SANDIA REPORT
SAND2021-11995
Printed September 2021

Sandia National Laboratories

Cyber-Physical Risks for Advanced Reactors

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Milestone No. M2CT-21SN1104023

Ray Fasano, Andrew Hahn, Jacob James, Christopher Lamb, Alexandria Haddad

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

Cybersecurity for industrial control systems is an important consideration that advance reactor designers will need to consider. How cyber risk is managed is the subject of on-going research and debate in the nuclear industry. This report seeks to identify potential cyber risks for advance reactors. Identified risks are divided into absorbed risk and licensee managed risk to clearly show how cyber risks for advance reactors can potentially be transferred. Absorbed risks are risks that originate external to the licensee but may unknowingly propagate into the plant. Insights include (1) the need for unification of safety, physical security, and cybersecurity risk assessment frameworks to ensure optimal coordination of risk, (2) a quantitative risk assessment methodology in conjunction with qualitative assessments may be useful in efficiently and sufficiently managing cyber risks, and (3) cyber risk management techniques should align with a risked informed regulatory framework for advance reactors.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the hard work and commitment of all contributors to the project. In particular, we would like to acknowledge the strong support and leadership of Rebecca Onuschak at the Department of Energy. Chris Spirito, Katya Le Blanc (INL) and Lon Dawson, Michael Rowland (SNL) are to be commended for their programmatic and technical guidance.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page left blank

EXECUTIVE SUMMARY

Advanced Reactors (ARs), also referred to as small modular reactors (SMRs) and generation IV reactors[1], are the nuclear energy of the future. In general, they are projected to be less expensive to build and operate because of their smaller size, output, and advanced safety features, providing a reduced financial risk to investors and to public safety. If successful ARs will eventually replace the current generation II fleet, providing clean energy integration with the grid, while working synergistically with renewable energy sources such as solar and wind. ARs have the potential to dramatically increase productivity of clean energy grid assets, improve energy grid resilience, and provide a clear path to meeting U.S. clean energy goals and promises. This will be accomplished by providing a robust baseload when energy from renewable sources is not available due to inherent intermittency or natural disasters.

Novel cyber-physical risks, from the perspective of the nuclear industry, must be thoroughly explored and understood prior to the implementation of ARs. Development of ARs without an effective cyber-physical risk management framework is possible but will greatly reduce the ability of AR designers to leverage cutting-edge digital technologies. Although costly in the short-term the long-term benefits of developing cyber-physical risk assessment methodologies for ARs can potentially lead to cost reductions, increased productivity, and resilience of power generating assets under cyber-attack. Reduction of cyber-physical risks should be included in the overall risk reduction strategy. Deployment of fleets of ARs represents an exponential increase in logistical complexity over generation II reactors due to the projected number of active reactor cores. Such a scenario may require the use of advanced computational tools and secure wireless infrastructure. To ensure future scalability, cyber-physical risk assessment methodologies may become a fundamental requirement in the future. This topic is an active area of research and regulatory concern.

Cyber-physical risks to ARs are identified in this report and limitations in the current cyber risk guidance for nuclear power plants (NPPs) are discussed. The cyber-physical risks covered in this report are established into two categories, absorbed risk and licensee managed risk. Under these categories the cyber-physical risks covered are supply chain, autonomous operation, AR OT architecture, off-site security operations, on-site security operations, and remote operation. These cyber-physical risks track the origination of the risks and denote how they are relevant for ARs. Other cyber-physical risks of interest not covered in this report are also enumerated in Table 1-1. Absorbed risks are those that originate externally to a licensee but may unknowingly be absorbed into the AR design. Ideally risks would be evenly distributed amongst the licensee and vendors however, when it comes to safety and security this is not true. In nuclear power it is the licensee's responsibility to ensure the safety and security of the plant up to and including the plants design basis.

Table 1-1. Cyber-Physical Risks for ARs

Absorbed Risk	Licensee Managed Risk
Autonomous Operation	On-site/Off-site Security Operations

[1] Definitions: https://world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-power-reactors/small-nuclear-power-reactors.aspx (SMRs) & https://world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-power-reactors/generation-iv-nuclear-reactors.aspx (Generation IV reactors)

Absorbed Risk
Advanced reactor OT architecture
External cellular networks
Security & network management tools (i.e. SolarWinds),
Software and control system development environments (i.e. Studio 5000, Sematic Step 7, etc..)
Supply Chain

Licensee Managed Risk
Remote operation
On-site hydrogen production
Virtual and augmented reality for operations and maintenance
Unmanned aerial vehicles
On-site special nuclear material inventory

Current U.S. guidance for securing reactor systems is based on identifying individual critical digital assets (CDAs) and administering all applicable security controls [1, 2]. By not taking system risk into consideration and instead focusing on asset-based mitigation, this methodology places a heavy burden on the licensee as they take on the responsibility of deeming when control measures are sufficient across their site. As the industry shifts towards establishing a risk informed framework, all stakeholders involved in AR operations will be better prepared to mitigate cyber-physical threats [3].

Risk responsibility remains a challenge for AR ownership. Under U.S. light water reactor (LWR) guidance, the plant owner is responsible for identifying and modifying risk within their reactor systems. While this methodology is extensive given the high number of CDAs identified, it is within the operator's ability to maintain and apply controls for the current nuclear fleet. AR owners and operators, however, will not have the capability to assess and manage cybersecurity risks in systems that have been designed and manufactured by an external party. The approach of CDA identification and assessment may not be feasible as the plant owners will have a limited knowledge of the system and licensee's limited capability to modify or access the risk of a system or component. Thus, in-order to discuss cyber-physical risks for ARs it is also pertinent to understand current cyber risk management framework for NPPs.

This report has identified limitations in the current practice of cyber physical risk management for ARs and identified important cyber physical risks that AR designers should carefully consider. Methodologies like Hazard and Consequence Analysis for Digital Systems (HAZCADS) provide the initial technical basis for quantitatively identifying and analyzing risks associated with CDAs. However, fundamental investments need to be made to improve current cyber-physical risk methodologies and unify security and safety-based analyses.

On-going challenges regarding cyber-physical risk methodologies include organizational and funding divisions between cybersecurity, physical security, and safety analysis groups for ARs. Cross-cutting research will need to be continuously pursued to ensure that cybersecurity, physical security, and safety analysis groups are working towards a common goal, as well as scalable and modular solutions. If ARs repeat the same mistakes as generation II reactors, U.S. based ARs may not be economically competitive or manageable at scale. AR designers should be encouraged to continue to innovate and regulators need to have the necessary tools and resources to properly assess associated risks. Coordinating the relationship between DoE NE, U.S. NRC, the National Laboratories, and

the nuclear industry will need enduring resolve to ensuring stress points in relationships are mitigated and an optimal balance in the public-private partnership maintained.

ACRONYMS

AI	Artificial Intelligence
AR	Advance Reactor
BoP	Balance of Plant
CAS	Central Alarm Station
CDA	Critical Digital Asset
CFR	Code of Federal Regulations
CS	Critical System
CSAT	Cyber Security Assessment Team
DBT	Design Basis Threat
DEG	Digital Engineering Guide
DiD	Defense in Depth
DRAM	Digital Reliability Analysis Methodology
DT	Digital Twin
EPRI	Electric Power Research Institute
FSAR	Final Safety Analysis Report
HAZCADS	Hazard and Consequence Analysis for Digital Systems
I&C	Instrumentation & Control
IC	Integrated Circuit
ICS	Industrial Control System
IEC	International Electrotechnical Commission
INL	Idaho National Laboratories
IP	Internet Protocol
ISMS	Information Security Management System
ISO	International Organization for Standardization
IT	Informational Technology
LWR	Light-water reactor
MCR	Main Control Room
ML	Machine Learning
NEI	Nuclear Energy Institute
NIST	National Institute of Standards and Technology
NPP	Nuclear Power Plant
NRC	Nuclear Regulatory Commission
OT	Operational Technology
PC	Personal Computer
PCB	Printed Circuit Board
PIP	package-in-package

PLC	Programmable Logic Controller
POP	package-on-package
PPS	Physical Protection System
PRA	Probabilistic Risk Assessment
RMF	Risk Management Framework
SAS	Secondary Alarm Station
SD	Secure Digital
SMR	Small Modular Reactor
SNL	Sandia National Laboratories
SP	Special Publication
STPA	Systems Theoretic Process Analysis
TAM	Technical Assessment Methodology
TCP/IP	Transmission Control Protocol over Internet Protocol
TTPs	Techniques, Tactics, and Procedures
USB	Universal Serial Bus

1. INTRODUCTION

This report fulfills milestone report Cyber-Physical Risks for Advanced Reactors (M2CT-21SN1104023) under work package M3CT-21SN110402. This work was sponsored by the Department of Energy's Office of Nuclear Energy (DoE-NE).

Advanced reactors (ARs) are needed to replace the generation II fleet that will largely be decommission in the 2030s and 2040s to enable the United States (U.S.) clean energy goals and retain world leadership in nuclear energy. Generally, ARs are significantly smaller in both size and electrical output, relative to generation II reactors, to reduce financial risk due to construction, simplify integration with the grid, compliment renewable energy, and create manufacturing economies of scale. Consequently, due to reduced operational revenue and a competitive energy market ARs will need to leverage enabling technologies. Sufficiently and efficiently managing cyber-physical risk will be critical for ARs to utilize current and future enabling technologies in the AR's operational technology (OT) architecture. The enabling technologies include, but are not limited to, digital twins, autonomous control systems, predictive maintenance, remote operation infrastructure, virtual & augmented reality, and cloud infrastructure. These technologies not only have the potential to dramatically increase productivity of energy grid assets but improve energy grid resilience.

Cyber-physical security risk management for advanced reactors is an active area of research and regulatory concern. This report seeks to identify emerging cyber-physical risks for ARs and establish two risk categories. These risk categories are absorbed risk and licensee managed risk, clearly delineating the cyber-physical risk distribution between stakeholders. Specific cyber-physical risks identified in this report are not exhaustive but are highly relevant for ARs. These risks are based on the ARs surveyed in the Advance Reactor Operational Technology Architecture Categorization (M2CT-21SN1104024) report, submitted in conjunction with this report.

Based on the distribution of cyber-physical risks and the lack of a quantitative cyber-physical risk assessment methodology significant residual cyber risk may remain within an ARs OT architecture. Future research should investigate quantitative methods for cyber-physical risk assessment and the establishment of regulatory cyber-physical risk acceptance criteria. The unification of safety and security analysis of accident scenarios that explore a spectrum of hazards, especially degradation scenarios for passive safety systems, should be prioritized. Such unification will require a methodology gap analysis and a flexible modeling and simulation stack that incorporates both safety and security considerations.

2. CONCEPTS AND DEFINITIONS

The following sub-sections introduce fundamental concepts and definitions used throughout this report.

2.1. Cyber-Physical Risks

Adversaries can use any combination of cyber-physical tactics, techniques, and procedures (TTPs) in an attack. Cyber TTPs may include initial access, privilege escalation, and command and control, while physical TTPs may include automatic assault rifles, explosives, and insider knowledge. Thus, it is necessary to identify risks as cyber-physical rather than just cyber or physical. Cyber-physical attacks can also be referred to as hybrid attack scenarios, leading to a spectrum of possible attack scenarios, Figure 2-1. Although this report focuses primarily on cyber based risks it is important to recognize the spectrum of hazards presented by all possible attack scenarios. Furthermore, cyber-attacks on the physical protection system (PPS) are highly relevant and is often overlooked due to the organization division between physical and cybersecurity. As well as the analysis methodologies used by physical and cybersecurity groups. Fundamentally cyber-physical risks are the probability of cyber-physical initiated events multiplied by the consequences of that event.

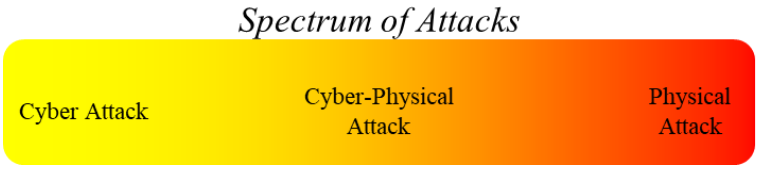

Figure 2-1. Spectrum of possible attacks

Division of physical and cybersecurity groups creates a division of responsibility and a high probability of vulnerabilities along the fault lines. This was historically demonstrated during the security breach at the National Nuclear Security Administration's Y-12 National Security Complex [4]. A key finding from the Y-12 incident indicated that a "bifurcated line of contractor accountability and responsibility" led to conflicting priorities regarding the site's PPS. Although the Y-12 incident did not include a cyber element it is logical to extrapolate potential vulnerabilities originating from the division between physical and cybersecurity groups at a plant and research groups. Unlike generation II reactors advanced reactors (ARs) will most likely not be able to support staffing of an on-site response force and separate security operation centers for both physical and cybersecurity. Suggesting that the traditional domain specific approach for nuclear security will not be sufficient for ARs. Further supporting the assertion of a spectrum of attacks.

2.2. Risk Definitions

Several definitions of risk exist, definitions are evaluated from the National Institute of Standards and Technology (NIST) Special Publication (SP) 800-37 Risk Management Framework for Information Systems and Organizations and International Organization for Standardization (ISO) 27000:2018 Information Technology – Security Techniques – Information Security Management Systems [5, 6]. The definitions established by both references overlap considerably however, for clarity and a more complete set the ISO/IEC 27000:2018 (originally from ISO Guide 73:2009) definitions are used in this report, see Table 1-1.

