Case Study: Defining and Measuring Protection Signal Transfer Speed, Latency, and Reliability Within Digital Trip Circuits

Matt Ross and John Bettler, *Commonwealth Edison*

Andrew Sprenger, *Puget Sound Energy*

Jesse Silva, *Southern California Edison*

Austin Wade, David Dolezilek, Mauricio Silveira, and Rodrigo Abboud,
Schweitzer Engineering Laboratories, Inc.

Revised edition released October 2022

Presented at the
49th Annual Western Protective Relay Conference, October 2022

Originally presented at the
75th Annual Conference for Protective Relay Engineers, March 2022

Case Study: Defining and Measuring Protection Signal Transfer Speed, Latency, and Reliability Within Digital Trip Circuits

Matt Ross and John Bettler, *Commonwealth Edison*
Andrew Sprenger, *Puget Sound Energy*
Jesse Silva, *Southern California Edison*
Austin Wade, David Dolezilek, Mauricio Silveira, and Rodrigo Abboud, *Schweitzer Engineering Laboratories, Inc.*

***Abstract*—The recent Newton-Evans study of the North American market for substation, automation, and integration systems reveals that 56 percent of respondents plan to use digital trip circuits to replace their legacy analog hardwired trip circuits. However, it is important to recognize that the top three protection system failures identified by the Electric Reliability Organization similarly affect digital trip circuits. It is also important to anticipate and reduce communications errors, hardware malfunctions, and logic and setting errors, which has accounted for 60 percent of misoperations since 2011. This paper includes the standardized methods to evaluate, test, and measure digital trip circuit designs, and it includes test results to compare speed and reliability of typical designs.**

To understand fault-clearing time, various trip-circuit latencies must be evaluated. The trip circuit time is the sum of the protective relay's decision time, the relay's physical output contact time assertion, and any additional timing required by auxiliary trip circuits and digital communications paths. The trip-circuit time can be estimated using theoretical information from the sources of the different components. However, it is good practice to measure the trip circuit time before placing the system in operation, especially if the trip circuit is composed of nondeterministic technologies, such as Ethernet-based devices and protocols. This paper includes measurements to verify fault-clearing time calculations.

When applied with knowledge guided by experience, serial and Ethernet process bus communications-based digital trip circuits can reduce physical wiring and provide signal transfer supervision, and therefore, safety. Recent technology advances, including faster phasor and time-domain protection algorithms, better zero-crossing detection algorithms, faster open-phase detection, and high-speed output contacts can improve protective relay decision time. Reduced latency and jitter of serial and Ethernet communications trip signals improve the speed and reliability of digital trip circuits.

This paper includes the IEEE and the IEC definitions of latency to describe, stage, and test the speed of several trip circuit designs. Recent activity in the working group tasked with improving IEEE 1646-2004 (*Communication Delivery Time Performance Requirements for Electric Power Substation Automation*) includes several specific latency descriptions. Latencies of several trip circuit designs are physically measured using the white-box and black-box testing of several topologies for both analog and digital trip circuits, including hardwire input/output (I/O), direct, and networked communications cables among relays; auxiliary-tripping devices; merging units (MUs); process interface devices; and breaker control units from multiple suppliers. Measurements gathered during lab testing are included to provide evidence for comparing the various speeds of the technologies and verifying the veracity of associated calculations. IEC 60870 reliability analysis and reliability-centered maintenance strategies are used to compare technologies based on availability, outage time, and operations and maintenance costs. Analysis and maintenance strategies also help eliminate undetected faults and single points of failure, which lead to unplanned repairs.

Results from these tests provide evidence of trip circuit behavior and guidance for engineers to evaluate the impact of device-to-device communications, digital versus electromechanical lockouts, and digital versus analog auxiliary trip circuits with respect to speed and performance.

I. Introduction

Fifty-six percent of respondents to the Newton-Evans study of the North American market for substation, automation, and integration systems plan to replace their legacy hardwired input/output (I/O) systems [1]. Though it is not clear if respondents plan to use IEC 61850 Generic Object-Oriented Substation Event (GOOSE) or Mirrored Bits communications, there is a lot of interest in using both for digital trip circuit designs [2]. However, this report and others that have similar content for international utilities reveal that very little Ethernet has actually been deployed in substations for station bus communications, and almost none is in service for process bus communications and digital trip circuits. As designers contemplate Ethernet for process bus communications, it is important to define and measure protection signal transfer speed, latency, and reliability within digital trip circuits.

Digital secondary systems (DSSs) and process bus communications that replace all or part of traditional trip circuit copper wiring with digital messaging over communications cables must satisfy fault avoidance and tolerance for inherent conditions. Process bus installations experience wind, rain, snow, electrical storms, fluctuations in temperature and humidity, and natural disasters [3]. Digital trip circuit designs must understand and mitigate all of these physical and environmental concerns. As new technologies are adopted, designers must anticipate cyber attacks and other external threats that can affect digital messaging and information sharing that may impact digital process communications.

According to [4], communications-assisted applications need to satisfy numerous service-level specifications, including

many that are related to the performance of the underlying application, such as using a trip signal transfer via GOOSE messaging that must satisfy IEC 61850 Class TT6 (<3 milliseconds). Reference [5] provides advice on network engineering and commissioning. Section 5.3.17 in [5] describes network testing, which recommends that a network's communications design is verified, and that the network performance is tested for requirement compliance during both factory and site acceptance testing [4]. Service-level agreements (SLAs) define the expected performance of the service (e.g., speed and availability) that the service provider offers and agrees to meet. Relays and intelligent electronic devices (IEDs) within a system contain all the test points in the system or have the ability to make all speed measurements, so SLAs are confirmed by operational latency and jitter tests. The founding basis of SLAs that is provided to end users includes communications, acceptance criteria, and other metrics (referenced in related international standards). Mission-critical digital trip circuits that satisfy protection have strict latency, jitter, and availability requirements that are addressed by SLAs. Ongoing fulfillment of these metrics or key performance indicators is necessary for the safe and reliable operation of digital trip circuits. Testing must be performed during design and commissioning to confirm a scenario and must be monitored while in service to understand performance degradation and associated risk. Large deviations must prompt root cause analysis and service improvements.

The Power System Relaying Committee (PSRC) is sponsoring the renewal of [6], which "defines power systems communication delivery time performance requirements to be exchanged within and external to substation integrated protection, control, metering, and data acquisition systems." Ongoing work is addressing newer technologies, including packet switched Ethernet, and creating more precise descriptions of latencies and jitter.

