Hidden in Plain Sight: Anticipating and Avoiding Hidden Failures in Communications-Assisted Protection

Allan McDonald
Aerospace Consultant

Anna Dolezilek
Role-Based Certification Consultant

David Dolezilek
Schweitzer Engineering Laboratories, Inc.

Presented at the
47th Annual Western Protective Relay Conference
Virtual Format
October 20–22, 2020

Hidden in Plain Sight: Anticipating and Avoiding Hidden Failures in Communications-Assisted Protection

Allan McDonald, *Aerospace Consultant*
Anna Dolezilek, *Role-Based Certification Consultant*
David Dolezilek, *Schweitzer Engineering Laboratories, Inc.*

***Abstract*—Before digital trip circuits even became a trend, the Electric Reliability Organization (ERO) documented human error as a root cause or contributing cause of many annual misoperations in traditional protection systems. The three most common failure modes are expected to become even more prevalent with the inclusion of multiple time-synchronized intelligent devices and the use of Ethernet communications within the trip circuit path.**

As utilities modify their use of Ethernet from station bus operator control commands to perform process bus automatic trip commands, energy control system designs must change to be much more resilient and fault-tolerant. An important part of this is recognizing the appropriateness of technology and the true severity of failure. This paper demonstrates the importance of this by drawing parallels between the roles and responsibilities of the ECS design team and real-world examples of severe technological failures in the aerospace, aviation, and healthcare industries. Deadly, catastrophic failures resulting from the inappropriate use of technology by subject matter experts include the space shuttle *Challenger* O-ring and *Columbia* insulative foam and the Boeing 737 MAX angle of attack sensor. Deadly, catastrophic failures resulting in the incorrect categorization of failure severity include the heat of the *Columbia*'s wing upon reentry and the strength of the 737 MAX override of pilot control. While the design failures of the *Challenger* disaster and 737 MAX illustrate misuse and abuse of failure condition analysis, the *Columbia* and 737 MAX disasters also illustrate a breach of the public trust. NASA disregarded likely successful repair or rescue, and the FAA approved the 737 MAX for continued service.

A challenge to digital trip circuit design exists when designers make the dangerous assumption that station bus and process bus communications require the same level of service. This paper emphasizes that public safety demands that designers perform appropriate failure analysis and require resilience appropriate to the severity of a station bus or process bus failure. Process bus communications require completely different service level agreements and key process indicators to understand digital trip circuit communications in real time. For example, the hidden failure modes of IEC 62439-3 Parallel Redundancy Protocol may not adversely affect commanded control on the station bus but will jeopardize peer-to-peer trip commands on the process bus that need to meet NERC N – 1, N – 1 – 1, or N – 2 resilience.

Carefully and appropriately adopting new technology represents a positive change that will expand the knowledge of engineers and technicians as they learn new skills. Better, more fully thought-out decisions prevent small faults from cascading into larger failures, ultimately leading to better systems.

I. Introduction

The role of the electric power energy delivery system (EDS) is to deliver electricity through the power grid to all points of consumption. EDS components are collectively called the primary system and include components within the energy delivery process, including generators, transmission lines, and circuit breakers.

According to the National Institute of Standards and Technology (NIST), a more resilient EDS is one that is better able to sustain and recover from adverse events like severe weather. A more reliable EDS is one with fewer and shorter power interruptions. Resilience is defined by Presidential Policy Directive 21: Critical Infrastructure Security and Resilience (PPD-21) as the "ability to prepare for and adapt to changing conditions and withstand and recover rapidly from disruptions" [1]. An adequately resilient EDS should be designed with this breadth of resilience in mind so that it can withstand and recover from a variety of potential threats, such as accidents, adverse natural conditions, and deliberate attacks. Therefore, the methods of monitoring and controlling the EDS need to be resilient as well.

The EDS is monitored, protected, and controlled by the energy control system (ECS), aptly termed the secondary system. ECS components consist of modern technologies such as intelligent electronic devices (IEDs), controllers, and communications devices. These components are designed to detect and mitigate faults in the EDS.

ECS devices exchange protection, interlocking, command and control, and engineering access signals via digital messages on one or more communications networks. These networks are designed using various topologies, including direct, multidrop, star, and ring, to support protocol message and device interface requirements. Care must be taken in network design to ensure appropriate messaging and topology is dedicated to detecting and mitigating faults, as well as responding to operator commands and requests for fault records.

The terms "station bus" and "process bus" were coined decades ago when multidrop bus topologies were popular in industrial control system (ICS) designs. Physical bus topologies do not provide the resilience and availability required for an

ECS, but the terminology persists as a way to describe different but overlapping groups of communications applications used in power system IEDs.

Protocols that send operator control commands and transmit and receive system information use human-to-machine (H2M) connections to networked IEDs on the station bus. Engineering access, metering, and monitoring, as well as supervisory control and data acquisition (SCADA), are accomplished by means of automated and human-activated client-server communications.

Process bus communications are machine-to-machine (M2M) connections and protocols that exchange input/output (I/O) process information between IEDs and process instrumentation and control devices, including data-acquisition devices, instrument transformers, and controllers [2]. Process bus communications systems use information in digital messages passed among intelligent devices to replace low-level energy over copper wires and send raw analog samples, status, alarms, and trip and control signals.

The North American Electric Reliability Corporation (NERC) is an international regulatory authority that works to improve the reliability and security of the EDS. NERC promotes the development and enforcement of reliability standards to support EDS quality requirements [3]. One of these standards, titled System Performance Under Normal Conditions (TLP-001-1), discusses system reliability following the loss of a single or several EDS components. As designated in TLP-001-1, continued performance after loss of a single component is known as an N – 1 contingency, continued performance after loss of two components *consecutively* is N – 1 – 1, and continued performance after loss of two components *simultaneously* is N – 2 [4].

North American EDS system dependability now mandates a momentary outage duration of less than one minute. To minimize the duration of an outage, the ECS that controls the EDS must quickly and reliably initiate use of hot-standby primary equipment. This means that ECS components must also meet N – 1, N – 1 – 1, or N – 2 contingency standards because the corresponding momentary ECS outage duration must be much shorter. Any outage of the ECS communications system must be 15 milliseconds or less so that it can isolate faulted primary equipment and automatically dispatch replacement equipment. Ensuring resilience of each means that faulted equipment in the EDS needs to be restored to service as soon as possible. However, to be prepared for subsequent events, the faulted ECS component must first be restored to service in the same 15 milliseconds. Perhaps even more important than designing for fault avoidance and fault recovery, though, is designing for awareness and alarming of faulted components. Without this awareness, failures remain hidden, and otherwise avoidable initiating events can result in cascading failures.

In this paper, we discuss the importance of the various roles of members of the digital process bus ECS design team and emphasize their accountability to anticipate and avoid hidden failures.

II. Roles and Responsibilities of the ECS Design Team

Critical systems protect the safety of people directly (e.g., by controlling air traffic) and indirectly (e.g., by controlling stock trading). In most cases, these critical systems rely on a constant and reliable source of electric power from a local EDS or a utility intertie. The role of the ECS is to anticipate and mitigate system faults to ensure critical and noncritical systems alike have access to reliable electric energy.

Inherent failures are unintended mechanical, electrical, or other functional failures of system components to perform their intended services, either due to natural causes or errors in planning and execution of design and manufacture. It is not possible to completely remove all inherent natural and human-caused threats to performance, and no design can eliminate all vulnerabilities. The responsibility of the ECS design team is thus to understand that inherent failures that are both natural (e.g., lightning) and human-caused (e.g., premature mechanical failure under stress) must be quantified, anticipated, and mitigated.

Extraneous failures occur when system components fail to perform their intended services due to intentional human-caused events that induce component failure. Extraneous failures can be caused by legitimate actions made in error but are primarily caused by malicious actions, such as terrorist attacks. Whether the intent of the action is legitimate or malicious, the effects of the extraneous failure on the system are the same. The ECS design team must anticipate threats that can be thwarted and threats that are not preventable, while also recognizing that they cannot anticipate all threats. For both inherent and extraneous threats, the design team is responsible for building detection, isolation, and recovery into a system that is also redundant and resilient, while meeting or exceeding the determined level of availability.

The ECS design team is also responsible for understanding and supporting all categories of ECS operations and resilience, including process bus designs or combinations of station bus, interlocking, and engineering access designs. This paper addresses concerns about unintentional and dangerous mistakes hidden in plain sight when ECS design teams implement digital trip circuits based on IEC 61850 process bus protocols.

According to the Relay Trip Circuit Design Working Group of the IEEE Power System Relaying Committee, the trip circuit supervises trip coils and operates circuit breakers and switches at high speed. Examples include trip circuit improved security via a breaker failure function and immediate retrip of the protected breaker and improved resilience via a separate breaker failure relay and alternate trip coils. Digital secondary systems that replace part of the trip circuit copper wiring with communications cables must satisfy fault avoidance and tolerance for inherent conditions, including fluctuations in temperature, wind, ice, rain, snow, electrical storms, humidity, altitude, and earthquakes, for both wiring and cables [5]. The role of the ECS design team is to mitigate all these physical and environmental concerns, as well as anticipate new concerns that

may threaten digital process communications, such as cyberattacks and other extraneous threats that can affect digital messaging and information sharing.

The ECS design team is responsible for thoroughly understanding the criticality of the system, the effectiveness of their design, and the impact that their design choices have on the EDS. The ECS design team and ECS owner are also responsible for deciding which vulnerabilities are acceptable and which must be mitigated based on cost, schedule, and performance. Then, after minimizing risk, the ECS design team must make an informed decision about acceptable vulnerabilities and associated risk in the final design. This decision must be well understood by all members of the team and well documented.

