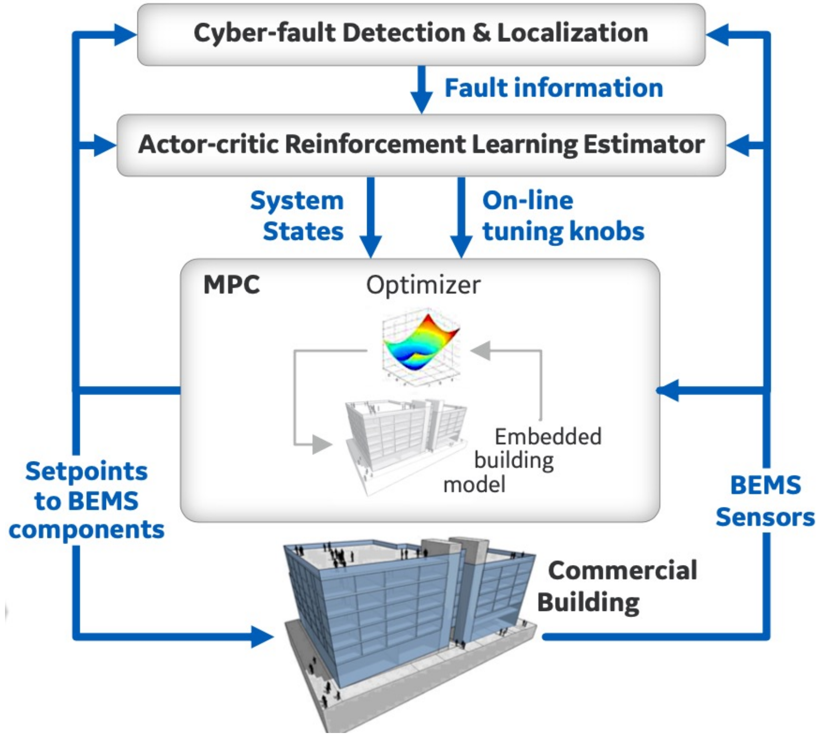
Adaptive Cyber-Physical Resilience for Building Control Systems



General Electric Research, Pacific Northwest National Laboratory
Mustafa Dokucu, Principal Engineer, Controls and Optimization, dokucu@ge.com
Craig Bakker, Research Scientist, Applied Statistics and Computational Modeling, craig.bakker@pnnl.gov

Project Summary

Timeline:

Start date: 04/2020

Planned end date: 03/2023

Key Milestones

1. Requirements Document; 06/2020
2. 90% accuracy for cyber-attack detection; 04/2021
3. 99% accuracy for cyber-attack detection; 04/2022

Budget:

Total Project $ to Date:

- DOE: $1,485,794
- Cost Share: $588,027

Total Project $:

- DOE: $2,973,087
- Cost Share: $1,093,534

Key Partners:

General Electric Research
Pacific Northwest National Laboratory

Project Outcome:

Automated Fault Detection and Diagnosis (AFDD) algorithms that detect and isolate cyber-attacks and system faults (99% accuracy).

Model Predictive Control (MPC) architecture with Reinforcement Learning (RL) based estimation to enable a cyber-resilient building energy management system.

Team

Technical Manager – **Erika Gupta**
Project Manager – **Jason Conley**

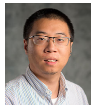
Pacific Northwest NATIONAL LABORATORY

- Dr. Draguna Vrabie: advanced control methods

- Dr. Craig Bakker: data driven modeling, cyberphysical security

- Mr. Andrew August: software integration and performance analysis

- Dr. Shant Mahserejian: data science

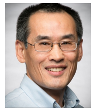


- Dr. Sen Huang: building modeling and simulation

- Dr. Soumya Vasisht: experimental design

- Dr. Mustafa Dokucu: cybersecurity, model-based estimation, advanced controls

- Dr. Subhrajit Roychowdhury: Reinforcement Learning, cybersecurity, optimal control



- Dr. Weizhong Yan: cybersecurity, machine-learning, fault detection and isolation

- Dr. Abhay Harpale: cybersecurity, machine-learning, fault detection and isolation

- Dr. Annarita Giani: experimental design, data science

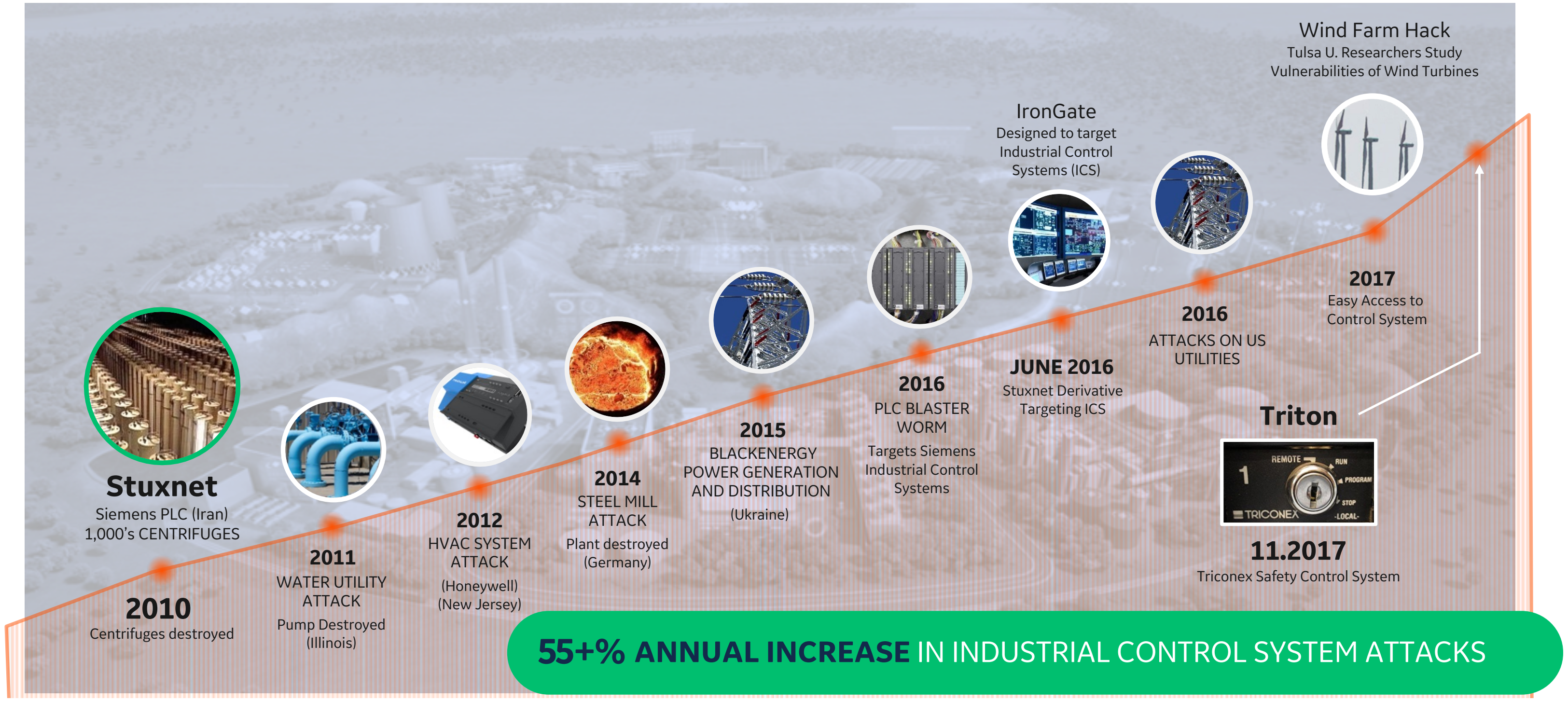
- Dr. Reza Ghaemi: model-predictive controls, building modeling and simulation

