Complete IEC 61850 Protection and Control System Cybersecurity Is So Much More Than Device Features Based on IEC 62351 and IEC 62443

David Dolezilek, Dennis Gammel, and William Fernandes
Schweitzer Engineering Laboratories, Inc.

Revised edition released July 2019

Originally presented at the
10th Annual Protection, Automation and Control World Conference, June 2019

Complete IEC 61850 Protection and Control System Cybersecurity Is So Much More Than Device Features Based on IEC 62351 and IEC 62443

David Dolezilek, Dennis Gammel, and William Fernandes, *Schweitzer Engineering Laboratories, Inc.*

***Abstract*—The Purdue Model has become a useful reference for energy control system (ECS) architectures as a mission-critical subset of industrial control systems. This model provides a method to identify and define the multiple distinct segments of the ECS network based on their information technology (IT) and operational technology (OT) characteristics. This model has driven the development of ISA99 and IEC 62443 defense-in-depth strategies, where cybersecurity features are distributed among multiple levels of the control system. This defense-in-depth strategy provides complete cybersecurity instead of the inadequate device-level features called out in IEC 62351 and IEC 62443 Part 4. Recent failures of the technology these device-level features are based on illustrate that they often create new, unintended vulnerabilities much worse than the challenges they were intended to mitigate.**

In this paper, the Purdue Model is used to design defense-in-depth cybersecurity methods to implement human-to-machine and machine-to-machine digital communications within an ECS communications network. The ECS communications architecture is divided into multiple appropriate levels with unique requirements and features from the process up through the station and finally to the control center. Using these levels, it is possible to appropriately identify interacting cyber defense technologies, the levels at which they should be deployed, and which devices they belong to (IEC 62443 Part 3) instead of the arbitrary defense-in-breadth strategy of requesting that every device include every cyber defense technology (IEC 62443 Part 4).

I. Momentary and Sustained Outages in Energy Delivery and Energy Control Systems

The word "cyber," originating from the Greek word meaning "skilled steering or guidance," has taken on the modern meaning of using digital communications within and among intelligent devices to perform information gathering and commanded control. Information technology (IT) systems include networked communications among computers, business systems, and the internet. Operational technology (OT) systems include networked communications among industrial control system (ICS) devices performing automatic safety, operational, and monitoring processes. Energy control systems (ECSs, often referred to as secondary systems) are a specialized type of ICS that use cyber methods within and between intelligent electronic devices (IEDs) to perform protection, monitoring, and control of energy delivery systems (EDSs, often referred to as primary systems). Machine-to-machine communications, such as Mirrored Bits communications and IEC Generic Object-Oriented Substation Event (GOOSE) messages, perform automatic detection and reaction to EDS events. The same OT ECS communications support the gathering of IT information from OT devices, such as equipment monitoring and metering values, and the commanded control of the ECS via operator workstations. Therefore, the OT ECS components must be robust and resilient enough to perform the automatic machine-to-machine communications necessary to detect and isolate faults and to safely restore energy delivery in the EDS. Cyber availability for resilience requires that components and networks be both reliable and dependable.

EDS reliability is often defined as the ability of the primary system (i.e., process components, including generators, transmission lines, and breakers) to deliver electricity to all points of consumption and satisfy customer quantity and quality requirements.

North American Electric Reliability Corporation (NERC) transmission availability data system definitions [1] include outage indices used by the IEEE to define EDS reliability. Dispatching and planned operational outages do not affect reliability indices. Automatic operation of an energy switching device to cause a primary system element to change from an in-service state to a not-in-service state makes it unavailable.

These automatic operations are typically in reaction to the detection and isolation of a power system fault. The duration of the outage is related to the time necessary to automatically react, reconfigure the power system, and restore the flow of energy to the consumers. Communications-assisted protection systems within the ECS enable faster and better automatic operations and shorter outages.

A power system momentary outage is defined as an automatic outage with a duration of less than one minute. A power system sustained outage is defined as an automatic outage with a duration of a minute or longer [1].

ECS reliability is often defined as the ability of the secondary system components (including protective relays, controllers, and communications switches) to deliver information to all points of consumption and satisfy the OT requirements for information quantity and quality.

Communications reliability of the secondary system is measured by the same outage indices as the primary system. As with the primary system, secondary system reliability is not affected by planned and operational outages but rather only by automatic outages that result from the operation of a communications switching device causing an element to change from an in-service state to a not-in-service state. These automatic operations are typically in reaction to the detection

and isolation of a communications system fault. The duration of the outage is related to the time necessary to automatically react, reconfigure the communications system, and restore the flow of information to the OT information consumers. High-speed communication of protection signals through the communications system enables faster and better communications-assisted decision making to support faster automatic operations in the protection system and shorter outage durations.

IEC 60834 [2] requires that transmission and receipt of protection control signals be less than 10 milliseconds and IEC 61850 [3] requires that they be less than 3 milliseconds. IEEE 1646-2004 [4] requires that protection information shared between IEDs within a substation be within one-fourth of a cycle (~4 milliseconds in a 60 Hz system) and information shared externally be within 8 to 12 milliseconds.