Table 2-1. Definitions regarding risk

Term	Primary Reference	Definition
Risk	OMB A-130	A measure of the extent to which an entity is threatened by a potential circumstance or event, and typically is a function of: (i) the adverse impact, or magnitude of harm, that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence.
	ISO Guide 73:2009	Effect of uncertainty on objectives Note 1 to entry: An effect is a deviation from the expected — positive or negative. Note 2 to entry: Uncertainty is the state, even partial, of deficiency of information related to, understanding or knowledge of, an event, its consequence, or likelihood. Note 3 to entry: Risk is often characterized by reference to potential "events" (as defined in ISO Guide 73:2009, 3.5.1.3) and "consequences" (as defined in ISO Guide 73:2009, 3.6.1.3), or a combination of these. Note 4 to entry: Risk is often expressed in terms of a combination of the consequences of an event (including changes in circumstances) and the associated "likelihood" (as defined in ISO Guide 73:2009, 3.6.1.1) of occurrence. Note 5 to entry: In the context of information security management systems, information security risks can be expressed as effect of uncertainty on information security objectives. Note 6 to entry: Information security risk is associated with the potential that threats will exploit vulnerabilities of an information asset or group of information assets and thereby cause harm to an organization.
Risk Acceptance	ISO Guide 73:2009	Informed decision to take a particular risk Note 1 to entry: Risk acceptance can occur without risk treatment or during the process of risk treatment. Note 2 to entry: Accepted risks are subject to monitoring and review.
Risk Analysis	ISO Guide 73:2009	Process to comprehend the nature of risk and to determine the level of risk Note 1 to entry: Risk analysis provides the basis for risk evaluation and decisions about risk treatment. Note 2 to entry: Risk analysis includes risk estimation.
Risk Assessment	NIST SP 800-30	The process of identifying risks to organizational operations (including mission, functions, image, reputation), organizational Assets, individuals, other organizations, and the Nation, resulting from the operation of a system.
	ISO Guide 73:2009	Overall process of risk identification, risk analysis, and risk evaluation
Risk Criteria	ISO Guide 73:2009	Terms of reference against which the significance of risk is evaluated Note 1 to entry: Risk criteria are based on organizational objectives, and external context and internal context. Note 2 to entry: Risk criteria can be derived from standards, laws, policies and other requirements.

Term	Primary Reference	Definition
Risk Evaluation	ISO Guide 73:2009	Process of comparing the results of risk analysis with risk criteria to determine whether the risk and/or its magnitude is acceptable or tolerable Note 1 to entry: Risk evaluation assists in the decision about risk treatment.
Risk Identification	ISO Guide 73:2009	Process of finding, recognizing and describing risks Note 1 to entry: Risk identification involves the identification of risk sources, events, their causes and their potential consequences. Note 2 to entry: Risk identification can involve historical data, theoretical analysis, informed and expert opinions, and stakeholders' needs.
Risk Management	OMB A-130	The program and supporting processes to manage risk to agency operations (including mission, functions, image, reputation), agency assets, individuals, other organizations, and the Nation, and includes: establishing the context for risk-related activities; assessing risk; responding to risk once determined; and monitoring risk over time.
	ISO Guide 73:2009	Coordinated activities to direct and control an organization with regard to risk
Risk Management Process	ISO Guide 73:2009	Systematic application of management policies, procedures and practices to the activities of communicating, consulting, establishing the context and identifying, analysing, evaluating, treating, monitoring and reviewing risk Note 1 to entry: ISO/IEC 27005 uses the term "process" to describe risk management overall. The elements within the risk management process are referred to as "activities".
Risk Owner	ISO Guide 73:2009	Person or entity with the accountability and authority to manage a risk
Risk Treatment	ISO Guide 73:2009	Process to modify risk Note 1 to entry: Risk treatment can involve: • avoiding the risk by deciding not to start or continue with the activity that gives rise to the risk; • taking or increasing risk in order to pursue an opportunity; • removing the risk source; • changing the likelihood; • changing the consequences; • sharing the risk with another party or parties (including contracts and risk financing); • retaining the risk by informed choice. Note 2 to entry: Risk treatments that deal with negative consequences are sometimes referred to as "risk mitigation", "risk elimination", "risk prevention" and "risk reduction". Note 3 to entry: Risk treatment can create new risks or modify existing risks.

2.3. Operational Technology Architecture

The term operational technology (OT) Architecture is defined in the Advance Reactor Operational Technology Architecture Categorization report but is included in this report for completeness [7].

Architecture is defined as "the complex or carefully designed structure of something,"[1] and is used heavily in the design and construction of buildings. The term architecture, however, is gaining adoption in cloud computing infrastructure, software, and network design. A security architect is now a key position within an organization's cybersecurity team. The hardware and software used to monitor or control NPP system functions, is the definition of OT. For the purposes of this report the term OT architecture is defined as outlined in this section.

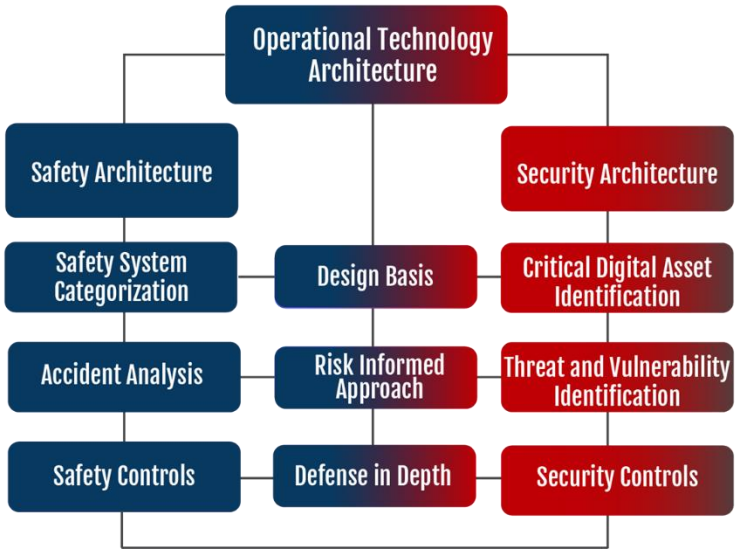

Figure 2-2. Visual representation of OT architecture definition

The Purdue Architecture is the first, best known, application of an architecture-based approach for OT infrastructure, as seen in Figure 2-3. This architecture provided designers with a framework to design an OT network by dividing system functions into five distinct levels. This concept was expanded to a security architecture that assigns unique security and assurance requirements to each level. The use of firewalls or data diodes between levels or sensitive parts of the network were used to control access and the flow of information. However, as the complexity of systems scale a unified OT architecture is needed to properly coordinated safety and security requirements. It is important to note that the digital systems used within a PPS are included within the definition of OT architecture.

[1] https://www.google.com/search?q=architecture+definition&rlz=1C5GCEM_enUS946US953&oq=architecture+definition&aqs=chrome.0.0i433i512j0i512l4j0i10i512j0i512l4.5619j0j15&sourceid=chrome&ie=UTF-8 .

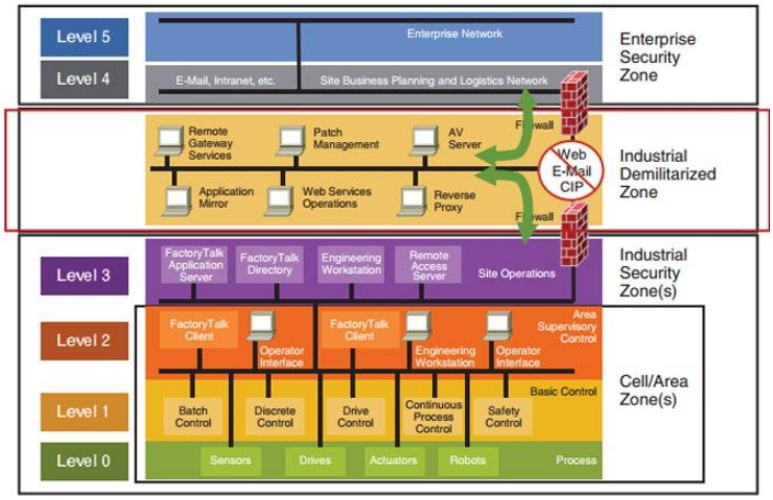

Figure 2-3. Five levels of the Purdue Architecture [8]

3. CURRENT APPROACHES IN RISK MANAGEMENT

This section will discuss U.S. Nuclear Regulatory Commission (NRC) regulations and domestic and international standards relevant to cyber physical risk.

3.1. U.S. NRC Regulatory Approach to Cyber-Physical Risk

10 Code of Federal Regulation (CFR) Part 73 (NRC 10 CFR 73) Physical Protection of Plants and Materials is the primary regulation for NPP security. Cyber and physical security requirements are primarily derived from NRC 10 CFR 73.54 and 73.55 respectively [9].

> NRC 10 CFR 73.55 Requirements for physical protection of licensed activities in nuclear power reactors against radiological sabotage states "(a) Introduction. (1) By March 31, 2010, each nuclear power reactor licensee, licensed under 10 CFR part 50, shall implement the requirements of this section through its Commission-approved Physical Security Plan, Training and Qualification Plan, Safeguards Contingency Plan, and Cyber Security Plan referred to collectively hereafter as "security plans.' ' Current applicants for an operating license under 10 CFR part 50, or combined license under 10 CFR part 52 who have submitted their applications to the Commission prior to the effective date of this rule must amend their applications to include security plans consistent with this section."
>
> NRC 10 CFR 73.54 Protection of digital computer and communication systems and networks states "(a) Each licensee subject to the requirements of this section shall provide high assurance that digital computer and communication systems and networks are adequately protected against cyber-attacks, up to and including the design basis threat as described in § 73.1."

NRC 10 CFR 73.55 states that licensees are responsible for creating and implementing a Physical Security Plan, Training and Qualification Plan, Safeguards Contingency Plan, and the Cyber Security Plan (CSP), collectively referred to as security plans. The intended purpose of these plans is best stated by NRC 10 CFR 73.55 "(b) General performance objective and requirements. (1) The licensee shall establish and maintain a physical protection program, to include a security organization, which will have as its objective to provide high assurance that activities involving special nuclear material are not inimical to the common defense and security and do not constitute an unreasonable risk to the public health and safety." The definition of an "unreasonable risk" is not defined quantitatively in the regulation and is determined by U.S. NRC during a licensing review.

Since this report is primarily focused on cyber risks for ARs the CSP discussed further to show U.S. NRCs current approach to cyber risk. Guidance for preparing a CSP is provided by U.S. NRC Regulation Guide (RG) 5.71. "Appendix A" of NRC RG 5.71 provides a CSP template, which outlines the procedure for identifying components as critical digital assets (CDAs) based on their role within Critical Systems (CS), and mitigating vulnerabilities by application of security controls. The security controls listed in NRC RG 5.71 are based on NIST SP 800-82 Guide to Industrial Control System (ICS) Security and NIST SP 800-53 Security and Privacy Controls for Information Systems and Organizations [10, 11]. Further guidance for application and assessment of security controls is provided by Nuclear Energy Institute (NEI) 13-10 Cyber Security Control Assessments. The relationship between the regulatory requirements and guidance as licensees work toward preparing their CSP is shown in Figure 3-1.

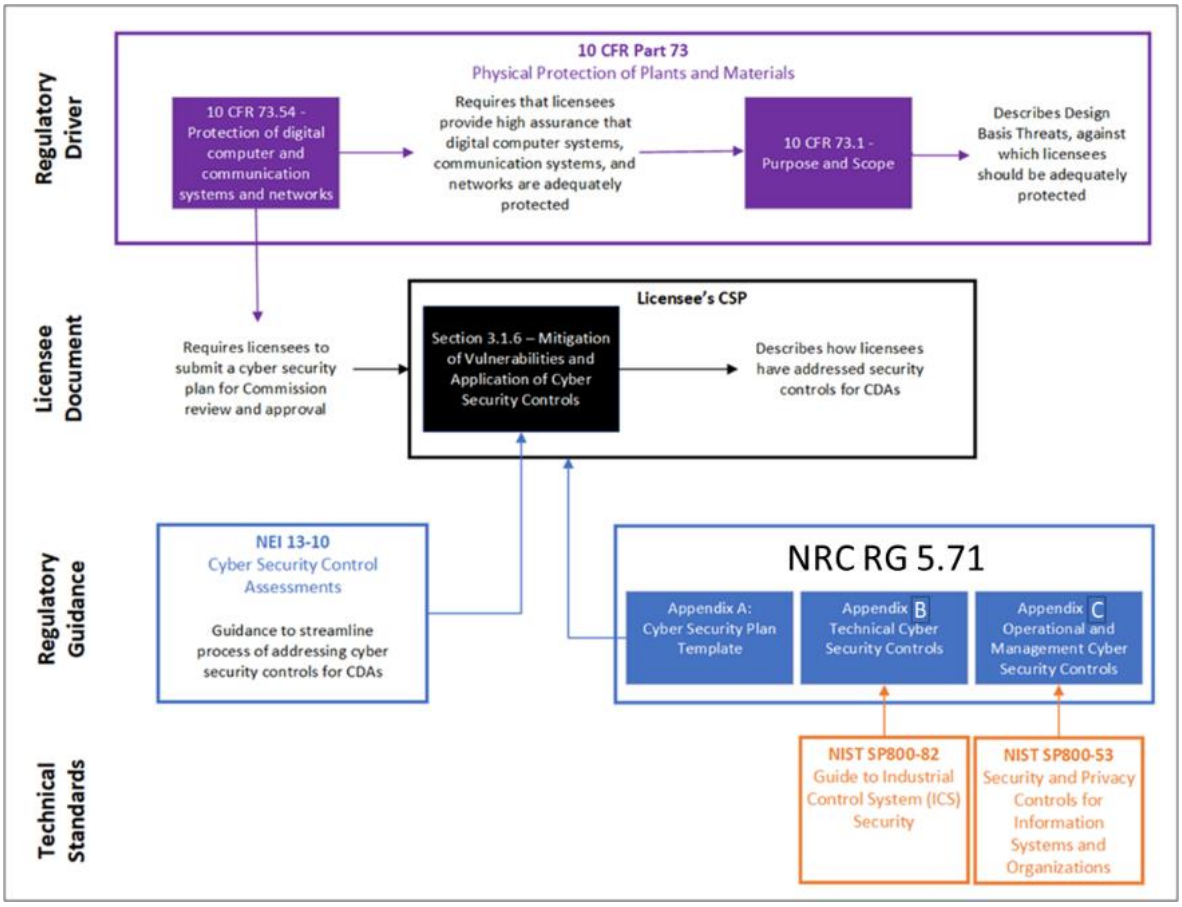

Figure 3-1: U.S. regulation and guidance for U.S. NPP cybersecurity

U.S. guidance for cyber-physical risk relies on risk modification, by identifying CDAs and applying a set of controls applied. NEI 13-10 refined this approach by providing criteria against which CDAs may be screened based on their impact to safety or security systems. An initial attempt at a consequence-based approach, which increased the efficiency of security controls. There is no regulatory guidance however, regarding the sufficiency of security control measures applied to a CDA. Without adequately understanding the sufficiency of a security control or the probability that a vulnerability being compromised, quantification of cyber-physical risk is impossible. Therefore, reduction of cyber risk is purely reliant on risk modification of CDAs and the overall cyber risk level remains unknown. This contrasts with NRC's approach to safety risks, which are quantified and modified using iterative design modifications until safety risk criteria are met.