Strong, cross-functional team with expertise for the desired technology

Challenge

Cyber attacks on Industrial Control System are increasing rapidly…

impacting customer security, safety, availability…



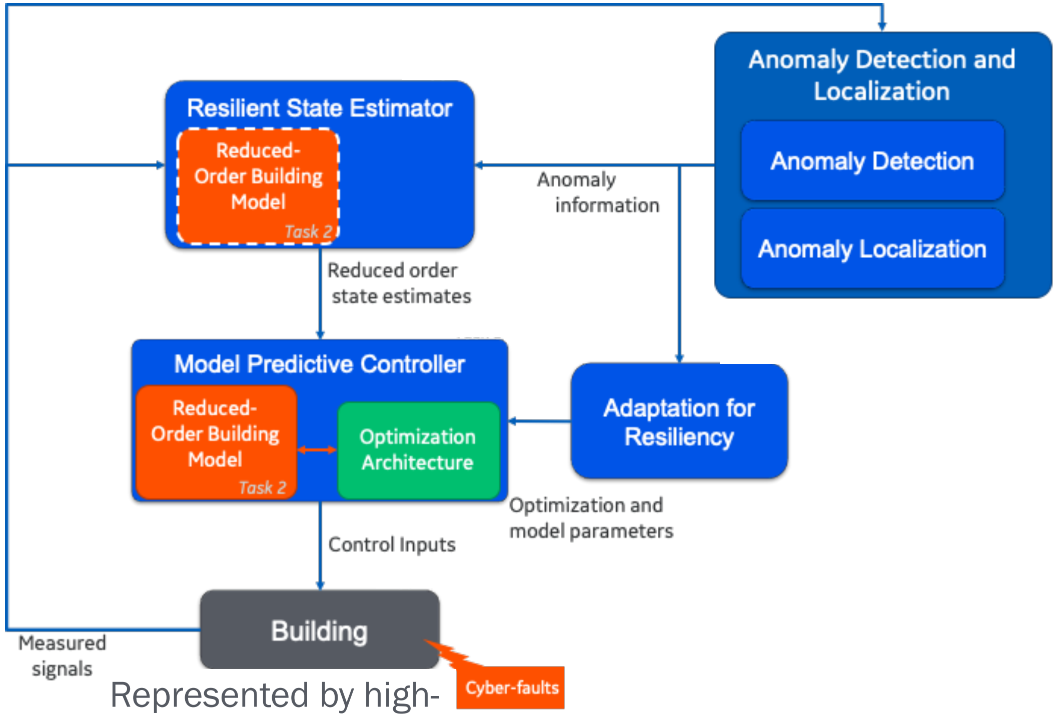
- Grid-interactive building control systems increase attack and fault scenarios and likelihoods
- IT/OT layer protection is necessary but not sufficient
- Protection at cyber-physical layer is also necessary
- Cohesive protection will ensure optimal energy savings/occupant comfort together with smart building technologies

Cybersecure advanced-controls will enable energy savings of ~30%

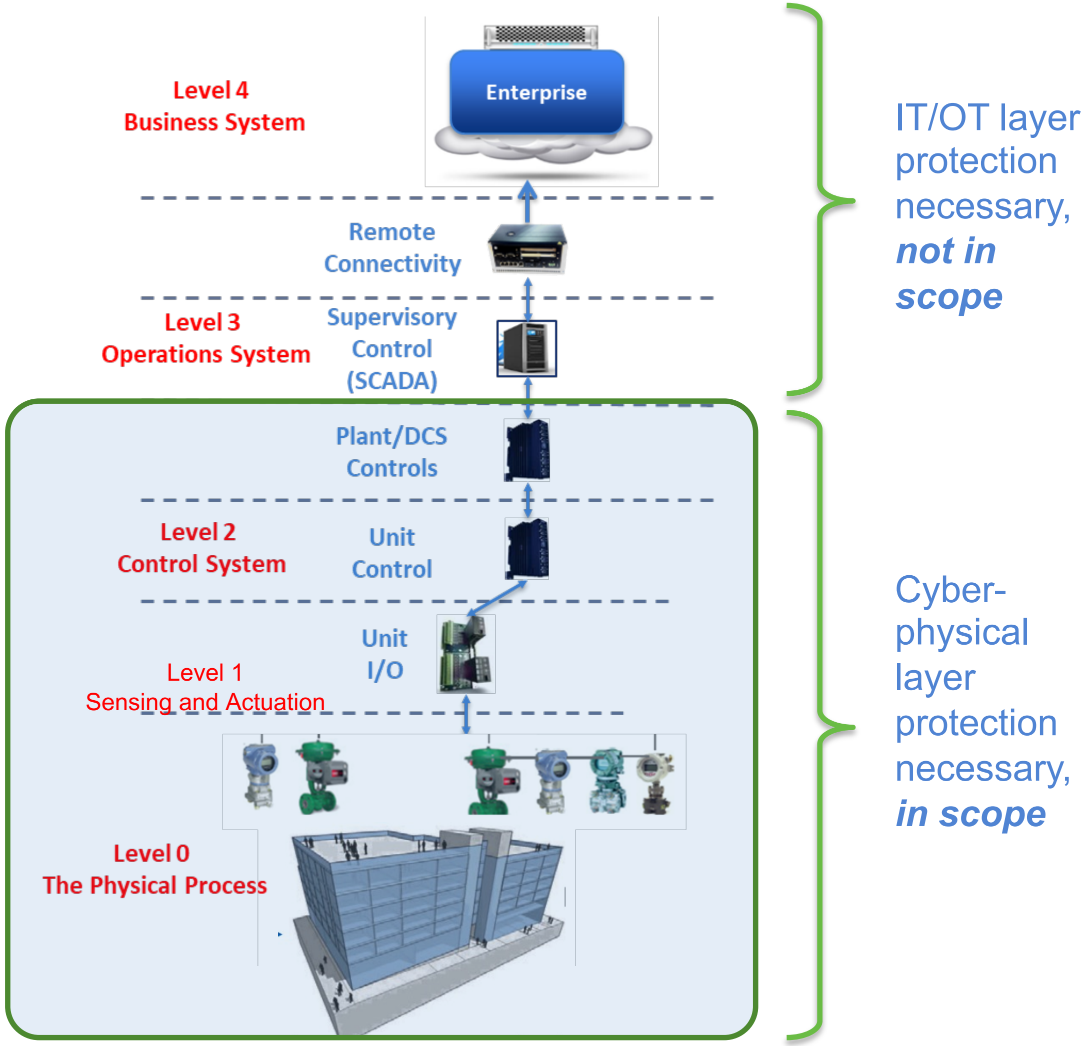
Approach - *Overall*

Integrated Technology for Cyber-Physical Building Control Systems

- Is the system abnormal or normal?
- Pinpoint root cause of the abnormality; system fault or cyber-attack?
- Make building control system more resilient to the detected/localized cyber-fault



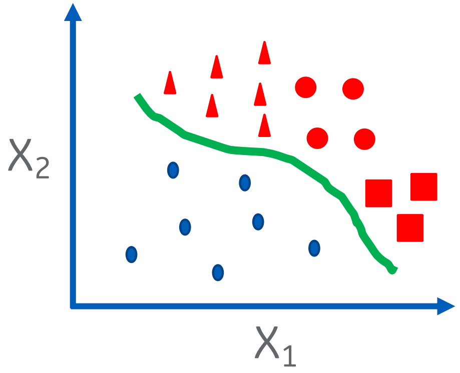
Represented by high-fidelity emulator



Approach – detection and localization

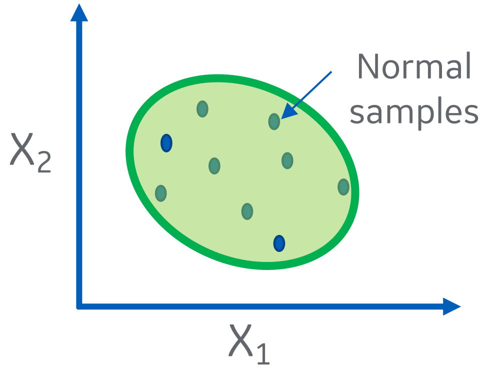
Anomaly Detection (AD) Design Strategy -

Supervised AD



VS.