III. NERC Failure Avoidance, Tolerance, and Acceptance

The simplest concept of fault tolerance is to design a system in which the failure of any single component would not cause the system to stop performing its intended function. This is referred to as N – 1 tolerance, where *N* represents the total number of components. Critical EDS networks that require N – 1 availability need an ECS that will immediately detect a failure of any single critical primary system component and automatically mitigate that failure to maintain the intended flow of energy to all points of consumption.

Since it is critical that the trip circuit perform its function of detecting and isolating a fault, it is the responsibility of the process bus ECS design team to understand whether the system requires NERC N – 1, N – 1 – 1, or N – 2 performance, or other appropriate requirements, and the impact that their design choices have on this performance. Further, it is essential that the design team understands the application and behavior of the chosen technology to prevent a small and seemingly insignificant failure from creating a chain reaction that could lead to a critical failure. An undetected failure of an ECS communications system component can defeat an EDS N – 1 design and cause an outage or can reduce N – 1 – 1 and N – 2 designs to N – 1 capability, without any operator being aware of this potentially dangerous change.

EDS and ECS fault tolerance are maximized with dual-primary resilient designs that recover quickly from a fault. Dual-primary *diverse* resilient ECS designs offer the greatest performance statistically because they use different technologies to avoid common mode failure. However, some ECS owners avoid diverse systems to minimize training and parts needed to maintain the system, which is elevated when operating different technologies. Statistically, the next best performance is achieved by dual-primary redundant resilient systems, followed by dual-primary redundant systems without resilience. Less effective fault-tolerant designs include a design that uses one IED and one controller with dual-trip circuit communications systems between them. This paper focuses on this choice because of its common usage and delves into its many hidden failure modes.

A. Applying IEC 62439 to ECS Design

Within the scope of the trip circuit communications, international standard IEC 62439-1 describes numerous technologies to improve the availability of Ethernet communications via redundancy and replication [6].

ECS design teams that choose a single group of IEDs and controllers and focus only on the communications system are responsible for understanding and mitigating the associated technical challenges. Indeed, implementing a nondiverse system dramatically reduces hardware reliability and increases the risk of inherent failure modes. Specifically, having Ethernet local area networks (LANs) made of collections of the same information technology (IT) communications devices—although this may be an easy solution to sketch without much thought of performance—dramatically reduces ECS reliability because of both inherent failure modes of devices with a lower mean time between failure and nondiversity of system design. ECS design teams should use specific operational technology (OT) Ethernet communications devices, IEDs, and controllers for appropriate reliability of the design. Although improving communications availability is essential to improving the reliability of communicating process bus trip signals, it may not be enough to provide appropriate ECS resilience to match the criticality of the EDS system, and a dual-primary design may be needed.

Section 5.1.1 of IEC 62439-1 addresses the resilience required in case of a failure in a high-availability automation network. It describes how industrial systems (e.g., ECSs) rely on network recovery time being shorter than what is known as the *grace time*: the duration of time a system can tolerate degradation. Methods differ on how to handle resilience, but their key performance factor is the *recovery time*: the duration of time necessary to restore operation after a disruption. The preferred IEC 62439-1 method for creating high-availability communications networks is resilience through recoverability, whereby faults are detected and isolated, and network traffic is rerouted without human interaction. This is accomplished by rapid resolution of the Rapid Spanning Tree Algorithm (RSTA)—as outlined in IEEE 802.1w, Rapid Reconfiguration of Spanning Tree—or software-defined networking (SDN), which detects and compensates for Ethernet faults in real time [6]. Resilience is measured by the speed at which the system reestablishes communications after the communications fault is detected and isolated. For an ECS, the acceptable grace time is less than 16 milliseconds for an EDS grace time of 1 minute [7].

The alternative counteraction methods of Parallel Redundancy Protocol (PRP) and High-Availability Seamless Redundancy (HSR) outlined in IEC 62439-3 are only defined as *repairable* and do not provide true resilience [7]. These replication protocols were developed for industrial processes where permanent human staff are available to detect and correct communications failures. These repairable methods act like a fuse, and if they fail and remain undetected, they cause a sustained outage until manually corrected. Because these

methods have no detection or compensation technology, any failure is indefinitely persistent.

Though not directly mentioned in IEC 62439, the method of publishing dual-primary IEC 61850 Generic Object-Oriented Substation Event (GOOSE) messages from each source IED through separate networks provides much better availability for trip signals. In addition, these GOOSE messages are supervised at each subscriber for performance based on standardized methods. With two messages containing unique signals representing field contacts, this method provides depth of information with detail about the health and behavior of the data source. Like visual depth perception gained from two independent eyes, information depth provides much more information about the source than replicating one signal. This method of using two encrypted messages with the same content to achieve information breadth, rather than one message replicated twice, is the core technology used to decode digital encryption methods, such as was done by the Allied Forces to decode the German Enigma machine in World War II.

Unfortunately, misapplication of Ethernet and manually corrected IEC 62439 techniques, such as replication via PRP and HSR, instead of compensation technologies, including resilient (and often redundant) RSTA and SDN, leads to dangerous failure modes that are hidden in plain sight. This is primarily due to the lack of standardized methods to detect and alarm failure as well as to take corrective action when a failure is present.

B. Understanding Differences Between High Availability and Fault Tolerance

In addition to technologies that do not provide fault detection, specific challenges evolve from misunderstanding the differences between high availability and fault tolerance. In [8], Paul Rubens discusses fault-tolerant design as applied to data centers. These concepts, which have been in use for decades within analog ECS design, are essential for the design of critical systems using digital signal processing to avoid and tolerate faults. Data center services are critical and rely heavily on a constant and reliable source of electric power. Rubens's discussion is useful to ECS design team members with an IT background, as they are often called upon to help design for the protection and control of an EDS.

In general terms, high availability means that a system will experience minimal outages, while fault tolerance means there will be *no* outages. Designs that automatically detect faults and mitigate them—for example, by applying a redundant or hot-standby component—without an outage are fault-tolerant. High-availability designs allow systems to operate to failure and then correct the outage as quickly as possible. Fault-tolerant designs demand more resources to implement, so ECS design teams are responsible for understanding the criticality of system components and subsystems and then deploying a combination of fault-tolerant and high-availability technology to meet the system's unique needs.

On occasion, when an ECS design team increases fault tolerance with dual-primary systems, resilience, or a combination of both, there can be a tendency to offset the increased cost by using inferior IT communications devices. This is a misguided choice that is often prompted by statistical analysis that demonstrates that, in a well-designed resilient system, the availability of an individual device is no longer critical. However, this cost-saving practice is quickly offset by reduced product life span and increased maintenance, and it can result in a less available system overall.

Rubens also points out that fault tolerance makes it more difficult to detect some ECS component failures when they do not lead to systemic failure [8]. For example, there is no standardized fault detection of failures with PRP and HSR replication methods, so they require unique, customized strategies to test and monitor the system in each ECS installation. However, this is rarely done because it can be quite costly, so these failure modes remain hidden in plain sight.

IEC 61850 process bus ECS design teams are responsible for ensuring that their systems are not only fit for the *use* of Ethernet messaging but also fit for the *purpose* of reliably and dependably transferring mission-critical trip signals.

IV. Lessons Learned in Aerospace, Aviation, and ECS

The similarity of critical systems and subsystems in aerospace, healthcare, and electric power delivery has led to decades of crossover technology sharing. The advent of digital communications systems has revealed a crossover of design failures as well.

In "Landing on the Hudson River: Lessons for Health Care," pilot Jeff Skiles discusses safety issues affecting healthcare and aviation that apply to ECS design as well [9]. The article concludes with a compelling explanation that computer and IT specialists often inadvertently make design choices that would otherwise satisfy a business system performance but that result in failure when applied to a mission-critical system. Poorly performing communications methods that do not impact business functions hide potential failure modes that could jeopardize a mission-critical system. The article points out that these poor design choices are often made when the design team is not familiar with the behavior and requirements associated with the critical nature of the underlying system,

It further points out that the engineering term "preoccupation with failure" should not be considered negative because the term describes how highly skilled professionals are continuously vigilant for opportunities to prevent catastrophic failure. Effective designers prevent big problems by being preoccupied with preventing small faults that could cascade into large failures.

In aviation, an accident is almost always the culmination of a chain of smaller errors, which means the accident will not happen if one of the small errors is fixed, breaking the chain. Skiles explains that breaking the chain is accomplished by developing "barriers to error," including procedures and checklists that are predictable and effective based on root-cause analysis of previous failures [9]. The most compelling barriers to error are the design team members themselves, namely their accountability and interpersonal communications.

Inappropriate ICS design choices have led to well-documented unintended events causing death, injury, or

ecological disasters. Like several modern ECS failures, the *Challenger* shuttle accident was caused by active dismissal of a simple—even low-probability—but important condition that became the initiating event of a cascading failure.

In *Truth, Lies, and O-Rings: Inside the Space Shuttle* Challenger *Disaster*, aerospace consultant Allan McDonald explains numerous engineering obstacles associated with the space shuttle *Challenger* disaster [10]. A particularly stark reality was the revocation of responsibility from the design team to decide when a design vulnerability associated with an "unlikely or a low-probability condition" was considered too remote for concern. On January 28, 1986, the responsible and accountable design team expressed that, based on all available evidence, the system conditions were outside the safety margin. Other interested parties usurped the role of responsibility and decided to proceed with launch. Tragically, 73 seconds into its flight, *Challenger* broke up, resulting in the deaths of Francis R. Scobee, Michael J. Smith, Ronald McNair, Ellison Onizuka, Judith Resnik, Gregory Jarvis, and Christa McAuliffe.