Digital message signal transfer time latency is not directly measurable in IEDs, because most IEDs do not time-stamp when messages enter and leave the IED Ethernet interface [7]. To validate the transit time message delivery duration between IEDs, there are sophisticated and expensive test tools that are used to measure various latencies. However, with knowledge of Ethernet switch and IED processing, important latencies are easily measured with the appropriate accuracy using relays and controllers. It is important to be aware that message analyzer tools running on nondeterministic operating systems, such as Microsoft Windows, create inaccurate time stamps with falsely reported resolution [8].

II. Trip Circuit Designs

Historically, trip circuits consist of a direct current (dc) supply, a protective relay trip contact, and a circuit breaker trip coil. To interrupt the primary circuit, the protective relay closes its trip contact. The dc supply then flows through copper trip circuit wires to the breaker where it then energizes the trip coil, moving the solenoid, tripping the breaker latch, and allowing the stored energy in the breaker mechanism to open the primary contacts and interrupt the flow of current.

Electromechanical (EM) and solid-state relays typically have a limited number of output contacts, so when an application requires a relay to trip multiple breakers, such as breaker failure or bus protection trips, the trip circuit schemes use indirect tripping using diodes or auxiliary relay contact replication for multibreaker tripping [3].

Microprocessor relays typically do not have the same constraint on output contacts and often have enough output contacts to provide direct tripping to multiple breakers and perform ancillary functions, as shown in Fig. 1. These multiple output contacts are isolated, allowing breakers or trip coils to use different supplies from different control power circuits and battery systems without the use of diodes. Direct tripping also eliminates the delay caused by intermediate devices used for trip contact multiplying or lockout relaying, reducing the total time to interrupt a fault. This reduced fault-clearing time can be beneficial to systems to preserve system stability for sensitive and momentary voltage disturbances [9].

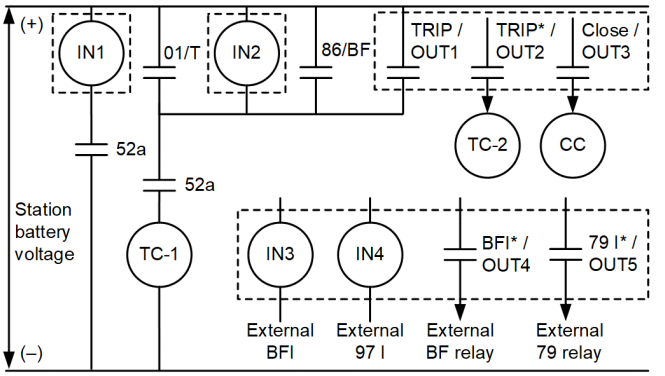

Fig. 1. Drawing of a typical microprocessor-based trip circuit.

Subsequently, microprocessor relays led to the development of communications-based protection signal exchange, such as Mirrored Bits communications or IEC 61850 GOOSE, and the creation of a new paradigm for trip circuit design among DSS devices connected to the process-level primary equipment—often referred to as the process bus.

III. Station Bus and Process Bus

A DSS architecture is usually divided into several categories, including process, bay, and station. Information exchange is often described by two communication buses: process and station [10]. According to Fig. 2, the process bus communications is the middleware interface between the process and bay levels. The station bus interface is the middleware interface between the bay and station levels. The

process and the station bus can also be bound together, or merged, if required. An example of this is downloading and retrieving settings of merging units (MUs) via a single interface that also serves process bus communications or getting data from the MU for human-machine interface (HMI) and supervisory control and data acquisition (SCADA) purposes via the single, merged communications interface.

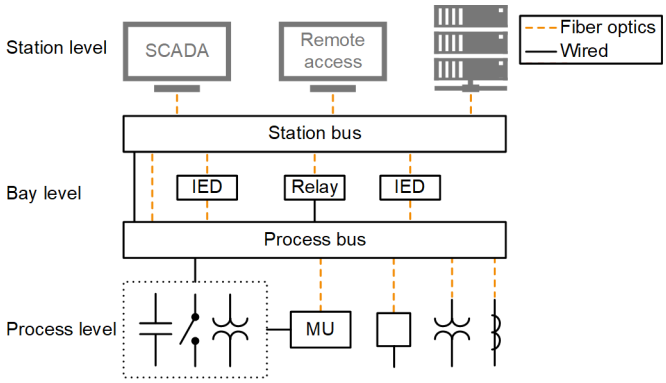

Fig. 2. DSS architecture.

A. Process Bus Overview

The process bus is composed of connections and protocols used to transmit and receive process-level signals via digital messages. Process-level analog signals can be transferred through traditional wired cables using traditional current transformers (CTs) and potential transformers (PTs) via low-energy analog sensors or Rogowski coil low-power CTs. Digitized process-level analog signals can be transferred using fiber optics and protocol technologies, such as IEC 61850-9-2 and IEC 61869 Sampled Values (SV) and time-domain technology protocol [11]. Digital signals are transferred via GOOSE and MIRRORED BITS communications, and time is transferred via Precision Time Protocol (PTP) and IRIG. The data communication flow starts with the process-level devices. The high-voltage equipment I/O are sampled by remote process interface units, including traditional CTs, PTs, digital CTs, digital PTs, wired I/O contacts, MUs, and digital sensors. The data gathering is usually done near the high-voltage equipment sources in the substation yard. Though resilient and robust protective relays are frequently installed in the yard, the power system protection and control (P&C) is performed at the bay level in the substation control house most often. Therefore, the main goal of the process bus is to serve as the communication among devices in the process, but more often, it serves as the communication between devices in the process and bay levels. The process bus communication links require high speed and high performance to achieve protection signal class rates [5], but noncritical communication data paths can coexist in the process bus.

B. Station Bus Overview

The station bus, in turn, is composed of connections and protocols, such as IEC 61850 Manufacturing Message Specification, Distributed Network Protocol, File Transfer Protocol, Telnet, PTP, and IRIG, that transmit and receive the system, engineering, and configuration information and time distribution as well as send operator commands to networked IEDs or relays [12]. In modern substations, IEDs are usually connected to the station bus through Ethernet connections, but serial communications are also used to connect the bay level to the station level. In addition, the station bus can have critical messages, such as trip and interlock signals and time-synchronization signals required for the process bus. Therefore, the performance requirements of the process bus and station bus with GOOSE are equally important and need to be engineered to accommodate high-speed protection signals.

IV. MEASURING SPEED AND LATENCY

The requirements for delivering signals for electric power substation automation system can be found in [6]. While this standard is presently being revised as of 2022, it defines communication delivery times of information to be transferred within and external to the substation integrated protection, control, and data acquisition system. This section uses definitions from [6] to define metrics and test procedures to measure the speed and latency for the substation automation system.