Fault clearing applications define the worst-case protection signal exchange time to be 20 milliseconds, even in the presence of a fault in the secondary system. Implementation of IEC 61850 GOOSE protection signal multicasting includes an instantaneous GOOSE publication after detection of a power system fault followed by a minimum of four additional GOOSE messages at various intervals. This repetitive burst of messages containing the power system fault information is done to increase the likelihood that the signal, within each of the repetitive messages, will get through the OT communications network. An outage in the communications system must be sufficiently short to enable at least one GOOSE message to reach its destination. A typical GOOSE message burst ends 16 milliseconds after detection of the power system fault in order to accomplish protection operation within the 20-millisecond worst-case operation time. Therefore, the duration of a momentary outage of the communications network must be less than 16 milliseconds for the secondary system to correctly serve the operation of the primary system.

A secondary system communications network momentary outage of less than 16 milliseconds will not cause a failure of the protection system to automatically operate the primary system.

A sustained outage in the secondary system communications network may cause a failure of the protection system to automatically operate the primary system. The risk of a sustained outage and the consequences must be understood in order to accept the risk or change the secondary system to mitigate the vulnerability.

Considering the lack of monitoring of communications system outage durations (except within newer software-defined networks), most indices are calculated based on interruptions of protection signal message receipt experienced by ECS components. Key process indicators include the total number of protection signal delivery outage events, the duration of each event, the accumulated durations of events, and the duration of the longest outage.

II. N-1 Availability of EDSs and ECSs

As described in [5], the EDS is often networked to improve service capability and availability. Networks of components enable devices and EDS sections to be manually removed from service for planned maintenance or automatically forced out of service to clear a fault without disrupting energy delivery. Therefore, networking enables the system to experience an outage and still perform. Designing for availability by using redundancy provides an alternative service when a component fails to serve its intended purpose, thus providing N-1 reliability. Primary systems are composed of an undetermined quantity of components, referred to as *N*; N-1 reliability means that one of the devices can fail to serve its purpose and the system will continue to function. That is, no single point of failure alone can cause a loss of service.

However, it is important to note that redundancy does not address the removal of the fault, and while the first fault exists the system will be in a N-0 state and no longer fulfill the design requirement of N-1. Therefore, redundancy may enable automatic operation to limit the effect of the fault to a momentary outage and the system will continue to function in the N-0 state. However, a second fault will result in a sustained outage that may cause loss of service until the outage is detected, isolated, and resolved by human interaction. Design for duplication means that two N-0 networks work independently and failure in one will result in a sustained outage that will remain until detected and corrected by human intervention. However, the second independent network will function in an N-0 state and provide service until it experiences an automatic outage.

Design for availability using resiliency is defined as the ability to automatically detect and isolate each failure and react to restore service, mitigate the initial fault, and return the system to its N-1 state after a brief outage.

The ECS is essential to automatically operating the EDS equipment to clear faults. The ECS uses OT as a tool to protect, monitor, and control the EDS. Since the resiliency of the primary system relies on the availability of the ECS, the ECS must be more reliable and available than the primary system it is tasked with keeping in service. It is particularly important that the secondary system be fully functional during the time that an automatic operation in the primary system has reduced the EDS reliability to N-0. In this condition it is essential that automatic ECS outages be momentary to ensure that the ECS will be available to support the return of the EDS to service. Ethernet-based OT requires that the Ethernet network be fault tolerant and remain in a N-1 condition even in the presence of an Ethernet fault. Like the primary system, networking of the Ethernet communications permits sections to be removed from service for planned maintenance or forced out of service to clear a fault without disrupting information delivery.

Protection system redundancy is best achieved with two independent and resilient systems, such as dual-primary protective relays communicating using robust dual-primary LANs. In this way, when there is a fault present in the System A, System B remains in service and the System A automatic outage is momentary after which it returns to service. To accomplish this, the communications network must be resilient and automatically detect, isolate, and reconfigure around a communications failure in order to preserve the

operation of the System B protection functions. This becomes an N-1 requirement for the secondary system during an automatic outage of the secondary system to ensure service to the primary equipment. This requires dual-primary protection and independent LANs with resilient reconfiguration that create only momentary outages of less than 16 milliseconds.

An alternative is singular protection devices dually connected to one or two independent LANs with resilient reconfiguration that creates only momentary outages of less than 16 milliseconds.

IEC 62439 Part 1 describes numerous technologies to improve the availability of Ethernet communications [6]. IEC 62439 Part 5.1.1, Resilience in Case of Failure, describes how industrial systems such as an EDSs rely on the correct function of the automation system. Industrial systems tolerate a degradation of the automation system for only a short time, called the grace time. The network recovery time should be shorter than the grace time since the application typically needs to perform additional tasks (related to protocol and data handling, waiting for the next scheduled communication cycle, and so on) before the ICS is back to the fully operational state. Therefore, the automatic outage should be momentary in duration.

ECS and ICS automation systems may contain redundancy to cope with a single component failure, but mission-critical designs require resiliency. Methods differ on how to handle resiliency, but their key performance factor is the recovery time (i.e., the time needed to restore operation after the occurrence of a disruption). If the recovery time exceeds the grace time of the industrial system, protection mechanisms initiate a (safe) shutdown, which may cause significant loss of production and plant operational availability.

III. ICSs, the Purdue Model, and Defense in Depth

The first ICSs were standalone systems, physically isolated and disconnected from external networks such as the business systems in the enterprise. The risk of cyberattacks was minimal and organizations were mainly focused on implementing physical security controls. Security by obscurity was commonly used as a defensive strategy to protect isolated ICS networks.