When visible and known vulnerabilities are dismissed by individuals unprepared for the role of mission-critical design assessment, unintended consequences result in hidden failures that reduce availability, reliability, and dependability.

Other observations from Skiles that affect healthcare and ECS design alike include those outlined in Table I [9].

TABLE I
RECOMMENDATIONS FOR SUCCESS APPLIED TO
IEC 61850 PROCESS BUS DIGITAL TRIP CIRCUIT DESIGN

Observation	Application to ECS Communications
There is no complete record of annual failures or near misses.	Due to resilient Ethernet design and differing design criteria (e.g., SCADA commands are 3,000 times slower than trip commands), it is not possible to know the number of failures that occur. Also, due to the relative obscurity of digital process bus systems, there is no statistically significant installed base from which to draw conclusions.
Failures are often due to computer system faults, misapplied technology, technicians with incomplete training, or a combination of all three.	ECS station bus and process bus communications networks are often built with IT devices and processes that are designed to operate to failure before being replaced. Often, technicians remove IT devices from service to install software patches without realizing that the device is part of a mission-critical control system and not simply a business system.
Technicians are not specifically trained or educated in healthcare or aviation.	Ethernet and IT specialists are recruited for the ECS design team but are not necessarily ECS or EDS specialists and thus have limited knowledge of the impact their design choices may have on the safety of people and equipment.
Many companies have no requirements for proof of healthcare or aviation competency when recruiting design team members.	ECS design team members may mistake awareness of IT and industrial IT requirements for competency in mission-critical OT requirements. ECS and EDS competency for IT ECS members is tested only rarely, if ever.

Aviation rules and regulations are often said to be "written in blood," meaning that root-cause analysis of deadly accidents has led to precautions and checklists to prevent reoccurrence of design flaws. While ECS station bus communications are often not this critical, it is necessary to presume that process bus trip circuits are always this critical.

On January 15, 2009, U.S. Airways Flight 1549 experienced a very low-probability condition. Dual-simultaneous bird strikes were the initiating event of a cascading failure that occurred two minutes after liftoff from LaGuardia Airport. The Airbus A320 captain Chesley "Sully" Sullenberger and first officer Jeffrey Skiles safely landed the aircraft on the Hudson river after the loss of its redundant engines.

Ten seconds after the plane ran into the flock of Canada geese, copilot Skiles began running through the emergency checklist for dual-engine loss. However, this checklist was only partially helpful because it was designed for problems at cruising altitude, when pilots have far more time to cope, and not the unlikely or a low-probability condition of dual-engine loss—an N – 2 condition—at takeoff.

The National Transportation Safety Board investigation of the event, discussed in [11], praised the pilots' quick thinking, and their recommendations included the following:

- Create a checklist for low-altitude dual-engine failures.
- Reevaluate how engines are designed and tested for bird strikes.
- Reconsider the brace position of the new type of seat in the Airbus A320, since it may have contributed to the shoulder fractures of two passengers.

Another dramatic ECS case-in-point is the fault in Sao Paulo, Brazil, which started at 2:58 p.m. on November 1, 2019, and remained energized for 110 seconds because a redundant breaker design failed to clear the fault. The fault inception was an N – 2 condition created by two simultaneous two-phase short circuits caused by vegetation. Viral videos of the fault filmed by local residents show how both transmission and distribution circuits were involved. After the primary circuit breaker failed to open, the backup breaker did not operate within the expected time coordination because the breaker failure protection failed. The primary breaker experienced an inherent mechanical failure, and the breaker failure protection may have experienced an extraneous man-made failure. A possible cause of this could have been that the breaker failure trip circuit had previously been disabled for testing and was accidentally left out of service. In this case, due to the characteristics of traditional trip circuit design, the unintended trip circuit outage was not detected, alarmed, or reconfigured, and it was unavailable when needed to open the breaker.

V. UNSKILLED, UNAWARE, AND MAYBE EVEN DANGEROUS

The Dunning-Kruger effect explains why some people, including ECS design team members, often do very poor work without realizing it. David Dunning cowrote a paper with colleague Justin Kruger titled "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," which describes why so many of us who are unskilled are also wholly unaware of our own lack of skills [12].

Flaws in IT Ethernet technology may not prevent it from being fit for the *use* of moving Ethernet packets, but it does

prevent it from being fit for the *purpose* of transferring protection signals within Ethernet packets in a trip circuit. For example, when technicians feel that Ethernet design flaws that are considered acceptable for IT systems are an unlikely or low-probability condition, they promote designs that are inappropriate for Ethernet-based transfer trip circuit signals. Dunning points out that some industries that support system infrastructure, such as IT, often employ technicians skilled in general-purpose methods who are overconfident that their design will adequately support any purpose. They may, however, fail to realize that on a control system design team, they are also accountable for their role of satisfying the criticality of the underlying aviation, healthcare, and electric power systems.

Also, in the interest of following social norms, people rarely contradict an individual with real or perceived authority, even when they can see fault in the person's actions. Then, when technology is misapplied where a fault should be predicted and mitigated, the system is allowed to drive to failure. More often than not, this is then blamed on bad luck rather than lack of preparation.

For example, in as many as 16 deadly aviation accidents, the crew knew the plane was going to crash, but flight recorder data show that they deferred to the pilot's mistaken confidence and authority as the aircraft proceeded to drive to failure, resulting in a crash [12]. When technical solutions are presented with authority, audience members often follow social norms and will assume that the speaker is telling the truth, unless they have clear evidence to the contrary. However, presumption of truth without evidence often leads to failure.

For example, the IEEE Code of Ethics points out that the operators at Chernobyl bypassed emergency cooling and protection systems during a test because they succumbed to the social norms of accepting instructions from an authority figure that were in conflict with established safety set points [13]. It further provides the example that it is paramount to protect public safety, so it is improper to consider safety as holding the same weight as other design goals and, thus, being subject to the same tradeoffs.

If, for example, an engineer involved in a design believes that the design is unreliable, he or she is under obligation to call attention to the problem. Then, if the engineer's manager still wants to retain the existing design but does not provide technically persuasive arguments, the design engineer is obligated to pursue the idea further and bring it to the attention of a higher level of supervision. Depending on the seriousness of the issue and the level of certainty the engineer has, it may even be necessary to contact a regulatory group outside the company, such as the Federal Aviation Administration. Section 7.8 of [13] states that engineers must promptly disclose factors that might endanger the public or the environment. Although issues of this magnitude are rare, it is always the role and responsibility of an engineer to safeguard public safety at all costs.

Further, people are most convincing when they are convinced they have the correct answer, even if it is not grounded in research or proven data but rather in their feeling of confidence in its truth. In a recent interview for a *Bloomberg Opinion* column, Dunning points out that people are much more likely to believe that fake and incorrect things are true rather than to believe that a true thing is fake [14]. People are too easily swayed to believe what is not true, so it is the responsibility of ECS design team members to invariably choose true science over social science. People can reduce their vulnerability to false information simply by evaluating the data and the source. Mission-critical designs require that every team member be a fact checker and review the evidence before making a decision.

People do not realize how much work it takes to produce an evidence-based analysis of the behavior of a technology. For example, behavior of IT and OT Ethernet switch networks requires a statistically significant number of fault tests of every component. This research and data gathering has been shown to take months of 24-hour-per-day automated testing to analyze multiple topologies for a ten-switch Ethernet network. Technicians who are not directly accountable for the EDS or ECS often do not adequately apply critical thought when promoting new technologies. Adapting new technologies to mission-critical applications like trip circuit design requires rational and clear thought, in addition to a full understanding of the logical connections among ideas and methods. It is necessary, in fact, to approach any new technology and method with scientific skepticism.

Dunning and Kruger claim that ignorance of the scientific method is so profound because people who make rash decisions fail to acknowledge that scientists collect data to aid in their decision-making [12]. For example, when promoting IEC 62439-3 PRP and HSR replication techniques, technicians often rely on social truth rather than scientific truth. They reference a document (e.g., a standard or magazine) rather than a scientific experiment that would immediately demonstrate undetected failure modes.

Well-intentioned audiences often assume the common shortcoming of basing their beliefs on what other people say. Dunning points out that this method of relying on social proof is the same reasoning used by those who choose to believe in the supernatural, ghosts, karma, and miracles. However, when dealing with potential misinformation, people should insist on scientific proof and evidence, since data have the final authority.

In *Controlling Technology: Ethics and the Responsible Engineer*, Stephen H. Unger points out that engineering managers who request personnel do not often take the time to ask for proof of specific qualifications [15]. His example explains that engineers sometimes accept assignments they are unqualified for because they do not realize that they are unqualified. This is due in part to not fully understanding what the job entails and the fact that some engineers can have an inaccurate perception of their experience level when applying their skills to a new specialty area. He points out that it is becoming more commonplace for managers and engineers alike to blame others when they discover a skills gap rather than ensuring team members' skills adequately prepare them for the

responsibilities demanded of a specialized, mission-critical application before they are hired.