Semi-supervised AD



What is it? To formulate attack detection as a binary classification problem
Pros: Tends to be more accurate for detecting the simulated attacks
Cons:
- Requires both normal and attack samples available;
- May not generalize well to novel attacks.

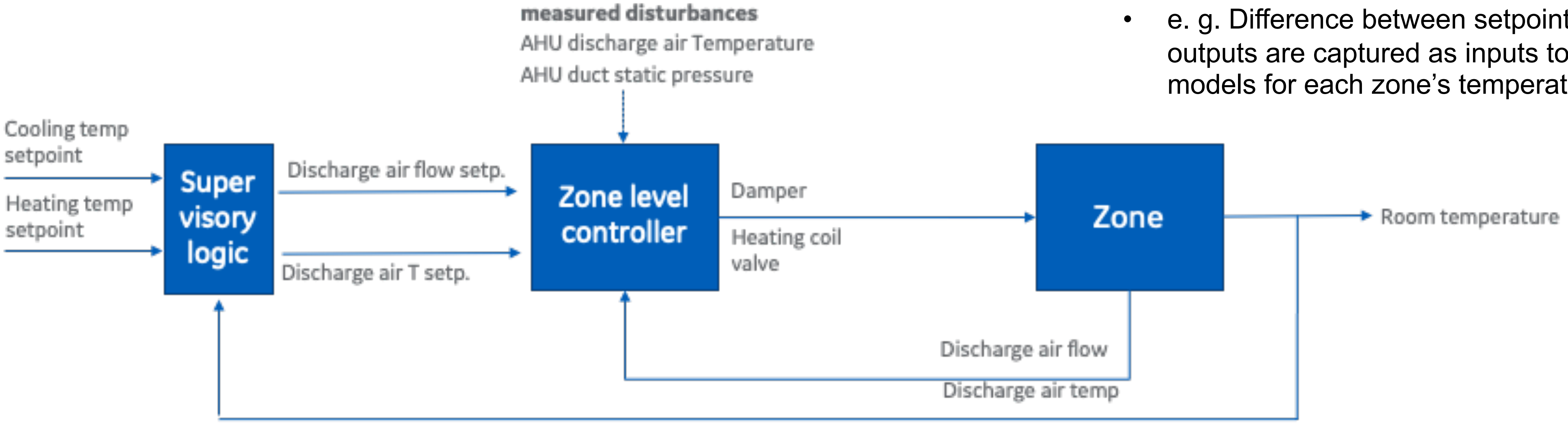
What is it? To characterize the system normal behavior (normality) first and then to perform attack detection by monitoring any deviation from the normality.
Pros:
- Only normal data is needed
- Can handle novel attacks

Cons: false alarms might be high

Approach - *Detection*

Normality modeling strategy - data-driven vs. physics-guided

Simplified Zone/VAV control diagram

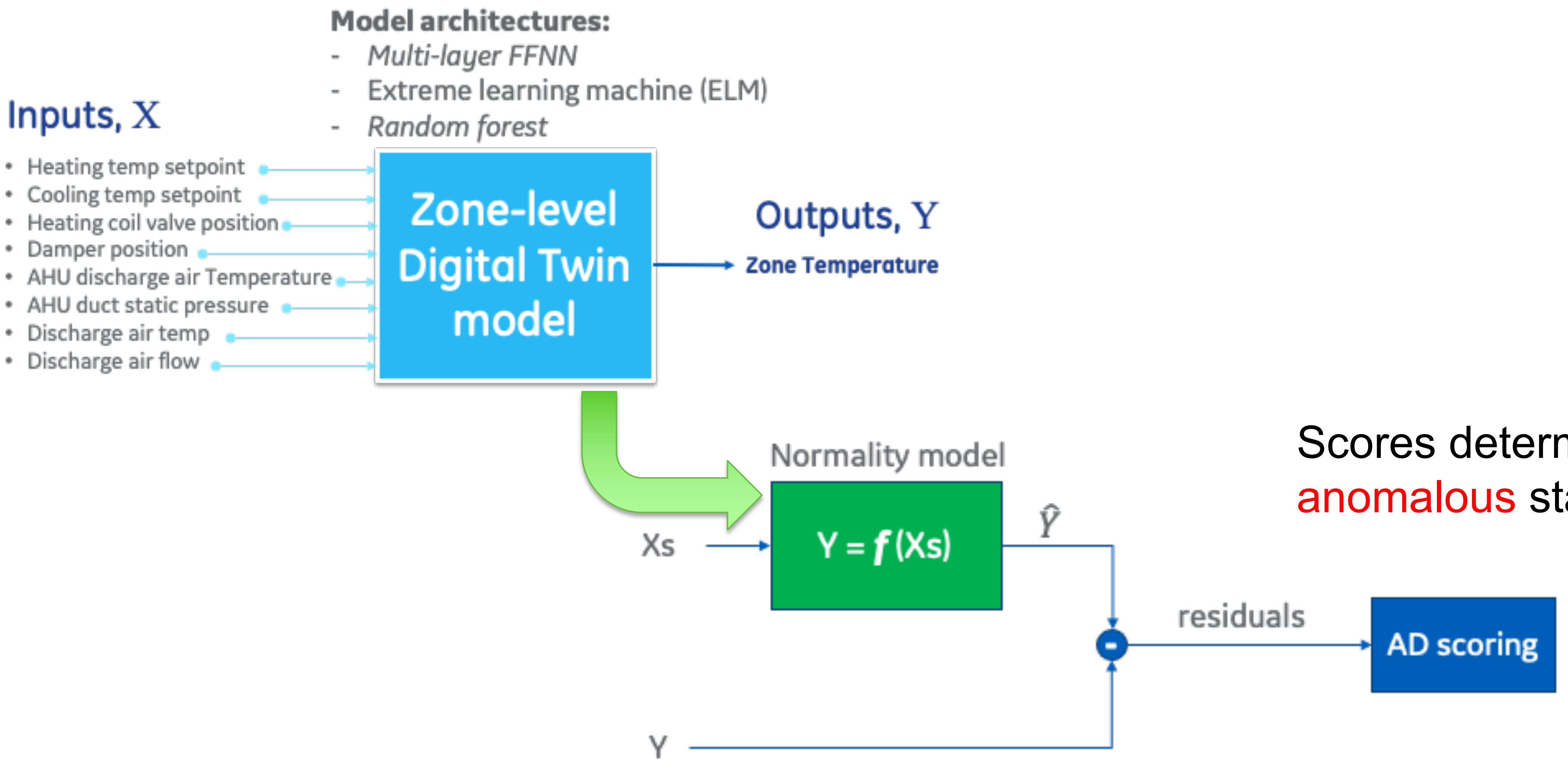


Inputs and outputs are selected based on the cascade control structures

- e. g. Difference between setpoints and controlled outputs are captured as inputs to the normality models for each zone's temperature