The needs to increase productivity, reduce operational costs, and access real-time information led asset owners to integrate the corporate and ICS networks with common communications protocols and open standards. While this replaced the diverse and unique proprietary solutions, it inadvertently exposed ICSs to a broader range of threats and introduced risks that did not exist in isolated ICSs. The security-by-obscurity philosophy became obsolete as a defensive strategy for ICSs due largely to the convergence of networks.

The perimeter defense emerged as a security strategy for ICSs. Firewalls strategically placed at the edge of the network were used to inspect and filter traffic and protect the network against external attacks; however, this offered no restrictions or traffic inspection on the network itself, thus still allowing internal attacks.

Some organizations, tried to implement the processes, methods, and techniques of the IT world to incorrectly secure ICSs, ignoring the special characteristics and restrictions of ICSs, some of which are as follows:

- Many ICS devices (e.g., protective relays) have the capacity to execute only their intended task and cannot implement demanding security controls like authentication or encryption.
- Firewalls and intrusion detection systems may introduce latency that can negatively affect the real-time communications and determinism of certain tasks.
- ICS processes have extremely high uptime requirements with no time for maintenance, patching, or other security-related activities. This makes nearly impossible the implementation of black-listing technologies.
- In case of emergency, operators must interact quickly and precisely with the ICS. The use of complex passwords, for example, may create delays that can mean the difference between life and death.

None of the security strategies mentioned so far offered the correct levels of security for ICSs, so it became necessary to create a methodical approach to compartmentalizing the applications and features within the ICS and securing them against internal and external attacks.

This was accomplished by the Purdue Enterprise Reference Architecture (also known as the Purdue Model) which was developed during the 1990s by Theodore J. Williams and the Purdue University Consortium for Computer Integrated Manufacturing.

Initially, the Purdue research did not take security or safety into account [7]. Its intended purpose was to improve factory ICS efficiency and reduce costs, with automation based on cyber processing and communications methods. The use of the Purdue Model for safety and cybersecurity came later. Initial research on cybersecurity using the Purdue Model was based on IT influences and misrepresented attributes of OT as challenges or threats (e.g., distributed and embedded devices, real-time control). However, these and other OT system attributes can be leveraged for whitelisting, baselining, monitoring, threat detection, and other means of securing OT systems.

Throughout the years, the Purdue Model has been used and adapted to different industries. In fact, the International Society of Automation's ISA99 framework is based on the Purdue Model and is used to describe the basic functions, composition, and levels of an ICS [8].

The ISA99 framework was developed by the Standards and Practices Committee 99 (SP99), but it is now aligned with IEC 62443 [9], which organizes a series of standards into four groups that address a wide range of topics related to ICS security.

The Purdue Model divides the ICS architecture into three zones and six levels, as shown in Fig. 1.

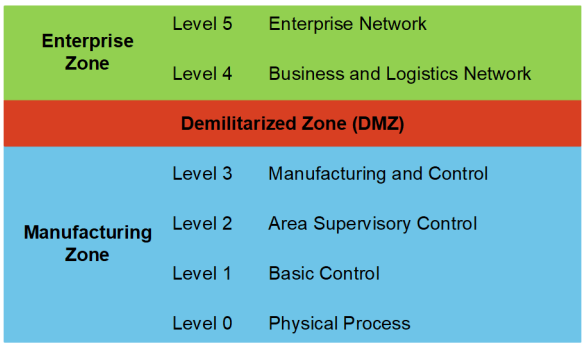

Fig. 1. The Purdue Model

The enterprise zone is not part of the ICS, but it obtains data from it to support business systems. This zone is composed of the enterprise network and the business and logistics network.

The enterprise network (Level 5) uses the data from lower levels to determine the status of production, inventory, and demand. The business and logistic network (Level 4) uses IT systems, not only to send production statistics to the enterprise network but also to distribute business information to some of the systems in the manufacturing zone. Database servers and file servers are normally included in this level.

The DMZ separates the systems in the enterprise zone from the systems in the manufacturing zone, preventing direct communication and allowing secure connection between these systems. In this zone, it is common to find web servers and database replication servers.

The DMZ was not included in the original Purdue Model, but it was added at a later stage to incorporate the resiliency requirements of the ECS. Also, some component categories were moved to more appropriate levels based on the work of ECS security standards like NERC Critical Infrastructure Protection and the National Institute of Standards and Technology (NIST) Cybersecurity Framework.

The manufacturing zone is where actual physical processes exists, and it can be subdivided in four different levels, as shown in Fig. 1.

Manufacturing and control (Level 3) provides monitoring and control functions (e.g., quality checks, sequence of events recording, and alarm monitoring) by allowing operators to interact with the ICS using human-machine interfaces (HMIs).

The area supervisory control (Level 2) focuses on specific parts of the ICS and provides monitoring and management through HMIs, programmable logic controllers (PLCs), and other systems.

The basic control (Level 1) includes equipment and systems that directly interact with the physical processes of the ICS. Devices like PLCs and protective relays are normally found in this level.

The actual physical processes are executed at Level 0. Devices in this level send analog data to devices in higher levels; the information is subsequently processed to determine the necessary controls to be issued to the devices in Level 0. Examples of these devices are circuit breakers and measurement transformers.

The Purdue Model shows that each level of the ICS has different performance and security requirements. Considering these factors, it makes sense to adopt a multilayer security approach, where the security of the ICS relies on layer upon layer of different controls specifically designed to cope with the requirements and limitations of each part of the system.

The Purdue Model has been adapted by the United States Department of Homeland Security in conjunction with some research organizations to create a flexible defense-in-depth model to secure ICSs without affecting their performance.

