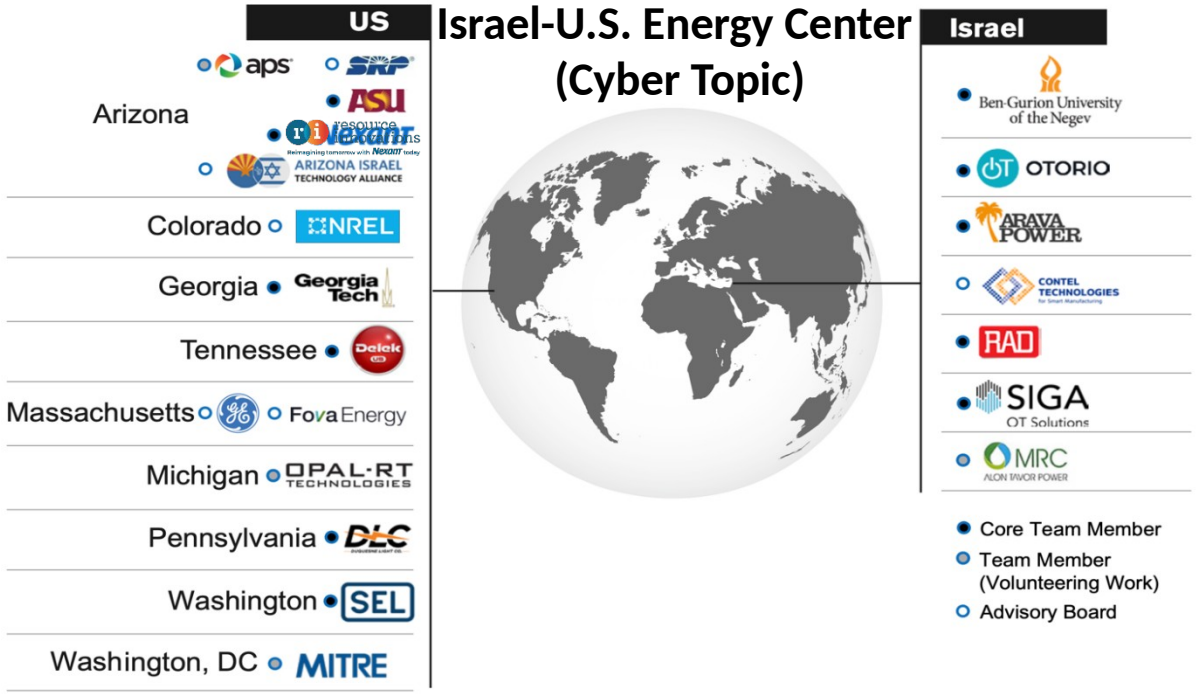
Comprehensive Cybersecurity Technology for Critical Power Infrastructure AI-Based Centralized Defense and Edge Resilience



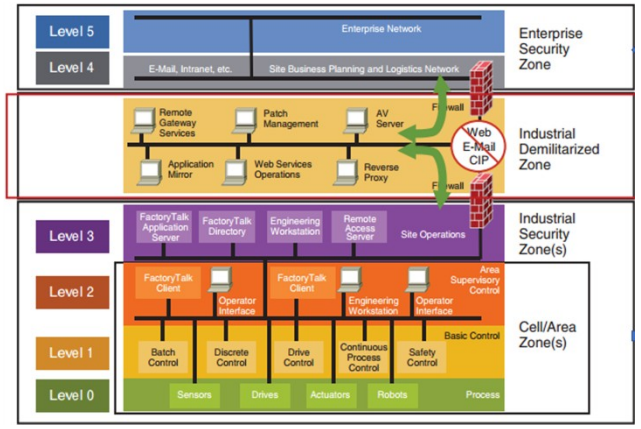
Prepared for

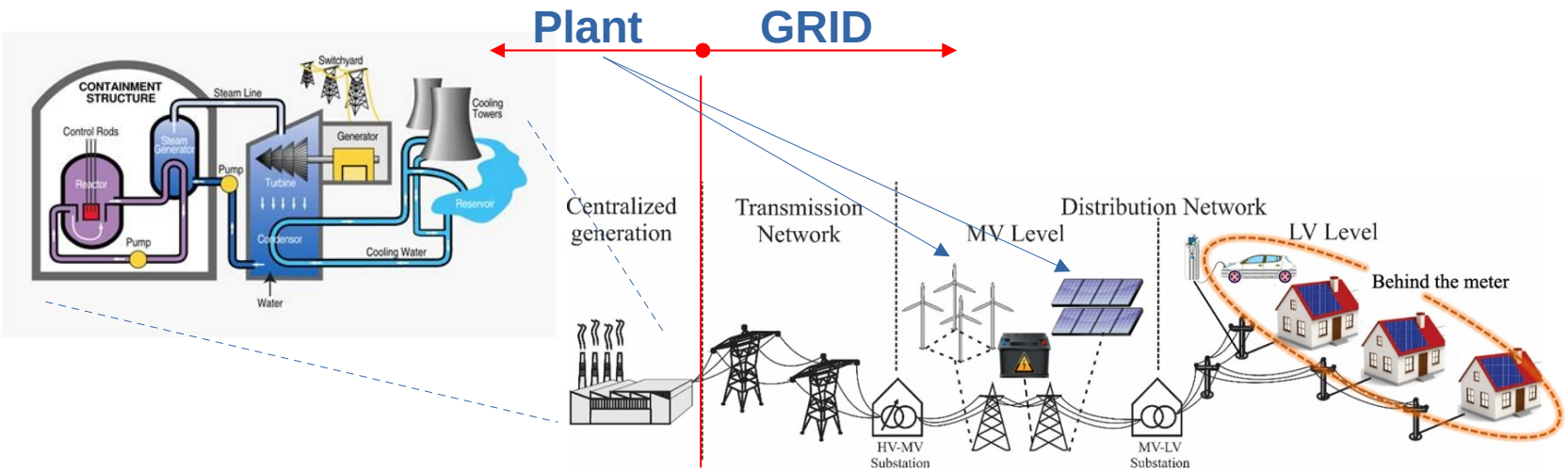
Eitan Yudilevich, Eynan Lichterman, and **Tal Fischelovitch**