Indeed, although freezing temperatures on a Florida launch pad is an unlikely or a low-probability condition, individuals with a misguided perception of their decision-making skills made a choice that ultimately cascaded into a catastrophic failure that resulted in the deaths of seven people. Although the N – 2 condition of dual-primary failure is an unlikely or a low-probability condition in a protection system, and there is no way of knowing how often it happens, it did happen in the presence of a fault as recently as December 2019 in Brazil. And, although the N – 2 condition of dual bird strikes is an unlikely or a low-probability condition during takeoff, it happened in 2009 and cascaded into the emergency water landing of a commercial airliner. Social proof labeled this event the "Miracle on the Hudson," but scientific proof documents a very well-designed aircraft skillfully and patiently navigated by two pilots with extensive experience and training who practiced great concentration, discipline, critical thinking, and interpersonal communications. It was not a miracle; it was a feat of extreme skill.

VI. Role-Based Access (RBA) and Accountability

Operators and engineers are given permission to use the in-service ECS to control the EDS and to modify the ECS, based on the job functions associated with their role in the utility. This RBA to the ECS is identified by the system owner and managed in real time by the ECS by controlling user authorization and permissions based on their job functions. This RBA control (RBAC) includes role assignment, role authorization, and permission authorization, which are defined in Table II.

TABLE II
RBAC Job Functions

Function	Task
Role assignment	Identify each necessary role and the associated tasks and functions necessary for a team member to accomplish each role.
Role authorization	Authorize each team member to perform one or more roles based on job function. Verify all team members have appropriate education and training for their role in ECS operations, as well as the associated interfaces and tools. Ensure team members can provide proof that they understand the ECS interface and the impact that their commands will have on EDS operations. *Team members are only assigned the roles for which they are authorized.*
Permission authorization	Associate each role with commands necessary to perform duties associated with that role. Assign permission for team members with role to execute associated commands using the ECS interface. *RBAC provides permission to execute each function only if the permission is granted for a team member's presently active role.*

Similarly, designers and engineers accept the responsibility to create an ECS to control the EDS and meet the associated design criteria, including availability, resilience, security, dependability, and speed. This role-based accountability for designing the ECS is no less important than RBAC for operating the ECS. However, skill verification is often done open-loop and without appropriate supervision, which leads to design flaws.

The use of nonspecific tools, including information processing software and Ethernet communications networks, in very specific applications (e.g., an ECS) requires role-based knowledge and certification. The ECS owner should manage role-based *accountability* based on design team member job functions using role-based accountability control. Similar to RBAC, role-based accountability methods include role assignment, role authorization, and permission authorization, with each of these functions fulfilling the same tasks outlined in Table II. However, whereas RBAC is defined in terms of permission and responsibility to perform functions to operate the ECS in real time, role-based accountability control is defined in terms of permission and responsibility to perform functions to design and build the ECS in advance, which is then, in turn, used to control the EDS. Examples of role responsibilities assigned as a part of role-based accountability are as follows:

- Identify communications-assisted applications.
- Document data server and destination.
- Select protocols.
- Identify client and associated restrictions.
- Create data-set mapping and matrix.
- Create last-mile cyber restrictions.
- Deploy deny-by-default security to prevent nonengineered communications.
- Create SDN whitelisting and traditional Ethernet blacklisting.
- Configure interface to a security information and event manager.
- Design packet flow.
- Document LAN requirements.
- Design LAN to meet ECS criteria for availability, resilience, security, dependability, and speed.
- Document the data flow design.

Adopting role-based accountability for members of an ECS design team makes team member capabilities, assignments, and decisions much more transparent. It reveals important decisions that have not yet been assigned to an accountable decision maker and promotes collaboration and peer review of important decisions made by others. This is especially helpful because it allows the design team to hold team members accountable for tasks associated with specific categories of criticality, including performance and resilience.

VII. Resilient EDS Design

A. Definitions and Standards Ensuring Quality of EDS Design

An EDS is a network of redundant and hot-standby primary system components built to satisfy quantity and quality requirements of each point of consumption. One role of an ECS is to automatically detect faults and then isolate them via automatic operation of energy-switching devices. This transitions a failed primary system element from an in-service

state to a not-in-service state. The role of an ECS is to then improve resilience by quickly restoring the flow of energy after fault isolation by automatically transitioning a nonfaulted redundant or hot-standby primary system element from a not-in-service state to an in-service state.

One measure of EDS performance is the duration (if any) required to perform automatic EDS network reconfiguration, which determines the duration of the power outage. Designers of the ECS must provide a design that will fulfill its role of controlling the EDS, which entails the following:

- Immediate detection of faulted EDS component.
- Automatic isolation of faulted EDS component.
- Automatic energization of alternate EDS component.
- Automatic restoration of energy flow to all points of consumption.
- As a condition of N – 1, N – 1 – 1, or N – 2 EDS state, immediate and automatic restoration of N – 1, N – 1 – 1, or N – 2 status of ECS.

According to NERC, electrical faults occur within an EDS when the flow of energy is interrupted by an outside disturbance (e.g., lightning) or by an internal equipment failure. Within the ECS, the role of the protection system is to detect faults and to quickly relay a request to open a circuit within the EDS to isolate the fault. In this role, the protection system senses changes in EDS currents, voltages, traveling waves, or other physical quantities resulting from the electrical fault. It is critical that the ECS quickly detect faults and then automatically communicate commands to controllable devices in the EDS. Very high currents associated with faults can be deadly, as well as very destructive to network equipment, so speed and availability requirements are defined for these systems as well [16].

The presently approved NERC definition of a protection system, illustrated in Fig. 1, includes the following devices:

- Protective relays.
- Associated communications systems.
- Voltage and current sensing devices.
- Station batteries and dc control circuitry [17].

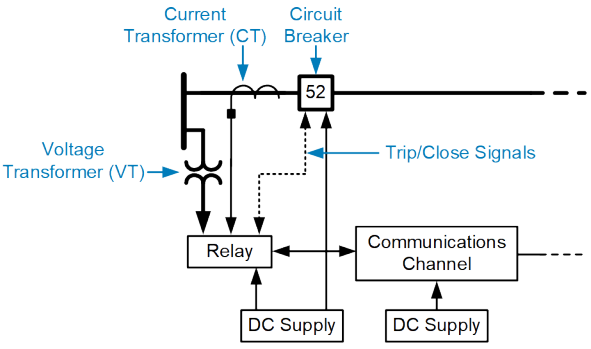

Fig. 1. Current NERC Definition of Protection System

The proposed new definition better represents modern system designs and emphasizes communications circuitry among the devices within the process. This system is shown in Fig. 2 and includes the following elements:

- Protective relays.
- Communications systems necessary for correct operation of protective functions.
- Voltage and current sensing inputs to protective relays and associated circuitry from the voltage and current sensing devices.
- Station dc supply.
- Control circuitry associated with protective functions from the station dc supply through the trip coil(s) of the circuit breakers or other interrupting devices [17].

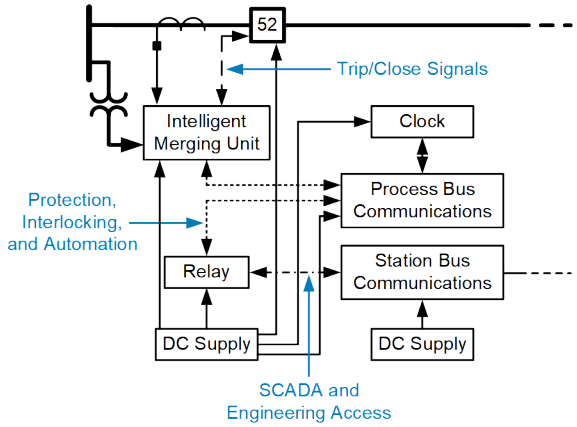

Fig. 2. Proposed New NERC Definition of Protection System

According to NERC, to coordinate among systems interconnected in the EDS, the role of ECS designers at transmission and distribution companies includes creating and sharing details of their protection systems with interconnected EDS utilities. Design elements related to the digital communications within the ECS to be shared to improve coordination ECSs include the following:

- CT and VT/capacitively coupled voltage transformer (CCVT) configurations.
- Documentation showing the function of all protective functions.
- Communications-assisted schemes [18].

NERC also describes the steps of a maintenance program to include one or more of the following activities:

1. **Verify** – Determine that the component is functioning correctly.
2. **Monitor** – Observe the routine in-service operation of the component.
3. **Test** – Apply signals to a component to observe functional performance or output behavior, or to diagnose problems.
4. **Inspect** – Examine for signs of component failure, reduced performance, or degradation.
5. **Calibrate** – Adjust the operating threshold or measurement accuracy of a measuring element to meet the intended performance requirement [19].

All five of these steps apply to ECS devices that create, consume, and communicate protection signals as part of electric fault mitigation. The role-based accountability of the ECS designer is to provide for fast and effective digital communications within the ECS, specifically to perform protective control actions.

In the absence of a modern IEEE digital communications trip circuit design guide, ECS requirements for process bus trip communications should at minimum be similar to those for remedial action scheme (RAS) tripping based on digital communications. Energy Coordinating Council requirements for RAS designs include the following:

- Logic should be designed so that loss of channel, noise, or other channel failure will not result in a false operation of the scheme.
- All channels and channel equipment should be monitored and alarmed.
- Any part of the RAS that has lost redundancy or duplication, due to failure of another component, must provide an alarm, since failure of that equipment would create a sustained outage.
- All channels and channel equipment should be monitored and alarmed to the dispatch center so that timely diagnostic and repair action will take place upon failure.
- Communications channels used for sending and receiving logic or other information between local and remote sites and/or transfer trip devices must meet at least the same criteria as for other relaying protection communication channels [20].