The defense-in-depth model divides the ICS into seven different levels, each one with specific security and technical requirements and limitations. Reference [10] describes using standards to segregate the components of an ECS as a specialized ICS. The ECS defense-in-depth strategy is illustrated in Fig. 2.

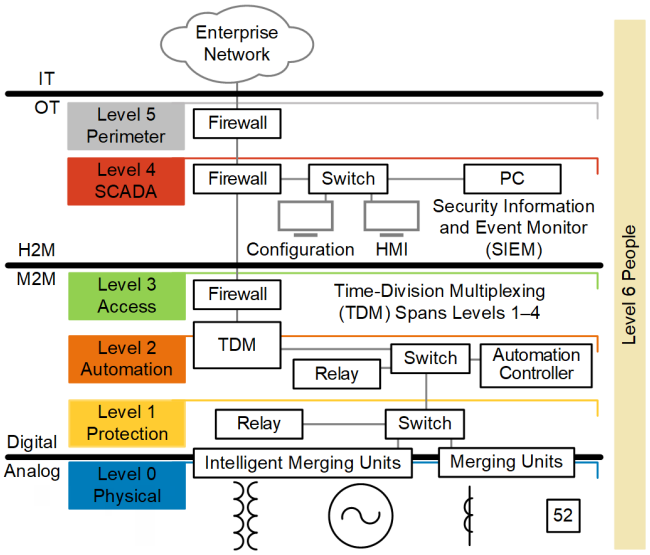

Fig. 2. Defense-in-Depth Levels Diagram

A simplified description of each level is provided as follows:

- Level 0: Digital and analog data are sent to higher levels while controls are received to ensure the system is safe and stable. Physical security controls are required in this level, including closed-circuit television, physical barriers, and alarms.
- Level 1: The information from Level 0 is processed in this level to determine the necessary controls to issue. The integrity of the devices may be determined by baselining and periodic baseline verifications.
- Level 2: Monitoring, automation processes, and further controls reside in this level. Communication filtering and processing in this level may prevent certain attacks, like denial of service.
- Level 3: This level provides internal segmentation to the ICS, separating the machine-to-machine levels from the human-to-machine levels. This level ensures that only authorized communications are exchanged between upper and lower levels.

- Level 4: Data are concentrated in this level for analysis and monitoring. Encryption may be used to reduce the risks introduced by general purpose devices.
- Level 5: This level provides physical and logical separation between the ICS and the enterprise network. Physical security controls may be implemented, as well as virtual private networks to provide confidentiality, integrity, and encryption.
- Level 6: This level is technically not part of the ICS. It includes policies, procedures, risk analysis, and other human-based tools used to secure the ICS.

The focus of this paper is cyber resilience and the security of the ECS to maintain availability to protect and control the EDS. Security is critical at all levels of a control system, but each level may have a different focus for its security, as represented in Fig. 3 [10].

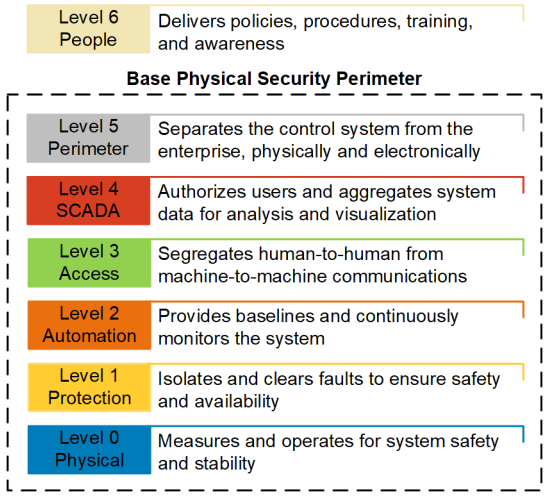

Fig. 3. Security Focus of Each Defense-in-Depth Level

IV. Counteracting and Compensating Technologies for Communication Network Faults

Whenever operation depends on the correct function of the automation network, it may become necessary to increase the availability of the network through counteraction or compensation. Counteraction in the ECS is the act of adding technology to a system component to nullify the effects of some previous choice. The simplest and least expensive way to increase ECS availability and reduce maintenance is to use components with a demonstrated low failure rate. As an alternative, the IEC 62439 standard considers the use of IT components with a high failure rate in the OT ECS and counteracting this choice with communications protocols that introduce redundancy.

Compensation in the ECS is the act of adding technology to one system component to intentionally avoid the adverse effects it would have on other components. Robust communications security systems are frequently patched or updated and use significant processing and memory to provide popular methods of obscuring information and preventing intrusion. The simplest and least expensive way to increase availability and reduce maintenance to the ECS is to compensate by adding a firewall with robust security that shields the protective relays and other IEDs.

The preferred IEC 62439 method for creating high-availability communications networks, described in IEC 62439 Part 1, is resiliency via recoverability, whereby faults are detected and isolated, and network traffic is rerouted without human interaction. After system reliability is reduced by a failure, the IEEE 802.1w Spanning Tree Algorithm (STA) can detect the failure and react to return the system to a greater level of reliability. Rapid STA (RSTA) deploys the same logic processing but with more efficient wait times and state transitions in order to be appropriate for OT. Resiliency is measured by the speed with which the system reestablishes communications after the communications fault is detected and isolated. Correctly implemented OT-class RSTA networks recover after a momentary outage of less than 16 milliseconds and return the system to an N-1 state. However, IT-class RSTA behavior is usually far too slow for OT resiliency needs.