Reference [6] specifies the communication delivery time performance requirements for mission-critical applications. Therefore, the terminology, test procedures, and measurements must be very precisely defined.

The validity of the results and comparisons drawn from an experiment must be supported by the validity of the measurements made. For each test, including speed and latency, the resolution and inaccuracy of the test measurement results must be defined and understood by the end user.

A. The Definition of Latency

Latency is a duration of time. In other words, it is a time delay in a device or system that can be measured as a period of time. Latency measurements are a duration of time measured by a piece of test equipment or the numerical difference between the time stamp of the start of a process and time stamp of the end of a process, or operational stop. Therefore, the accuracy of the latency measurement is related to the accuracy of the clock being used for time measurements and the ability of the IED to synchronize to the reference clock and precisely start and stop the time measurement appropriately. The difference in the IED time-stamp calculation of latency is directly related to the accuracy of the IED to precisely time-stamp values related to the actual start and stop of the time duration.

Operational latency is the time delay between an input (cause) and the desired output (effect) of an operation. It is also known as the aggregate of operations. When measured by test equipment, the operational start must accurately begin the time measurement, and the operational stop must accurately complete the time measurement. When measured as the difference between time stamps in two different IEDs, each device must be synchronized to the same time source and have the ability to accurately apply time stamps for the start and the stop of the test for the latency measurement to be meaningful. Operational latency of a system may be the combined aggregate

of time delays between an input (cause) and the desired output (effect) where the desired output is the input to a second duration, and so on. Operational latency of a workflow can be defined as the sum of time durations and operational latencies of operations within a workflow. Variations in latency are called jitter.

B. Accuracy of Timing Sources

Time accuracy of a measurement is related to the ability of the device creating time stamps to both synchronize to a system time reference and apply an accurate and absolute time stamp to each event. The accuracy of a device to synchronize its clock and detect and time-stamp an operational start or stop should be understood by the test designer to evaluate if the proposed test plan is appropriate. SV are published every 208 microseconds based on [13] for a power system frequency of 60 Hz. Dwell time in a typical Ethernet switch, the time for a message to pass through the hardware at line speed, is 10 microseconds, and GOOSE should travel through a network in under 1 millisecond and should cause a trip within several milliseconds. Test methods must have the appropriate time accuracy to match the purpose of the test.

The required time accuracy of a measurement is related to the nature of the operation and the purpose of the test. Like any engineering project, the scope, cost, and schedule of a test procedure must be understood to design an appropriate test. Laboratory testing with inexpensive and available automation devices acting as test equipment, with inaccuracy up to 0.5 milliseconds, is often adequate to measure the transmission and application time of GOOSE transfer for circuit breaker tripping. Other test cases, including power system faults and the transmission of SV, may require more precise test equipment and associated verification of their calibration.

C. Deterministic Delivery

Deterministic Ethernet message delivery, which includes the transfer of the associated Ethernet packets, is the ability to consistently deliver messages across a specified communications channel with predictable timing from beginning to end or operational latency. The variation between delivery times is called jitter or delay variation. Deterministic means no jitter or delay variation in the delivery latency; however, interaction among system components often creates variations of packet delivery times, also known as packet delay variation (PDV). Designers must be aware of the typical delivery time and the maximum PDV of a proposed communications system to evaluate if the underlying application is resilient enough to withstand the maximum PDV. If not, either the application or the communications system needs to be redesigned. IEC 61850 Protection Class 2 or 3 messages have delivery performance requirements for time and availability, regardless of quantity, frequency, or network configuration [5].

D. Latency, Accuracy, and Determinism in Digital Trip Circuits

Tripping schemes for energy delivery systems have typical protection signal transfer times and maximum signal transfer times, regardless of the signal transfer technology used. For digital trip circuits, the operational times are mission-critical. The applications may suffer due to a larger PDV, and the maximum must be specified as part of the acceptance criteria for the Ethernet communications design.

Communications network latency is the elapsed time between the first bit of an outgoing message (from the source device exiting the physical interface [PHY] and entering the network) and the same bit entering the PHY of the receiving device after passing through the communications network if the complete message is received successfully. While in service, latency may include network reconvergence and delays associated with bandwidth saturation and message prioritization.

Network jitter is the variance between network latency time duration among numerous delivery instances measured from the same message over the same network conditions.

E. IED Contribution to Latency and Jitter

The IEDs at each end of the communication channel in digital trip circuits must be included in the overall channel analysis for deterministic communication. When transferring data, the IEDs participate in or complete the intended process. To guarantee the application is deterministic, the complete process must be considered. When analyzing digital trip circuits, we must account for the delays due to IED communications processing, message encoding and decoding, and the IED process intervals. We must also understand that both the publishing IED and the subscribing IED contribute to these times [14].

Information transfer operational latency may be defined as the sum of operational latencies within the workflow of the following:

- Data change detection in the publishing IED.
- Data change verification and data set generation in the publishing IED (may include strategic delay to coordinate message delivery and reception).
- Message creation and encoding, publication onto physical communication interface at publisher.
- Message transfer across communications media.
- Message receipt onto physical communication interface at the subscriber IED.
- Message verification, and decoding by the subscriber.
- Data parsing and mapping into virtual data placeholders by the subscriber.
- Operation on the data by the subscriber.

Operational latency measurements that are defined to include an input or output within the processing of a device triggering or triggered by a signal message cannot be measured by a black-box test and cannot be precisely consistent among multiple test fixtures. True operational latency must be tested using the actual devices and communications paths fully configured to perform protection, control, monitoring, and communications. However, similar devices, communications networks, and network loading may be considered adequate to gain representative results.

Device data processing latencies may be influenced by the following:

- Microprocessor operations
- I/O buffers
- Interdependencies between subroutines

Device communications interface latencies may be influenced by the following:

- Connectors
- Microcontrollers
- Transceivers
- Oscillators

Device contact I/O latencies are influenced by the design and components of each device interface and their environment. Device contact I/O latencies may be influenced by the following:

- Shock
- Vibration
- Humidity
- Temperature
- Thermal expansion

Additionally, these environmental influences may change over time due to stress and corrosion that ultimately affect contact reliability.