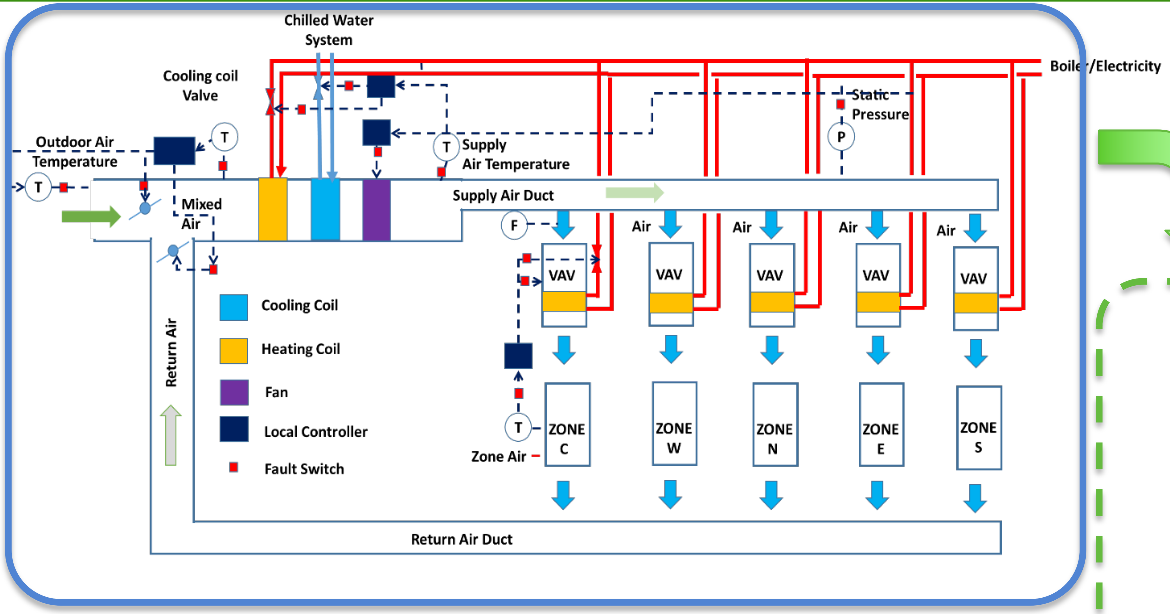
Data-driven (black-box) → physics-guided normality modeling

Approach - *Detection*



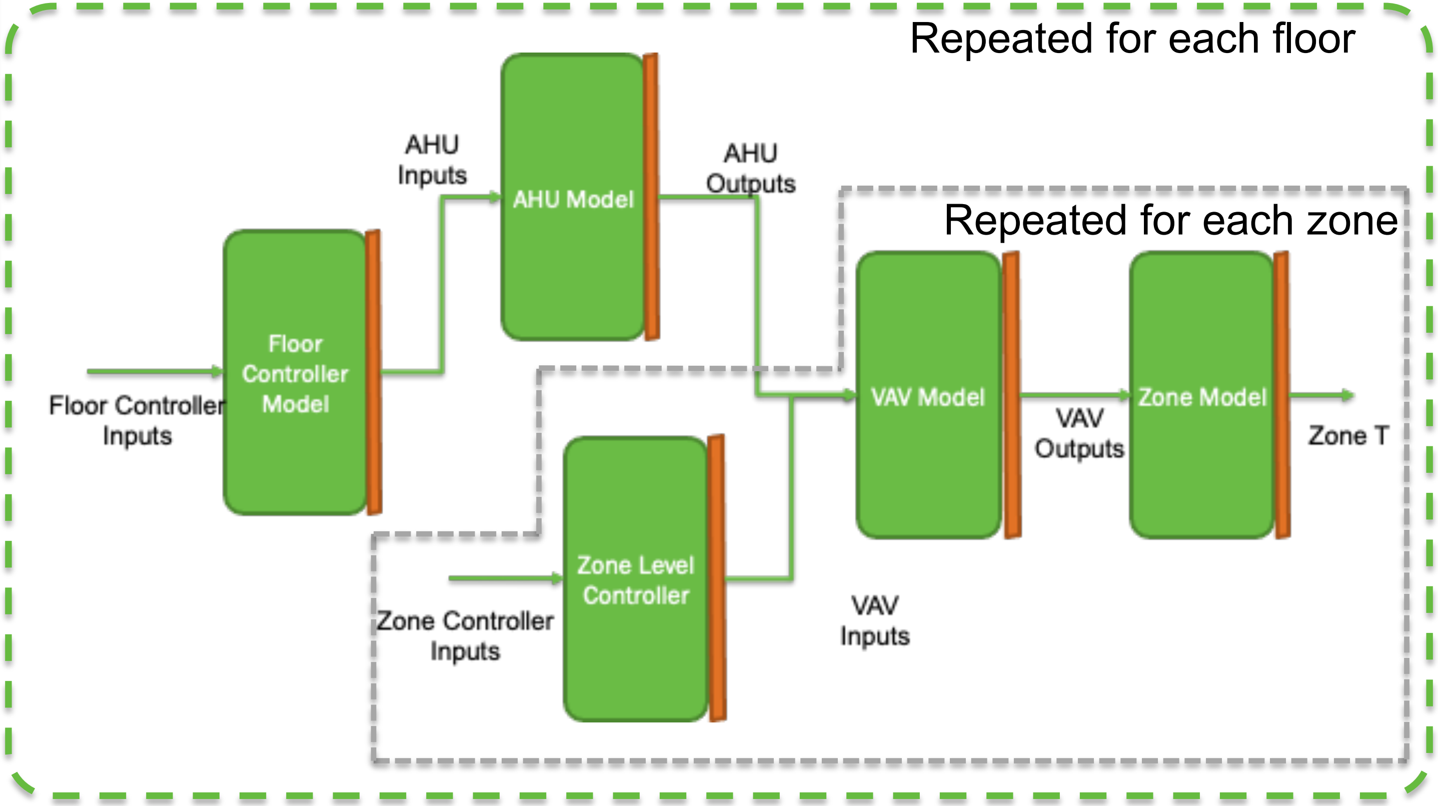
Scores determine anomalous state

Approach - Localization



Building Control System Blueprint

Residual location is used for isolating the attack/fault



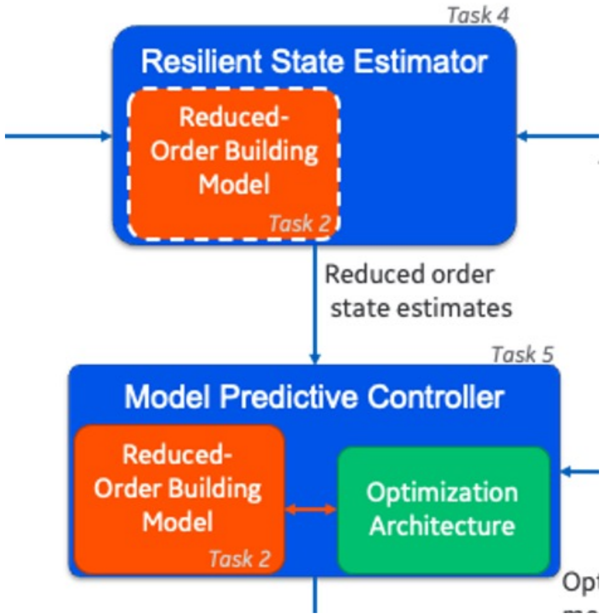
Building Control System ML Model

Every model consists of a characteristic part (suitable for transfer learning) and an individualized part (fine-tuned Digital Twin)