3.2. NIST SP 800-37

To address the growing concern of cyber-attacks on critical infrastructure a Risk Management Framework (RMF) is established in NIST SP 800-37 Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy. The RMF applies across both Internet Technology (IT) and OT systems and establishes a formal methodology to secure critical digital systems. The framework documents an organization wide implementation

strategy in a tiered system, Figure 3-2. Level 1 addresses risk from an organizational perspective, frames the risk and provides context for the risk management activities for the organization. This directs overarching business and mission decisions with respect to risk, and effects the information system architecture development in the other levels.

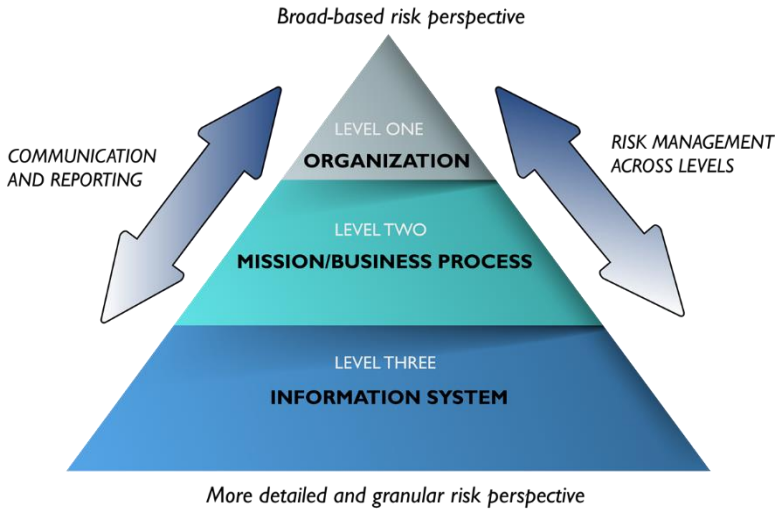

Figure 3-2. Organizational-wide risk management approach (NIST SP 800-37)

Level 2 addresses the risk from a mission or business perspective, informed by the organizational decisions made at level 1. This level considers the mission/business process needs, prioritizes those processes and their security requirements, and establishes a system architecture that meets these needs. Informed by the activities in level 1 and 2, the 3rd level seeks to address risk from an information system perspective. The allocation and implementation of security controls is done in Level 3. The management, assessment, and continued monitoring of these controls are also conducted in Level 3 of the RMF.

Clear reporting and communication paths between organizational levels enables effective allocation of resources and ensures that identified risks are addressed at each level. The granularity of risk controls is the most detailed at Level 3 of the RMF. Levels closer to the actual technical implementation can respond to minor requirement and implementation changes, without effecting the organization. The RMF is divided into seven major steps that form a continuous cycle:

I. **Prepare:** Establish the context and priorities for managing the security and privacy risks from a system-level perspective.

II. **Categorize:** The system and the information in contact with that system are categorized based on the impact of loss.

III. **Select:** An initial set of risk controls are selected and configured as necessary to reduce risk to an acceptable level.

IV. **Implement:** The selected controls are implemented within the system and its operational environment.

V. **Assess:** Determine if the controls are implemented correctly, operate as expected, and satisfy the security requirements.

VI. **Authorize:** Make a determination whether the system or controls fall within an acceptable risk level based on the security requirements.

VII. **Monitor:** Continuously assess the effectiveness of the risk control and security posture of the system.

The RMF outlines a broad and highly flexible system for organizations to frame, assess, respond, and monitor risk. It seeks to reduce the implementation cost, improve efficiency, and produce a better cybersecurity implementation across information systems. The RMF attempts to address the cybersecurity risks present in the full system life cycle from implementation to disposal. Common control methods are leveraged as much as possible to reduce development cost and time. The process is iterative and continuous to capture new risks and the changing structure of information systems. Providing the technical justification for an iterative OT architectures design approach that focuses on optimizing safety and security requirements. Applied thoroughly it can help organizations path find in the ambiguous territory of cyber risk. Ultimately this methods effectiveness relies on the rigor to which the organization applies the RMF and the technical awareness and skill of that organizations staff at each level.

3.3. ISO/IEC 27005

The international community has also sought to address cybersecurity and risk management in the released ISO/IEC 27005: Information security risk management standard. The standard describes an Information Security Management System (ISMS) that is applicable to the organization, down to any discrete parts of the organization such as departments, locations, or services. This standard modifies the risk management process specified in ISO 31000 to be more applicable to cybersecurity, as well as becoming iterative and continuous. Figure 3-3 illustrates the risk management process, starting with contextualizing the risk.

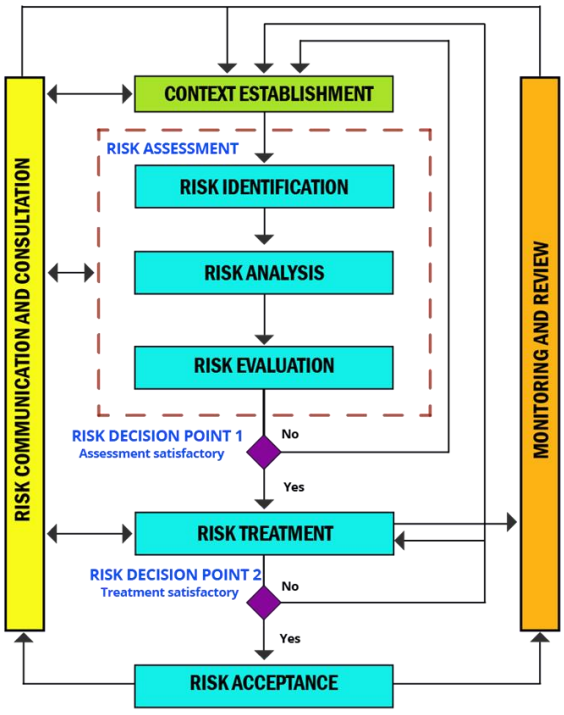

Figure 3-3. Illustration of an information security risk management process (ISO/IEC 27005)

Context establishes a scope and boundary to the risk management activity. The purpose of the risk management, supporting an ISMS, legal compliance, etc., inform the context as well. Risks identified at any other stage of the process are communicated to the context step as the starting point of the iterative approach. Context also defines the risk management approach, how risk will be evaluated, the impact of a security breach, and the acceptable level of risk.

From the gathered information, a risk assessment can be completed for the system under analysis. In the first step of identification, loss scenarios, assets at risk, and threats and vulnerabilities of the system are identified. Consequences of each loss scenario and the failure modes of the assets are also identified. Identification information informs risk analysis, which determines the likelihood of each scenario. The likelihood of a scenario is then scaled by consequence to determine the risk. The level of risk is then compared to the risk evaluation criteria and acceptable level of risk established in the context phase. Risk evaluation will inform whether the risk assessment is satisfactory or requires further iteration.

After the risk assessment is deemed to have satisfactorily captured all the risks and categorized them correctly, those risks can be addressed. Risk treatment takes the identified risks in previous steps and applies methods to attempt to reduce the risk to an acceptable limit. These methods are risk modification, risk retention, risk avoidance, and risk sharing. Risk modification seeks to manage the risk by introducing or altering controls so that the residual risk can be reassessed as acceptable. Risk retention accepts the risk if it is below the acceptable level of risk. Risk avoidance simply avoids the problem, removing the high-risk activity or device. Risk sharing partitions the risk with another entity that may be more equipped to address that risk, such as hiring a security partner. Dividing risk responsibility can bring new risks or change the nature of other risks. Once the risk treatment options are determined, the final step is risk acceptance.

After considering the risks and their treatment residual risk will remain, this must be evaluated to determine if the residual risk is acceptable. Risk acceptance enables stakeholders to authorize the implementation of additional risk treatments to further reduce or quantify residual risk. After all risk treatments are implemented, the implementation of risk control and external factors will be monitored to ensure the level of risk remains at an acceptable level over time. If new threats, vulnerabilities, or initial risk values change a new risk assessment with new context is conducted. During every stage, information security risk communication and consultation interact with the activity. Communication and consultation help ensure a more rounded, accurate, and effective risk management system that is responsive to wide range of expertise and developing threat concerns.

Risk management is a vital strategy to reduce the risk of informational and operational technologies. Implemented with rigor and skilled staff, both ISO 27005 and NIST SP 800-37 will reduce the risk in any organizations IT and OT systems. The problem isn't the processes themselves, it's a lack of tools and knowledge available to organizations to sufficiently identify risk and protect themselves from skilled and persistent adversaries. The frameworks can only ever be as good as the tools and knowledge used to implement them. Currently, air gapped networks, application of security controls, and onsite security are the primary cyber defense for NPP control systems. As proven by attacks like STUXNET significant residual risk may remain in I&C systems despite application of strict security controls [12]. Cyber mitigations degrade with time, and adversaries continue to erode network isolation strategies. New tools and knowledge must be developed to successfully apply risk management frameworks to address modern adversaries using both qualitative and quantitative assessments.

4. LIMITATIONS OF CURRENT PRACTICE

Understanding cyber risk given the high uncertainties associated with ARs is a major challenge. While regulation drives towards risk-informed approaches to security, there are many challenges that make security risk difficult to quantify. Additionally, the existing framework for LWRs requires the plant licensee to assess and modify risk to their systems. With AR technology, however, the role of risk responsibility between the owner and system manufacturer is not well-established. System assessment and modification may require expertise that the licensee does not have, meaning the risk responsibility may need to be shared with vendors. As the industry moves toward guidance that provides licensees the ability to identify risk within digital systems based on existing safety evaluations (i.e. consequence based), the path forward for risk analysis and evaluation remains less clear.

U.S. NRC currently does not cite any standard for cybersecurity risk management in NRC 10 CFR 73. Cybersecurity guidance for U.S. NPPs rely on NIST SP 800-82, controls placed at the component and network level, and NIST SP 800-53, organizational controls. As covered in Section 3.3, ISO/IEC 27005 guidance is split up to address risk assessment, risk treatment, and risk acceptance, monitoring, and review. To understand where gaps may exist in U.S. guidance with respect to cybersecurity risk management at NPPs, the analysis in Table 4-1 has identified ISO 27005 clauses that do not have an associated clause under U.S. NRC cybersecurity regulation and guidance.

Table 4-1: ISO 27005 clauses that <u>do not</u> have associated guidance within U.S. NRC cybersecurity regulation and guidance

ISO 27005 Clause	Category	Description
8.3.1	Risk Assessment Methodologies	Defines qualitative and quantitative risk analysis and weighs the advantages of each as well as when each method should be implemented. Discusses the roles that consequence and likelihood play in a risk assessment.
8.3.3	Assessment of Incident Likelihood	Assuming a plant has a comprehensive list of assets, vulnerabilities, and identified threats, assess the likelihood of given scenarios either by qualitative or quantitative methods. Likelihood analyses should leverage experience, available statistics, threat sources, and the effectiveness of existing controls.
9.4	Risk Avoidance	Defining the threshold at which the risk is considered too high for an activity to take place.
9.5	Risk Sharing	Sharing risk responsibility with an external party which can be done by insurance (consequence coverage) or a partnership with a party who would monitor information systems and take action to stop an attack.

ISO 27005 Clause	Category	Description
12.2	Risk Management Monitoring, Review and Improvement	The information security risk management program should be reviewed to identify if there are areas in which risk is overlooked and if appropriate decisions are made in response to the risk. Risk assessment inputs and criteria should also be reviewed to ensure they are valid under present circumstances.