The alternative IEC 62439 counteraction methods of Parallel Redundancy Protocol (PRP) and High-Availability Seamless Redundancy (HSR), described in IEC 62439 Part 3, are defined as repairable and thus provide no resiliency. These protocols were developed for industrial processes with permanent human staff to detect and correct communications failures. These repairable methods act like a fuse and create a sustained outage. The failure is persistent because the methods have no detection or correction intelligence and human intervention is required to detect and repair faults. These methods result in sustained outages of indefinite duration and reduce the secondary system to an N-0 state.

In lieu of a method to detect PRP or HSR failures, some technology providers have invented methods to automatically detect the loss of a link that the protocols are active on. However, these methods have become proprietary and are not interoperable due to the lack of a standardized method. Manual IT methods exist to poll network devices, learn link statuses, and then manually react to Ethernet faults. However, these methods require that detection be successful and followed by human intervention to poll for statuses, plan mitigations, and manually change settings to implement the mitigation. Each of these failures will result in a sustained outage of infinite duration and cause the EDS to be in an N-0 state.

HSR is not appropriate for ECSs due to its poor performance and lack of resiliency, but the largest vulnerability is the fact that it is not interoperable with Ethernet.

PRP systems can be improved to mitigate the vulnerabilities of the repairable design by combining PRP with IEEE 802.1w RSTA as prescribed in IEC 62439 Part 1. Correctly implemented RSTA OT networks will recover after a momentary outage of less than 16 milliseconds and return the system to an N-1 state.

OT-based software-defined networking (SDN) is a packet-switching technology that gives unprecedented control over network traffic and failover speeds. Instead of counteracting poor design choices, OT SDN works autonomously or provides compensation techniques to existing technologies to increase the reliability and security of the overall system [11].

V. NIST Threat Sources and Examples

NIST describes a threat source as "the intent and method targeted at the intentional exploitation of a vulnerability or a situation and method that may accidentally trigger a vulnerability," [12]. Categories include natural, technical, operational, environmental, human, and physical threats.

Recent publicized EDS failures provide specific examples of failures due to each threat category.

- Natural: Hurricane Maria caused dramatic and long-term EDS outages [13].
- Technical: Faulty ECS equipment contributed to a historic blackout during Hurricane Maria [14].
- Operational: A lack of mitigation planning created chaos in Puerto Rico after Hurricane Maria [15].
- Environmental: After lines were repaired following Hurricane Maria, falling tree causes another wide-spread blackout [16].
- Human: A nonmalicious insider (a utility test technician) caused the 2008 Florida blackout [17]. The nonmalicious outsider that caused the 2013 Super Bowl blackout was a contracted commissioning technician [18]. Malicious insider energy traders created congestion for profit, which caused a blackout when a forest fire caused a line to go out of service [19]. Malicious outsiders caused outages in the Ukraine in 2015 and 2016 [20].
- Physical: On February 15, 2019 a Transport Layer Security (TLS) 1.3 vulnerability was discovered that enabled hackers to eavesdrop on encrypted traffic [21]; before the guidelines were published, the vulnerability in the recommended version of TLS had been weaponized.

TLS is a cryptographic protocol used for internet communications and online transactions. Like other technologies, this cryptography method is "perishable" and must be replaced when new processors make it obsolete or when a vulnerability is found. Unfortunately, some IT designers promote it for use in OT devices. TLS is an example of the unintended consequences of an inappropriate deployment of technology creating the need to physically modify an in-service device. In March 2018, TLS 1.3 was finalized and the Internet Engineering Task Force (IETF) published it as RFC 8446 in August 2018. In December 2018, the comment period closed on NIST Special Publication 800-52 Revision 2 [22]. This documents states that all government TLS servers and clients must upgrade to TLS 1.3 by January 1, 2024. As noted, in February 2019 a TLS 1.3 vulnerability was weaponized by hackers to eavesdrop on encrypted traffic before the guidelines were even published [21], thus requiring the removal of each affected device from service for repair.

A second simple physical threat is the forget-and-flood feature in every Ethernet switch chip. When a switch does not know the destination of a received frame, it floods, or sends the frame to all ports except the port it was received on. Malicious and nonmalicious uses of this feature may physically prohibit the delivery of ECS protection messages via sustained bandwidth saturation.

VI. IEC 62443 Defense in Depth

According to Reference [23], "The ISA/IEC 62443 series of standards, developed by the ISA99 committee as American National Standards and adopted globally by the International Electrotechnical Commission (IEC), is designed to provide a flexible framework to address and mitigate current and future security vulnerabilities."

The NIST Risk Management Framework (RMF) and Cybersecurity Framework (CSF) itemize controls to implement a program and reference international standards, such as IEC 62351 and IEC 62443, that provide details. The RMF core defines five main functions: identify, protect, detect, respond, and recover. Coincidentally, these are the same steps used to manage the EDS. The RMF and CSF provide actionable information to choose the correct implementation of the related technical standards.

ISA/IEC 62443 Part 3-3: System Security Requirements and Security Levels provides detailed technical control system requirements and defines the requirements for control system capability security levels. These levels reflect the Purdue Model levels for device function and capability to describe the appropriate security technologies to be deployed in devices at each level.

ISA/IEC 62443 Part 4-1-2018: Secure Product Development Lifecycle Requirements defines a secure product development lifecycle. This lifecycle includes security requirements definitions, secure design, secure implementation (including coding guidelines), verification and validation, defect management, patch management, and product end-of-life guidelines.