F. Additional Considerations for Time, Speed, and Latency Measurements

- Operational start and operational stop test measurements must be time-stamped to support operational latency measurements and calculations. The time accuracy needs to be appropriate for the purpose of the test.
- Test measurements require that the device time stamps be precisely accurate to the absolute time of the operational I/O event. Accuracy must be calibrated and provided as a measure of possible error. Calibration by triggering all test devices from a single event and comparing the resultant time stamp may be adequate for laboratory and field test calibration. Calibration of purpose-built, high-precision test equipment must be provided by and confirmed by the manufacturer.
- Test measurements of the same event, made by more than one device, require that the devices be precisely time-synchronized. This clock time accuracy must be confirmed by the manufacturer and provided as a measure of possible error. Time accuracy of device time synchronization and time stamping must be confirmed to be appropriate for each specific test. The operational time-stamp error is related to the device clock time-stamp error and the device event time-stamp error.
- It is necessary to understand the time-stamp accuracy and error and understand clock time accuracy and error when combining test measurements from more than one device.
- Some P&C operational latency measurements for process bus and digital trip circuits may require microsecond accuracy and appropriately small error.
- Operational time-stamp accuracy and errors affect the precision of recording operational starts, stops, and latency measurements. This influences the applications that they are suitable for because of the differences in speed and precision of GOOSE, PTP, SV, and input and output contacts.
- Black-box test measurements are done by a device separate from the devices under test (DUT).
- White-box test measurements are made by the DUT.

Fig. 3 is the time requirement with the transit time illustrated as a network of communications devices between the IEDs or in the middle of the communications path. Various middle boxes are used in process bus and station bus communications for serial time-division multiplexing and packetized Ethernet data message transfer [4] [15]. Reference [16] describes the Rapid Spanning Tree Protocol (RSTP) in IEEE 802.1w as a recoverable technology that automatically detects Ethernet faults and restores data flow. Software-defined networking (SDN) is also compatible with the IEEE 802.1w methods of packetized Ethernet.

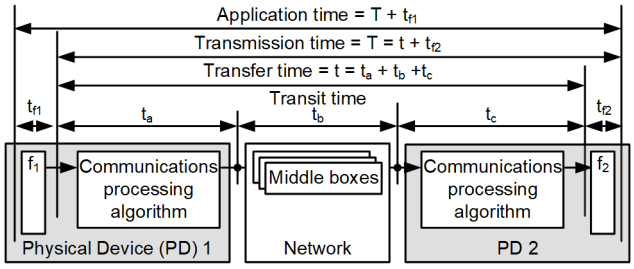

Fig. 3. IEC 61850-5 transfer time.

G. "How Much Additional Time Will You Accept?"

At the end of the last century, the Electric Power Research Institute and IEEE were working to document Ethernet message delivery and application times. As stated in [15], considerations of digital messaging included that data are being updated and transmitted at the analog-to-digital conversion sampling rate of the sending end device for use by a digital relay, and thus, any message delay affects the protection performance. Each message between Protection Device 1 and Protection Device 2 and each message from a protection device to a digital breaker control collocated with the switchgear to replace copper wire connections from a digital relay to a circuit breaker need to be delivered in less than 2 milliseconds. As this interface may be used for tripping, and since any local-area network (LAN) will be slower than a copper wire connection, the question became "HOW much additional time will you accept?" [15]. The document further anticipated that in a substation LAN environment, the protective relay could issue a trip command on the LAN, which is received by each of the breaker controllers for breakers being commanded to trip, without the need for any auxiliary-tripping relays. Since these auxiliary relays, and their operating times, could be eliminated in a LAN environment, the relay engineers agreed to a

requirement that such messages be delivered—application to application—in 4 milliseconds or less [15]. At the time, high-speed EM auxiliary-tripping relays that operated in 4 milliseconds were available.

This became one of the earliest descriptions of application-to-application protection signal transfer time: 4 milliseconds or less between a microprocessor relay to a microprocessor-based breaker across a LAN. The 4-millisecond time requirement is defined as $t_a + t_b + t_c$ in Fig. 3. This is the application-to-application time. It starts after PD 1 detected a fault and issued a trip output (f_1) to the communications processor in the IED. PD 2 is located at the breaker control. Its communications processor delivers the trip signal by the end of time (t_c) to the breaker trip controls, to trip the circuit breaker. This is perhaps the earliest documentation of the application-to-application time of an Ethernet message from a digital relay (PD 1) to a digital circuit breaker (PD 2). Twenty-two years after this document was written, digitally controlled circuit breakers are not commonplace, so it is not appropriate to eliminate the auxiliary relay in this example and consider PD 2 as part of the breaker. Instead, we use a digital breaker control as PD 2, which is hardwired to the breaker. PD 2 can be a logical device in a multifunction digital relay or programmable automation controller and translate the digital protection signal within the digital message into a physical breaker contact.

Assuming that PD 1 is the sending IED, the function, which is identified as latency time (t_{f1}), reflects the elapsed time for the information generated to be forwarded to the communications processing algorithm. These functions in an IED include performing the analog-to-digital conversion of hardwired field signals received on the IED data acquisition interface and values calculated by the IED protection and automation logic.

As illustrated in Fig. 4, processing hardwired contact inputs as part of f_1 input processing consists of four basic steps:

- The field wiring signal detects an open or close signal on the contact input.
- The input creates an energized or de-energized optoisolated input signal.
- The IED performs debounce timing on input signals.
- The associated internal logic variable states are updated to reflect signal status.

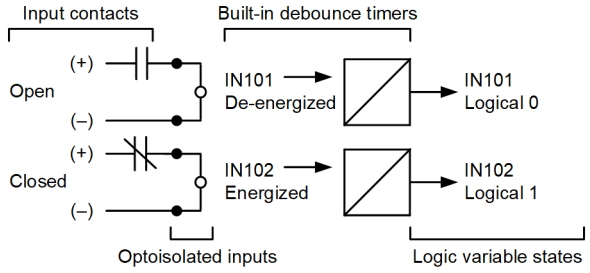

Fig. 4. Contact input function (f_1) of PD 1.

The communications processing algorithm in the sending IED PD 1 from Fig. 3, identified as latency time t_a, consists of four basic steps, including:

- Generating data sets that represent internal analog and logic variable state values as part of the communications application.
- Encoding message content, which includes generating header and trailer and performing a cyclic redundancy check (CRC).
- Processing the assembled message through the communications stack, which includes converting the message information components into an Ethernet or serial message.
- Presenting the message to a PHY.

The communication processing algorithm in the receiving IED, identified as latency time t_c in PD 2, has the following basic steps:

- Receive the message from the PHY.
- Disassemble the serial or Ethernet message components, check the CRC, and process the complete message through the communications stack.
- Decode message content.
- Process data set contents and identify logic variable and analog states as part of communications application.

Processing hardwired contact outputs, identified as part of f_2 output processing and as latency time t_{f2}, reflect the elapsed time between moving the newly calculated or communicated changes to digital variables and analog states into the IED logic data objects. This, in turn, triggers changes to the associated contact outputs. As illustrated in Fig. 5, processing hardwired contact outputs, as part of f_2, consists of three basic steps:

1. Updating associated variable states to reflect internal logic state.
2. Energizing or de-energizing output coils based on variable state (set or unset).
3. Opening or closing output contacts to reflect changes to the coil energization state.

