ADVANCED REACTOR SAFEGUARDS & SECURITY

# Use of AI/ML to Assess Cyber Vulnerabilities in Diverse Defense-In-Depth (D3) Advanced Reactor Designs and Impact on Common Cause Failures

*ARSS Spring Program Meeting April 28-30, 2026*


PRESENTED BY

Fleur de Peralta, Project Manager

Dr. William Hutton, Principal Investigator



PNNL-SA-222544



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.


Sandia National Laboratories | U.S. DEPARTMENT OF ENERGY

# Research Objectives


- Use of artificial intelligence and machine learning (AI/ML) methods to develop a decision-making analysis framework to assess vulnerabilities in Diversity of Defense-In-Depth (D3) advanced reactor designs and Common Cause Failure (CCF)
  - Modern programmable digital devices contain **multiple hardware and software components** and often **complicated supply chains**.
  - Vendor diversity in components is no guarantee of diversity and may hide CCF issues.

- Use of open-source AI/ML models to automate the analysis of data sets reduces human error, improves time-efficiency and produces more complete and correct analysis results.

# Approach


- Identify Digital Devices and software used in advanced reactor designs
  - Collaborate directly with industry stakeholders (developers, utilities, NEI, etc.)
  - Identify integration of these devices into advanced reactor design and safety systems
    - Diversity of Defense-In-Depth
    - Common Cause Failure Scenarios
- Evaluate HBOM (Hardware Bill-of-Materials) and SBOM (Software Bill-of-Materials) extraction methods
- Evaluate artificial intelligence methods (Large-Language Models) and Machine Learning approaches
- Evaluate suitability of AI/ML methods of D3 Analysis and CCF
- Document results of research and feasibility of approach

# Why SBOMs and HBOMs?


- Software Bill of Materials (SBOMs) identify various components used to build the software (open source or proprietary components).
- Hardware Bill of Materials (HBOMs) identify physical hardware parts, components and firmware in a device.
- Transparency enables organizations to identify vulnerable software and hardware components that are then identified in databases and security advisories of known vulnerabilities
- Provides insights on source of components and enables visibility to the software supply chain traceability and dependencies

## Challenges with Approach


- SBOMs and HBOMs may be viewed as proprietary or a competitive advantage by vendors and system designers, leading to repeating work to generate HBOMs and SBOMs for standard industry devices.

- Without HBOMs and SBOMs, programmable digital devices may be opaque to subject matter experts, leading to increased uncertainty on risk despite best efforts of D3 and CCF analysis.

# Cybersecurity Scenarios, Partial

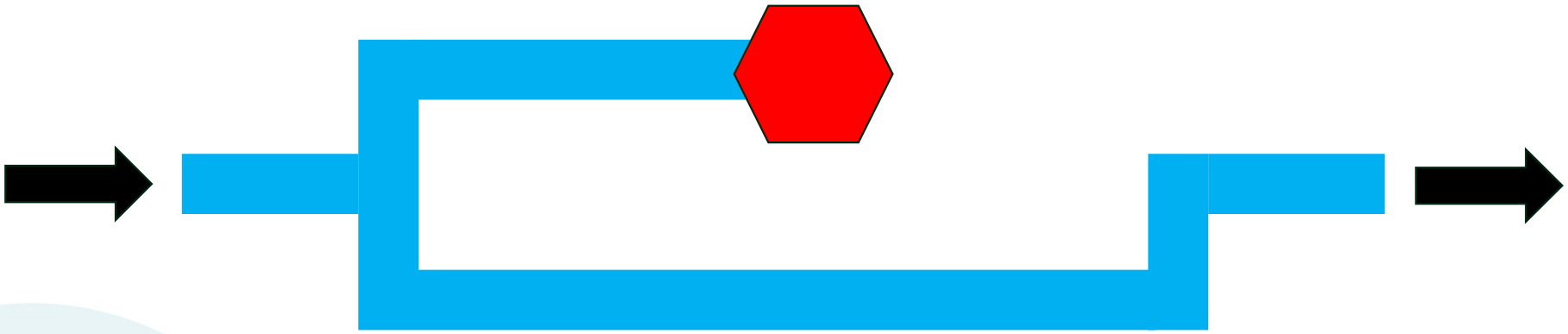
Traditional **fault tolerant design** pattern based on Lamport's *Byzantine Generals* problem (1982)