An appropriate use of IEC 62443 Part 3-3 and Part 4-1 illustrates where safety and security technologies should be deployed among OT devices to maximize impact and reduce unintended vulnerabilities.

VII. Unintended Consequences of IEC 62443 Part 4-2 Defense in Breadth

Defense in breadth is the concept of putting every possible security feature in every single device, ignoring the appropriate deployment based on the Purdue Model levels. This strategy is often promoted for devices with very short lifespans deployed in locations that cannot be protected by a defense-in-depth system. An example is a remote wireless sensor that publishes data but has no control with an end of life that coincides with the expiration of the internal encryption. Data encryption converts data so that it can only be read by people or devices with access to a secret decryption key or password. This creates the challenge of making sure that all clients speaking to a single OT device data server are updated to the same patch version at the same time, and that all of the OT data servers talking to each OT client are similarly updated with the same patch version simultaneously in order to maintain communications. In addition to the challenge of patching firmware when the encryption expires, malware hidden inside encrypted traffic cannot be seen or stopped by most security technologies. This is not appropriate for the millions of devices deployed in ICS and ECS systems.

As an example, consider the earlier example of TLS 1.3. When considering the defect management portion of an OT device lifecycle, patch management creates a vulnerability and a high cost to the OT ECS. Laptops and smart phones are often upgraded automatically or on demand without consideration of a loss of use. OT devices require preplanned removal from service, which requires a planned EDS outage, ECS outage, or both. Also, personnel need to be onsite to apply and test the new TLS patch. Therefore, encryption should be deployed in a Level 3 device (shown as a firewall in Fig. 2) that shields the Level 2 and Level 1 devices on the protected LAN.

The cost of patch activity for utilities can be around 5,000 euros per ECS device, including several hours of personnel time to travel, patch, and test. Additional expenses are incurred if an EDS outage is also required to safely remove the ECS device from service.

Recent activities for industrial IT and internet of things (IoT) devices have prompted IEC 62443 Part 4-2 to promote deploying Level 3 security features in Level 1 and 2 devices. While this may be acceptable for IoT devices deployed outside a firewall and with a two-year lifecycle, this is very problematic for OT devices. Thus, NIST RMF and CSF strategies recommend IEC 62443 Part 3 defense in depth and not Part 4 endpoint security, like TLS. Instead of a strength, these security technologies become a very large vulnerability when deployed in Level 1 and 2 devices. Each change forces an unwanted outage to patch the security firmware. A typical large utility with 12,000 OT devices would be faced with over 1,000 days of effort at a cost of 60M euros.

If TLS is deployed in a Level 3 device, like the firewall in Fig. 2, none of the protection and automation devices need to be patched. Only one device per substation needs be patched. This firewall device can be safely removed from service while personnel are onsite without affecting the safety and protection of the EDS.

Security is not a goal that can be met but is rather an ongoing process. New vulnerabilities are discovered every day, existing threats evolve, and people make mistakes. Recent activities illustrate that newly added security technologies can become attack vectors to otherwise isolated OT systems

OT designs must prevent attacks and also anticipate that they will happen nonetheless. Attackers are often more skilled and motivated than defenders. IEC 62443 Part 3 explains defense-in-depth methods to compartmentalize devices and minimize what needs to be defended. This also minimizes the loss when a device is compromised.

VIII. Case Study of Large Utility's Evaluation of OT Device Access Control

Recently, a large European utility performed a thorough evaluation of the available international standards (IEC 62351, IEC 62443, and IEEE 1686) for ICS and ECS cybersecurity. These documents address the malicious human insider and outsider threats to the system. However, they do not adequately address nonmalicious human threats. More importantly, they do not address the other five threat categories at all. Based on these concerns, the utility IT and OT staff decided to adopt IEC 62443 Part 3 defense in depth, with security controls distributed among the six levels of the ECS. Due to the frequent disruption of OT devices caused by localized encryption and authentication, they decided not to adopt IEC 62443 Part 4-2. The following summary of their evaluation process presents a thorough description of the use of defense-in-depth security.

For system security, access control includes the authentication, authorization, and audit of an entity (subject) needing a resource (object), such as a settings file [24]. An access control list (ACL) contains a list of subjects, objects, and permissions. An ACL is also a resource under access control that can be viewed by certain subjects and modified by a subset thereof. For example, an administrator can use the ACL to set privileges as to who can access what resources, the time they can be accessed, and to what level of access [24].

Access control technology and architectures are congruent with IT enterprise networks and devices but are not congruent with ICS, ECS, or OT networks. IT system differences make these access control architectures unsuited to OT or control system networks. IT systems include devices that are accessed and updated on a regular basis (e.g., corporate servers, network appliances, laptops, and desktop computers). IT networks are also dynamic in operation and reactive in terms of user needs for resources and security. However, control system networks at many organizations today contain ten times as many devices (and growing) as their respective enterprise networks. These controls systems ideally communicate on closed restricted networks.

Administration of OT devices is different than that of IT devices in that they are commissioned and then updated on, at most, a yearly basis. Control system networks have access restricted to fewer individuals on infrequent and controlled intervals. Where IT networks are centered on information confidentiality, OT networks are centered on information availability.