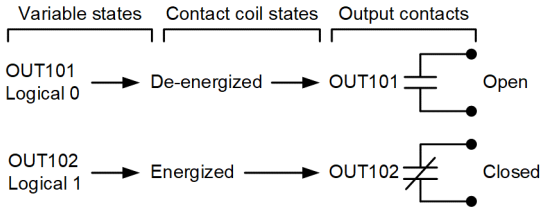

Fig. 5. Contact output function (f_2) of PD 2.

Although the start and stop of each of these operational latencies can be described precisely, most IEDs do not time-stamp the transition from t_{f1} to t_a or from t_c to t_{f2}. The previously mentioned transitions occur at varying times within the executable code that processes the IEDs messages, logic, and hardwired interfaces. Changing the IED firmware to provide a time stamp at these transitions from t_{f1} to t_a or from t_c to t_{f2} may delay the data processing. The delay is caused because the IEDs perform different tasks in different sequences at each process interval, and therefore, is not able to provide a

consistent start or stop time. In summary, changing the IED application firmware to provide test points within the sequential processing may affect the underlying application and its operational latency. This concept is known as the observer effect, in which the underlying system is disturbed by the act of observation.

The imprecision of transition timing from t_{f1} to t_a or from t_c to t_{f2} is illustrated in Fig. 3 as the undefined space between the functions and the communications processing algorithms in PD 1 and PD 2.

Internal processing latency durations are provided in product documentation. For devices that sequentially perform data acquisitions and generation logic, f_1 and f_2, as shown in Fig. 6, the physical input and logic variable data transitions, contact inputs, contact outputs, and protection logic transitions are all given the same time stamp at the end of the process interval. The exception is the high-accuracy signal and time stamp recorded in the high-resolution oscillography.

Each processing interval

1. Monitor field inputs
2. Perform analog filtering
3. Perform A/D
4. Perform digital filtering
5. Calculate phasor angle and magnitudes
6. Scan digital inputs
7. Create high resolution oscillography
8. Accept update from incoming communications processor algorithm (ICPA)
9. Run protection logic 8 times per cycle
10. Run user logic 8 times per cycle
11. Calculate logic variable states
12. Perform metering
13. Store information
14. Update logic variable states and Sequential Events Recorder (SER)
15. Update virtual outputs and energize physical outputs
16. Update outgoing communications processor algorithm (OCPA)

Fig. 6. Overview of typical sequence of data processing performed each IED process interval.

During the f_1 and f_2 functional process intervals (as illustrated in Fig. 3), it is not known when state transitions occur, even though they are time-stamped at the completion of the cycle. In fact, based on the sequential processing, some transitions are not detected or calculated until the following cycle. Therefore, using a time stamp of a phasor-based device that runs the process interval every 1/8th of a power system cycle creates time stamps with inaccuracy between 0 milliseconds and the duration of a full process interval (2.08 milliseconds at 60 Hz and 2.5 milliseconds at 50 Hz).

Using time stamps of a time-domain-based device that runs the process interval every 2 milliseconds regardless of power system frequency creates time stamps with inaccuracy between 0 milliseconds and the duration of a full process interval (or in the worst-case scenario, 2 milliseconds). Time-based devices that process edge transitions remove the inaccuracy involved in input debounce. Time-domain logic devices, like the ones used for research to support this paper, have even more precise and accurate time stamping, as illustrated in Fig. 7.

For testing in support of the paper, a high-precision output trigger was verified to change the state of the contact output with a median time of 350 nanoseconds after a precise top of second event. Although it is important to understand the precision of the test and recording devices, for these tests, this inaccuracy of 350 nanoseconds may be considered negligible.

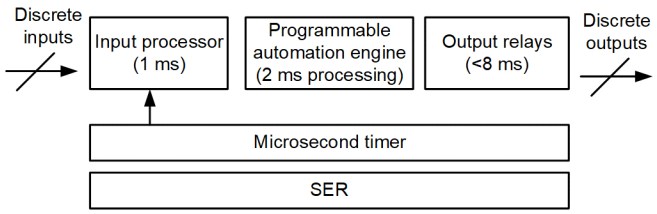

Fig. 7. Block diagram of microsecond timing of discrete inputs in time-domain IED with a 2-millisecond process interval.

The second test device recording contact inputs with high precision is a time-domain microprocessor controller with microsecond SER resolution. Using this precise trigger, the time-domain test device illustrated in Fig. 7 was verified to detect the input, verify the time, and create an SER with a time stamp that has the median time of 9 microseconds after a precise top of the second event. Therefore, the test recording device is verified to have a typical contact input detection inaccuracy of 9 microseconds for the rising edge, 54 microseconds for the falling edge, and one microsecond resolution in the SER. Since both devices are time-synchronized with high accuracy, the time-domain test device is confirmed to record state changes to contact inputs with inaccuracy between 9 and 54 microseconds.

However, the microsecond accuracy provided by the recording device is done using raw Relay Word bit information that detects and time-stamps the input assertion with microsecond accuracy, which can overload the memory when receiving multiple input signals. The use of filtered bit information better reflects real-world applications, and through similar precision tests, it was found that the error in the time stamp of such input assertions due to processing cycles is, at most, 500 microseconds. This is satisfactory for most use cases of event recording.

The SER of a sender IED logic variable transition of a data object is delayed until the completion of the process interval in which it is detected. When these data objects are in the data set of a message processed in the communications processing algorithm triggered at the end of the process interval, the message is published several microseconds later. This interaction of f_1 and the communications processor algorithm makes it possible for the SER of the data object logic variable transition to act as a proxy for the time stamp of the publication of the digital signal message itself.

For digital trip circuits, application speed and latency timing include all aspects of the act of communicating protection signals and protection signal digital messages as well as the generation or consumption of the signals or digital signal messages.

It is often appropriate to consider that these time stamps of virtual and physical output transitions, recorded as SER in f_1, are directly related to, and act as a proxy for, the operation time of publications from the OCPA in Fig. 8.

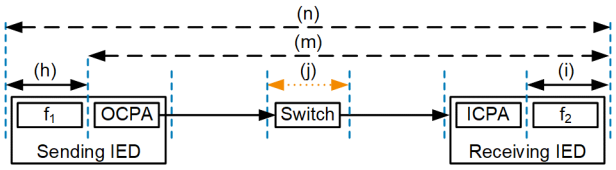

Fig. 8. Overview of various digital control systems with precise operational latency (start, stop, and duration).