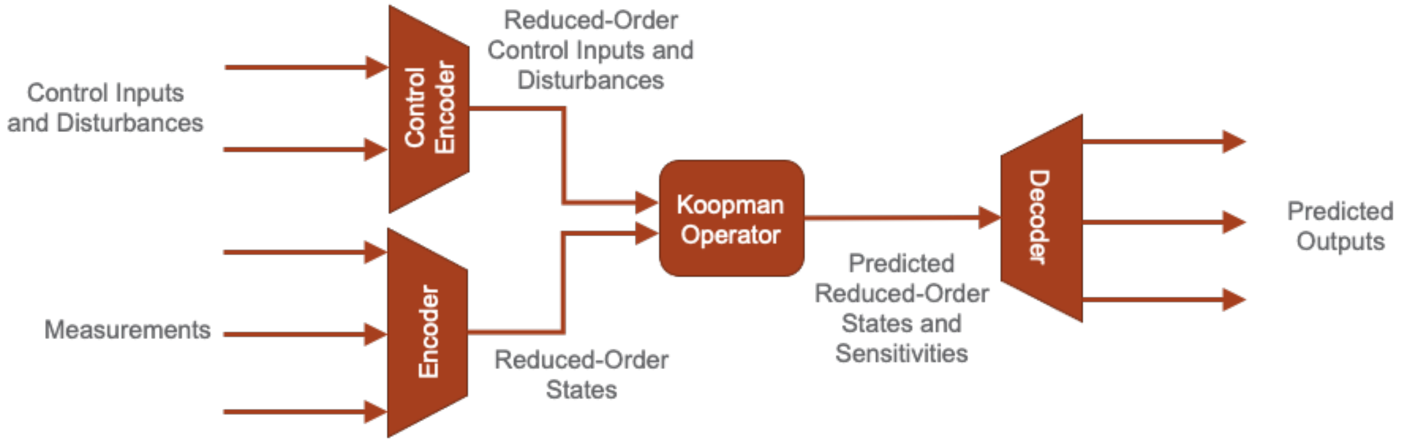
Approach – Reduced Order Modeling

Computationally efficient reduced-order model is the enabler for

- Resilient state estimator
- Model predictive controller

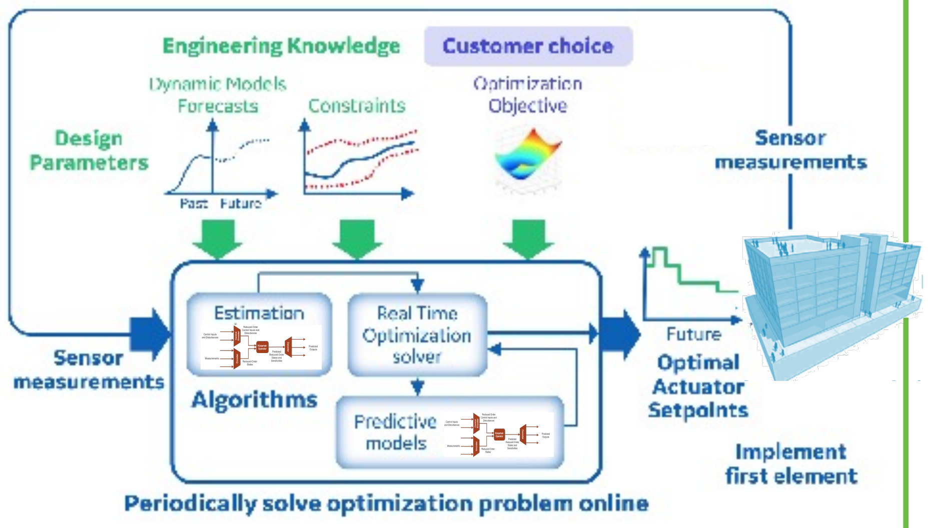


- Deep learning-based autoencoders for dimension reduction
- Koopman operator-based approach to learn the dynamics
- Trajectory prediction and linearized dynamics provided to MPC