As an example, a utility with approximately 1,500 employees will have approximately twice the number of computers, phones, and similar devices assigned to those individuals as part of the organization's IT enterprise system. Table I depicts the number of OT devices and IT appliances such as routers, firewalls, servers, and other Ethernet packet-forwarding devices in their respective OT and IT systems. While users constantly interact with their human-to-machine devices, access to the more numerous OT devices is ideally never and at most less than once a year.

Table I illustrates that the number of employees out of the 1,500 that can access the OT devices directly is a very small percentage of the organization's employees. Only about half that number are allowed access to the even greater number of utility substation assets (i.e., the relays).

TABLE I
TYPICAL USER ACCESS FOR IT AND OT SYSTEMS (FOR 1,500 TOTAL USERS)

Type	Cyber Asset Type	User Access	Total Assets	Users With Access (out of 1,500)
IT	Computers and phones	Constant	3,000	1,500
	IT appliances	Once per week	750	25
OT	All OT devices	< Once per year	12,000	40
	Relays (subset of OT devices)	< Once per year	8,000	20

Continuous access is neither necessary nor acceptable in OT networks. Human access to devices in an OT network is typically a scheduled event, planned and assessed for system impact. Unscheduled authorized human access to devices in an OT network is for emergency or mitigation events, which typically require greater scrutiny and analysis after human interaction with the ECS. For this reason, a different and simpler access control architecture that aligns with the defense-in-depth approach is necessary.

While users in IT systems accept and expect daily access to resources and systems, there is a tendency to work around or minimize the effectiveness of the access control mechanisms for OT devices and networks. The reasons for this are as follows:

- Access controls and credentials in OT networks are typically in addition to the enterprise network credentials that users must first present.
- OT access control systems are, in some cases, segregated from enterprise access control systems for incompatibility reasons and/or because they should not be integrated for best-practice security reasons.
- Because users rarely need or use their OT credentials, they often forget them.
- Because users rarely need or use OT credentials, shared access control credentials are incorporated, creating an issue for repudiation and revocation.
- Many OT sites are remote locations with no enterprise network access.
- Encrypted communications can conceal malicious or malformed messages and hinder detection and reaction.

An access control architecture for OT must provide simple-to-use security that does not hinder the intended OT applications. Effective OT security cannot rely on unwarranted, unnecessary, and often misunderstood trust in the most targeted and compromised organization asset, the human user. Access control for OT must ensure the following capabilities:

- Single-sign-on user access.
- Multifactor authentication without IEDs needing extra complexity and code.
- A focus on organizational trust rather than individual trust.
- Representation of IEDs (relays) as a system resource rather than an interactive device.
- Human-to-machine authentication, authorization, and accountability on OT devices.
- Machine-to-machine authentication and authorization for whitelisting, baselining, and monitoring communications between OT devices.
- Time-to-live accessibility (no 24/7 access availability).
- No need to manage roles and users for IEDs (access granted as planned or necessary).
- No complicated directory service query protocols such as Remote Authentication Dial-In User Service (RADIUS) or Lightweight Directory Access Protocol (LDAP) in IEDs.
- Certificate- or token-based access to minimize number of secrets on IEDs.
- Interoperability with existing central management servers, such as Active Directory, Citrix, HP, Google, or Oracle.
- Interoperability with configuration and monitoring software.
- Mutual authentication between objects and subjects.
- Protection against eavesdropping and replay attacks.
- Simplification and minimized number of managed credentials and public keys (necessary for revocation or rotation).
- System support of role-based access control, with authentication, authorization, and accountability independent of directory service protocol.
- IEDs that can support any user management system and protocol—past, present, or future—without implementing and maintaining a single protocol. Example protocols include RADIUS, LDAP, TACACS+, OpenOTP, and 2FA.

IX. CONCLUSION

The multilayer approach of defense in depth allows asset owners to implement the correct security controls in each layer of the ICS without degrading its performance. It allows the use of common standard protocols and it protects the ICS from internal and external attackers without hiding or obscuring the network. For all these reasons, defense in depth is the correct approach to properly securing modern ICSs against malicious and nonmalicious cyber attacks.

IEC 62443 Part 3 provides an appropriate and useful defense-in-depth strategy for OT networks. Based on work in the ISA99 and Purdue models for ICSs, the defense-in-depth strategy provides levels of appropriate security and prevents insecurity. Such insecurity is often the byproduct of unintended consequences resulting from vulnerabilities such as frequent

firmware patches in protective relays that have internal encryption and TLS based on IEC 62443 Part 4-2.

An access control architecture suited for OT must align and complement the defense-in-depth security approach. The access control architecture must be simple in both implementation and usability to realize the required capabilities described in this paper. Simplicity and security are achieved by removing the trust in the most targeted and compromised organization asset, the human user, and placing it solely with the organization. OT resource permissions in this new architecture should be nonpersistent and provided for only a limited window of time by a second party, taking separation of duty controls in account.

X. References

[1] North American Electric Reliability Corporation, "Transmission Availability Data System Definitions," January 2013. Available: https://www.nerc.com/.

[2] IEC 60834, Teleprotection Equipment of Power Systems—Performance and Testing, 1999.

[3] IEC 61850, Communication Networks and Systems for Power Utility Automation, 2019.

[4] IEEE Standard 1646-2004, IEEE Standard Communication Delivery Time Performance Requirements for Electric Power Substation Automation.

[5] CIGRE JWG 34/35.11, "Protection Using Telecommunications," 2001.

[6] IEC 62439, Industrial Communication Networks—High Availability Automation Networks, 2016.