BIRD

Aug. 25, 2022

CPS Devices in Energy Infrastructure





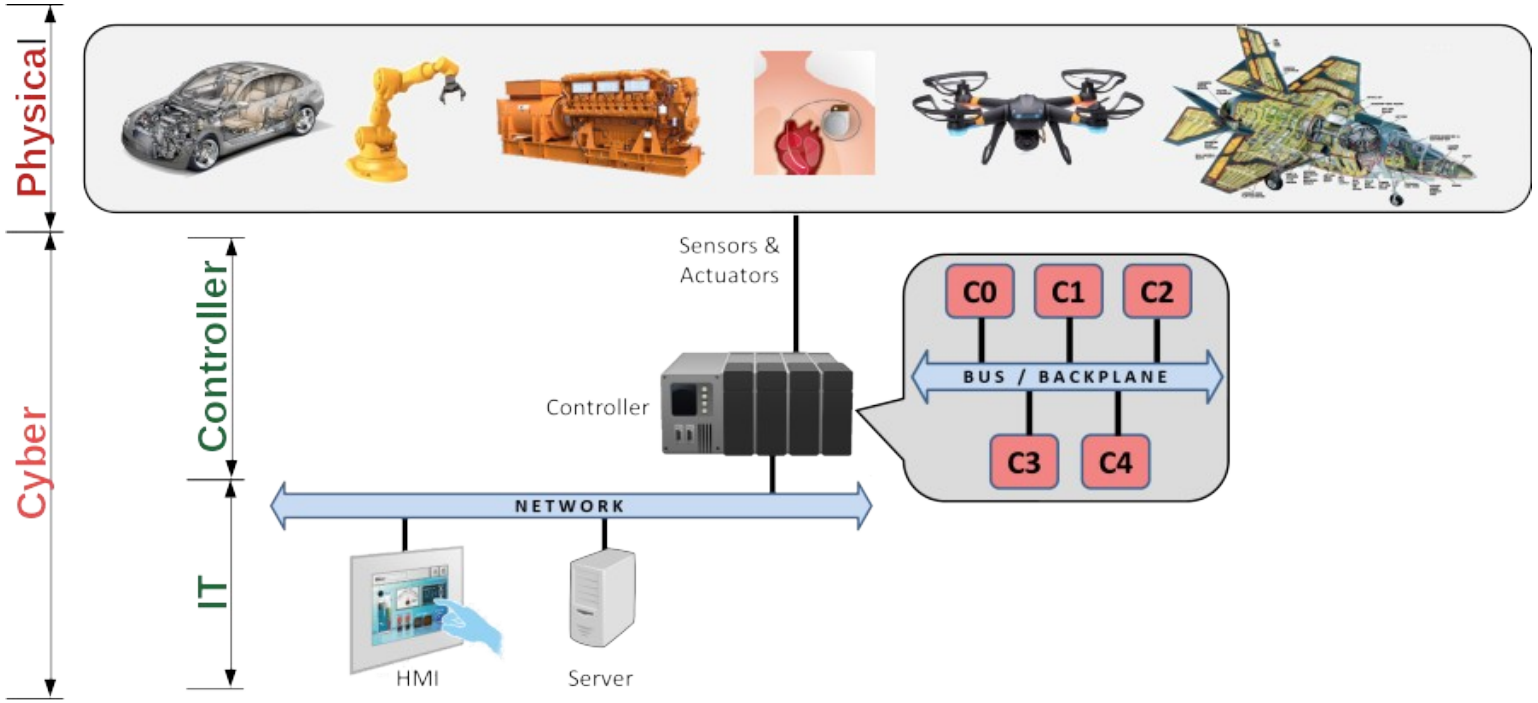
Power Transmission & Distribution GRID

- Small number of component/device type, large number or replication; rather homogeneous
- Underlying process: events propagation

Power Generation Plants

- Large & diverse component/device types, low number of replication; heterogeneous
- Underlying process: varies

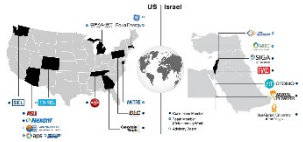
Device Level Security: Robustness from the Ground Up



- Effect of Compromised Device:
  - Compromised Devices → '**Insider Threats**'
  - **Lie** to monitors – doing one thing, reporting another (e.g. Stuxnet)
  - Transport-Layer/communication **encryption** generally **irrelevant** (data is generally ephemeral) – protecting the attacker. Although, **authentication** is **relevant**.

- Cyber Attack Resilience
  - Relying on CPS (controller) properties to tolerate direct cyber attack
  - Agnostic to the specificity of the attack (malware)
  - Complementary multi-factor authentication for firmware update, help to complete the security posture.