Following the process model for risk management in ISO 27005 from Figure 3-3, plant licensees should have processes in place to identify, analyze, and treat security risk. U.S. guidance for securing reactor systems is based on identifying individual digital components and administering all applicable security controls. By not taking system risk into consideration and instead focusing on asset-based mitigation, this methodology places a heavy burden on the licensee as they take on the responsibility of deeming when control measures are sufficient across their site. As the industry shifts towards establishing risk informed methodologies, all stakeholders involved in advanced reactor operations will be better prepared to mitigate cyber threats. A U.S. standard like ISO 27005 is needed, NIST SP 800-37 provides a foundation but needs refinement. Standards that can apply risk assessment specifically OT networks and the unique considerations as outlined in NIST SP 800-82 for industrial control systems would then be the logical next step in the standard development.

4.1. Risk Identification

U.S. NPPs rely on Probabilistic Risk Assessment (PRA) as the vehicle for assessing safety risk for scenarios leading to unacceptable consequences. PRA models consist of event trees (which evaluate scenarios from an initiating event that can lead to a particular consequence) and fault trees (which evaluate a system's probability of success or failure by reducing the system to basic events of component failures) to quantify the overall system risk. Plant systems rely on redundancy and diversity to ensure that single failures do not compromise the function of the system. PRA quantifies the overall risk of component or system failure leading to some consequence, which can be successfully mitigated by employing reductant systems. For example, a random pump failure may not be detrimental to the function of an emergency core cooling system. Assuming there is a backup pump or a redundant train to which the system can align. Current security risk identification methodologies take advantage of existing plant PRA models to identify digital controllers and components of interest.

These methods are the Digital Engineering Guide (DEG), Technical Assessment Methodology (TAM), Digital Reliability Analysis Methodology (DRAM), Electromagnetic Compatibility (EMC) Assessment Methodology (EMCAM), Human Factors Analysis Methodology (HFAM) and HAZard and Consequence Analysis for Digital Systems (HAZCADS), led by the Electric Power Research Institute (EPRI). Specifically, under DEG guidance, HAZCADS provides a method for identifying hazards and associated unsafe component actions, and TAM maps hazards to cybersecurity controls. Given a plant's PRA model, HAZCADS can identify not only the controllers whose failure can lead to an unsafe system state, but also the variety of ways which that controller may fail or be manipulated. While powerful, this approach has its limitations to the plants that would be expected to perform the analysis. HAZCADS relies on Systems Theoretic Process Analysis (STPA), a hazard analysis method that requires deep understanding of the system to consider all possible unsafe

control actions. Plant owners cannot realistically expect that their systems engineers to become experts in STPA and develop and maintain HAZCADS models for their systems. Nor can they support retaining STPA experts on staff who have the level of understanding for all systems necessary to perform a successful HAZCADS analysis.

It may not be necessary for licensees to perform comprehensive HAZCADS analyses on all systems, but the following features, at a minimum, are critical to risk identification:

- Produce a list of systems and subsystems necessary to maintain a plant function (context establishment phase)
- Identify digital components or groups of components which, if compromised, can lead to an unsafe system state
- Determine the consequence of failure for identified digital components

Identifying assets using a top-down approach allows plant owners to consider all CDAs that are required to maintain a plant function.

Cyber risk identification should begin in the OT architecture design phase. Communication between stakeholders is imperative to ensure that risk factors do not go unidentified. Regulators and plant owners need to clearly communicate their unacceptable consequences to the system manufacture and vendor to ensure that all risk sources that can compromise critical plant functions are identified. For advance reactors that seek a path towards licensing not utilizing a PRA process would require the current HAZCADS methodology to be modified. Indicating that more research is required in the domain of risk identification to address the gaps presented by STPA and systems without a PRA analysis.

4.2. Risk Analysis and Evaluation

A one-to-one mapping between the probabilistic analysis in PRA and security risk does not exist due to the inability to quantify the likelihood of a cyber-physical attack. Fault trees for safety analysis rely on historical data and testing to assign failure probabilities to components. This capability is currently not feasible in the security space due to lack of long-term available attack data and the rapidly changing landscape of cyberattacks. The inability for analysts to apply probabilistic assessments to security means that assessments are informed by consequence, not pure risk (i.e. NEI 13-10). Current security analysis methods are qualitative, meaning all attacks on all components are equally likely; the probability of an attack is assumed to be 100%.

There are many benefits for industry stakeholders to adopt a quantitative assessment approach to security risk. If security analysts are able to quantify risk to a system, they will be better positioned to apply control measures that are not only sufficient to mitigate threats to an asset, but are also efficient, ensuring that operators are not endlessly applying controls to all applicable CDAs. Regulators who carry the responsibility of providing guidance to licensees will be able to clearly state the requirements for a CDA that can be considered secure. Under guidance driven by quantitative analysis methods, licensees may be able to take credit for existing programs such as physical security protections around security zones and CDAs. As well as advanced safety features.

To move towards quantitative security assessments, the industry needs to assist plants with assigning likelihood of compromise to the components considered in the risk identification step. This requires system owners and security analysts to have a clear understanding of their control systems, the

ability to mitigate known vulnerabilities, the ability to assess impact of vulnerabilities, and the effect of controls on the overall system.

Additionally, stakeholders should consider potential threat vectors. Threat modeling tools, such as attack trees, can be used to provide analysts with an understanding of the steps required to perform an attack and the feasibility of attack scenarios under consideration. Since penetration production systems is not possible, cyber-physical testbeds or digital twins may be needed to conduct performance testing.

4.3. Risk Treatment and Acceptance

Risk treatment options discussed in ISO 27005 are shown in Figure 4-1. Risk modification and retention are the current U.S. standard for NPPs. Risk avoidance, in the context of reactor safety and security, refers to the suspension of reactor operations and is not normally considered as an option. Risk sharing refers to the transfer of some amount of risk to an external party and will be a necessity for advanced reactor risk management practices. If a licensee is unable to assess, monitor, or modify the risk of a given system due to unfamiliarity with the system, the risk will need to be managed and modified through the vendor or manufacturer of the system.

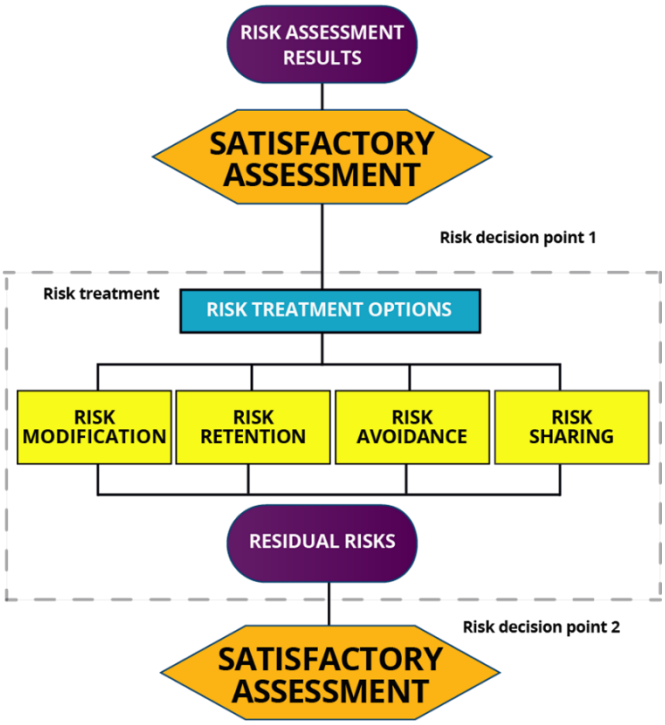

Figure 4-1: Risk treatment and decisions under ISO 27005 (ISO/IEC 27005)

Risk responsibility remains a challenge for AR ownership. Under U.S. LWR guidance, the plant owner is responsible for identifying and modifying risk within their reactor systems. While this methodology is extensive given the high number of CDAs identified, it is within the operator's ability to maintain and apply controls for the current nuclear fleet. AR owners and operators, however, will not have the capability to assess and manage cybersecurity risks in systems that have

been designed and manufactured by an external party. The approach of CDA identification and assessment may not be feasible as the plant owners will have a limited knowledge of the system and licensee's limited capability to modify or access the risk of a system or component. System developers would be required to take on the risk responsibility. Therefore, in this report cyber-physical risk has been clearly delineated between absorbed risk and licensee managed risk. This approach demonstrates an acceptable way to show the distribution of cyber-physical risks for ARs.

5. ABSORBED RISK

In all commerce there is inherent risk. To the customer, that risk takes the form of the purchase item being defective, being poor quality, or a service not rendered. The vendor's reputation allows them to convince customers to purchase their product, their risk is ultimately reputational harm. It's in the best interest for vendors to improve their reputation and gain more trust with the consumer. So, they may offer warranties or contract agreements to replace defective items or guarantee customer satisfaction. This effectively transfers some of the risk the customer has in the purchase onto the vendor. This works in low consequence transactions, the most either party is putting at risk is purely financial.

This transfer of risk is reversed when it concerns safety, and high consequence systems. The risk is mostly absorbed by the customer who is placing safety and security trust in a vendor's product. This is especially true in NPPs; they acquire the risks associated with the vendors and suppliers of equipment in the facility. Cyber risk for delivered components is normally cover using a contract that stipulates cybersecurity procurement provisions. It has been shown that a significant amount of residual risk is still present despite these cybersecurity procurement provisions [13]. This implied trust of vendors, suppliers, and equipment can expose a significant attack surface for NPPs. This section will focus on the cyber-physical risks absorbed by licensees due to interaction with external vendors and suppliers.

5.1. Supply Chain

The supply chain is the procurement and life cycle of hardware, software, and services that are utilized in a facility. Risks are present in every point of the supply chain, from manufacturing to maintenance services. To capture and mitigate all risk factors is difficult due to the evolving nature of technology and services that support NPPs. Therefore, each general sector of supply chain will be covered, beginning with the component level, and building in complexity.

It is important to understand the entirety of the supply chain to effectively allocate the resources necessary to mitigate risks within it. Risks within the supply chain will be dependent on the context of the application. The supply chain of safety equipment and software will require stringent controls, while some ancillary systems may not demand such rigor, linking back to a risk-informed approach using a cyber-physical risk assessment. The depth of the risks within the supply chain are complex and interrelated, risk in one area will be used to exploit the risk in another. Threat actors are varied and numerous, their operational resources, and technical skills should not be underestimated [14, 15].

5.1.1. Hardware Supply Chain

Design, manufacture, and delivery of components are encompassed in the hardware supply chain. In each step of the process a plethora of risks exist, some of which have solutions, but many that will require extensive research to remediate. The exploit paths into hardware requires sophisticated adversaries, but the detection difficultly of these exploits provides significant incentive. To fully evaluate the risk presented in the hardware supply chain we will need to dissect each component from the chip foundry to the delivery of a completed control system.

Consider that a facility purchases a piece of equipment from a domestic manufacturer that has outsourced the manufacturing of some critical modules like the motherboard. The motherboards are made by a separate manufacturer in an adversarial nation. The motherboard designs are modified to include an extra component that can exfiltrate information and exert command and control over the

motherboard and the equipment it is installed in. This is the most extreme form of compromise of a hardware supply chain but it this exact scenario has already played out with SuperMicro [16].

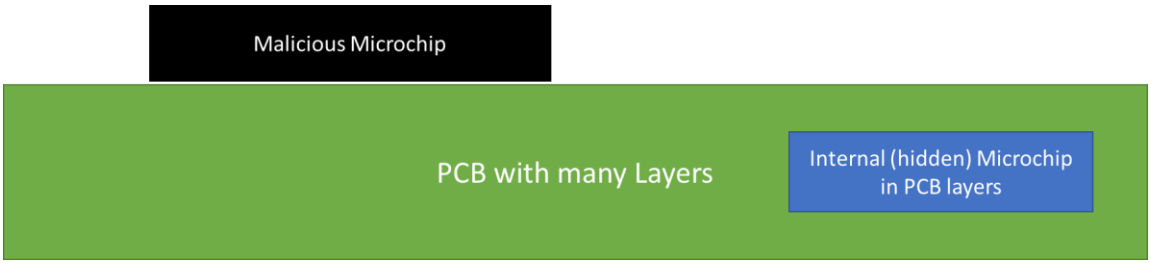

Figure 5-1. Layered Printed Circuit Board (PCB) with added malicious microchips

These types of hardware supply chain attacks can be impossible to detect, as the level of obfuscation can extend from embedding devices in between layers of a printed circuit board (PCB) [Figure 5-1] [17] to modifying the Integrated Circuits (ICs) in the chips [18]. In highly complex circuit board layouts, even an additional chip may be difficult to identify with human inspection. Embedded devices within the PCB would require x-ray inspection to visually identify. Altered PCBs could be identified through signal signature-based detection [19]. The traces and components on the PCB have characteristic resistance, capacitance, and induction that can form a signal response signature of the design [20]. This has the advantage of providing greater information of the device's health and security in-situ. Though this will require a reference design to compare to and does not protect against the possibility of a corrupted reference.

Hardware trojans are not limited to the board level of devices; the concern extends to the silicon of each microchip within devices. It is common now to stack silicon wafers inside encapsulated chips to improve component density and compact designs [17]. Placing a malicious trojan silicon wafer in the package-in-package (PIP), or package-on-package (POP) stack within a chip is entirely possible[21]. A trojan chip hidden within an encapsulated package may even escape x-ray inspection. The great variety of exploit paths and depth to which exploits can operate and be obfuscated make it impossible at present to verify that within each microchip there is no malicious logic implanted within [22].