Single channel (water) is vulnerable to breaks and blockages

# Cybersecurity Scenarios, Partial

Two (or more) channels provide redundancy

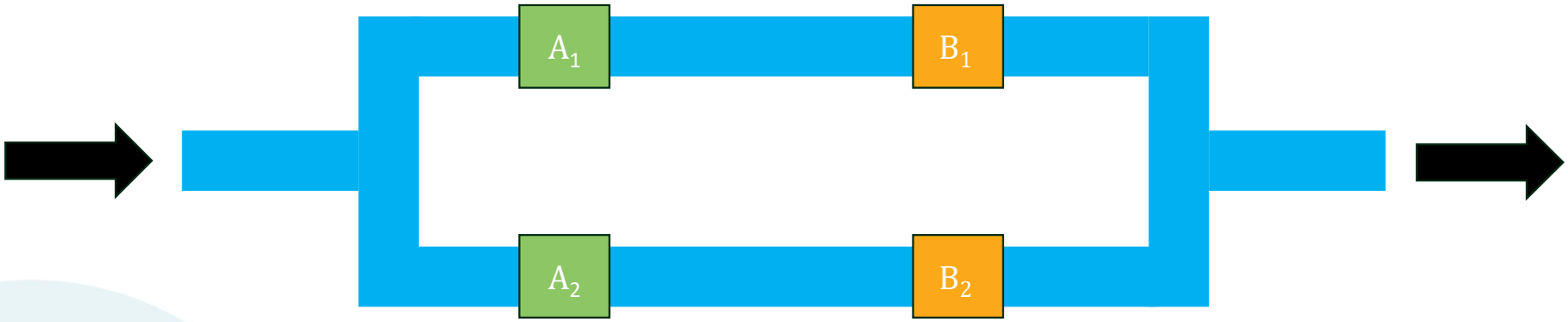
# Cybersecurity Scenarios, Partial



Two channels provide redundancy

# Cybersecurity Scenarios, Partial

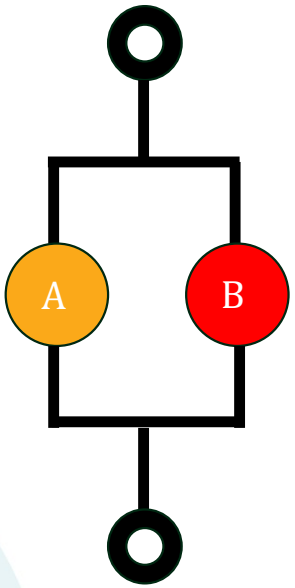
The more useful **two-channel, four sensor** design pattern



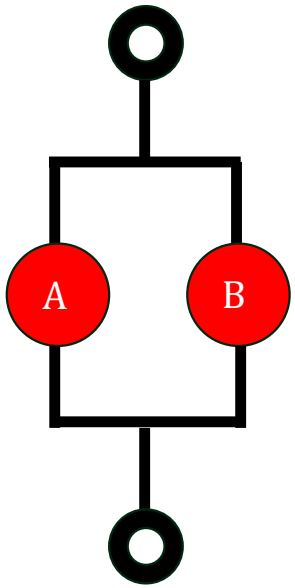

3n+1 sensors (4) can tolerate 1 failure (n)