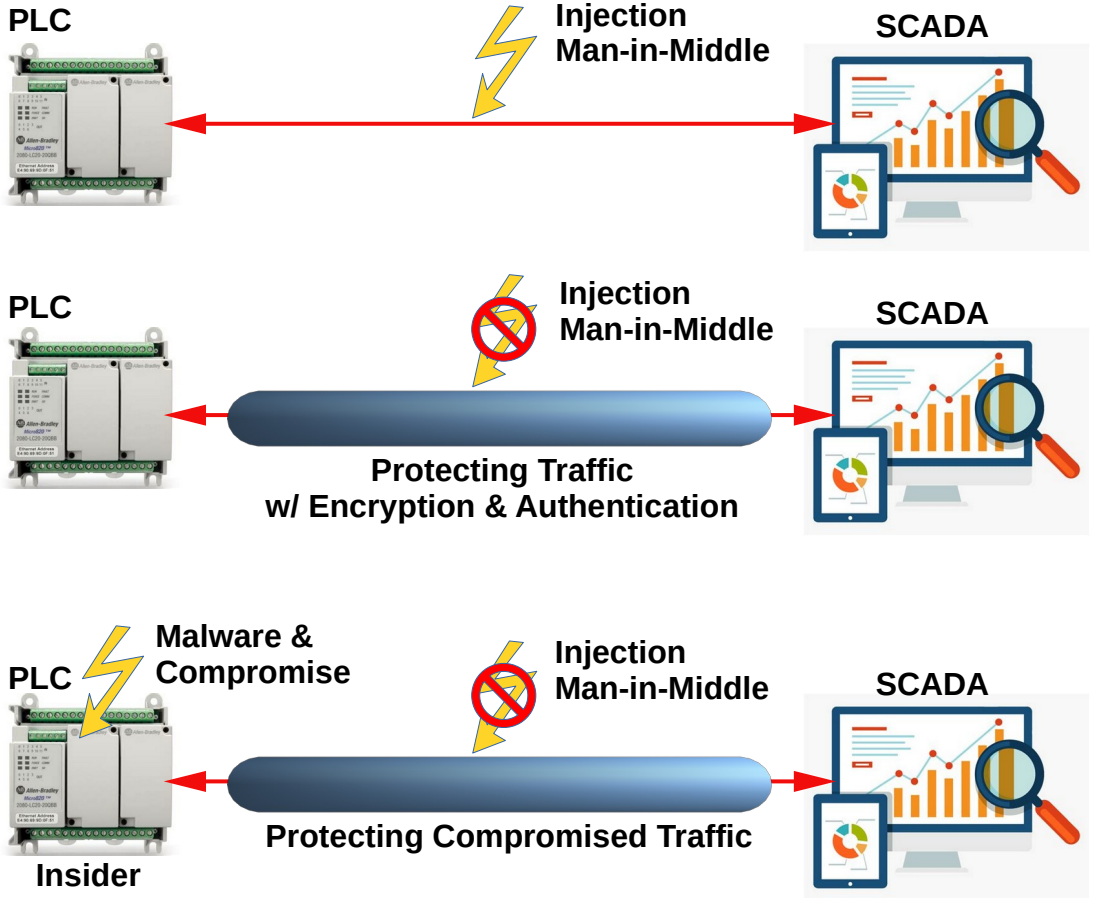
Device Level Security is Paramount



- **Compromised Devices act as Insider threat**
- **Communication & Network security:**
  - Authentication is always beneficial.
  - Encryption reduces attacker's ease of tracking & spoofing,

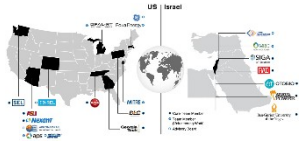
    but also complicates defender's monitoring efforts.
  - Effective against data injection & Man-in-the-Middle.
  - **Not** effective against **Insider Threats.**

Robust & Timely Defense and Resilience Cyber Physical Systems' Device is Highly Desired



Georgia Tech

CPS stability concerns



Cyber **Physical** **Systems**

Supporting
Is a "means to an end"

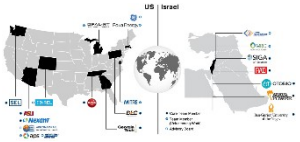
Primary Concern
"Stability & Integrity"

Cyber-attack resilient solutions should be primarily defined and motivated by physical requirements

The goal is for the physical subsystem to be stable, and not necessarily the cyber subsystems

Georgia Tech

CPS controller properties



Periodicity

- Continuous observe and control loop (scan cycle, usually ~1-300 Hz)

Periodicity tolerates some loss of inputs or data

- Sensitive to latency variations
- Not performing open-ended, general-purpose tasks like IT

Inertia

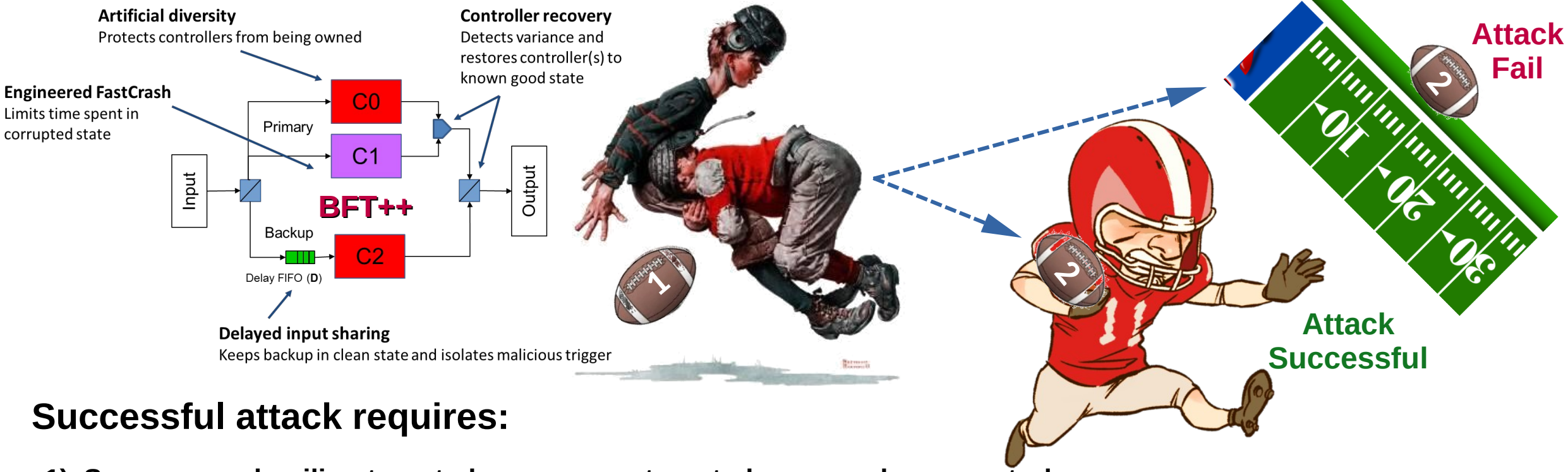
- Physical systems have inertia
- Effect: can tolerate some bad cycles and still maintain stability
- Missed output
- Wrong output (sensor blip, etc.)
- In context of cyber attack:

Inertia provides some natural (output) fault tolerance

- Not immediately uncorrectable
- How long is system-dependent

Periodicity and Inertia enable BFT++

Cyber Attack Invariant



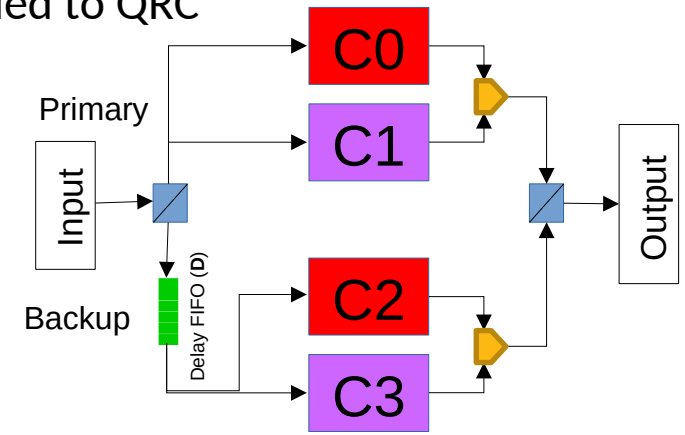
Successful attack requires:

1) **Success on derailing targeted program --> targeted program loses control**
   - Defense: avoid any bugs and flaws (formal methods, protection techniques)
2) **Success on capturing control --> attacker controls program execution**
   - Defense: avoid any predictability, cross-check with diversified version

ONR's RHIMES (BFT++)

BFT++ v1 (Vanilla) – NRL

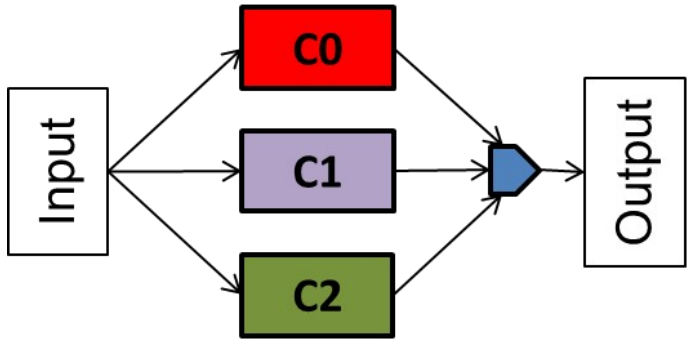
Original BFT++ applied to QRC (QRC++)



BFT++ v2 – Georgia Tech

More **robust**, more **costly,** protection is diversified

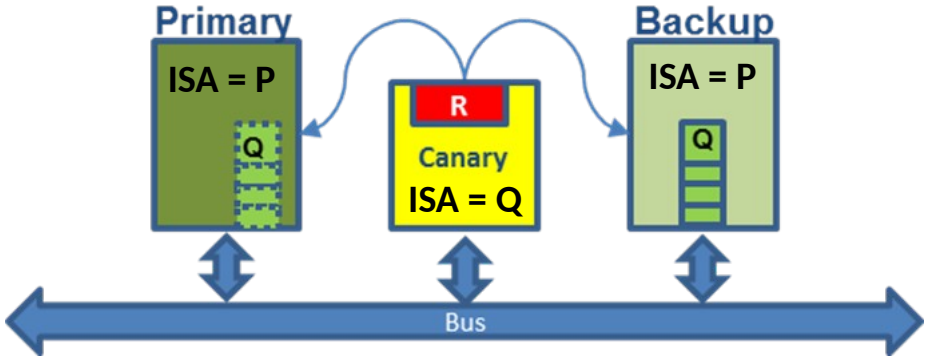
Diversified protection sets provides comprehensive protection with reasonable distributed overhead



BFT++ v4 (Rum Raisin)

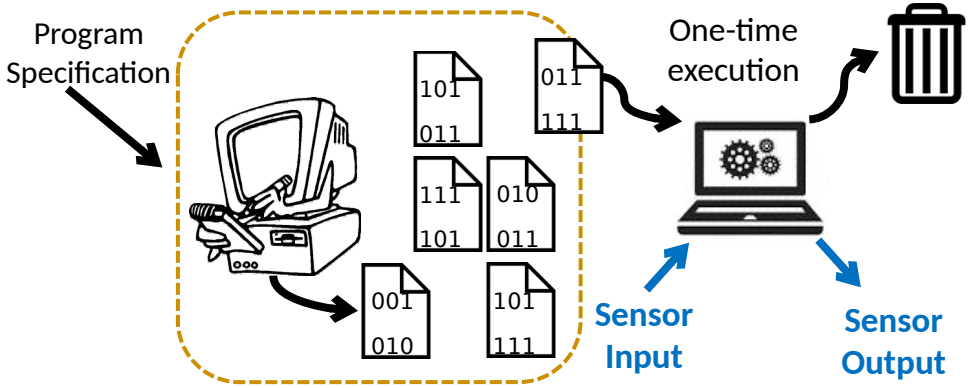
Variation of the original: ISA diversity

Q = delay queue
R = recovery trigger logic



BFT++ v3 – Columbia University

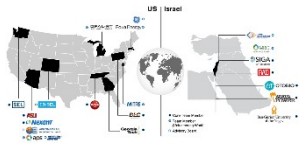
Lightweight, **probabilistic** guarantee, needs **no redundancy**