Currently there is not a comprehensive way to ensure the security of hardware, from chip, PCB, to assembled device [23]. Some method to validate the security of components and devices must be researched and a standard must be established to fully secure hardware from this level of supply chain attack. The threat of hardware trojan is an active area of research, but mitigations lag the development of new technology. Detecting trojans on 2D IC design have many potential detection mechanisms, but trojans in the 3D IC designs such as POP and PIP negate these [24, 25]. Though this is a highly advanced attack path and requires significant compromise of manufacturing, globalization of the hardware supply chain provides ample opportunity for threat actors.

The best solutions to date are to ensure the security of IC's is to manufacture them in highly secure and monitored foundries and provide tracking and validation. There are a number of methods to attempt to secure IC production. Blockchain-based security attempts to reduce the possibility of counterfeit chips by providing a non-replicable signature within the chip [26]. Machine learning is being applied to challenge the circuity of the chips to identify hardware trojans [27]. These methods are developing and are not at a stage to deploy commercially. Ultimately these methods will still rely

heavily on the security of the software they operate on and the software the ICs and PCBs are designed on.

5.1.2. Software Supply Chain

Computer system architecture can be abstracted into notional layers, from the hardware to the user interface. In a typical PC the applications we interact with are a layer that is supported by and interfaces with the operating system. The operating system is interfaced with the hardware through the firmware. Programmable Logic Controllers (PLCs) software is generated from a human-machine interface, which defines its operation [28]. The PLC's firmware bridges the gap between the software and hardware layers and is often incorporated with the software layer. When the PLC is programmed its software and firmware data are generated by the engineering application. The firmware and software data are transmitted through the layers of the engineering workstation to the hardware that physically connects the workstation to the PLC. At each layer along the data path that the PLC program traverses there exists supply chain risk and vulnerability, Figure 5-2. This section will focus on the supply chain risks of the software layers above the hardware, as discussed in the previous section.

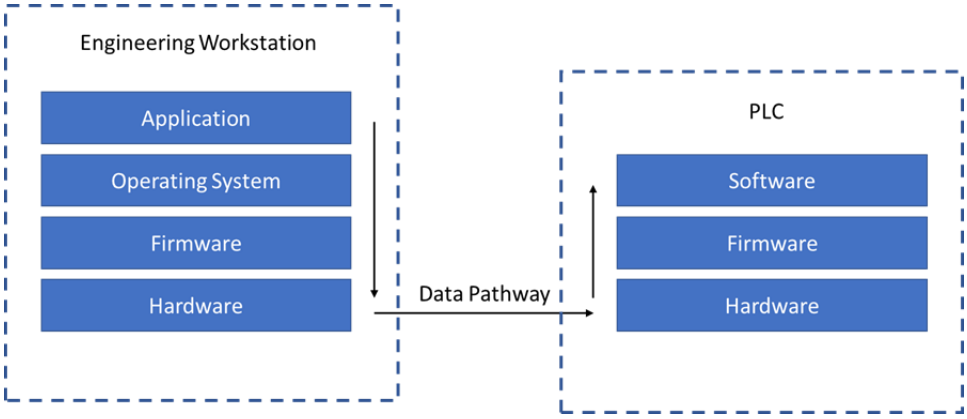

Figure 5-2. Data pathway for PLC software across system layers

Consider the scenario of a plant purchasing a new PLC for their facility that will be connected to the control system network. The PLC requires software to configure it which is supplied by the manufacture of the controller. During the development of the vendor supplied software, an open source library is included to perform a ubiquitous function. The opensource library contains malicious code that lies in wait until it detects deployment in a production environment. Once it has deployed in a production environment it moves laterally within a facilities network to identify and attack high consequence targets.

Software development is now a highly collaborative effort. It not only requires teams of developers, but often relies on opensource code libraries to reduce development time. A threat actor may add malware to these repositories [29]. These added exploits will often be obfuscated with encryption or different encoding to prevent detection. If the threat actor cannot become a contributor to the open repository, they may make their own repository that is one character different from the real one. Known as typo-squatting, the threat actors relies on a developer to make a typo and include it in their code base [30]. For qualified safety systems with highly controlled development environments

it is unlikely that an opensource piece of code would be used. The issue is: What code in the overall software ecosystem has contact with the control system network?

Ubiquitous code that enables hardware interfacing or simplifies complex tasks is incorporated into OT devices from 3rd party vendors. The code bases for these devices can be massive especially when considering the chipset firmware and discrete device drivers that must be included. Typically, 3rd party libraries are purchased out of necessity to reduce development burden and cost. Recently the NicheStack Transmission Control Protocol over Internet Protocol (TCP/IP) stack, a library which facilitates network communications and is commonly used in OT devices for critical infrastructure applications, was found to have 14 critical vulnerabilities [31]. These types of vulnerabilities can be exploited by other devices on the network. Code used in both the OT devices and anything that has access to control networks should to be vetted for latent vulnerabilities and hidden malware.

Evaluating each library would be a costly, intensive, and continuous process as updates for each library would require re-evaluation. The evaluation must be rapid; human evaluation of these libraries would not be able to keep pace with important updates. Machine Learning (ML) and Deep Learning detection methods may provide the tools to rapidly evaluate firmware [32]. While the research effort on ML vulnerability identification is quickly expanding, these methods are in the early stages of research [33]. The variety of code languages and chipset architectures used in an ICS network add to the difficultly of developing a comprehensive evaluation toolset. This technology is a double edge sword, the information ML vulnerability identification toolsets generate can also be used to exploit the systems under test. Machine Learning is not perfect, with sufficient knowledge adversaries can design their malware to evade detection [34]. Likewise, software must be placed under scrutiny to be evaluated, which may not happen when software comes from trusted sources.

For highly configurable controllers, the configuration utilities typically reside on engineering workstations. Part of keeping these workstations secure involves keeping their software up to date. Software both related and unrelated to the controller configuration utilities is updated regularly to patch vulnerabilities and fix bugs. If an attacker infiltrates the update servers of one of the programs used on these workstations a malicious update file can distributed via trusted channels. The malicious update looks legitimate and has a valid signature, without deep introspection it may be impossible to detect that this update contains a threat.

Numerous high consequence, high profile attacks have taken place in the software update sector of the software supply chain such as NotPetya and the SolarWinds compromise [35] [36]. These update server hijacks have already affected cybersecurity in nuclear power. In 2014 a computer connected to the business network in the Japanese Monju fast breeder reactor control room was compromised via an update server [37]. A media player, 'GOM Player', had its update server hijacked and employees following the general cyber-security guidance of updating software inadvertently installed the 'Gh0st RAT' trojan. This may represent the most concerning intrusion vector for malware into sensitive systems, as the software source is believed to be trustworthy. These types of attacks can be highly obfuscated which enables a long persistence, and thus a long data gathering period for adversaries to build a more dangerous attack. This attack method requires an advanced threat actor with significant resources and patience. Most attackers choose faster, easier methods that bypass a facilities security boundary such as contractors and service providers.

5.1.3. Service Supply Chain

Often facilities will contract out service providers for a variety of incidental operations. These contractors may bring in equipment that does not have a strict chain of custody or is simply exposed to connections with the open internet. Information shared with contractors is also not strictly controlled. Information of plant designs, system information, control systems, networks, if extracted from contractors is a viable reconnaissance vector for threat actors.

Assume, for example, a diagnostic tool that interfaces with PLCs is brought with contractors to setup a PLC on site. This tool has been connected off-site to a laptop connected to the internet. A virus that seeks computers with the drivers for this diagnostic tool injects itself on to the diagnostic tool when it is connected. This infected tool is then used on PLC's in the plant, as part of an attack campaign. This scenario assumes no malicious intent of the contractor (trusted insider), they were simply an easier target for the threat actor, as was the case in the Target breach in 2013 [38].

It is reasonable to conclude that the cyber-physical security of facilities is only as secure as the least secure contractor brought on site. With threat actors increased focus on the energy sector, the cyber-security of contractors must be scrutinized [39]. Threat actors have identified that contractors and suppliers are viable attack vectors into secure networks [40]. This is compounded by the potential for a malicious contractor (insider threat). A single trusted insider could be paid by a threat actor to bring a Micro SD card onsite with a USB adaptor. The USB adaptor presents no threat and can make it passed barriers and checks being that its innocuous. The Micro SD is small, even in a pocket it could be easily missed by security checkpoints. Together the USB adaptor and Micro SD can be used to deliver a cyber threat into the control system.

5.2. Autonomous Operation

Autonomous systems are a growing field of study and offer considerable benefits but have unique cyber-physical risks. These systems are based on ML and Artificial Intelligence (AI), which fundamentally are complex algorithm sets that are programmed by large training sets of data [41]. Autonomous control systems for NPPs are an active area of research that is rapidly evolving. Consequently, it is difficult to describe an exact system, but a notional model of the system functions can be analyzed to assess the core cyber-physical risks, Figure 5-3. An autonomous system can be described by the following decision process: detection, prediction, strategy selection, and strategy execution [42]. Detection pulls in data from sensors to inform the ML and update the prediction system with the current system condition. The prediction system provides probable system responses with control strategies and allows the ML to select the best strategy. The strategy is executed, and the decision process starts again.

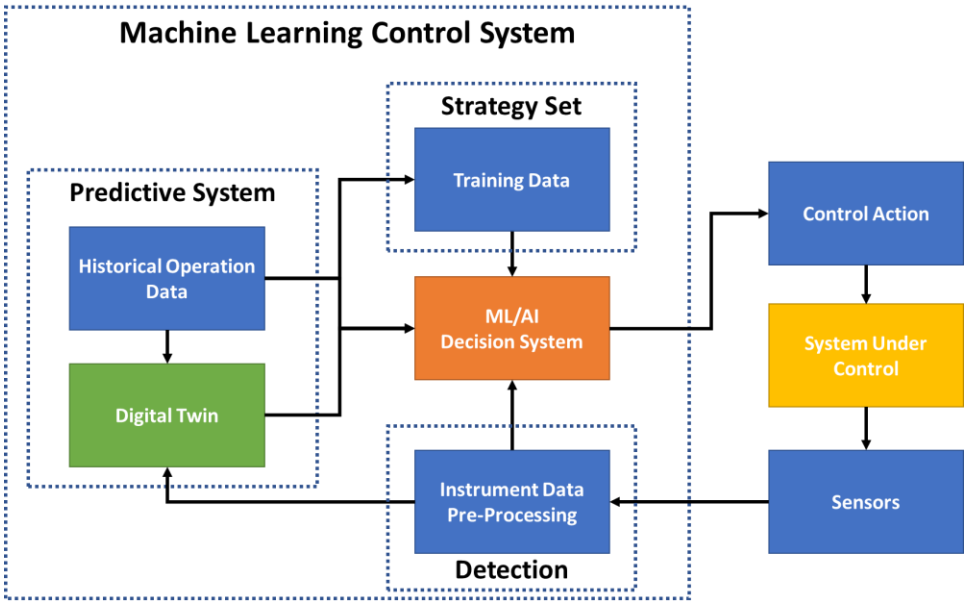

Figure 5-3. Notional machine learning system block diagram.

Training sets define the boundaries of acceptable system operation. Strategies are derived from these boundaries and the available control actions. To aid the system's predictive capabilities a reference model or Digital Twin (DT) of the plant physics is used to analyze probable future conditions and the possible consequence of control actions [43]. If the DT or training sets are corrupted the algorithm may drive the reactor into an unsafe condition [27]. These training sets and DTs will need to be highly protected and methods to validate their fidelity will need to be considered.

To continue to respond correctly to plant conditions, ML and AI need historical data. The DT will also be reliant on historical data to calibrate to the current condition of the fuel. The data can be in the form of operational data from another plant or the historical data from the plant under control. If an adversary were to corrupt or change the training data, the system would not respond correctly and possibly put the plant in an unsafe condition. It has been shown that malware could replace a significant number of nodes in a neural network while retaining the functionality of the program and being undetectable [44]. Corrupting training data would be difficult to detect since the data sets are large and the data syntax would not be broken. Thus, training data would require protection and validation methods to ensure that it is not corrupted or altered.

The cyber-physical risk presented by autonomous systems can be compounded by the consolidation of control systems. Ideally from a control perspective, a single autonomous controller should be responsible for the entire plant. In a generation II NPP, control is distributed across many PLCs. Which, while inefficient, increases system security. With consolidated control, a single compromise may be sufficient to cause harm. Autonomous architectures will need to include a defense-in-depth approach to ensure systems functions are sufficiently isolated and diverse such that the addition of autonomous control does not increase cyber-physical risk.

Autonomous control is a tantalizing concept for designers and researchers, it presents as an enabling technology for remote deployment of advanced micro reactors [45]. It has the potential to reduce the operation and maintenance staffing requirements of facilities especially sites with many SMRs. The transportation, defense, and aerospace industries have and continue to invest heavily in autonomous systems and use them extensively in safety critical systems. This has occasionally

resulted in disaster, such as in the case of the Boeing 737 MAX [46]. While the nuclear industry can learn from the lessons of other industries, they cannot endure the risks others have taken. Autonomous systems will certainly be an important piece in the advancement of nuclear power, the question is when. Autonomy requires complexity in domains outside of current nuclear operational experience and opens new attack vectors that have not been fully explored. It will require extensive development, research, and operational experience to fully understand the cyber risk presented by autonomous systems.