B. Electric Reliability Organization (ERO) Enterprise EDS Failure Severity

Similar to other technical industries, such as the aviation industry, EDS and ECS failure conditions are categorized based on severity. NERC and seven regional entities make up the ERO, which is tasked with creating a highly reliable and secure North American EDS. As part of their mission to "assure the effective and efficient reduction of risks to the reliability and security of the grid," the ERO recommends mitigating risks by means of consequence-informed planning and operation, as well as resilient system design [21]. To that end, they have categorized failure severity of disruption events occurring on the EDS based on various consequences, which are rated as follows [22]:

1. Loss of communications from SCADA remote terminal units or loss of automatic generation control communications.
2. Unintended island of 1,000 MW to 4,999 MW or loss of a greater than 300 MW load for more than 15 minutes.
3. Unintended island of 5,000 MW to 10,000 MW or loss of load or generation of greater than 2,000 MW.
4. Unintended island of greater than 10,000 MW or loss of load or generation between 5,001 MW and 9,999 MW.
5. Unintended loss of load or generation greater than 10,000 MW.

In an effort to correlate EDS and ECS components with outages, as well as analyze the effects of the risk attitude in design and operation, the Purdue Laboratory for Advancing Sustainable Critical Infrastructure studied major EDS outages witnessed by different states in the continental U.S. between January 2000 and July 2016 [23]. The summary of outages related to the five ERO severity categories are summarized in Table III.

TABLE III
ERO SEVERITY CATEGORIES

NERC Severity Category	MW Loss	Number of Events	Cause
5	≥10,000	9	Intentional attacks, vandalism, severe weather
4	Between 5,000 and 10,000	6	Intentional attacks, vandalism, severe weather
3	Between 2,000 and 5,000	15	Intentional attacks, vandalism, severe weather
2	Between 300 and 2,000	264	Sabotage, intentional attacks, vandalism, severe weather, wildfire
<2	<300	527	Intentional attacks, vandalism, severe weather, wildfire, earthquake
Not categorized	Not available	712	Intentional attacks, vandalism, severe weather, wildfire, earthquake

As shown in Fig. 3, the three most common causes of misoperations in the EDS are directly related to the ECS and are responsible for 60 percent of misoperations since 2011 [24]. ECS designers are accountable to understand and mitigate all failure modes including the top three: errors in settings, logic, and/or design; relay failures or malfunctions; and communications failures.

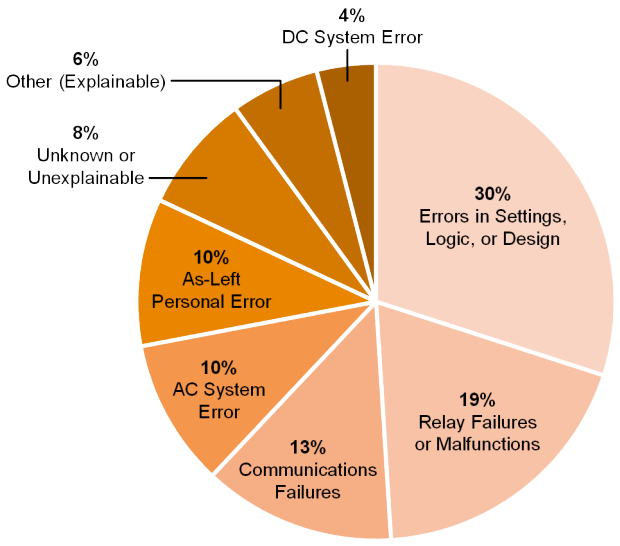

Fig. 3. Common Causes of Misoperations in the EDS

The Newton-Evans study of the North American market for substation automation and integration systems reveals that 56 percent of respondents plan to replace their legacy hardwired I/O. The study does not ask if this includes trip

circuit wiring or if the replacement will use the many IEC 61850 protocols for process bus design. However, this report, similar to those reflecting international respondents, does illustrate that, though it is becoming more popular, respondents are implementing very little Ethernet, even less IEC 61850, and even fewer, or no, process bus trip circuit installations at present [24].

Though presently trending downward, as Fig. 4 illustrates, the ERO results of the annual number of ECS misoperations as a result of human error remains large.

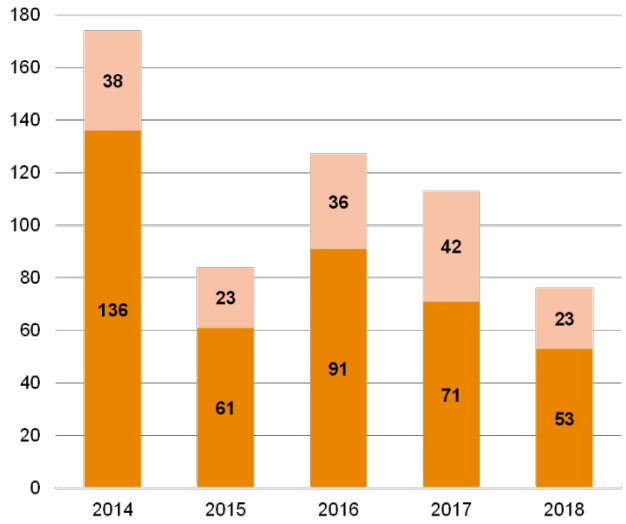

Fig. 4. ECS Misoperations Due to Human Error

VIII. The "Good" Bad Examples of American Space Shuttle Tragedies

Failure to consider that unlikely or low-probability events could cause a failure is irresponsible, especially when technology exists to detect and eliminate that failure mode. Assessment of this issue relative to both the *Challenger* launch disaster in 1986 and the breakup of the space shuttle *Columbia* in 2003 is appropriate because both were triggered by avoidable initiating events. The *Challenger* breakup was caused by an unlikely, low-probability condition that was never properly analyzed because it was considered by some to be too remote to worry about.

Trying to identify how a process, technology, or device like this can fail involves three areas of concern: a) known knowns, b) known unknowns, and c) unknown unknowns.

A. Known Knowns

Known knowns are by far the most prevalent areas of concern and receive the most attention, which is as it should be because they are most likely to cause a failure. The space shuttle solid rocket motor field joint and O-ring sealing system had numerous known knowns to account for. The engineering team inspected and controlled the surface finish for the O-ring sealing surfaces, dimensional cross section of the O-rings, O-ring grooves, case and O-ring diameters, and metal thicknesses and properties. X-ray and pressure tests were performed prior to assembly, and a high safety factor was applied to the design. The higher-than-typical safety factor was used to accommodate stress greater than the design margins to cover any unforeseen contingencies and to provide an added safety factor for operation.

For example, a common challenge to the design of ECS communications is inappropriate network design, such as provisioning Ethernet bandwidth based on IT practices of message size, rather than OT requirements of maximum message latency, are known knowns that will persistently cause latency of trip signals. Though it is often undetected, this flaw will constantly exist by design and be present when a fault occurs, and it may very well provoke a cascading failure.

B. Known Unknowns

Known unknowns are conditions that are known to eventually cause a failure, but there is not enough data or analytical capability to determine where that failure point is. Known unknowns were the exact reason why Morton Thiokol engineers recommended not launching the *Challenger* in temperatures below 11.6°C (53°F). The engineers knew that at some lower temperature the O-rings would fail to seal, but they did not know where that point was, and they did not have the tools, much less the time, to determine that point analytically. They had witnessed a concern due to a low temperature of 53°F one year earlier and did not feel comfortable going below that temperature. The key to the most successful design is to obtain the necessary data or develop the necessary tools to assess the boundaries of known unknowns or, if this is not possible, to provide an additional safety factor or operating constraint to cover the unknown.

Undetected failures of certain technologies, including PRP and HSR, and ECS outages longer than 15 seconds are known unknowns that will inevitably happen in the presence of a fault and cause a cascading failure. However, it is not possible to know their frequency or their existence. These must be anticipated, acknowledged, and openly discussed between the ECS design team and the ECS owner and mitigated to the furthest extent the project cost, schedule, and performance requirements allow.

The space shuttle *Columbia* is a "good" bad example illustrating the deadly consequences of poor decision-making and incorrect risk assessment when faced with a witnessed known failure condition. During failure analysis while the *Columbia* was in orbit, Boeing engineers explained that the foam strike during launch was 600 times larger than any previously tested damage [25].

In spite of this, and partly based on the history of noncatastrophic foam strikes on four previous launches, National Aeronautics and Space Administration (NASA) managers chose to reject the options of repair and rescue and proceeded with the *Columbia* landing as planned. The crew was never advised of the damage or of the decisions being made about the mission with inadequate regard for their welfare and public safety. On February 1, 2003, just prior to 8:00 a.m. local time, the *Columbia* disintegrated at 18 times the speed of sound, 38 miles above Dallas, Texas.

In a study conducted after the *Columbia* disaster, NASA was told to assume that correct actions had been taken to learn the extent of the damage in either of two ways: images of foam striking the shuttle, which were enhanced by available technology, and a spacewalk by the crew members prepared for this activity [26].

The results of the study determined that, with this information, either a repair or rescue conducted in orbit would likely have been successful. The repair could have been made with using existing materials onboard. The launch schedule of the shuttle *Atlantis* could have been safely accelerated, including all safety checks, and reached orbit with a five-day window of time, beginning February 10th, to rendezvous with *Columbia* before the crew's consumables ran out.

Why, then, with the crew and public safety at risk, did the known unknown fact that the foam strike was dramatically outside of test data remain hidden in plain sight?