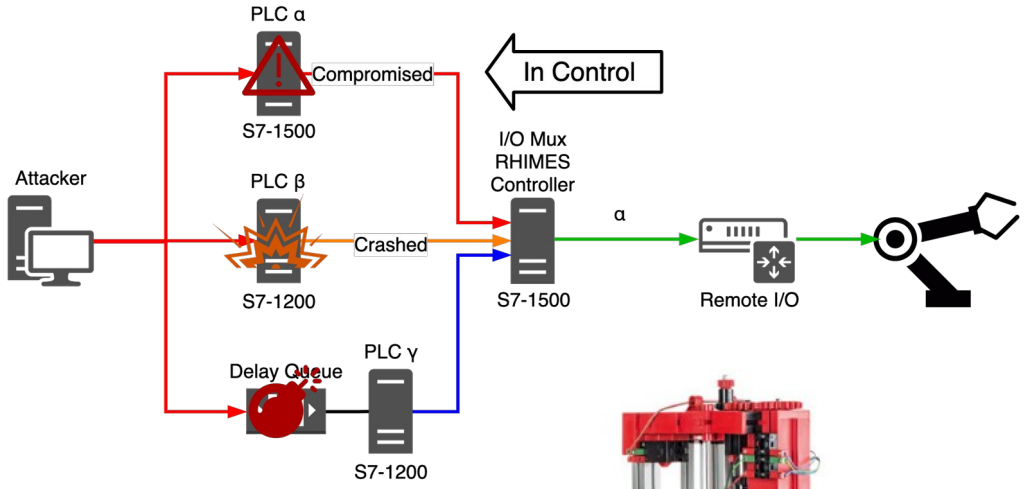
Georgia Tech

MITRE's RHIMES Laboratory Experiment

contributed by: Matt Mickelson, MITRE



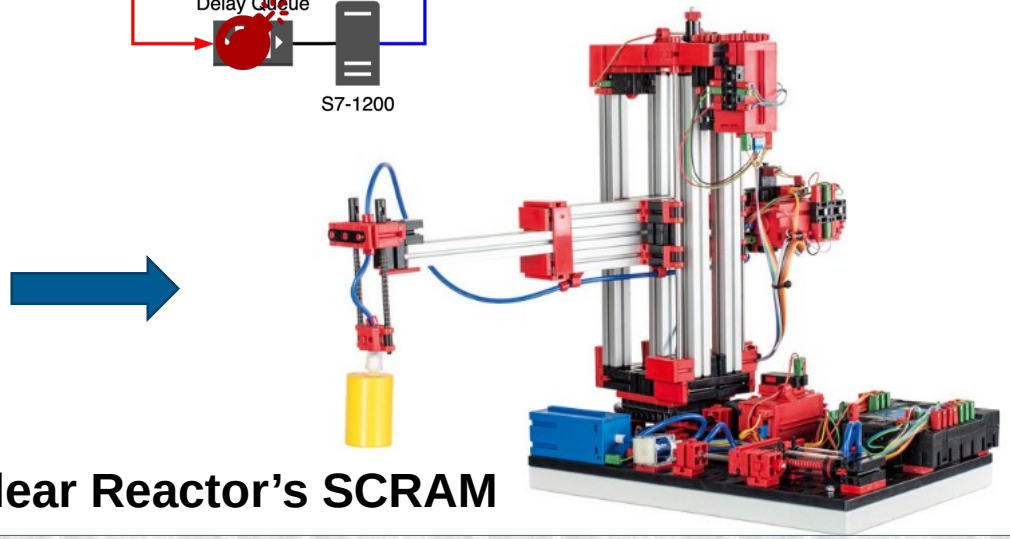
Hypothesis: The time it takes to detect a crash and switch to a hot backup PLC is less than the time it takes to lose a "puck" due to inertia of the gripper losing grip.



Hypothesis Confirmed

Full recovery is acquired if the first 2 PLCs can be rebooted and reassume control.

MITRE

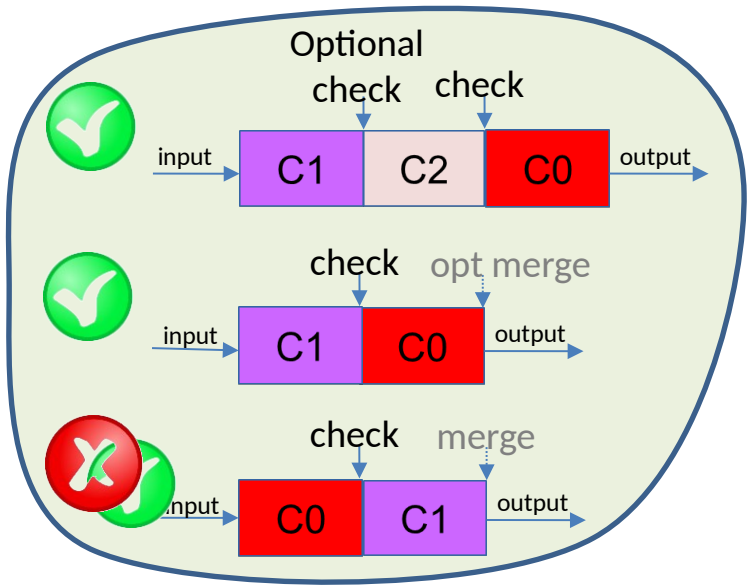


Emulating Nuclear Reactor's SCRAM

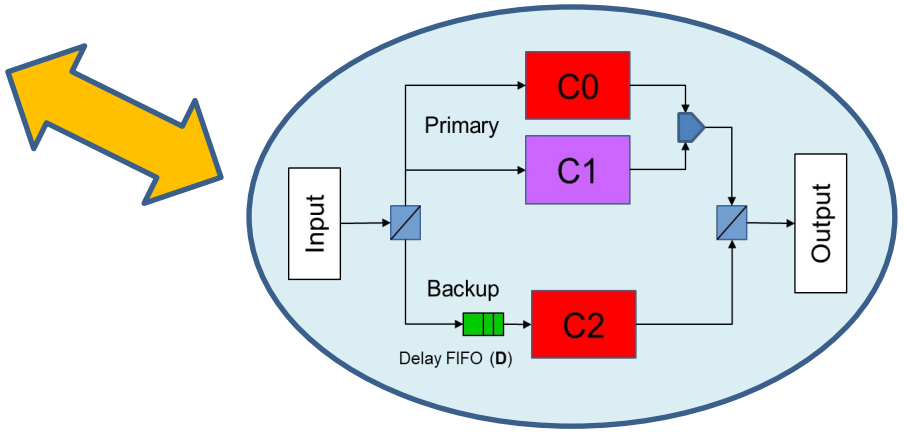
Demonstrated continuous operation despite repeated cyber exploit