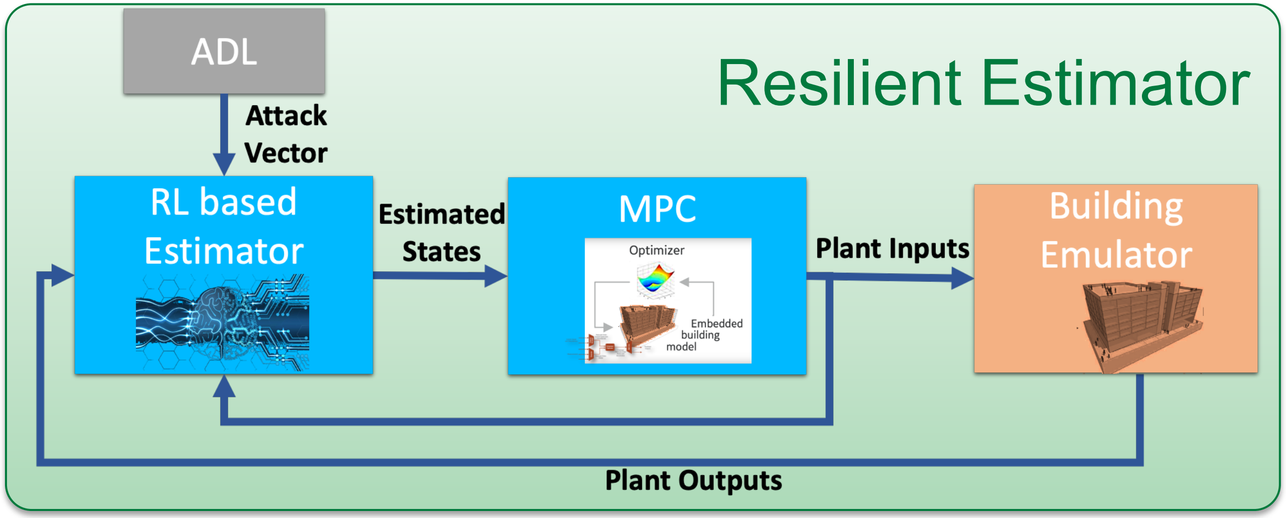
Approach – Resilient Control Layer

Resilient Control



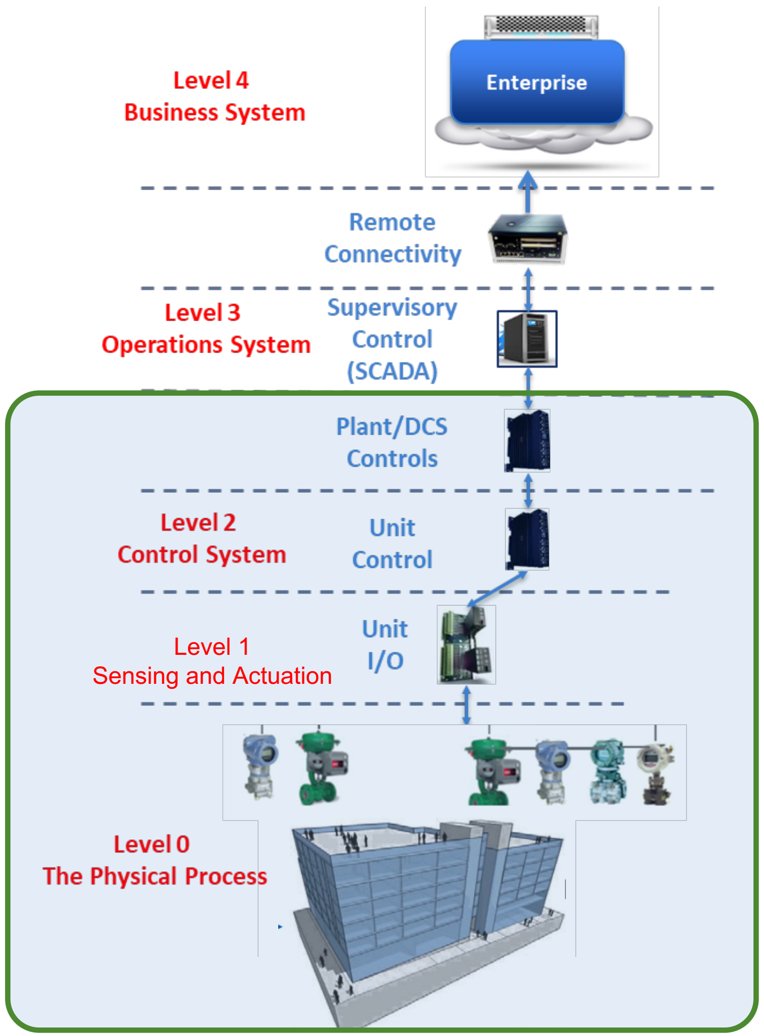
- Hybrid (linear/nonlinear) Model Predictive Controller (MPC)
- Utilizes the reduced-order model for predictions
- QP optimization problem solved at each time step
- Objective function adapted for cyber-resiliency

Resilient Estimator



- Estimates current status of the building system to the MPC
- Virtual estimation of attacked sensors
- Reinforcement Learning based policies trained with reduced-order model

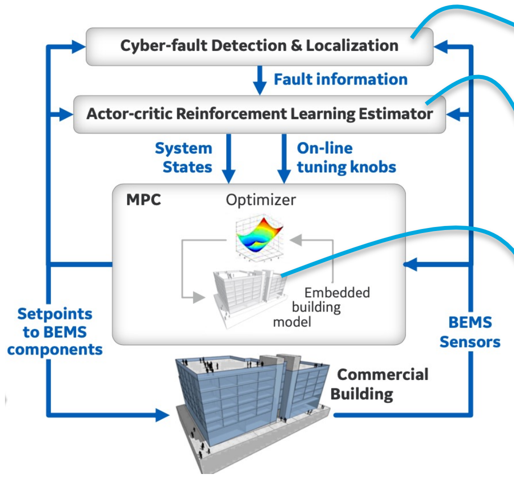
Impact



Existing cybersecurity solutions:

- Threat detection at IT/OT layer
- Utilize network information only
- Treat system as black-box
- No resilience if attack is successful

Technology being developed: Security at the Cyber-physical layer



Threat detection and isolation at the cyber-physical layer

Resilience if an attack is successful

Leverage physical understanding

Integrated solution will enable ~30% energy savings brought on by advanced building control systems robustly (MPC)

Progress – Attack detection and localization

- Project is approaching mid stage
- Budget Period 1 go/no-go decision point: above **90%** detection/localization accuracy
  - Detection accuracy – **93.5%**(True Positive Rate) **0.1%** (False Positive Rate)

Overall detection performance

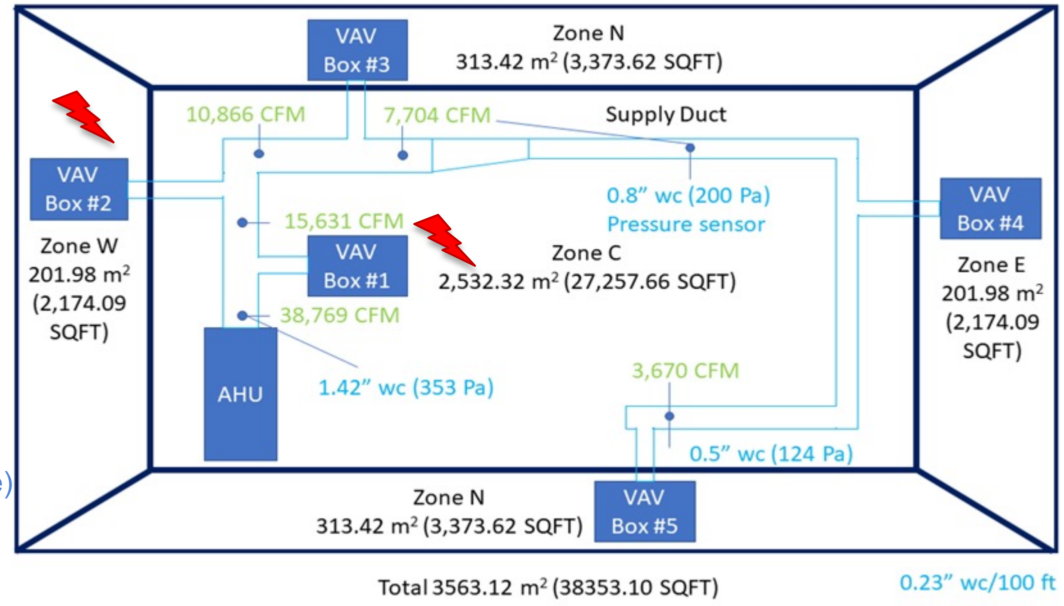
N: no attack A: attack		Prediction				
		N	A			
TRUE	N	**66,592**	68	FPR =	**0.1**	%
	A	3	**43**	TPR =	**93.5**	%

  - Localization accuracy – **100.0%**(True Positive Rate) **0.03%** (False Positive Rate)