[7] J. Brodsky, "What Was That Purdue Model Stuff, Anyway?" *SCADASEC Magazine*, March 2018. Available: http://scadamag.infracritical.com/index.php/2018/03/01/purdue-model-history/.

[8] The International Society of Automation, "ISA99, Industrial Automation and Control Systems Security." Available: https://www.isa.org/templates/two-column.aspx?pageid=124560.

[9] IEC 62443, Industrial Communication Networks—Network and System Security, 2009.

[10] J. Smith, N. Kipp, D. Gammel, and T. Watkins, "Defense-in-Depth Security for Industrial Control Systems," proceedings of the EEA Conference & Exhibition, Wellington, New Zealand, June 2016.

[11] D. J. Dolezilek, "Using Software-Defined Network Technology to Precisely and Reliably Transport Process Bus Ethernet Messages," proceedings of the 14th International Conference on Developments in Power System Protection, Belfast, UK, March 2018.

[12] Computer Security Resource Center, "Threat Source," *National Institute of Standards and Technology*. Available: https://csrc.nist.gov/glossary/term/threat-source.

[13] R. Bernal, "Hurricane Maria Knocks Out Power in All of Puerto Rico," *The Hill*, September 2017. Available: http://thehill.com/latino/351583-hurricane-maria-knocks-out-power-in-all-of-puerto-rico.

[14] E. L. Vélez, "Lack of Maintenance, Faulty Equipment Caused Historic Blackout," *Caribbean Business*, February 2017. Available: http://caribbeanbusiness.com/lack-of-maintenance-faulty-equipment-caused-historic-blackout/.

[15] F. Robles and D. Acosta, "Puerto Rico Cancels Whitefish Energy Contract to Rebuild Power Lines," *The New York Times*, October 2017. Available: https://www.nytimes.com/2017/10/29/us/whitefish-cancel-puerto-rico.html.

[16] N. Acevedo, "Puerto Rico: Single Fallen Tree on Power Line Leaves 900K Without Power," *NBC News*, April 2018. Available: https://www.nbcnews.com/storyline/puerto-rico-crisis/puerto-rico-fallen-tree-power-line-leaves-900k-without-power-n865506.

[17] K. Semple, "Florida Blackouts Affect One Million Across State," *The New York Times*, February 2008. Available: https://www.nytimes.com/2008/02/27/us/27florida.html.

[18] NOLA.com, "Super Bowl 2013 Blackout Caused by Faulty Relay Equipment, Entergy Says," *NOLA Media Group*, February 2013. Available: http://www.nola.com/superbowl/index.ssf/2013/02/super_bowl_blackout_caused_by.html.

[19] J. Roberts, "Enron Traders Caught on Tape," *CBS News*, June 2004. Available: https://www.cbsnews.com/news/enron-traders-caught-on-tape/.

[20] D. Goodin, "Found: 'Crash Override' Malware That Triggered Ukrainian Power Outage," *Ars Technica*, June 2017. Available: https://arstechnica.com/information-technology/2017/06/crash-override-malware-may-sabotage-electric-grids-but-its-no-stuxnet/.

[21] R. Millman, "TLS 1.3 Vulnerability Enables Hackers to Eavesdrop on Encrypted Traffic," *SC Media UK*, February 2019. Available: https://www.scmagazineuk.com/tls-13-vulnerability-enables-hackers-eavesdrop-encrypted-traffic/article/1525916.

[22] K. McKay and D. Cooper, "(DRAFT) NIST Special Publication 800-52 Revision 2: Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations," *National Institute of Standards and Technology*, November 2017. Available: https://csrc.nist.gov/.

[23] Automation.com, "ISA Announces ISA/IEC 62443-4-2-2018 Standard," *International Society of Automation*, September 2018. Available: https://www.automation.com/library/resources/isa-announces-isaiec-62443-4-2-2018-standard.

[24] Technopedia, "Access Control," 2019. Available: https://www.techopedia.com/definition/5831/access-control.

XI. Biographies

David Dolezilek is the international technical director at Schweitzer Engineering Laboratories, Inc. and has three decades of experience in electric power protection, automation, communication, and control. He leads a team that develops and implements innovative solutions to intricate power system challenges and teaches numerous topics as adjunct faculty. David is a patented inventor, has authored dozens of technical papers, and continues to research first principles of mission-critical technologies. Through his work, he has created methods to specify, design, and measure service level specifications for digital communication of signals, including class, source, destination, bandwidth, speed, latency, jitter, and acceptable loss. As a result, he helped coin the term operational technology (OT) to explain the difference in performance and security requirements of Ethernet for mission-critical applications versus IT applications. David is a founding member of the DNP3 Technical Committee (IEEE 1815), a founding member of UCA2, and a founding member of both IEC 61850 Technical Committee 57 and IEC 62351 for security. He is a member of the IEEE, the IEEE Reliability Society, and several CIGRE working groups.

Dennis Gammel is a graduate of the University of Idaho with a BS in Applied Mathematics and has been actively working in the computing and communications industries since 1996. His career experience includes network security design, CS network architecture, embedded product development, ASIC simulation, and firmware design with RTOS application development. Mr. Gammel is presently a research and development director at Schweitzer Engineering Laboratories, Inc. (SEL), responsible for security technology designed for and implemented in SEL product lines. He has been with SEL since March 2005 and carries with him over 20 years of secure firmware and network engineering experience.

William Fernandes is an application engineer with Schweitzer Engineering Laboratories, Inc. He earned a master's degree from La Sapienza, Universita di Roma.