5.3. Advanced Reactor Operational Technology Architecture

Licensees of ARs will be system integrators and will have to work with numerous vendors to meet all regulatory and operational requirements. Specific architectural considerations for each AR separated by coolant type are outlined in the Advance Reactor Operational Technology Architecture Categorization report [7]. A key observation of the architecture categorization report is that safety and security requirements need to be coordinated and an iterative approach to OT architecture taken to reduce cyber-physical risk in the design phase [47-49].

Cyber risk for delivered systems are currently covered in contracts that stipulate cybersecurity procurement provisions. It has been shown however, that a significant amount of residual risk is still present despite these cybersecurity procurement provisions [13]. The residual risk is assumed by the licensee when the delivered systems are used, and the cybersecurity procurement provisions are not enforced or quantified. For ARs the degree of standardization across all reactors, of a reference design, will be high to enable economies of scale and streamlined licensing. A high degree of standardization across reactors implies that the licensees will have less capability to implement unique cybersecurity controls, make changes to the OT architecture, and latent cyber-physical vulnerabilities will be systemic.

To illustrate the risk posed by latent cyber-physical vulnerabilities in an OT architecture, a comparison to airplanes can be made. Airplanes are systems with high public safety consequences and a high degree of architectural standardization for a reference design. In the U.S. an airplane vendor must get a new design approved by the Federal Aviation Administration. If the design is approved, airline companies can purchase the airplane and operate within the approved operational envelope. Assuming that the aircraft is operated correctly, and an accident occurs due to a fundamental architectural error the airplane design company is liable for damages. This is the case for the two Boeing 737 Max planes that crashed in 2018 and 2019, Boeing is responsible for $2.5 billion in related damages [50]. A significant portion of this settlement will go to the airlines that purchased the 737 and lost revenue, while the design was grounded. A cyber-physical attack that exploits a latent vulnerability in the architecture and causes an accident is different regarding liability. However, the ramifications of such a vulnerability would be the same.

If Boeing was an AR designer and an architectural flaw lead to an accident, Boeing would not be liable under the Price-Anderson Act. Under the Price-Anderson Act the airline would be liable. Thus, licensees of ARs cannot claim that OT architecture risk is transferred risk and the consequences of a latent AR OT architecture vulnerability are absorbed by the licensee. All reactors, of the same reference design, would be shut down until the architectural flaw is mitigated. The impact to the grid could be severe depending on the number of reactors in operation. Under a reformulation of the Price Anderson Act AR design companies may need to assume liability in accident scenarios with the root cause stemming from an architectural flaw. Assumption of liability may also incentivize mitigation of OT architecture cyber-physical risks in the design phase.

Emergent OT architecture cyber-physical risks include instrumentation & control (I&C) function consolidation, monolithic zones, and multi-unit operation. These risks stem from design requirements that reduce the complexity and cost of the manufacturing process, as well as costs associated with O&M. Advance reactor cores will most likely be assembled at a factory and shipped to the plant site. In this case the licensee will not have the ability to alter the OT architecture design associated with the primary loop and interconnecting systems that interface with the auxiliary plant. The same organizational dynamics are true for the auxiliary plant, which contains, at a minimum, energy conversion and heat transfer systems. These risks, although not exhaustive, will be coved in the following sub-sections.

5.3.1. Instrumentation & Control Function Consolidation

System function consolidation can be seen in many AR concepts and is driven by many different top-level design requirements. The primary diver of I&C function consolidation being the inclusion, and subsequent reliance on, passive safety systems. The advantages of consolidating I&C functions include, enhanced coordination between systems, reduced development and maintenance costs, and reduced need for operator action. A simple example is consolidating heating and cooling controllers into a unified architecture. Poor coordination between heating and cooling can cause the temperature to oscillate due to overshoot/undershoot. Given the system's dynamic response the control engineer may have to limit the temperature to a prespecified range to eliminate oscillation. If the controllers are combined however, the temperature range can be better coordinated. There are two different methods that can be used to consolidate distributed control systems, the first being executing the control logic on the same physical controller and the second being networking two independent controllers together so that information can be communicated, see Figure 5-4. From a security perspective, the consolidation of I&C functions increases risk due the increased level of control gained through malicious compromise. Compromise of the system yields more degrees of freedom the adversary can utilize, and the attack surface is large relative to two non-networked, independent controllers within different access levels. Unifying control structures typically also implies an increase in joint access, which a key administrative security control for NPPs.

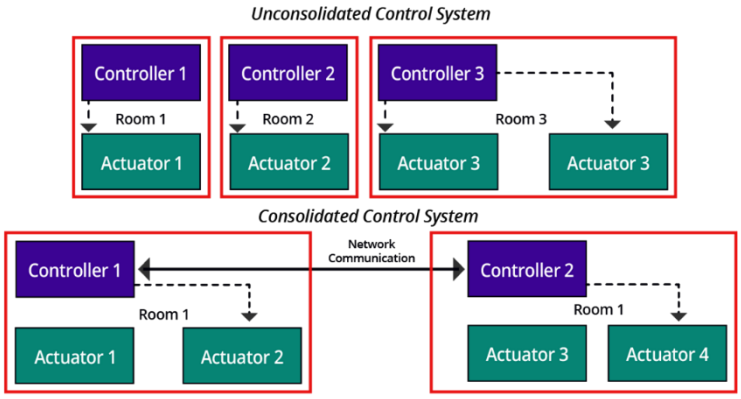

Figure 5-4. Example of an unconsolidated vs consolidate distributed control system

5.3.2. Monolithic Zones

The concept of security zones and security levels is refined in NST 47, as seen in Figure 5-5. A risk that has been identified with generation II reactor systems and can be repeated by ARs is using large monolithic zones within a security level. The genesis of monolithic zones can be mapped to the extensive requirements for safety and non-safety divisions. These requirements include, but are not limited to, fire safety, interdivisional requirements, safety requirements, and security requirements. For example, the use of a large monolithic zone within a security level simplifies the application of security controls due to security controls being common to all CDAs within a zone/level. Careful segmentation of zones as part of an iterative OT architecture design process could mitigate this issue but would increase cost and drastically decrease I&C function consolidation. Indicating a need for an optimization process that can coordinate conflicting requirements.

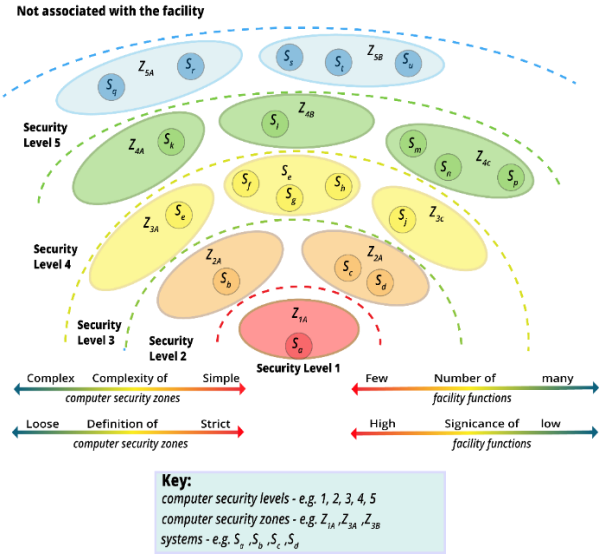

Figure 5-5. Visual depiction of the concept of security levels and security zones [51]

Large monolithic zones decrease the difficulty of pivoting within a security level, due to common security controls. From a security perspective when monolithic zones are used, communication requirements within a security zone and security level are equivalent. For the purposes of this report communication within a monolithic zone will be referred to as interlevel communication.

Separation between systems within the same security level are aligned with U.S. NRC's Highly-Integrated Control Rooms – Communication Issues (DI&C-ISG-04) [52]. The four principles that are used in interlevel communication are shown in Table 5-1.

Table 5-1. U.S. NRC's Highly-Integrated Control Rooms Communication Principles

Interdivisional Communication	Communications among different safety divisions or between a safety division and a non-safety entity

Command Prioritization	Selection of a particular command to send to an actuator when multiple and conflicting commands exist
Multidivisional Control and Display Stations	Use of operator workstations or displays that are associated with multiple safety divisions and/or with both safety and nonsafety functions
Digital System Network Configuration	The network or other interconnection of digital systems that might affect plant safety or conformance to plant safety analysis assumptions (interconnections among safety divisions or between safety and nonsafety divisions should also satisfy the guidance provided for interdivisional communications)

The primary premise for interdivisional communication is that safety functions are simple and isolated from information originating from within or external to its own safety division. A notable exception is division voting logic, which requires inputs from multiple safety divisions. Communication handshaking and interrupts from outside the systems own safety division is prohibited. Network communication should be highly formatted such that every message is identical. This is also true for shared, process, input, and output memory locations. Communication faults in non-safety systems should credible and not affect the integrity of the safety function. Error-detecting or error-correcting code should be used for vital communication such that invalid messages are always categorized recoverable or unrecoverable. The bandwidth of such systems should be tested such that there is always sufficient margin for safety functions to be performed.

Command prioritization requirements are regarding "priority modules" that can take safety and non-safety actuation commands and only send the command that has the highest priority. The actuated device then becomes a safety-related component. The priority module is also categorized as safety-related and must comply with U.S. NRC safety related requirements. The definitions of safety channel and safety division in DI&C-ISG-04 are relevant to the present analysis of interlevel communication requirements, so they are presented here for completeness.

A safety channel as used herein is a set of safety-related instruments and equipment, along with the associated software, that together generate a protective actuation or trip signal to initiate a single protective function. While an analog/hardwired system would have each functional circuit clearly assigned to only one channel, the processor and other components in a digital system may be assigned to multiple channels within a single division.

A safety division is the collection of all safety channels that are powered by a single power division. Different channels perform different functions. Different divisions perform the same set of functions, and are redundant to one another. Licensing typically credits redundancy among divisions. The voting logic that generates the final actuation signal to an item of plant equipment typically resides in one division and receives input from redundant channels in all divisions. For the purposes of this guidance, it is to be assumed that each of the actuation signals entering the voting logic that establishes the final actuation signal to an item of plant equipment is in a different division, regardless of the particular usage of the term "division" for a particular nuclear power plant.

By these definitions a priority module would be part of a safety channel, within a safety division. For multidivisional control and display stations the requirements stated in interdivisional communication and command prioritization are rolled up into station requirements. There are safety and non-safety stations with four informational scenarios of concern. These informational scenarios are represented in Table 5-2.

Table 5-2: Multidivionsal Control and Display Stations Informational Scenarios

Scenario 1	Nonsafety stations receiving information from one or more safety divisions
Scenario 2	Safety-related stations receiving information from other divisions
Scenario 3	Safety-related stations controlling the operation of equipment in other safety-related divisions
Scenario 4	Malfunctions and Spurious Actuations

Communicating between levels, intralevel communication, is much simpler in that only one-way communication is only permitted using deterministic, hardware-based network segmentation (i.e. data diodes). It can also be concluded that interlevel communication for non-safety systems have a lower set of requirements relative safety systems. These requirements are not specified in DI&C-ISG-04 and further research should be done to confirm communication requirements for zones with only non-safety systems. It may be logical to assume that non-safety systems can employ bidirectional communication between systems inside a monolithic zone architecture.

5.3.3. Multi-Unit Architectures

Multi-unit architectures for ARs are common and there are several different proposed Balance of Plant (BoP) layouts, Figure 5-6 [53]. NuScale has decided to take the shared main control room (MCR) approach, which is a significant departure from the traditional model used by U.S. generation II reactors. Due to the relatively small electrical outputs of many AR designs and desired minimal MCR staffing, it is unlikely that any ARs will use the traditional approach. For initial designs the shared MRC approach will be the easiest to implement due to a maximization of modularity. However, as AR production reaches a significant level a shared MCR and BoP model would be the most efficient option for a multi-unit configuration.

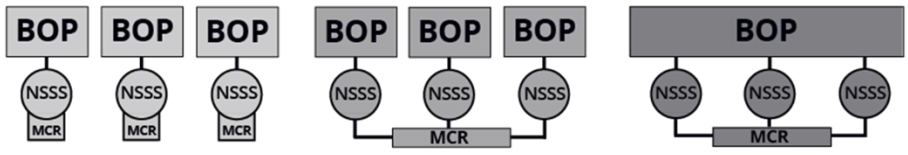

Figure 5-6. Leading concepts of multi-unit BoP architectures: Traditional Model(left), Shared MCR Model (Middle), Shared MCR and BOP Model (right)

From a cyber-physical risk perspective, the traditional approach would be the most resilient and the shared MRC and BoP would be the least resilience to a loss of business continuity scenario. A mutli-unit configuration with vital resources shared amongst nuclear steam supply systems has the potential to be more susceptible to common mode failure. The difference in safety between approaches is a significantly more difficult problem that requires multi-unit probabilistic risk assessment and can only be addressed on a case by case basis [54]. Multi-unit severe accident analysis can bound the worst-case scenario for a multi-unit configuration, but a security centric quantitative approach may also be needed to evaluate the impact of compromised shared systems. Again, referring to the need for quantitative cyber-physical risk assessment.

The departure from the traditional model also suggests that the attack surface maybe shifting from the nuclear island to the MCR or BOP. Since components within the BoP are considered non-safety, they are jointly regulated by the Federal Energy Regulatory Commission and U.S. NRC. Compromising the auxiliary plant of an AR can result in extended shutdown with low risk of radiological release, making the BoP an attractive target. Especially since security requirements for non-safety components are significantly less than safety components. Low risk of radiological release is important to minimize collateral damage, assuming this outcome is the adversary's goal. If an adversary's goal is to cause collateral damage it is unlikely that the BoP will be the primary or only target. U.S. NRC's guidance on a highly integrated MCR needs to be augmented to include additional guidance on highly integrated BoP for multi-unit AR architectures.