Whereas reporting known knowns is an effort to avoid doing harm, reporting known unknowns is an attempt to prevent harm from being done [13]. The degree of certainty and the magnitude of the danger should be considered when deciding the level of remediation to recommend, but the minimum obligation for engineers is to warn others of the danger.

C. Unknown Unknowns

Unknown unknowns are conditions that were never considered to even occur, much less to create a condition that could contribute to or cause a failure. The key to assessing unknown unknowns is truly by thinking outside the box to consider other factors that have not been previously considered as possible threats to the successful operation or use of a product or system. It is very difficult to quantify the improvement in reliability of the product or service, but it could well be an order of magnitude, or at the very least, it will always be better than any product that has not even considered such a possibility to exist.

It was, in fact, this third condition, an unknown unknown, that caused the *Challenger* disaster. Gaseous oxygen vapors from the external tank that were dumped near the launch service structure ultimately triggered the event. These vapors blew back onto the vehicle super-cooling the air around the shuttle, resulting in colder temperatures on the right side and bottom of the shuttle. This condition resulted in an O-ring temperature closer to –12.7°C (9°F) on the bottom field joint of the right-hand booster, when the other five joints were closer to the ambient temperature of 2.2°C (36°F) at liftoff. That is precisely why that joint failed and the other five did not. NASA never even considered this rare environmental condition, and therefore no engineering and design resources were allocated to analyze it, because it was so unlikely to ever occur, much less to be present on the very day and time when a space shuttle was being prepared for launch. The *Challenger* breakup is a very "good" bad example of what can happen by not considering everything that could create a condition leading to or contributing to a failure in any system, no matter how remote the possibility.

Unknown unknowns become a serious issue when engineers accept assignments for which they are not qualified [13]. This is a fundamental aspect of professional integrity, and deception or suppression of relevant information undermines the mutual confidence of designers and owners that is essential to technology-based systems. The exception is when a less than fully qualified engineer does accept roles and responsibilities that do not jeopardize public safety, and no deception is involved. There are also cases where engineers can gain experience and necessary qualifications by collaborating with others, under full disclosure, during a project.

IX. The Known Knowns of the Boeing 737 MAX Failure

Although all engineering tasks require application of rigor and knowledge, those that involve public safety, such as aviation and the ECS, require specific attention. A more recent "good" bad example with deadly consequences illustrates both bad design and an incorrect risk assessment of the consequences of a known failure condition.

The Federal Aviation Administration (FAA) allows different models of airplanes with similar design characteristics, such as different models of the Boeing 737, to share a common type certificate. New aircraft versions with common type certificates do not legally require new pilot certification, which lowers the cost to the airlines. The goal was to do this for the new Boeing 737 MAX jet.

To add the new, large engines of this new version of the jet, though, Boeing had to mount them more forward and higher on the wing than the previous model. Boeing developed a software system called the Maneuvering Characteristics Augmentation System (MCAS) as a temporary solution to compensate for the unusual lift associated with these new engines. This design was 8 percent more fuel efficient than any competitor, and the common type certificate required no expensive classroom or simulator time and only 2.5 hours of computer-based training for the pilots. The MCAS, which relies on an angle of attack (AoA) sensor, was provided as a compensation measure but was hidden and not disclosed to pilots.

The FAA defines failure as "a loss of function, or a malfunction, of a system or a part thereof" [27]. Each aircraft system failure condition is categorized with respect to severity based on a prediction of what will happen if the system fails. The FAA summarizes severity with respect to passengers, as follows:

- Minor – inconvenience to passengers.
- Major – discomfort to passengers.
- Hazardous – fatal to a small number of passengers.
- Catastrophic – fatal to all passengers.

These severity categories are used to specify the level of appropriate resilience and redundancy of the aviation systems.

The MCAS was categorized by the FAA as hazardous, which requires two or more levels of redundancy. However, the MCAS receives input from just one AoA sensor at a time, meaning it has no redundancy and thus cannot perform a

reasonability check with a second sensor when the primary is suspect. It cannot cope with sensor malfunction and does not satisfy the minimum requirements of a hazardous failure condition. The Boeing 737 MAX experienced nearly six times the allowed AoA sensor issues for a hazardous rating from 2014 to 2019. What is worse, though, the MCAS software is, in reality, four times more powerful than was documented. This increased power, which was hidden from but prone to interfere with pilots, raised the FAA failure condition severity rating from hazardous to catastrophic due to the even higher risk of failure.

The result is that 186 innocent people died when Lion Air Flight 610 crashed on October 29, 2018, and 157 when Ethiopian Airlines Flight 302 crashed on March 10, 2019 [28]. The passengers, like all of us, trust manufacturers, regulators, and engineers to be fully qualified for the roles they are assigned and fully trained in each specialized area of mission-critical systems in order to protect public safety.

The MAX failure is a "good" bad example of the cross purposes of business, technology, and safety. Boeing focused on efficiency instead of transparency, speed instead of rigor, and similarity instead of innovation. The FAA missed a clear opportunity to prevent the two crashes by failing to enforce its own safety regulations.

In the eight years after the MAX design was announced, Boeing stock quadrupled, its profits doubled, and its annual revenue grew by nearly 50 percent to 101 billion dollars. However, Boeing has lost over 25 billion dollars in market capitalization since the two crashes, and it may have to pay billions more to its suppliers and airline customers for costs related to the grounding [29].

The failure conditions of the 737 MAX were a result of known knowns, known unknowns, *and* unknown unknowns. However, the most egregious failures were a result of known unknowns and unknown unknowns becoming known and, yet, remaining hidden.

X. IEC 61850-Based Process Bus Trip Circuit

The use of process bus communications in the ECS including digital messaging, shared bandwidth Ethernet, reliance on precise time, and commercial versus utility grade devices creates many new known knowns, known unknowns, and unknown unknowns that must be addressed rather than remain hidden in plain sight.

Modern microprocessor-based IEDs often produce telecontrol, teleprotection, metering, protection, automation, and control signals that need to be delivered with mission-critical levels of dependability and security. Digital messaging defined by NIST includes protocols supported by standards developing organizations (SDOs)—including IEC 60870, IEC 61850, IEC 61158, and IEEE 1815, Standard for Electric Power Systems Communications-Distributed Network Protocol (DNP3)—and protocols supported by Standards Related Organizations (SROs)—including Mirrored Bits® communications [30].

H2M protocols send operator control commands and transmit and receive system information on the station bus. M2M connections and protocols exchange I/O process information between IEDs and process instrumentation and control devices. Interlocking, automation, and time-distribution signals are M2M messages that exist on both the station and process bus [2].

Each source device performs analog-to-digital conversion of the analog signals to create a pool of process-level, raw signal information [19]. Then, with each microprocessor operating cycle, the IEDs create processed signals via arithmetic and logic calculations. These local, raw, and calculated signals are used to make local decisions about the health and performance of the primary equipment and to perform local control and protection functions. When equipped with appropriate communications capabilities, each data consumer IED also receives remote, raw, and calculated values from other data producer IEDs, and the data consumers add these to the pool of local, raw, and calculated signals. Raw field signals and calculated quantities arrive at the receiver (data consumer) IED as contents of digital message payloads over various communications media. The process to convey data from the producer to the consumer after it is measured or calculated is as follows:

1. Data change detection and time-stamp creation in producer IED.
2. Message creation in producer IED.
3. Message publication in producer IED.
4. Message transfer across the communications media.
5. Message subscription in consumer IED.
6. Message verification and decoding in consumer IED.
7. Strategic delay in producer IED as appropriate to manage message delivery and reception from multiple sources.
8. Data parsing, processing, time-stamping, and mapping into virtual data placeholders in consumer IED.

Together, these eight steps result in a time latency associated with moving the payload from the data producer to the data consumer after it is available within the data producer. The precision of data alignment and latency compensation dictates what arithmetic and logic processes can be supported. Design teams must create service-level specifications (SLS) for each communications-assisted application describing the numerous performance requirements to adequately serve the underlying application [31]. As part of the project documentation, the design team must document their service level agreements (SLAs), which identify what metrics they plan to meet and how they will provide evidence of success. The IEC 61850-90-4 Technical Report describes network testing to verify process bus network performance and requires that an appropriate subset of the tests continue to monitor the network during operation. These tests detect and mitigate failures and ensure conformance to the SLAs that the design team has agreed to meet. Since relays that publish GOOSE and Sampled Value (SV) message data are unaware of multicast message delivery, signal exchange between them and data subscribers must be monitored by each subscriber.

To design a process bus system—with or without a digital trip circuit—the SLAs need to satisfy the international standards referenced by IEC 61850, as explained in [32].

Digital communications acceptance criteria have driven the design of SLA documentation and real-time monitoring to be provided to end users. The SLAs necessary for protection, automation, and trip circuits are very specific and numerous. The metrics needed to confirm satisfaction of an SLA are referred to as key performance indicators (KPIs). As an example of the role of the ECS team member responsible for these communications, a list of GOOSE exchange metrics are included in the following list for reference. Designers may choose to request stricter acceptance criteria, but, based on these standards, IEC 61850 GOOSE exchange SLAs for publishers must at minimum record KPIs in memory, display metrics in real time, and alarm failure to meet the following metrics thresholds:

- Protection signal exchange success rate greater than 99.99 percent.
- Device configuration and real-time details related to message publication.
- LAN signal transfer time between devices of less than 3 milliseconds.
- Signal transit via LAN of less than 1 millisecond.
- Maximum data delivery time, since last reset, between devices within a substation LAN of less than 0.25 cycles.
- Maximum data delivery time between devices, since last reset, external to a substation, across a wide-area network of between 8 and 12 milliseconds.