These outgoing messages may be either MIRRORED BITS communications, a time-domain protocol, or GOOSE. In Fig. 8, (h) and (i) represent the time duration of the f_1 and f_2 operation cycles, respectively. Latency (j) is the dwell time of the message as it passes through a switch. This is often 10 microseconds for modern Ethernet switches if there is no traffic inference. Latency (m) is the operational latency of the application to detect and time-stamp a change of state at the sender contact input and create a corresponding contact output at the receiver. While latency (n) is the full operational latency between the instant of the trigger signal change of state and corresponding contact output at the receiver.

Specifically, latency (h) represents the operational latency of the sender IED detecting a contact input as a signal value function (f_1), which includes the input contact, optoisolated input, debounce filter (if any), and conversion to logic variable status. The operational latency begins when the energy level in the input contact surpasses a detectable threshold in the sending IED and ends when the logic variable changes state in the sending IED.

Latency (i) represents the operational latency of the receiver IED function (f_2) to modify an output contact to represent a logical change in signal value. Logical 0, output contact coil de-energized, output contact is opened, and logical 1, output contact coil energized, output contact is closed. Latency (i) begins when the logic variable changes state in the receiving IED and ends when the output contact is done changing to the open or closed position.

Latency (j) represents the operational latency of the dwell time, which includes the signal message passing through a switch. It begins when the first bit of the message ingresses the associated PHY and ends when the first bit of the message egresses the associated PHY.

Latency (m) begins when the logic variable changes state in the sending IED and ends when the receiving IED output contact has completed changing to the open or closed position. This operational latency (m) time may include latency of communications devices between the IEDs if present in the data path. This latency, also known as transmission time, emulates the time duration of using energy to transmit an auxiliary trip signal output from an EM sending device to an EM receiving device and the associated output contact.

Latency (n) represents the time duration of the entire process of a field signal state change being detected and transferred from the sender and acted upon in the receiver. It begins when the input energy level changes on the contact input of the sending IED and ends when the receiving IED output contact has completed changing to the open or closed position. This operational latency (n), also known as application time, may include latency of communications devices between the IEDs if present in the data path.

V. LATENCY TESTING

The test case setups used to empirically determine some of the relevant latencies discussed in this paper are detailed in the following subsections. The goal is to compare the delays in the assertion of contact outputs of different devices through both a DSS and a traditional digital relay. A traditional RSTP switch was used for the station bus in both setups, while an SDN switch was used for the process bus in the DSS setup. However, the lack of intensive network traffic means that the dwell time is negligible through both switches (under 10 microseconds).

A. Test Case 1 Setup—Binary Contact Trigger

The target comparison was first made by means of a microprocessor relay that did not involve MUs, as shown in Fig. 9. An energized contact output closed at a specific time by a high-precision relay test set that served as the trigger for all subsequent signals in the setup. The signals went into the high-speed EM tripping relay (94X-HS), a contact input of a microprocessor relay (Relay A), and into the recorder to serve as the reference test initiate time stamp. Relay A then closed both a high-speed (HS) and a standard contact output based on this signal, both of which are operational stop (contact input) measurement points on the recorder. A second HS output is also used to activate the EM lockout relay (LOR), while a GOOSE message is sent to both a digital relay (Relay B) and a protection speed automation controller to trigger their own contact outputs. Additionally, Relay A sends a MIRRORED BITS communications message to Relay C and a time-domain trip message to the time-domain MU to trigger them to operate contact outputs. These output signals are also measured as operational stop (contact inputs) by the recorder. The test was repeated ten times, which illustrates the operational latencies measured for Test Case 1 setup while the resulting averages of the recorded time stamps were laid out in the graph seen in Fig. 10.

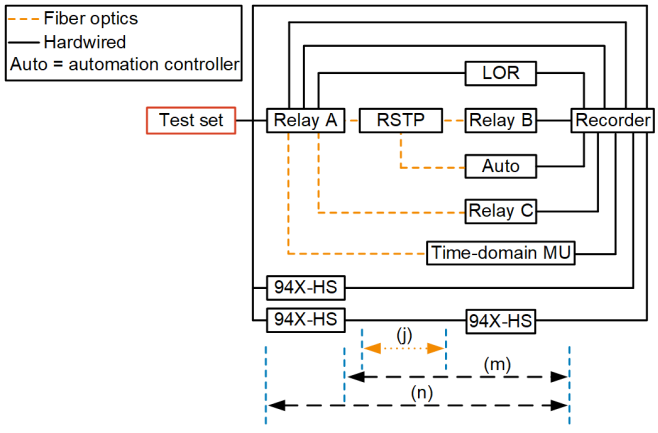

Fig. 9. Test Case 1—Traditional relay setup with a binary trigger and a timing diagram for test scenarios.

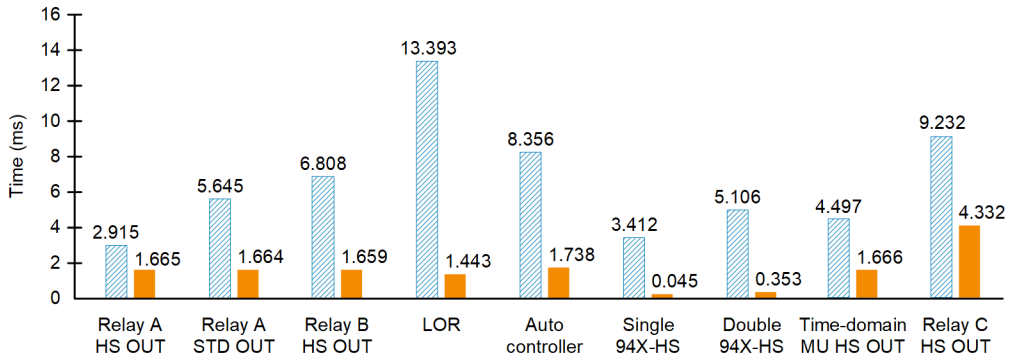

Fig. 10. Average operational latency (n) (blue stripes) and jitter (solid orange) for Test Case 1.

1) Scenario 1 Flow

The operational start is a contact output signal from the left test set to the microprocessor Relay A contact input. The signal is then converted to a GOOSE message and sent through the SDN switch. Relay B receives the message and converts it into a relay contact output. This output is recorded as an operational stop (contact input) by the recorder device's contact input.

2) Scenario 2 Flow

The operational start is a contact output signal from the left test set to the microprocessor Relay A contact input. The signal is then converted to a GOOSE message and sent through the SDN switch. The automation controller received the message and converts it into a controller contact output. This output is recorded as an operational stop (contact input) by the recorder device's contact input.