6. LICENSEE MANAGED RISK

In this section, licensee managed cyber-physical risks associated with off-site security operations, on-site security operations, and remote operation will be reviewed. Risks in these domains, although not exhaustive, are highly relevant for future AR cyber-physical risk management programs. Relative to generation II reactors it is expected that more cyber-physical risk will be absorbed by the licensee from external sources than internally managed. The high standardization proposed for fleets of AR based on a "reference design" is the primary driver of this trend. The relevant attack vectors for licensee managed risk are physical access, portable media & device connectivity, wired, and wireless communication. Based on relevant attack vectors for NPPs listed in NRC RG 5.71.

6.1. Off-Site Security Operations

The first and most significant barrier to current NPPs is the physical security of the site. Limiting access to sensitive areas is an obvious first line of defense. The PPS can be broken down into three fundamental functions, detect, delay, and respond, Figure 6-1 [55]. Detection is the discovery of an intrusion, such as seeing adversary on a camera or sensing a boundary crossing. Door locks, walls, barriers, and security personnel are meant to delay the actions of the adversary to damage the site. The delay provides time for the response force to neutralize the adversary. The PPS will function as intended if the detection and delay times enable the response force to arrive in time to neutralize the adversary.

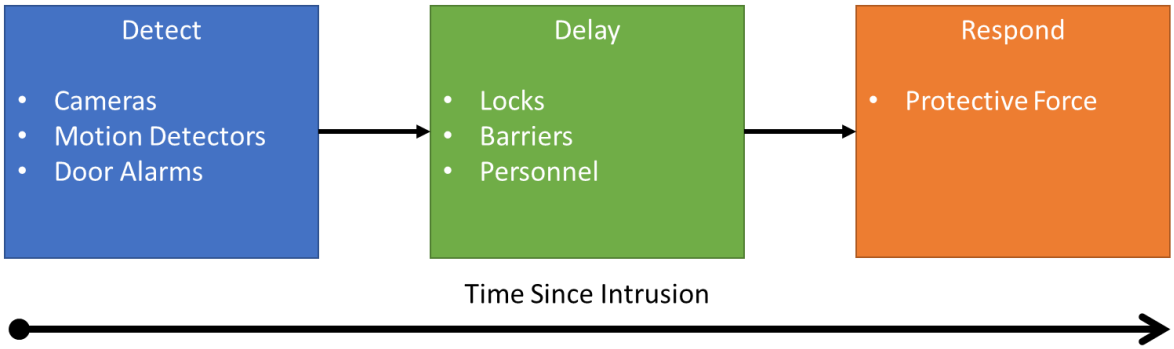

Figure 6-1. Physical Protection System functional diagram.

Digital PPSs have begun to replace analog systems, such as closed-circuit television being replaced with IP based cameras. Thus, cyber-security of these systems has become a concern [56]. Detect functions are largely digital systems and a cyber-attack could potentially degrade or negate these detection functions. The longer an adversary can delay their detection the more time they will have to complete their objectives.

Cybersecurity of the PPS is highly relevant for ARs that may be remotely operated or cannot economically support a large on-site security staff. The lack of an on-site response force will require additional delay mechanisms and changes to detection time will have greater significance [57]. The dependance on timely response greatly increases the cybersecurity assurance requirements of the PPS.

It is worth considering the scenario of an AR sited in a remote location. Assume the reactor is remotely operated, and security consists of intermittent patrols combined with remote monitoring

by a centralized security center. The site can be accessed by those with valid access cards and have scheduled access to the site. The central security center becomes compromised via their supply chain. The threat actor authorizes an access card enabling unauthorized access to the facility. The adversary now has access to the site and the security center believes this intrusion is an authorized event.

The core difference between the cybersecurity of control systems and physical security is that compromises of the physical security require a physical attack to elicit a consequence [58]. A cyber-attack on the control system may have severe consequences without an adversary being physically present, but a cyber-attack on the PPS could be motivated by an effort for an adversary to gain physical access to a facility. The consequences of a cyber-physical attack are potentially far greater than just a purely cyber-based or physical-based attack. Due to the PPS's reliance on digital systems an AR with off-site security operations will require higher PPS cybersecurity design assurance relative to ARs with on-site security. For off-site security a cyber-physical attack scenario is credible and will need to be included in the design basis threat (DBT) for the facility. The OT architecture of the PPS will also need to be included in the sites cyber-physical risk assessment to reduce cyber-physical risk in the design phase of the PPS.

6.2. On-Site Security Operations

In this section cyber-physical risks for on-site AR security operations will be covered. If a licensee decides to have an on-site security operation there are advantages and disadvantages to having the central alarm station (CAS), secondary alarm stations (SAS), CSAT, and the protective force on-site as part of the security plan. Advantages and disadvantages are outlined in Table 6-1. The effectiveness of on-site security, however, can only be determined on a case by case basis. The licensee's commitment to a rigorously defined and strictly enforced security plan will have a large impact of the NPPs security culture. Security culture is critical for maintaining security over an extended period.

Table 6-1. Potential advantages and disadvantages to having on-site security operations

Advantages	Disadvantages
A faster response time can be used in the PPS design. Detection and delay requirements can be decreased	Across a fleet of ARs on-site security is less efficient relative to off-site security
Decreases cybersecurity and assurance requirements	On-site security may not be sufficiently staffed to be effective
Security infrastructure can be easily changed if the DBT is reevaluated relative to off-site security	Decreased ability for ARs to leverage automation

Most advanced reactors will be categorized as category II facilities based on their proposed enrichment levels, storage of spent nuclear fuel in wet storage, and on-site inventory. See Table 6-2, for U.S. NRC special nuclear material classifications and definitions. The ARs surveyed in the Advanced Reactor Architectural Categorization report plan on using high-assay low enriched uranium (HALEU) fuel with enrichments between 5% and 20%. Notably exceptions are General Electric Hitachi's BWRX-300 and NuScale's SMR design. These design plan on using fuel with 5% or less enrichment. Enrichment levels and on-site fuel inventory will factor into NRCs future approach to on-site physical security requirements relating to NRC 10 CFR 73.55 [59].

Table 6-2. U.S. NRC Special Nuclear Material Classifications and Definitions

U.S. NRC's special nuclear material classifications	Definition
Category III is special nuclear material of low strategic significance	(1) Less than an amount of special nuclear material of moderate strategic significance as defined in paragraph (1) of the definition of strategic nuclear material of moderate strategic significance in this section, but more than 15 grams of uranium-235 (contained in uranium enriched to 20 percent or more in U–235 isotope) or 15 grams of uranium-233 or 15 grams of plutonium or the combination of 15 grams when computed by the equation, grams = (grams contained U–235) + (grams plutonium) + (grams U–233); or (2) Less than 10,000 grams but more than 1,000 grams of uranium-235 (contained in uranium enriched to 10 percent or more but less than 20 percent in the U–235 isotope); or (3) 10,000 grams or more of uranium-235 (contained in uranium enriched above natural but less than 10 percent in the U–235 isotope).
Category II is special nuclear material of moderate strategic significance or irradiated fuel	Special nuclear material of moderate strategic significance means: (1) Less than a formula quantity of strategic special nuclear material but more than 1,000 grams of uranium-235 (contained in uranium enriched to 20 percent or more in the U–235 isotope) or more than 500 grams of uranium-233 or plutonium, or in a combined quantity of

U.S. NRC's special nuclear material classifications	Definition
	more than 1,000 grams when computed by the equation, grams = (grams contained U–235) + 2 (grams U–233 + grams plutonium); or (2) 10,000 grams or more of uranium-235 (contained in uranium enriched to 10 percent or more but less than 20 percent in the U–235 isotope).
Category I is a formula quantity of strategic special nuclear material	Formula quantity means strategic special nuclear material in any combination in a quantity of 5,000 grams or more computed by the formula, grams = (grams contained U–235) + 2.5 (grams U–233 + grams plutonium). This class of material is sometimes referred to as a Category I quantity of material.

6.3. Remote Operation

In a remote operation architecture, the licensee has a MCR and other operation and maintenance buildings, such as the SAS, located off-site that can monitor and send manual commands to the control system. The intrinsic benefit of this setup is a dramatic cost reduction and increase in productivity can be realized for fleets of ARs. This report differentiates remote operation from autonomous operation based on the unique characteristics of each architecture and the assumed risk profile generated by each approach. Autonomous operation is an absorbed risk because it is unlikely that licensees will develop autonomous architectures and go through a U.S. NRC approval process. Autonomous architectures will most likely be purchased and used as a black box from the licensee's perspective. The R&D and domain specific expertise required to develop an autonomous architecture exceeds the capability and general scope of a licensee's business model as covered in Section 5.2 - Autonomous Operation. It is more likely that the licensee will maintain the infrastructure required to remotely operate a fleet of reactors. Therefore, for the purposes of this report remote operation is assumed to be a licensee managed risk.

Due to the prohibited cost of a direct connection using cabling, remote operation architectures will have to rely on wireless communication. Ultra-secure high reliability wireless schemes have been proposed that rely on Suite B Cryptography as defined by the National Security Administration. This type of encryption protocol can handle U.S. government information up to the secret level. Type-1 encryption is used to handle information up to the classified level. J. Cordaro et. al. have shown that for radiation monitoring I&C an ultra-secure, short range wireless setup can be established [60]. How this method can be expanded to work in long-range applications is still an on-going challenge [61]. It may be possible to use 5G and beyond-5G cellular networks with modifications to ensure not just a high reliability but an ultra-secure connection [62].

The risks associated with remote operation are centered around the cyber-physical risks to the physical connection between the operator and the NPP. If the connection between the NPP is lost the plant must be designed such that no external operator action is required given the entire spectrum of hazards that can lead to a significant increase risk of radiological release or theft of radiological material. It is unlikely that the assurance of an ultra-secure high reliability wireless network would warrant a differing fundamental design requirement, from a regulatory perspective. This is due to the fact that the number of ways a wireless connection can be eliminated through cyber-physical means is potentially large enough that it is justifiable to assume that loss of remote connection is a feasible plant state that should be incorporated into the design.

Furthermore, if remote operators can execute manual control actions the possibility of hijacking the connection is a plausible scenario. Design requirements in this case will need to focus on how to design the on-site control system such that a hijacked connection cannot lead to an accident scenario. Once accident scenarios due to a hijacked connection are addressed, control systems resilient to hijacked connections and additional security controls can be applied such that the risk to business continuity is minimized. Prevention of a hijacked connection is the first goal of the security architect. However, assuming the connection is hijacked allows additional security controls to be implemented to mitigate further propagation of the attack. The control system in the case of an actual or suspected hijacking will need the ability to successfully recognize, with a high degree of confidence, anomalies in manual commands. Extensive coordination with plant safety will be required to ensure that manual commands essential to ensure safety are not blocked due to false classification.

As covered in Section 6.1 Off-Site Security Operations, NPPs requiring remote operation will also most likely adopt an off-site security architecture, where the design of the PPS is heavily reliant on high assurance detection and delay systems. The response time of an off-site response force will factor significantly into the detection and delay requirements for the PPS. Coordinated research between safety and security engineers can be conducted to explore the plausibility of connecting the security and physical state of the plant. Such that safety and security control systems can communication to position the plant into a state of optimal protection in the event of a cyber-physical attack. Such a state would be defensive in nature and prioritize the safety and security goals of the site, over day to day operations and maintenance considerations. The interface between security and safety control systems should be considered for both autonomous and remote operation architectures for ARs as a potentially way to lower cyber-physical risk.

7. CONCLUSIONS

This report has identified limitations in the current practice of cyber physical risk management for ARs and identified important cyber physical risks that AR designers should carefully consider. The initial intention of this report was to take a detailed look at specific cyber risks faced by AR designs coming to market. This however was not possible due to challenges that could not be resolved. (1) AR designers contacted did not have detailed design information of an I&C system that could be analyzed or refused to collaborate. (2) Designers that were willing to collaborate had proprietary concerns regarding their design information that required NDAs and external approval.

To effectively study cyber risks detailed design information is required to conduct experiments or detailed analyses and obtain results. Detailed design information necessitates the existence of a finished design and partners that are willing to collaborate. To mitigate the lack of information, a comprehensive review of current guidance regarding NPP cyber physical risks and a review of risk categories pertinent to ARs was adopted. NIST 800-37 and particularly ISO 27005 provide an excellent risk management framework for information systems that can be leveraged in future AR guidance and domestic standards. Methodologies like the DEG, HAZCADS, TAM, and DRAM provide the initial technical basis for quantitatively identifying and analyzing cyber risks. However, fundamental investments need to be made to improve current cyber-physical risk methodologies and unify security and safety-based analyses. Missing from the broader discussion is a method to unify cyber and physical security-based analysis to reduce overall security costs for ARs.

On-going challenges regarding cyber-physical risk methodologies include organizational and funding divisions between cybersecurity, physical security, and safety analysis groups for ARs. Cross-cutting research will need to be continuously pursued to ensure that cybersecurity, physical security, and safety analysis groups are working towards a common goal, scalable approaches, and modular solutions. If ARs repeat the same mistakes as generation II reactors, U.S. based ARs may not be economically competitive or manageable at scale. AR designers should be encouraged to continue to innovate and regulators need to have the necessary tools and resources to properly assess associated risks. Coordinating the relationship between DoE NE, U.S. NRC, the National Laboratories, and the nuclear industry will need enduring resolve to ensuring stress points in relationships are mitigated and an optimal balance in the public-private partnership maintained.