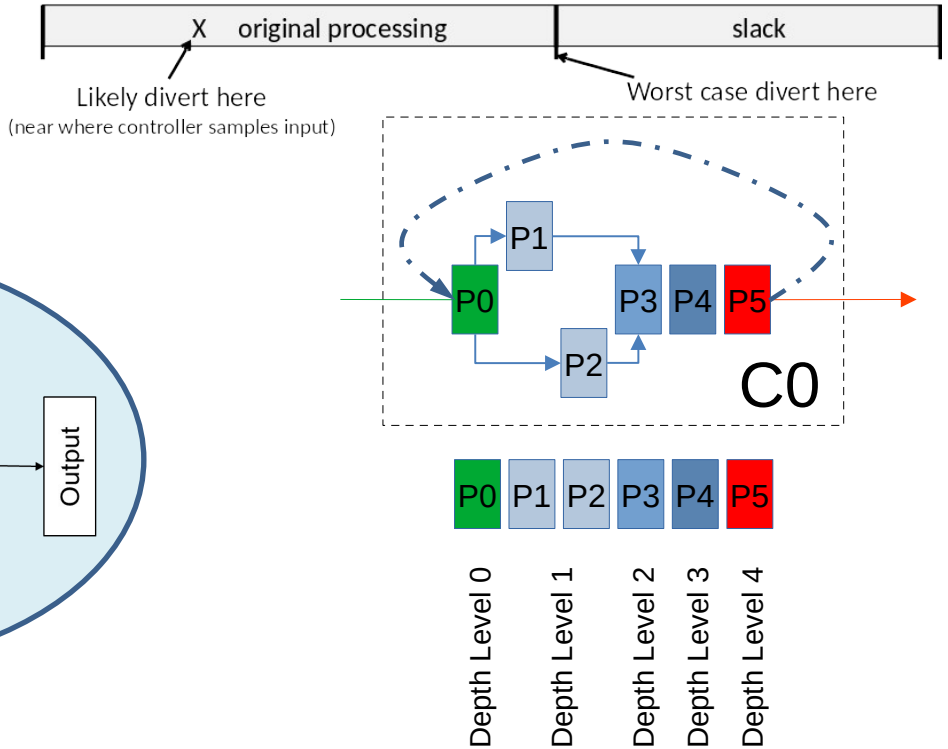
Georgia Tech

Diversified Redundancy on Single Processor
Parallel vs. Serial



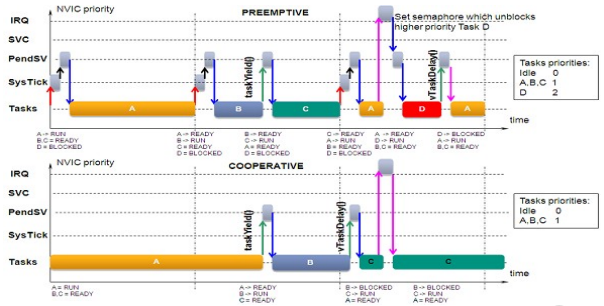
Legends:
- **C0** Original control program
- **C1** & **C2** Diversified programs
- If **check** fails, drop input



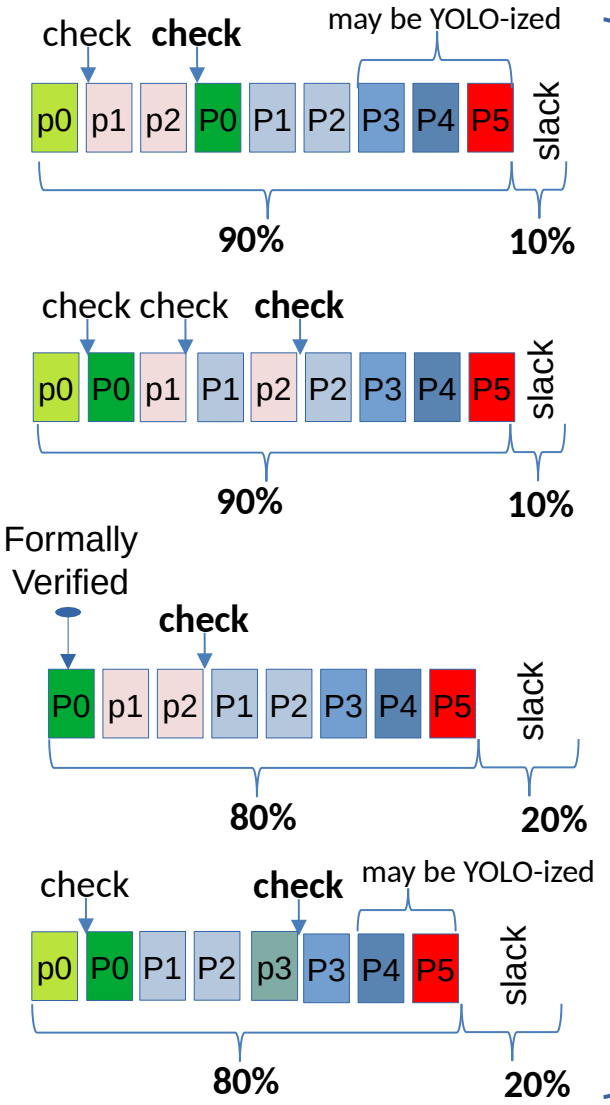


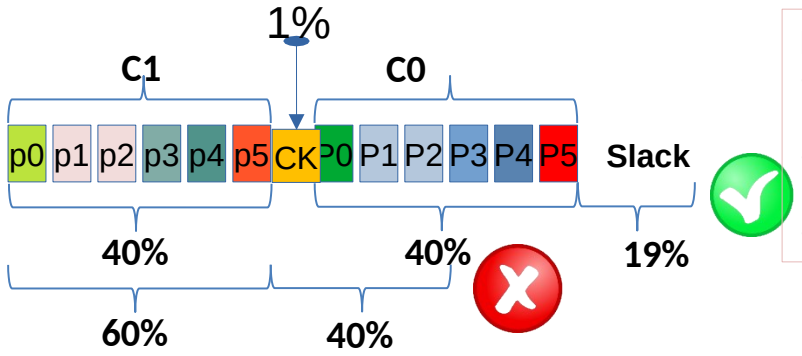
Design parameters:

- Slack availability
- Number of sensitive processes
  - Depth relative to input
- Pre-emptive vs. Co-operative scheduling



Diversified Redundancy on Single Processor
Serial in Finer Granularity

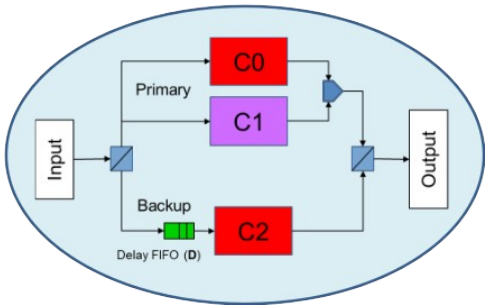




Legends:
- **C0** Original program
- **C1** Diversified program
- If **check** fails, drop input