To verify IEC 61850 GOOSE protocol, SLAs require that each subscribing relay uniquely monitor and validate each protection signal exchange. In the case of GOOSE messages, each exchange is supervised in real time to confirm its integrity before its contents are used for communications-assisted protection, automation, and control The IEC 61850 GOOSE KPIs and related SLAs for subscribers must at minimum record KPIs in memory, display status and metrics in real time, and alarm failure to meet threshold as appropriate for the following:

- Detect and display delayed GOOSE messages for each subscription.
- Detect and display lost, undelivered GOOSE messages for each subscription.
- Detect and display maximum quantity of packets lost in a single event, total aggregate quantity of packets lost, and maximum outage time as the duration of time for which GOOSE messages are not received for each GOOSE subscription since last reset.
- Create and store GOOSE message receipt reports containing message configuration information as well as message status, including priority tag, virtual LAN, state number, time-to-live (TTL) value, sequence number, and error code for each subscription. The TTL value is to be recalculated in real time and represents the expected duration before receipt of the next GOOSE message.
- Create, store, and display the TTL count for each subscription.
- Create, display, and store an out-of-sequence count for each subscription.
- Create, display, and store a decode error count for each subscription.

Persistent fulfillment of these SLAs in real time is required as evidence of the safe and reliable operation of process bus applications. The values act as KPIs to identify and illustrate performance degradation. Metrics that flag problematic behavior prompt root-cause analysis and service improvements.

SLAs and KPIs for SV, time distribution, and other process bus trip circuit elements are similar and are also useful to quantify in-service behavior in order to understand and mitigate risk.

XI. ECS Design Team Role-Based Accountability

Industry-specific technology providers, such as manufacturers of protection and control equipment and developers of ECS solutions, are also responsible for training users to protect and operate an ECS with their products. They are also capable of teaching users about the roles and responsibilities of designing an ECS. Providers that offer other peripheral services, such as testing and training, may be helpful but have no specific role in the design or operation of the ECS. Their responsibility, though, is to properly teach the use and application of their products.

Sophisticated communications components and information processing software are tools common across most industries, including healthcare, aviation, and electric power systems. Technology providers that produce information and engineering tools and Ethernet communications components and systems recognize that they are not subject matter experts (SMEs) in these industries and other fields where their products are used, but they are experts in developing and using their products. Role-based accountability includes knowledge, skills, and certification of the same. Technology providers need to help users become proficient in the use of their technology and in the appropriate use of the technology for designing a mission-critical solution.

Technology providers are not necessarily SMEs in mission-critical systems and thus do not have the capacity to verify a candidate's knowledge of the fundamentals of mission-critical systems. Often, technology providers are not even experts in the use of their own tools, so they rely on consumer and consultant SMEs familiar with using their technology to create solutions. These companies acknowledge that certification of this knowledge must be found elsewhere. Design team members must demonstrate the following skills before assuming the responsibility of designing a mission-critical system:

- Knowledge in the underlying mission-critical system.
- Proficiency in the use of the proposed technology.
- Proficiency in the use of the technology to perform specific project-related tasks.

Even today, when digital trip circuits are virtually nonexistent, ECS misoperations as a result of human error are large, as shown in Fig. 4. This number can be predicted to rise

as the industry migrates toward digital trip circuits (Fig. 2) and replaces low-level energy signals flowing over pairs of copper wires with signals in packetized Ethernet messages. Even for SMEs who understand IT Ethernet business and, perhaps, ECS station bus applications, OT Ethernet process bus protection applications require a completely different understanding to design a system of digital relays, intelligent merging units, and clocks that is energized by separate power supplies and exchanges information in messages on a network of switches and cables. The added technical complexity and the more severe service level requirements lead to more risk and potential for complications, which may contribute to increased minor, major, hazardous, or even catastrophic failures.

In light of this, it is of the utmost importance that the ECS design team fully understands digital trip circuit design and implementation so as to not exacerbate the top three human-induced failures illustrated in Fig. 3: hardware selection, protection settings and configuration errors, and communications settings and configuration errors.

The ECS design team is responsible for being aware of the consequences of design choices and the severity of a potential failure. Severity can be predicted by understanding the associated category of failure designated by the ERO and defined previously in Table III. Each category specifically documents the impact of known knowns associated with the complexity in ECSs including "failure to properly design, coordinate, commission, operate, maintain, prudently replace, and upgrade…[EDS and ECS] assets could negatively impact system resilience and result in more frequent and wider-spread outages initiated or exacerbated by protection and control system misoperations or failures" [21]. As ECS design teams migrate toward ECS designs similar to Fig. 2, care must be taken to address all failure modes associated with digital protection process bus technology based on IEC 61850.

Although it is not possible for most general-purpose technology providers to certify whether team members are knowledgeable in the underlying mission-critical system, they are increasingly obligated to provide skills certification of their proficiency in the technology being used and collaborate with SMEs to develop certification processes. Product certification tests knowledge of and ability to effectively use a specific technology. Role-based certification tests knowledge and ability to use the tool to perform tasks and duties associated with a specific role. As mentioned, general-purpose technology providers use SMEs—trusted and capable users and consultants—to create tests that audit necessary skills and competencies for certification of specific roles. Even though general-purpose technology providers recognize that they lack specific industry knowledge, they do recognize that they can encourage candidates to learn to ask the appropriate questions to learn industry-related essentials. Further, they should consistently evaluate the reliability and timeliness of their certification procedures.

To better understand this, consider the hypothetical example of creating an information analysis dashboard for a physician who queries a huge database of patient records for trends and predictions. A competent member of the design team for this database should not only be a proficient software user but also understand queries of healthcare-related databases. It is also essential for that same designer to learn and understand the underlying use of the dashboard, in addition to its availability and performance requirements. Fault tolerance of a dashboard that needs to be available ad hoc for a physician to make a diagnosis based on test results prior to a patient's office visit is completely different than one used constantly in real-time to make life-or-death decisions about a patient's care in the operating room.

A real-world example of how team members who are unqualified to apply their general experience to a mission-critical application can result in preventable failures can be found in one of the many tragedies that occurred in the aftermath of Hurricane Katrina [33]. Rain and flooding is a common occurrence during hurricane season in Gulf Coast states, so the Pendleton Methodist hospital in New Orleans decided to decline help in evacuating the hospital inpatient population because they were mistakenly confident that they could safely weather the storm. Although the hospital administrators were the most skilled in hospital operations, they relied on disaster planners with limited knowledge of hospital operations to plan for and safeguard their hospital during the storm. The overconfident disaster planning team misjudged the hospital infrastructure and declined to recommend evacuation because they had installed emergency generators. However, after the utility feed was lost, the emergency generators were inundated by floodwaters and went offline, resulting in a blackout that caused the death of at least one patient. This avoidable tragedy is another example of a failure hidden in plain sight, which would have been simple to mitigate with proper training, peer review, and quality control.

Recently, a knowledgeable utility ECS design engineer had to temporarily suspend work on a process bus design when a second engineer from a department that was not a stakeholder in the initial design tried to prohibit the digital communications process bus upgrade. The second engineer misunderstood the changes and had felt that the upgrade would affect his established methods to provide service to the primary equipment. The ECS design team, though, correctly decided that he was not impacted by the new technology or responsible or accountable for the design. Ultimately, they collectively applied the RACI matrix to resolve the problem, which is defined in Table IV [34].

TABLE IV
RACI MATRIX

Role	Responsibility
Responsible	Performs the work of completing the task. Each task has at least one responsible party.
Accountable	Delegates work and performs final review and approval. An individual may be both responsible and accountable, but there is only one person accountable for each task.
Consulted	Is recruited by the other team members for review and consultation. Consulted party is an SME and/or is a user who will be affected by the design.
Informed	Is not responsible for the project but is kept informed of its progress.

Using the RACI matrix, the design engineer's team not only made each team member's roles and responsibilities clear, but also, they were able to clearly show who was accountable for making specific design choices. In this case, the engineer confused by the project was identified as a consultant and therefore was not responsible or accountable. The other team members were able to negotiate and explain their plan without relinquishing their authority to make the change.

At its core, a RACI matrix helps set clear expectations about project roles and responsibilities. Having tasks clearly defined at the beginning of a project prohibits the conflict of having multiple people working on the same task or against one another. When using the RACI matrix, teams are able to encourage individuals to accept responsibility for their work and, in some cases, defer to others when they recognize a skills gap. It is a useful tool to depersonalize the process of selecting the right team members and assigning roles, responsibilities, and accountability more effectively.

XII. Conclusion

When using Ethernet M2M messaging techniques for a process bus trip circuit, it is important to consider all issues related to the design and performance of the underlying EDS infrastructure. ECS design teams need to understand not only how to use Ethernet GOOSE and SV messages but also how to design the system to immediately detect and mitigate Ethernet faults so that those messages reliably convey trip circuit signals.

An overview of trip circuit components and design requirements is provided in this paper as a resource for those needing to know the basics of the ECS protection system. The ECS design team must fully understand and accept their roles and responsibilities to build a safe and effective IEC 61850-based process bus trip circuit. The ECS design needs to perform the following functions:

- Detect and report when a system fault occurs.
- Analyze and report why it occurred.
- Determine whether the design criteria were met.
- Evaluate if the fault is likely to happen again.
- Determine if and how criteria should be changed.
- Evaluate if the ECS and EDS responded as expected.
- Assess whether the failure and resultant behavior was within the accepted risk assessment boundaries.
- Determine if it was exacerbated by design choices and, if so, determine whether these choices should be reconsidered.