REFERENCES

[1] *Regulation Guide 5.71 - CYBER SECURITY PROGRAMS FOR NUCLEAR FACILITIES,* U.S. Nuclear Regulatory Commission, 2010.

[2] *NEI 13-10 Cyber Security Control Assessments Rev. 6,* Nuclear Energy Institute, 2017.

[3] *10 Code of Federal Regulation Part 53 LICENSING AND REGULATION OF ADVANCED NUCLEAR REACTORS,* U.S. Nuclear Regulatory Commission, TBD.

[4] U.S. Department of Energy Office of Inspector General Office of Audits and Inspections, "Inquiry into the Security Breach at the National Nuclear Security Administration's Y-12 National Security Complex - DOE/IG-0868," August 2012.

[5] *NIST SP 800-37: Risk Management Framework for Information Systems and Organizations*, NIST, 2018.

[6] *ISO/IEC 27000:2018: Information security risk management*, ISO/IEC, 2018.

[7] R. Fasano, A. Hahn, A. Haddad, and C. Lamb, "Advance Reactor Operational Technology Architecture Categorization," DOE NE, 2021.

[8] D. Greenfield. "Is the Purdue Model Still Relevant?" Automation World. https://www.automationworld.com/factory/iiot/article/21132891/is-the-purdue-model-still-relevant (accessed Sep. 20, 2021, 2021).

[9] *10 Code of Federal Regulation PART 73 - PHYSICAL PROTECTION OF PLANTS AND MATERIALS,* U.S. Nuclear Regulatory Commission.

[10] *NIST SP 800-82: Guide to Industrial Control Systems (ICS) Security*, NIST, 2015.

[11] *NIST SP 800-53: Security and Privacy Controls for Information Systems and Organizations*, NIST, 2020.

[12] J. P. Farwell and R. Rohozinski, "Stuxnet and the Future of Cyber War," *Survival,* vol. 53, no. 1, pp. 23-40, 2011, doi: 10.1080/00396338.2011.555586.

[13] L. A. Dawson, Joshua Michael Daley, Kandy Phan, "Enabling Secure and Cost-Effective Nuclear Power Plant Wireless Communications (No. SAND2019-1239C)," presented at the 11th Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies, 2019.

[14] S. Eggers, "A novel approach for analyzing the nuclear supply chain cyber-attack surface," *Nuclear Engineering and Technology,* vol. 53, no. 3, pp. 879-887, 2020, doi: https://doi.org/10.1016/j.net.2020.08.021.

[15] S. L. Eggers, *The Nuclear Digital I&C System Supply Chain Cyber-Attack Surface* (Conference: ANS Annual Meeting 2020, Phoenix, AZ, 06/07/2020 - 06/11/2020). ; Idaho National Lab. (INL), Idaho Falls, ID (United States), 2020, p. Medium: ED.

[16] D. Mehta *et al.*, "The Big Hack Explained: Detection and Prevention of PCB Supply Chain Implants," *J. Emerg. Technol. Comput. Syst.,* vol. 16, no. 4, p. Article 42, 2020, doi: 10.1145/3401980.

[17] H. Kwak and T. Hubing, "An Overview of Advanced Electronic Packaging Technology," 2007.

[18] K. Basu *et al.*, "CAD-Base: An Attack Vector into the Electronics Supply Chain," *ACM Trans. Des. Autom. Electron. Syst.,* vol. 24, no. 4, p. Article 38, 2019, doi: 10.1145/3315574.

[19] S. Paley, T. Hoque, and S. Bhunia, "Active protection against PCB physical tampering," presented at the 17th International Symposium on Quality Electronic Design, 2016.

[20] J. Harrison, N. Asadizanjani, and M. Tehranipoor, "On malicious implants in PCBs throughout the supply chain," *Integration,* vol. 79, pp. 12-22, 2021, doi: https://doi.org/10.1016/j.vlsi.2021.03.002.
[21] Z. Zhang and Q. Yu, "Modeling Hardware Trojans in 3D ICs," in *2019 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 15-17 July 2019 2019, pp. 483-488, doi: 10.1109/ISVLSI.2019.00093.
[22] E. V. Levine, "The Die is Cast: Hardware Security is Not Assured," *Queue,* vol. 18, no. 4, p. Pages 60, 2020, doi: 10.1145/3424302.3431245.
[23] Y. Hayashi and S. Kawamura, "Survey of Hardware Trojan Threats and Detection," presented at the 2020 International Symposium on Electromagnetic Compatibility - EMC EUROPE, 2020.
[24] H. Li, Q. Liu, and J. Zhang, "A survey of hardware Trojan threat and defense," *Integration,* vol. 55, pp. 426-437, 2016/09/01/ 2016, doi: 10.1016/j.vlsi.2016.01.004.
[25] Z. Zhang, J. Dofe, P. Yellu, and Q. Yu, "Comprehensive Analysis on Hardware Trojans in 3D ICs: Characterization and Experimental Impact Assessment," *SN Computer Science,* vol. 1, no. 4, p. 233, 2020/07/16 2020, doi: 10.1007/s42979-020-00220-0.
[26] X. Xu, F. Rahman, B. Shakya, A. Vassilev, D. Forte, and M. Tehranipoor, "Electronics Supply Chain Integrity Enabled by Blockchain," *ACM Trans. Des. Autom. Electron. Syst.,* vol. 24, no. 3, p. Article 31, 2019, doi: 10.1145/3315571.
[27] Z. Huang, Q. Wang, Y. Chen, and X. Jiang, "A Survey on Machine Learning Against Hardware Trojan Attacks: Recent Advances and Challenges," *IEEE Access,* vol. 8, pp. 10796-10826, 2020, doi: 10.1109/access.2020.2965016.
[28] C. Schuett, J. Butts, and S. Dunlap, "An evaluation of modification attacks on programmable logic controllers," *International Journal of Critical Infrastructure Protection,* vol. 7, no. 1, pp. 61-68, 2014, doi: https://doi.org/10.1016/j.ijcip.2014.01.004.
[29] M. Ohm, H. Plate, A. Sykosch, and M. Meier, "Backstabber's Knife Collection: A Review of Open Source Software Supply Chain Attacks," Cham, 2020: Springer International Publishing, in Detection of Intrusions and Malware, and Vulnerability Assessment, pp. 23-43.
[30] D.-L. Vu, I. Pashchenko, F. Massacci, H. Plate, and A. Sabetta, "Typosquatting and Combosquatting Attacks on the Python Ecosystem," presented at the 2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), 2020.
[31] Forescout. "New Critical Operational Technology Vulnerabilities Found on NicheStack – Mitigation Advised." Forescout Research Labs. https://www.forescout.com/blog/new-critical-operational-technology-vulnerabilities-found-on-nichestack/ (accessed 2021).
[32] G. Lin, W. Xiao, L. Y. Zhang, S. Gao, Y. Tai, and J. Zhang, "Deep neural-based vulnerability discovery demystified: data, model and performance," *Neural Computing and Applications,* 2021, doi: 10.1007/s00521-021-05954-3.
[33] A. Qasem, P. Shirani, M. Debbabi, L. Wang, B. Lebel, and B. L. Agba, "Automatic Vulnerability Detection in Embedded Devices and Firmware: Survey and Layered Taxonomies," *ACM Comput. Surv.,* vol. 54, no. 2, p. Article 25, 2021, doi: 10.1145/3432893.
[34] V. Duddu, "A Survey of Adversarial Machine Learning in Cyber Warfare," *Defence Science Journal,* vol. 68, no. 4, 2018, doi: 10.14429/dsj.68.12371.
[35] D. Maynor, A. Nikolic, M. Olney, and Y. Younan. " The MeDoc Connection " Talos Intelligence. https://blog.talosintelligence.com/2017/07/the-medoc-connection.html (accessed Aug. 12th, 2021, 2021).

[36] CISA. (2021). *AA20-352A, Advanced Persistent Threat Compromise of Government Agencies, Critical Infrastructure, and Private Sector Organizations*. [Online] Available: https://us-cert.cisa.gov/ncas/alerts/aa20-352a
[37] M. Graham, "Context Threat Intelligence - The Monju Incident," ed: Context Threat Intelligence, 2014.
[38] M. Plachkinova and C. Maurer, "Teaching Case: Security Breach at Target," *Journal of Information Systems Education,* vol. 29, no. 1, pp. 11-20, 2018.
[39] C. Glenn, D. Sterbentz, and A. Wright, "Cyber Threat and Vulnerability Analysis of the U.S. Electric Sector,"; Idaho National Lab. (INL), Idaho Falls, ID (United States), INL/EXT-16-40692 United States 10.2172/1337873 INL English, 2016. [Online]. Available: https://www.osti.gov/servlets/purl/1337873
[40] (2018). *Alert (TA18-074A): Russian Government Cyber Activity Targeting Energy and Other Critical Infrastructure Sectors*.
[41] Y. Zeng, J. Liu, K. Sun, and L.-w. Hu, "Machine learning based system performance prediction model for reactor control," *Annals of Nuclear Energy,* vol. 113, pp. 270-278, 2018, doi: https://doi.org/10.1016/j.anucene.2017.11.014.
[42] C. Spirito, S. Aghara, C. Duffley, J. Strandburg, J. Coble, and F. Zhang, "Threat Assessment Methodology for Autonomous and Remote Operations for Advanced Reactors," Idaho National Laboratory, 2021.
[43] L. Lin *et al.*, "Development and assessment of a nearly autonomous management and control system for advanced reactors," *Annals of Nuclear Energy,* vol. 150, 2021, doi: https://doi.org/10.1016/j.anucene.2020.107861.
[44] Z. Wang, C. Liu, and X. Cui, "EvilModel: Hiding Malware Inside of Neural Network Models," presented at the 26th IEEE Symposium on Computers and Communications (ISCC 2021), 2021.
[45] R. T. Wood, B. R. Upadhyaya, and D. C. Floyd, "An autonomous control framework for advanced reactors," (in English), *Nuclear Engineering and Technology,* vol. 49, no. 5, pp. 896-904, Aug 2017, doi: 10.1016/j.net.2017.07.001.
[46] J. Herkert, J. Borenstein, and K. Miller, "The Boeing 737 MAX: Lessons for Engineering Ethics," *Science and Engineering Ethics,* vol. 26, no. 6, pp. 2957-2974, 2020/12/01 2020, doi: 10.1007/s11948-020-00252-y.
[47] *IEC 62645:2019: Nuclear power plants - Instrumentation, control and electrical power systems - Cybersecurity requirements*, International Electrotechnical Commission, Geneva, 2019.
[48] *IEC 62859:2016/AMD1:2019 : Amendment 1 - Nuclear power plants - Instrumentation and control systems - Requirements for coordinating safety and cybersecurity*, International Electrotechnical Commission, Geneva, 2019.
[49] *IEC 63096: Nuclear power plants – Instrumentation, control and electrical power systems – Security controls*, International Electrotechnical Commission, Geneva, 2020.
[50] D. Schaper, "Boeing To Pay $2.5 Billion Settlement Over Deadly 737 Max Crashes," in *NPC*, ed, 2021.
[51] International Atomic Energy Agency, "NST047 - Computer Security Techniques at Nuclear Facilities," Vienna, TBD.
[52] (2009). *Highly-Integrated Control Rooms—Communications Issues*.
[53] International Atomic Energy Agency, "Instrumentation and Control Systems for Advanced Small Modular Reactors," Vienna, 2017.
[54] M. Modarres, T. Zhou, and M. Massoud, "Advances in multi-unit nuclear power plant probabilistic risk assessment," *Reliability Engineering & System Safety,* vol. 157, pp. 87-100, 2017.

[55] M. L. Garcia, *The Design and Evaluation of Physical Protection Systems*. 2008.
[56] J. Clem, W. Atkins, and V. Urias, "Investigation of Cyber-Enabled Physical Attack Scenarios," 2015.
[57] A. S. Evans, J. M. Parks, S. Horowitz, L. Gilbert, and R. Whalen, "U.S. Domestic Small Modular Reactor Security by Design," 2021. [Online]. Available: https://gain.inl.gov/SiteAssets/2021-April_SafeguardsAndSecurityWorkshop/Reading/U.S.%20Domestic%20Small%20Modular%20Reactor%20Security%20by%20Design%20-%20SNL.pdf
[58] M. T. Rowland, J. Sladek, and C. Nickerson, *EVALUATION OF THE APPROPRIATENESS OF TRUST MODELS TO SPECIFY DEFENSIVE COMPUTER SECURITY ARCHITECTURES FOR PHYSICAL PROTECTION SYSTEMS* (Conference: Proposed for presentation at the International Conference on Nuclear Security held February 10-14, 2020 in Vienna, Vienna, Austria.). ; Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), 2019, p. Medium: ED; Size: 12 p.
[59] *10 Code of Federal Regulation PART 73.55 - Requirements for physical protection of licensed activities in nuclear power reactors against radiological sabotage*, U. S. N. R. Commission.
[60] J. V. Cordaro, D. Shull, M. Farrar, and G. Reeves, "Ultra secure high reliability wireless radiation monitoring system," *IEEE instrumentation & measurement magazine,* vol. 14, no. 6, pp. 14-18, 2011.
[61] *Application of Wireless Technologies in Nuclear Power Plant Instrumentation and Control Systems*. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2020.
[62] R. Kumar, "Achieving Ultra-high Reliability and Low-latency in Future Wireless Networks," New York University Tandon School of Engineering, 2020.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Technical Library	01977	sanddocs@sandia.gov

This page left blank

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.