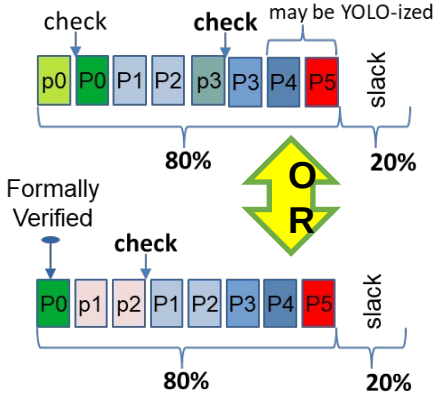
1001 ways to implement BFT++ concept w/ sub-process replication;
- **Diversified replication can co-exist w/ Formal-Methods, Protections & YOLO**
- **Engineering for sub-processes replication depends on:**
  - **Available Slack & Desired Slack,**
  - **Sub-processes' Depth Level,**
  - **& particular sub-process' properties**

Cyber-Attack Resilience for CPS – Part B



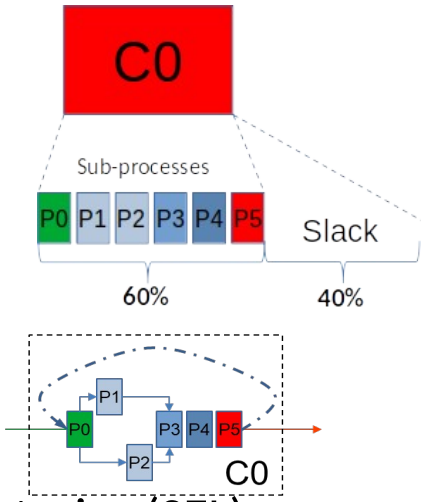
BFT++

equival ent



SubProcess BFT++

SubProcess BFT++ Engineering **Tool**

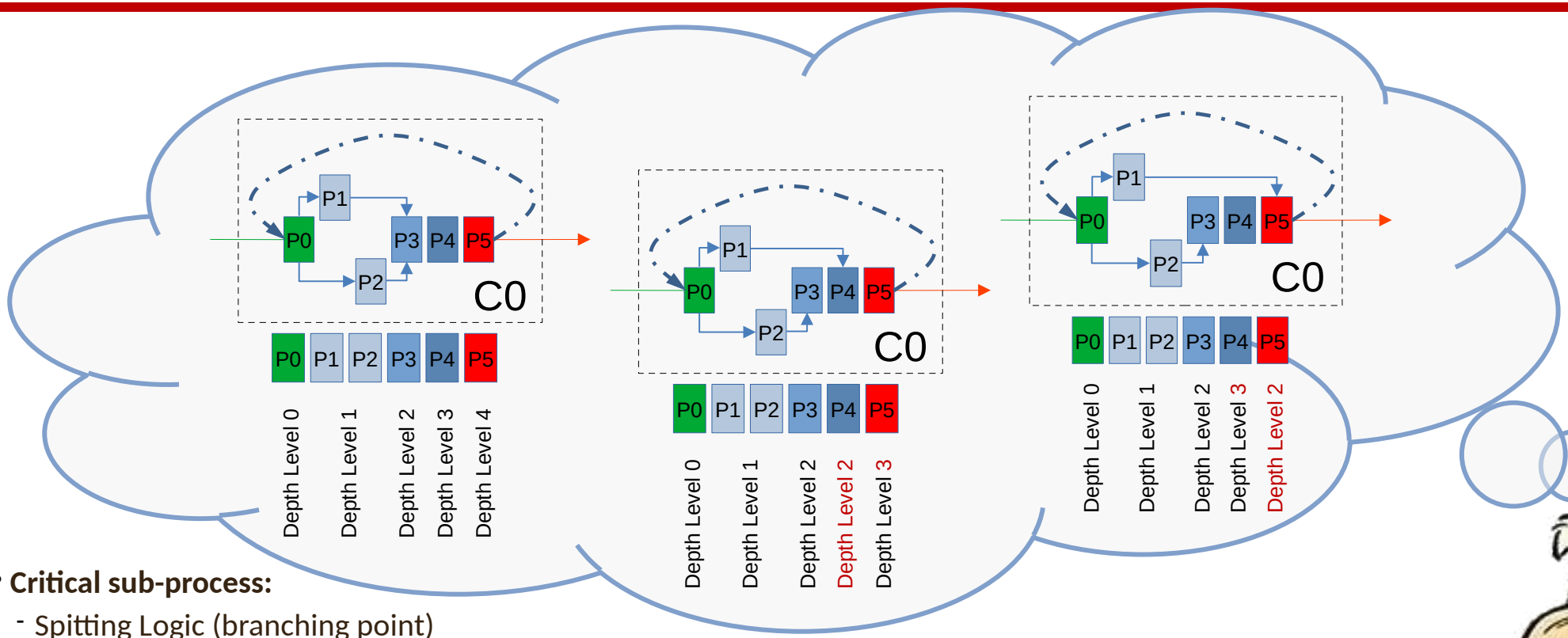


We plan to integrate the SubProcess BTF++ engineering tool into ~~Schweitzer Engineering Laboratories (SEL)~~ open source PLC design tools and environment.

Impact:

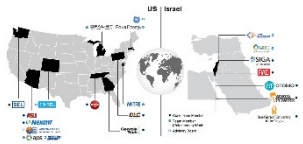
- Alleviate the need for redundant device in BFT++, providing cyber attack resilience for application which cannot afford device redundancy
- Significantly widen the applicability of BFT++ and resilience against direct cyber-attack
- BFT++ automatically isolate offending data, can be communicated to other system components, e.g. SCATOPSY, RAM$^{2.}$, to prevent repeat attack.
- Integration into ~~SEL~~ open source PLC (or Siemens) design environment for ease of deployment and dissemination.

Theory & Formulation to Develop



- **Critical sub-process:**
  - Spitting Logic (branching point)
  - Gating Logic (where 2 or more flow merge)
  - Sensitivity Level (minimum distance to inputs)
- **Need to develop Theory & Formulations**
  - Identification of key sub-process (Sensitivity Level, Gating & Splitting Logic)
  - Sub-Process BFT++ per-component mode selection/trade-offs: redundant-diversified, YOLO, or Formal methods

Cyber-Attack Resilience for CPS – Progress



Current status:

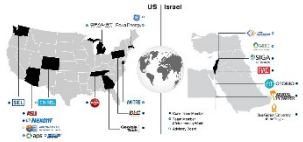
- Coordination meetings with SEL:
  - Initial Coordination meeting on April 5th 2022
  - Training for SEL RTAC development tools and environment on April 18th 2022
  - Meeting for further deep dive into SEL's:
    - Operating System implementation
    - Compiler and code generation process
    - Real time scheduler
    - Etc.