Overall localization performance

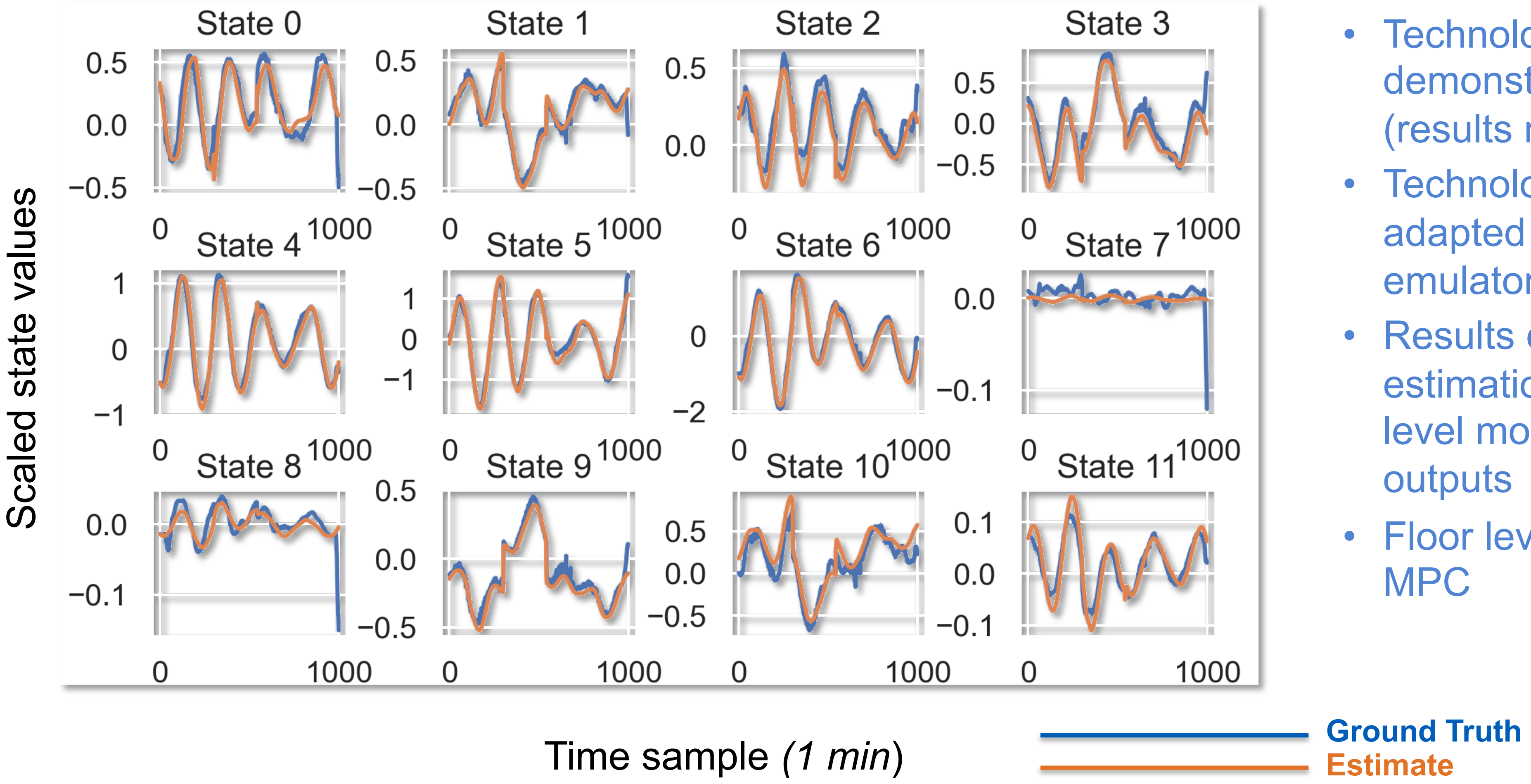
N: no attack A: attack		Prediction				
		N	A			
TRUE	N	**44,880**	13	FPR =	0.0290	%
	A	0	**30**	TPR =	100.0	%

Man-in-the-middle attacks on Temperature Sensors (Floors 1 and 2)



Developed technology can detect and localize cyber-attacks to desired accuracy
Attack sophistication will be increased to stress-test and improve the technology

Progress : RL based resilient estimator



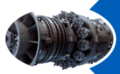
- Technology developed and demonstrated on a model system (results not shown here)
- Technology currently being adapted to PNNL building emulator model
- Results demonstrate satisfactory estimation of ROM states for floor level model with 13 inputs and 27 outputs
- Floor level estimator integrated to MPC

Developed technology can reconstruct states with desired accuracy

Stakeholder Engagement

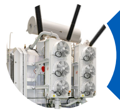
Digital Ghost is a cyber-physical protection framework that encompasses, detection, localization, and neutralization functionalities for industrial control systems

Digital Ghost is being productized

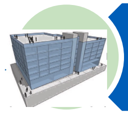


GE Gas Power

GE Renewables: Wind



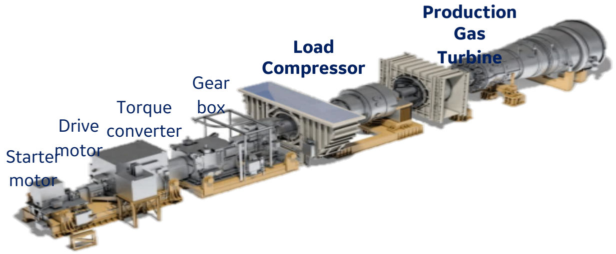
GE Renewables: Grid Solutions



Building Control Systems

Tailor Cyber-security solution for Building Control systems based on feedback/learnings from ongoing productization efforts

Digital Ghost is tested on industrial assets



Digital Ghost Gas Turbine Test Results

- ✓Prevented sophisticated attacks that were not detected by available cybersecurity solution
- ✓DG successfully detected the attacks & localized the sensors under attack
- ✓Neutralization able to provide replacement sensor values in closed-loop control

Stakeholder Engagement

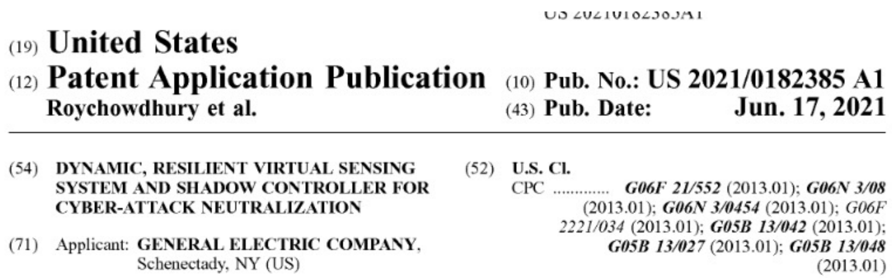
Publications/Presentations:

- ACC manuscript
  - “A resilient control architecture for building control systems”
- ASHRAE presentation, planned

Invention disclosures:

- *Roychowdhury, Subhrajit, Masoud Abbaszadeh, and Mustafa Tekin Dokucu. "Dynamic, resilient virtual sensing system and shadow controller for cyber-attack neutralization." U.S. Patent Application No. 16/710,051.*

- *Blueprint networks for anomaly localization in interconnected and interdependent systems (in process of filing)*



(19) **United States**
(12) **Patent Application Publication**
Roychowdhury et al.

(10) **Pub. No.: US 2021/0182385 A1**
(43) **Pub. Date: Jun. 17, 2021**

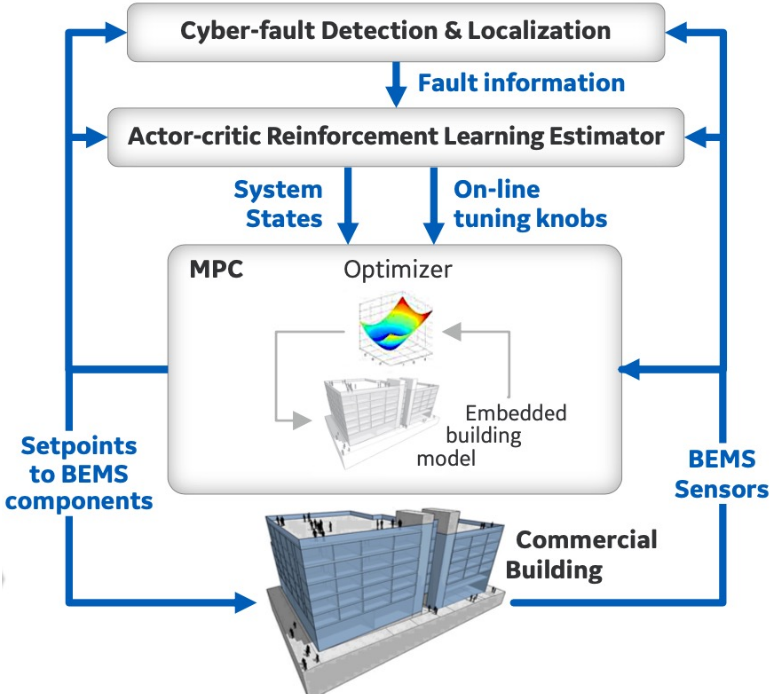
(54) **DYNAMIC, RESILIENT VIRTUAL SENSING SYSTEM AND SHADOW CONTROLLER FOR CYBER-ATTACK NEUTRALIZATION**