3) Scenario 3 Flow

The operational start is a contact output signal from the left test set to the microprocessor Relay A contact input. The signal is then sent to a contact output and energizes the trip coil of the LOR. The LOR then operates and converts the signal to the contact output recorded as an operational stop (contact input) on the recorder device's contact input.

4) Scenario 4 Flow

The operational start is a contact output signal from the left test set to the contact input to the left EM high-speed tripping relay (94X-HS). The relay converts the signal to a contact output received as contact input at the right 94X-HS. The relay converts the signal to the contact output recorded as an operational stop (contact input) on the recorder device's contact input.

5) Scenario 5 Flow

The operational start is a contact output signal from the left test set to the microprocessor Relay A contact input and is simultaneously recorded as the recorder device's contact input. The signal is then converted to a time-domain trip message and sent through the direct fiber-optic connection to the time-domain MU. The time-domain MU receives the message and converts it into a contact output. This output is recorded as an operational stop (contact input) by the recorder device's contact input.

6) Scenario 6 Flow

The operational start is a contact output signal from the left test set to the microprocessor Relay A contact input and is simultaneously recorded as the recorder device's contact input. The signal is then converted to a MIRRORED BITS communications message and sent through the direct fiber-optic connection to Relay C, which receives the message and converts it into a contact output. This output is recorded as an operational stop (contact input) by the recorder device's contact input.

Latency (j) for Scenarios 1 and 2 is 10 microseconds and potentially negligible for this testing. However, it is important to keep it in mind, because large and asynchronous network paths affect test and application results. Scenario 3 does not include a device latency between relays.

Latency (m) for Scenario 1 and 2 operational latency begins when the contact output of Relay A changes state and is recorded as a contact input in the recorder on the right. This SER time stamp acts as a proxy for a time stamp of the logic variable change of state in the sending Relay A acting as an intelligent MU (IMU) IED. For Scenarios 1 and 2, operational latency ends when the recorder time-stamps an SER of the associated contact input related to contact outputs of Relay B or when the automation controller contact output changes to the open or closed position. This operational latency (m) time includes latency (j) of 10 microseconds. This is a logical assumption because no other network traffic is present, which may otherwise prolong latency (j). The operational latency (m) is calculated as the difference between SER time stamps.

Latency (m) for Scenario 3 begins when the contact output of the left 94X-HS relay changes to an open or closed position and is recorded as a contact input in the right recorder. This SER time stamp represents the completion of the internal mechanical processing of the left 94X-HS relay. For Scenario 3, operational latency ends when the recorder time-stamps an SER of the associated contact input related to the contact output of the right 94X-HS relay. The operational latency (m) is calculated as the difference between the SER time stamps.

The operational latency (n) in Scenarios 1, 2, and 3 begins when the contact output of the test set changes state and is recorded as a contact input in the right recorder. The operational latency (n) in Scenarios 1, 2, and 3 ends when the recorder time-stamps an SER of the associated contact input related to contact outputs of Relay A. The automation controller, or the right 94X-HS relay, changes the contact output to an open or closed position. This operational latency (n) for Scenarios 1 and 2 includes a latency (j) of 10 microseconds. Again, this is a logical assumption because no other network traffic is present that may otherwise prolong latency (j). The operational latency (n) is calculated as the difference between SER time stamps.

As shown in Fig. 10, the average operational latency is lowest for Relay A directly tripping, while the EM lockout had the longest operational latency. As noted in [9], the ability to directly trip can improve tripping speeds. Our results also show that for applications where hardwired direct tripping is not possible, using digital trip circuits can be nearly as fast as a standard speed output in modern microprocessor relays. The operational jitter in Test Case 1 is very consistent for all scenarios, and the most jitter is attributed to the IED processing interval.

B. Test Case 2 Setup—Fault Applied to MU Trigger

Test Case 2 setup is then changed to explore similar behavior with different process bus solutions. The diagram in Fig. 11 details the two options used: a time-domain MU with a point-to-point connection to an MU and an SV MU connected to an SV relay by means of an SDN switch.

For both scenarios, the high-precision relay test set is used to inject currents into the MUs, and instantaneous overcurrent protection served as the trigger (instead of a binary input). To accurately time these triggers, the recorder measures a dc voltage signal, which is applied by the test set at the same time that the current is largely increased. The EM contact output on the test set cannot be used for this test because it takes approximately 6 milliseconds to assert, while the analog signal state change can initiate the trip process in under 1 millisecond. This test was also repeated ten times for each setup, and the results are displayed in Fig. 12.

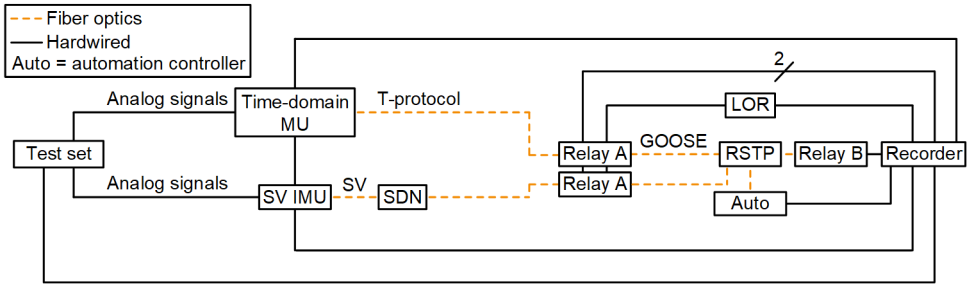

Fig. 11. Test Case 2—DSS setup with an overcurrent protection trigger.

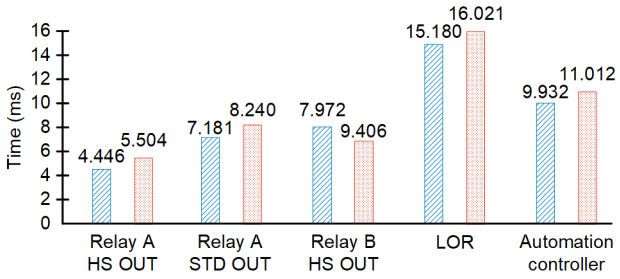

Fig. 12. Operational latency (n) for time-domain link technology (blue stripes) and SV (red dots) for Test Case 2.

1) Scenario 7 Flow

The operational start is a fault current signal from the test set as an analog input to an MU. The signal is then converted from an analog signal to a time-domain analog sample message sent directly to Relay A. Relay A detects the fault and converts the signal to a contact output recorded as an operational stop (contact input) by the recorder contact input. An SER of the operational start is triggered by the recorder device based on the transition of a dc square wave that is produced by the test set at the same instance as the fault current.