ECS design team members must understand their responsibility to deploy technology that is both useful and fit for its intended purpose. EDS and ECS systems are very specific and require highly specialized knowledge. Regulating organizations like NERC are very public about the fact that requirements for failure avoidance, tolerance, and acceptance are a matter of national security. Engineering has made wonderful strides in developing technology that harnesses the resources of the physical world to enable human society and innovations to thrive. But, with this control comes tremendous responsibility. Dramatic examples of dual-primary N – 2 failures in aerospace, aviation, healthcare, and electric power systems show that while unlikely or low-probability conditions may be too remote for mitigation, they are never too remote for concern and communication.

In the same way that real-time control of mission-critical systems requires individual role-based authentication control, there must be role-based accountability control of each design team member. This should be in the form of certification and experience. These design teams must understand and satisfy the resilience and fault-tolerance requirements of the EDS and ECS based on vulnerability analysis and risk assessment. Any gaps must be clearly discussed with the ECS owner and potential mitigation strategies considered in light of cost, schedule, and performance requirements.

Unfortunately, it is human nature for some to be overconfident in their skills, which can, in the worst of cases, lead to having team members who are unskilled, unaware, and maybe even dangerous. Examples discussed in this paper illustrate that these challenges are real and present dangers, and there is no way of knowing how many ECS faults exist, how often, and for what duration they remain hidden. It is essential that engineers identify and mitigate failure modes hidden in plain sight and to minimize human-caused and natural failure modes.

Preoccupation with failure should not be considered a detriment to progress but rather a method for highly skilled professionals to remain continuously vigilant for opportunities to prevent catastrophic failure. Effective designers prevent big problems by being preoccupied with preventing small faults that could inevitably cascade into large failures. Without appropriate design requirements, we are haunted by unknown unknowns. We do not know the boundaries of correct operation, and thus we do not know when operations are outside the safety boundaries.

XIII. References

[1] Interagency Security Committee, "Presidential Policy Directive 21 Implementation: An Interagency Security Committee White Paper," February 2015. Available: cisa.gov.

[2] D. Dolezilek and A. Kalra, "Comparison of Standardized Protocols Available to Replace Substation Copper Field Wiring With Digital Communications," proceedings of the 14th International Conference on Developments in Power System Protection, Belfast, UK, March 2018.

[3] North American Electric Reliability Corporation, "About NERC." Available: nerc.com.

[4] NERC Standard TPL-001-1 – System Performance Under Normal Conditions. Available: nerc.com.

[5] PSRC Relay Trip Circuit Design Working Group, "Summary of Relay Trip Circuit Design," *IEEE Xplore*, August 2002. Available: ieee.com.

[6] IEC 62439-1, Industrial Communication Networks – High Availability Automation Networks – Part 1: General Concepts and Calculation Methods, 2010.

[7] IEC 62439-3, Industrial Communication Networks – High Availability Automation Networks – Part 3: Parallel Redundancy Protocol (PRP) and High-Availability Seamless Redundancy (HSR), 2016.

[8] P. Rubens, "Understanding Fault Tolerance: Securing Your System," *Enterprise Storage Forum*, June 2019. Available: enterprisestorageforum.com.

[9] B. Silbaugh and J. Skiles, "Landing on the Hudson River: Lessons for Health Care," *Physician Executive*, Vol. 36, Issue 3, May–June 2010, pp. 22–29.

[10] A. J. McDonald, J. R. Hanson, *Truth, Lies, and O-Rings: Inside the Space Shuttle* Challenger *Disaster*, University Press of Florida, Gainesville, FL, 2009.

[11] A. Davies, "10 Years Later, Retrace the 'Miracle on the Hudson' Flight," *Wired*, January 2019. Available: wired.com.

[12] D. Dunning and J. Kruger, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology*, Vol. 77, Issue 6, January 2000, pp. 1121–1134.

[13] IEEE Board of Directors, "IEEE Code of Ethics," *IEEE Policies, Section 7 – Professional Activities*, June 2020. Available: ieee.org.

[14] B. Ritholtz, "Unskilled, Unaware and Maybe Even Dangerous: The Man Who Makes Up One Half of the Dunning-Kruger Effect Shares His Thoughts," *Bloomberg Opinion*, March 25, 2020. Available: bloomberg.com/opinion.

[15] S. H. Unger, *Controlling Technology: Ethics and the Responsible Engineer*, John Wiley & Sons, Inc., New York, NY, 1994.

[16] NERC System Protection and Control Subcommittee, "Reliability Fundamentals of System Protection," *North American Electric Reliability Corporation*, December 2010. Available: nerc.com.

[17] North American Electric Reliability Corporation, "Protection System Definition," [n.d.]. Available: nerc.com.

[18] North American Electric Reliability Corporation, "Power Plant and Transmission System Protection Coordination: Technical Reference Document Overview," *NERC Protection Coordination Webinar Series*, May 2010. Available: nerc.com.

[19] North American Electric Reliability Corporation, "Glossary of Terms Used in NERC Reliability Standards," June 2020. Available: nerc.com.

[20] WECC Relay Work Group, "Remedial Action Scheme Design Guide," *Western Electricity Coordinating Council*, January 2017. Available: wecc.org.

[21] North American Electric Reliability Corporation, "ERO Reliability Risk Priorities: RISC Recommendations to the NERC Board of Trustees," February 2018. Available: nerc.com.

[22] North American Electric Reliability Corporation, "State of Reliability," June 2019. Available: nerc.com.

[23] R. Nateghi, "Major Power Outage Risks in the U.S.," *Purdue University Laboratory for Advancing Sustainable Critical Infrastructure*, [n.d.]. Available: engineering.purdue.edu.

[24] Newton-Evans Research Company, "The World Market for Substation Automation and Integration Program in Electric Utilities: 2017–2020," November 2017. Available: newton-evans.com.

[25] P. Parker, D. Chao, I. Norman, and M. Dunham, "Orbiter Assessment of STS-107 ET Bipod Insulation Ramp Intact," January 2003. Available: nasa.gov.

[26] *Columbia* Accident Investigation Board, "Possibility of Rescue or Repair," *Spaceflight Now*, [n.d.]. Available: spaceflightnow.com.

[27] Federal Aviation Administration, "Certification Maintenancy Requirements," *Advisory Circular*, November 1994.

[28] "What Caused Two Devastating Crashes of the 737 MAX Airplane?" *USA Today*, February 5, 2020. Available: usatoday.com.

[29] D. Campbell, "Redline: The Many Human Errors That Brought Down the Boeing 737 MAX," *The Verge*, May 2, 2019. Available: theverge.com.

[30] D. Dolezilek, P. Lima, G. Rocha, A. Rufino, and W. Fernandes, "Comparing the Cost, Complexity, and Performance of Several In-Service Process Bus Merging Unit Solutions Based on IEC 61850," proceedings of the 15th International Conference on Developments in Power System Protection, Liverpool, UK, March 2020.

[31] A. Kalra, D. Dolezilek, G. Vielmini, T. Grigg, and L. D. Carpini, "Using Real-Time Testing Tools to Baseline the Performance of OT Networks for High-Speed Communications," proceedings of the 15th International Conference on Developments in Power System Protection, Liverpool, UK, March 2020.

[32] D. Dolezilek, N. Fischer, and R. Schloss, "Case Study: Dramatic Improvements in Teleprotection and Telecontrol Capabilities Via Synchronous Wide-Area Data Acquisition," proceedings of the 2nd Annual Protection, Automation and Control World Conference, Dublin, Ireland, June 2011.

[33] J. W. Siems, "Disaster Threat and the Dunning-Kruger Effect," *Homeland Security Affairs*, December 2016. Available: hsaj.org.

[34] B. Harned, "How to Clear Project Confusion With a RACI Chart [Template]," *Team Gantt*, September 2019. Available: teamgantt.com.

XIV. Biographies

Allan McDonald received a B.S. in chemical engineering from Montana State University in 1959 and an M.S. in engineering administration from the University of Utah in 1967. He retired in 2001 from ATK Thiokol Propulsion after a 42-year career with the company. He was the director of the Space Shuttle Solid Rocket Motor Project at the time of the *Challenger* accident and led the redesign of the solid rocket motors as vice president of engineering for space operations. He has several patents related to rocket propulsion, has published over 80 technical papers that have been presented in national and international conferences, and received numerous professional awards. He received an honorary doctorate in engineering from Montana State University in 1986, was selected as Montana State University`s centennial alumnus in 1987 by the National Association of State Universities and Land Grant Colleges, is a fellow member and a distinguished lecturer for the American Institute of Aeronautics and Astronautics, and is currently on the board of directors of Orbital Technologies Corporation in Madison, Wisconsin.

Anna Dolezilek has been working in the certification industry for over four years. She joined Tableau in 2017 as a program specialist, working directly with customers on their exam experience and supporting the exam development process. In 2019, Anna became an associate program manager and is now responsible for the certification program customer experience. In her several roles, she has gained experience in certification and working with subject matter experts for exam creation to validate skills and comprehension. She especially enjoys helping others communicate their skills through independently verified solutions.

David Dolezilek is a principal engineer at Schweitzer Engineering Laboratories, Inc., and has three decades of experience in electric power protection, automation, communication, and control. He 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 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.