(71) Applicant: **GENERAL ELECTRIC COMPANY**, Schenectady, NY (US)

(52) **U.S. Cl.**
CPC ***G06F 21/552*** (2013.01); ***G06N 3/08*** (2013.01); ***G06N 3/0454*** (2013.01); *G06F 2221/034* (2013.01); ***G05B 13/042*** (2013.01); ***G05B 13/027*** (2013.01); ***G05B 13/048*** (2013.01)

Remaining Project Work

Core algorithms for all layers have been developed in Budget Phase I



Near-term (Budget Phase II) remaining work

Detection/Localization: Broader operation (normal) conditions

Reduced Order Model:

Extend to full building, robustness

Consider more attack/fault scenarios, better separation of attacks/faults

MPC:

Rule-base for adapting MPC objective function for resiliency

Reinforcement Learning based adaptation

Integration of individual layers of technology

End Goal (Budget Phase III) work:

- ***99%*** accuracy in detection/localization
- Resiliency (at different levels) for ***all*** cyber-fault scenarios
- Architecture ready for deployment
- Successful separation of faults and cyber-attacks

Thank You

General Electric, Pacific Northwest National Laboratory

Mustafa Dokucu, dokucu@ge.com

Craig Bakker, craig.bakker@pnnl.gov

Acknowledgment: “This material is based upon work supported by the U.S. Department of Energy’s Office of Energy Efficiency and Renewable Energy (EERE) under the Building Technologies Office (BTO) Emerging Technologies (ET) Program and Grid Modernization Initiative (GMI), Award Number DE-EE0009151.”

REFERENCE SLIDES

Project Budget

Project Budget: total: **$4.1M, cost share GE:$1.1M, PNNL share: $0.7M, GE share: $3.4M**
Variances: none
Cost to Date: 56% of the budget has been spent
Additional Funding: no additional funding other than cost share

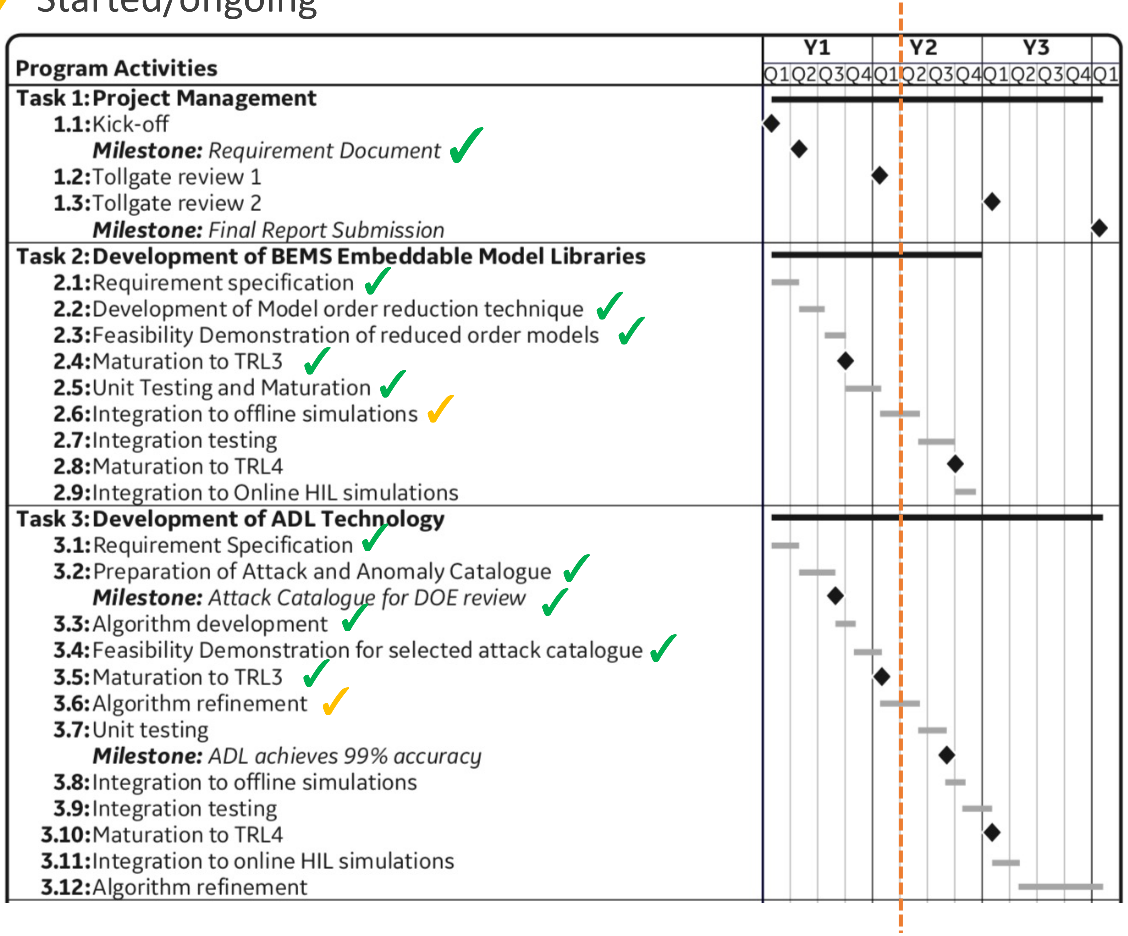
	Budget History							
	April– FY 2020 (past)		FY 2021 (current until Q3)		FY 2022 (planned)		FY 2023 (planned)	
	DOE	Cost-share	DOE	Cost-share	DOE	Cost-share	DOE	Cost-share
GE	944,993.70	336,006.30	423,543.77	226,523.53	733,233.23	260,711.42	213,365.89	758,65.36
PNNL	82,767.43		260,007.85		332,224.72			

Project Plan and Schedule

Project initiation: 04/2020, Project completion: 03/2023

✓Completed

✓ Started/ongoing



Task 1: Project Management

- MILESTONE 1.1 Requirements documents (M3), completed
- MILESTONE 1.2 BP1 review (M12), completed

Task 2: BEMS Embeddable Model Libraries (PNNL)

- Reduced order model (ROM) requirements defined
- ROM mathematically formulated and aligned with ACRE, ADL, & MPC
- ROM code written and tested on example-problem
- ROM methodology applied to the building emulator to create a floor-level ROM

Next Steps: Extend floor level model to building-level and integrate with other modules

Task 3: Attack Detection and Localization

- Attack detection and localization methodologies developed and tested on a high-fidelity emulator and achieved **>90% accuracy**
- Cyber-fault catalogue created (PNNL +GE)

Next Steps:

- More fault attack scenarios
- Algorithm refinement, fault vs. attack and improved accuracy

Project Plan and Schedule

Task 4: Development of ACRE Technology

- First version of resilient state estimator developed for a 3-state example system
- Tested in conjunction with the model predictive controller
- Resilient state estimator integrated with ROM of the example system
- MILESTONE 4.1 Feasibility demonstration (M12), established
- Application of the methodology to the floor-level ROM
- Next Steps:
  - Application of the methodology to the building-level ROM
  - Integration with MPC

Task 5: Development of Model Predictive Control

- Developed fast QP-based MPC and tested in conjunction with the resilient state estimator for a 3-state example system
- MILESTONE 5.1 MPC integrated with ROM (M12), established
- Next Steps:
  - Integration of the MPC with the ROM-based resilient estimator (ACRE)

Task 6: Overall Technology Integration

- First version of high-fidelity model developed (PNNL)
- Next version of high-fidelity model released (PNNL)
- MILESTONE 6.1 Deliver report on high-fidelity model selection (M12), established