# Cybersecurity Scenarios, Partial

Two ways Diversity and Defense-in-Depth ($D^3$) may introduce vulnerabilities

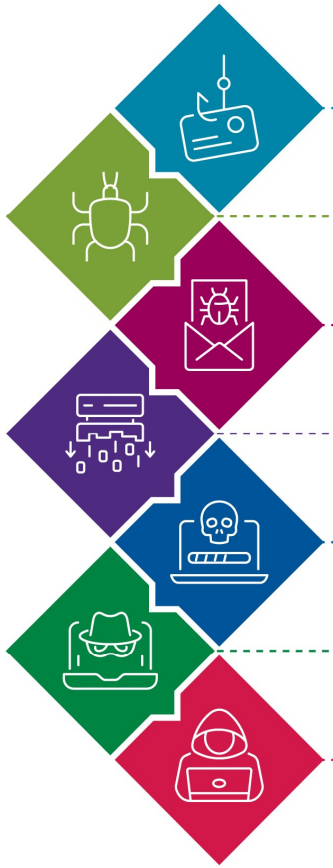



A. Objectively "better"




B. Hidden homogeneity (CCF)

# Diversity and Defense-in-Depth

**1. RECONNAISSANCE**
Harvesting email addresses, conference information, etc.

**2. WEAPONIZATION**
Coupling exploit with backdoor into deliverable payload

**3. DELIVERY**
Delivering weaponized bundle to the victim via email, web, USB, etc.

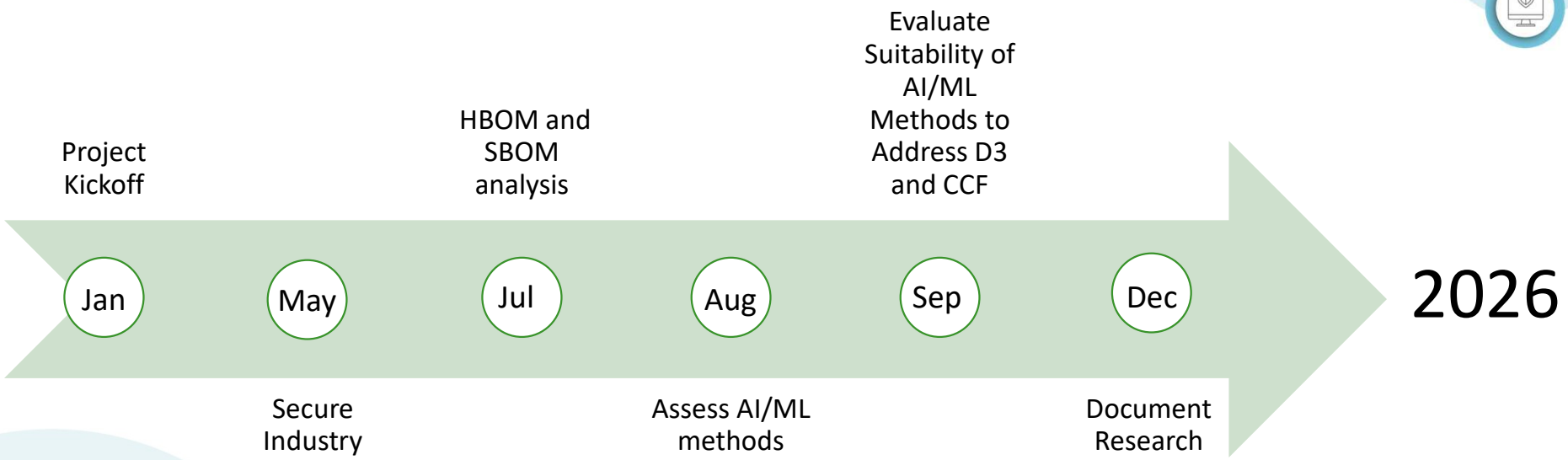
**4. EXPLOITATION**
Exploiting a vulnerability to execute code on victim's system

**5. INSTALLATION**
Installing malware on the asset

**6. COMMAND & CONTROL (C2)**
Command channel for remote manipulation of victim

**7. ACTIONS ON OBJECTIVES**
With "Hands on Keyboard" access, intruders accomplish their original goals

# Research Timeline



# Progress to Date


- Met with industry partners
  - Advanced reactor developers (Kairos)
  - Utilities deploying advanced reactors (Columbia Generating Station and X-Energy)
    - HBOM / SBOM extraction methods
    - Artificial intelligence models (LLMs)
    - Machine learning algorithms
  - Subject matter experts (NEI)
    - Nuclear safety systems
    - Computer science
- Drafted NDAs

# PNNL Team


**Fleur de Peralta, P.E.**
Nuclear Safety Systems and Cybersecurity
Email: fleurdeliza.deperalta@pnnl.gov

**Dr. William Hutton, CISSP**
Sr. Cybersecurity Researcher
Email: Will.Hutton@pnnl.gov

**Corrine Roth**
Data Scientist
Email: corinne.roth@pnnl.gov