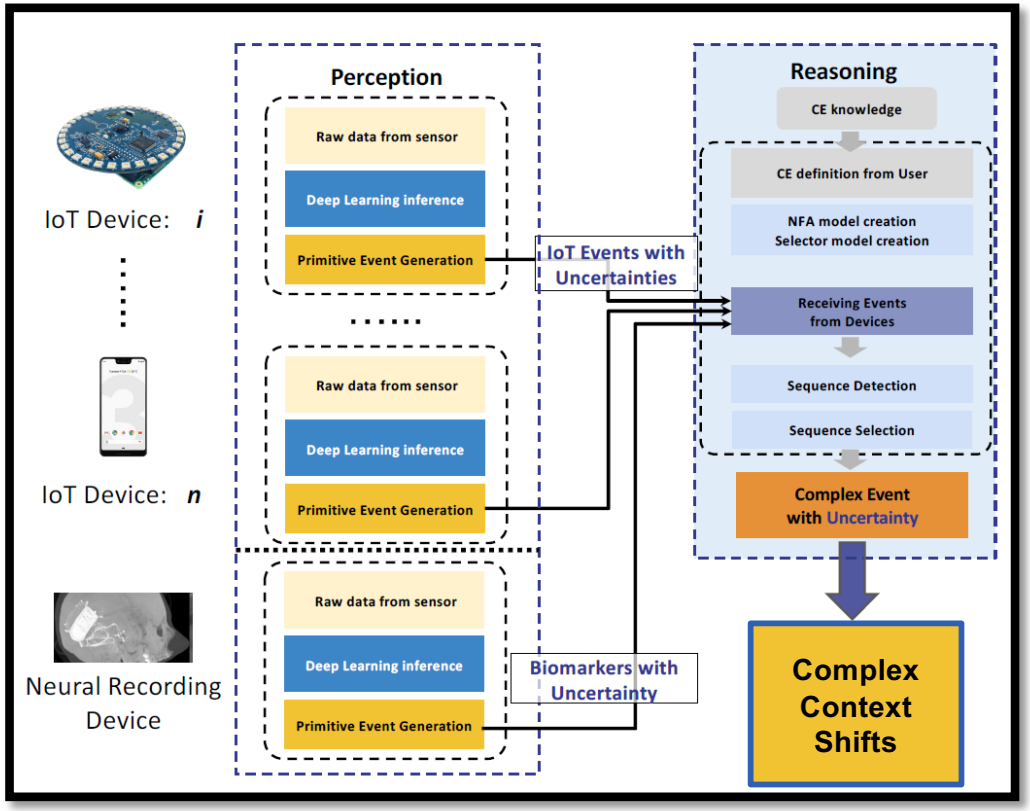

Future Directions + Concluding Thoughts
Enabling Research Infrastructure
Retrofitting Security for CPS
Cross-Domain Threats
Security of Engineered Systems (CPS)

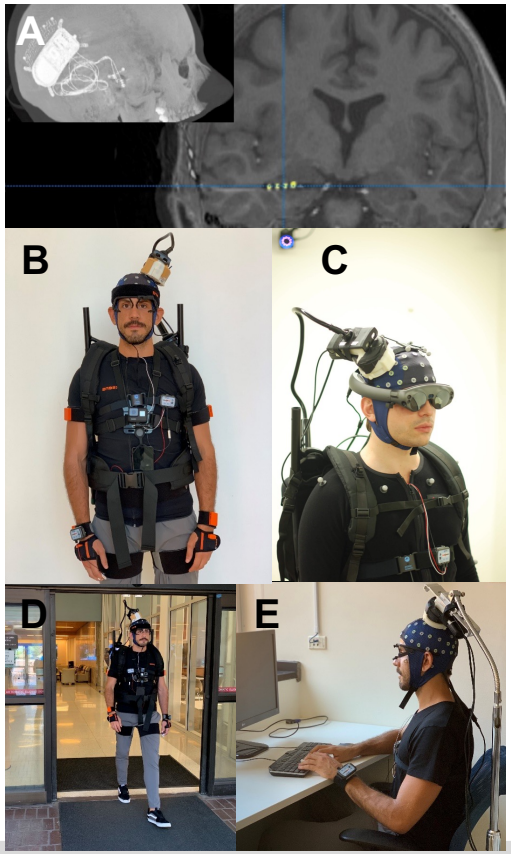
Other Projects: Formally Verifying PLC Code against Safety Properties



THE UNIVERSITY OF UTAH
College of Social and Behavioral Science

Tomorrow's Talk:
Trustworthy IoT-in-the-loop Neuroscience





Concluding Thoughts

- Attackers can opportunistically leverage **physical properties** of CPS—exploiting IT/OT convergence
- For legacy systems, we can opportunistically leverage physical properties to **retrofit security**
- We are actively building a **large ICS testbed** with real-world use cases across multiple domains to enable accessible ICS security research

Concluding Thoughts

- Attackers can opportunistically leverage **physical properties** of CPS—exploiting IT/OT convergence

- For legacy systems, we can opportunistically leverage physical properties to **retrofit security**

- We are actively building a **large ICS testbed** with real-world use cases across multiple domains to enable accessible ICS security research

- **Ongoing and Future Work**
  - *Human-in-the-loop CPS: Safety, security, and privacy guarantees*
  - *Secure Peripheral Abstractions in Cyber-Physical Systems*
  - *Neurosymbolic Architectures for Trustworthy Cyber-physical Systems*

Luis Garcia
la.garcia@utah.edu
https://iotrustlab.com