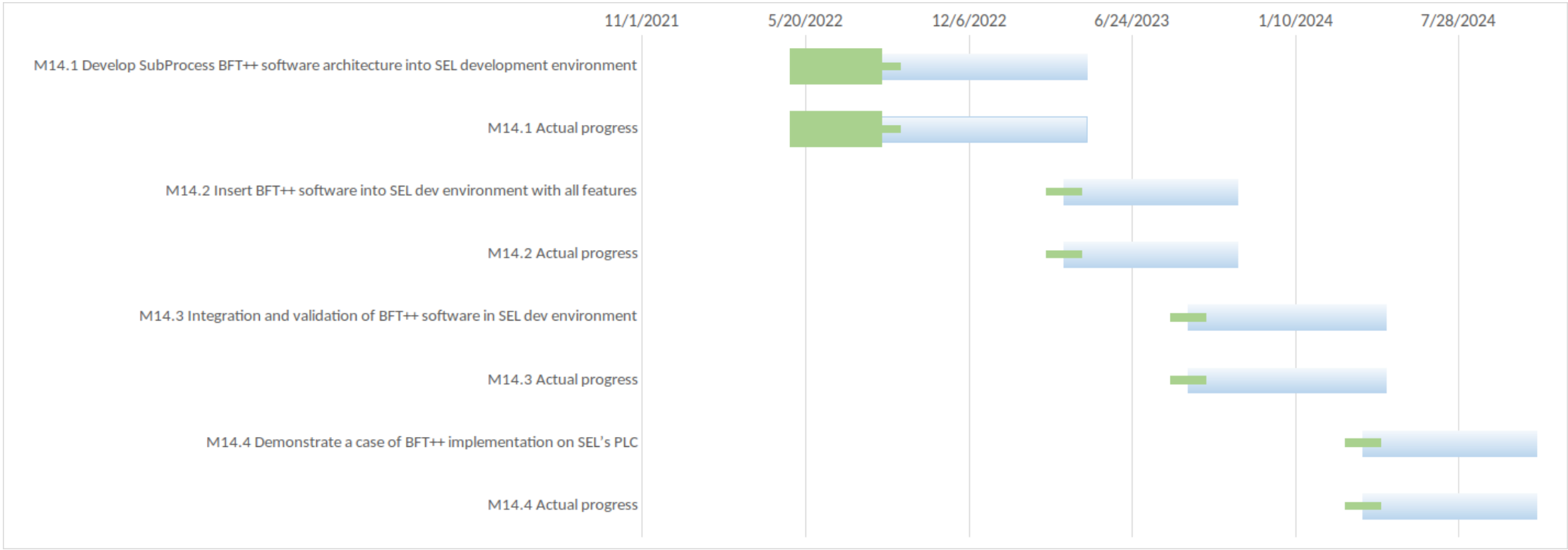
  SEL cannot support transition
- Changing Transition Plan:
  - Transition into open source OpenPLC
  - Siemens ???

- Started on May 1st 2022
- Initial research will use an open source PLC environment: **ClassicLadder**.
  - For experimentation platform and
  - For analyzing generated codes for PLCs
    - Understanding scheduling structure
    - Studying design trade offs for integrating sub-process BFT++
  - A Linux toolset, as oppose to Windoze
- Exploring OpenPLC for R&D (instead of ClassicLadder) and Transition
- Future: integration into OpenPLC ??? design tools and environment

Georgia Tech

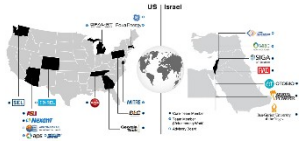
Schedule & Milestones:

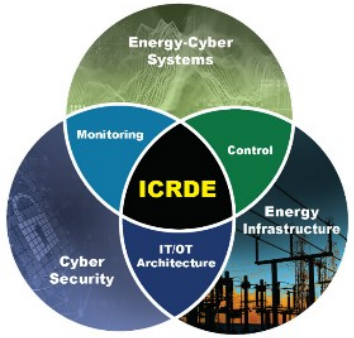




- We are starting in May 2022
- Team:
  - Dr. Sukarno Mertoguno
  - Interviewing a new PostDoc 8/31
  - M. Faraz Karim, Ph.D. student visa isn't completed in Pakistan, delayed arrival

Comprehensive Cybersecurity Technology for Critical Power Infrastructure AI-based Centralized Defence and Edge Resilience





Theme 1: Cyber-Physical System (CPS) Data Preparation

- **Project 1**: Physical Processes as Cybersecurity Assets
  - **T1: Realization of advanced energy management applications**
  - **T2**: Digital representation of physical processes and operational process modelling
    - **T3**: Data collection and aggregation
- **Project 2**: Knowledge Base Establishment and Representation Learning of Threats
  - **T4**: Multi-level threat intelligence knowledge base (MITRE)
    - **T5**: Adversarial mimicking (Theoretic)
    - **T6**: Threat hunting (Actively learning)

Theme 2: Data Analytics for Monitoring

- **Project 3**: Cyber Attacks Detection of IT/OT Architecture
  - **T7**: Malware threats mitigation (Smart enemy? Reduce probing and plan)
  - **T8**: Detect event mimicking attacks
    - **T10**: Multilayer anomaly detection (Uncertainty)
    - **T9**: False Data Injection (Physical Embedding)
  - **T11**: AI-based intrusion detection (Engage operator)
  - **T12**: Explainable cyber-AI analytics (Metric + Temporal)

Theme 3: Control and Validation

- **Project 4**: Cyber Resilience and Robustness for Control Actions
  - **T13**: Firmware verification (Edge)
  - **T14**: Attack tolerance (Insensitive to the attack)
  - **T15**: Self healing and auto-remediation (Optimization)
    - **T16**: RL control for CPS (Advise Sys Operator)
- **Project 5**: Future-Proof Architectures
  - **T17**: ICS security by design
  - **T18**: Hardware in the loop