2) Scenario 8 Flow

The operational start is a fault current signal from the test set as an analog input to an SV IMU. The signal is then converted from an analog signal to an SV (IEC 61850 SV) message. The message is then sent through the SDN Switch 1 to Relay A, which detects the fault and converts the signal to a contact output recorded as an operational stop (contact input) by the recorder contact input. Similar to Scenario 5, the operational start is triggered by the recorder based on the transition of a dc square wave that is produced by the test set at same instance as the fault current.

3) Scenario 9 Flow

The operational start is a fault current signal from the test set as an analog input to an SV IMU. The SV IMU detects the fault and converts the signal to an SV IMU contact output recorded as an operational stop (contact input) by the recorder contact input. The operational start is the same as Scenario 7 and 8.

Scenarios 7, 8, and 9 operational latency (n) begins when the analog output of the test set introduces fault current to the analog input of the MU and IMU at the same instant that it triggers a dc square wave output, which is recorded as a contact input in the recorder on the right. For Scenarios 8 and 9, operational latency (n) ends when the recorder time-stamps an SER of the associated contact input related to contact outputs of Relay A. For Scenario 8, operational latency (n) ends when the IMU triggers a contact output, which is recorded as a contact input in the recorder. This operational latency (n) for Scenario 9 does include latency (j) of 10 microseconds. The operational latency (n) is calculated as the difference between SER time stamps.

After comparing the results from Fig. 12 and with the previous results in Test Case 1 from Fig. 10, we can see that adding the process bus in Test Case 2 leads to a slightly longer operational latency. Similar to Test Case 1, the operational latencies for all of the scenarios in Test Case 2 are very consistent.

VI. Design for Reliability

Availability of the individual components of a system, such as a digital trip circuit, is a common metric for component reliability. However, it is perhaps more important that the system has available digital components and communication links to react to an energy delivery system (EDS) fault and can perform mitigation (e.g., trip a breaker as part of the energy control system [ECS]). Design for reliability is not unique to the power delivery industry and many tools are available to support improving availability of trip devices and communications [17].

According to the North American Electric Reliability Corporation (NERC), it is critical that the DSS quickly detect faults and then automatically communicate commands to controllable devices in the ECS. Trip circuits perform critical functions of signaling breakers to isolate faults. ECS design teams for DSS with a process bus have the responsibility of understanding if their system requires NERC N-1, N-1-1, or N-2 performance. They must also understand any other appropriate requirements and how these requirements and design choices can impact the system performance. Failure to fully understand the application and chosen technology can lead to a critical failure occurring from a chain reaction of multiple small and seemingly insignificant failures. “An undetected failure of an ECS communications system component is dangerous and 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” [1].

A. Value Engineering

The value engineering process from ASTM E1699-14 is used to improve operations, reduce costs, and substitute current materials and methods with ones that are less expensive while preserving or improving functionality, reliability, and serviceability based on performance-based specifications [18]. This is a useful framework for replacing mature and well understood field hardwire engineering methods with process-level digital devices and digital communications.

B. Limited Vulnerability Design (LVD)

The U.S. Department of the Army’s LVD is used to ensure maximum functionality and performance of the ECS, including digital trip circuits, by preventing or reacting to natural and manmade failure events [19]. These practices create designs that protect against malicious attacks intended to interrupt the system and include the following steps:

- Identify and investigate design gaps.
- Recognize vulnerabilities associated with design gaps.
- Recognize risks associated with vulnerabilities.
- Limit vulnerability based on cost, schedule, and performance design choice.

C. Electronic Safety-Related Systems

Reference [20], IEC 61508, *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems*, provides guidance to improve the designs based on programmable electronic devices, including the following:

- Reduce the probability of the systems remaining in service failing.
- Differentiate between dangerous detected (DD) failures and dangerous undetected (DU) failures.
- Avoid creating a new system that is less reliable than the one it replaced. For example, DD digital message failure with immediate detection is more reliable than copper wires installed with DU failure modes.

- Use design tools and processes to make evidence-based decisions. For example, many IEEE and IEC standards recommend using fault tree analysis (FTA) as a tool to evaluate and compare the availability of different designs.

D. Reliability and Availability Using FTA

FTA is used to understand, measure, and compare systems based on the probabilities of component failures and their interdependencies [21]. For digital trip circuits, FTA is used for reliability and availability analysis that is based on failures that are measured by product quality. Using the device rate of failure and unavailability, systems and their components are evaluated based on the probability that they become unavailable to perform functions vital to system operation. The failure rate of a device, commonly expressed as the mean time between failures (MTBF), is the number of failures expected over a period of time. Availability and unavailability are usually expressed as probabilities [22]. Reliability and availability failure rates should be based on results of field data for proper evaluation.

1) Availability

The time to detect and repair a failure, mean time to repair (MTTR), divided by the MTBF provides a unitless value representing unavailability [23]. Wiring terminations fail over time, and [24] predicts that ten times the failure rate of the newly tested wiring is a realistic field failure rate, yielding a 5,000-year MTBF per wiring point. The failure data for terminal block connections [25] predict an MTBF of more than 4,400 years. Electric utility practice does not generally include automated detection and reporting of wiring failures, so the average detection time influencing the MTTR is conservatively predicted to be half of the time between periodic manual testing. Assuming an average testing interval of two years and a period of two days to make the repair, the unavailability of copper contact wiring with an MTBF of 5,000 years is 200.

The average unavailability of a fiber-optic connection to support digital messaging, as an alternative to copper wiring, has a typical value of 1.1 [26]. This unavailability is low, and is therefore much more available than copper contact wiring with an unavailability of 200, due to the quality of manufacture of the fiber-optic system used and the immediate detection of a failed fiber link, which automatically prompts corrective action.

2) Fault Tolerance

Fault tolerance is the ability of a system to perform its intended function even after the failure of any single component. This is often referred to as N-1 or N-2 tolerance. N is the total of all components required for a system to perform a function, and N-1 and N-2 are the number of components that can fail without the system failing. Trip circuits are critical systems that require a minimum fault tolerance of N-1. However, even with N-1 fault tolerance, the system must immediately detect and self-announce failures so that system operations can repair the failure before it becomes an N-1-1 failure [1].

3) Using the FTA Method to Evaluate Cyber-Attack Tolerance

Using the same approach as event tree analysis, attack tree analysis (ATA) uses the probabilities of a device's lack of cyber-attack tolerance that can lead to a system failure. ATA is used to understand, measure, and compare systems based on the probabilities of devices succumbing to cyber attacks and related component interdependencies [12]. "ATA provides an event tree method to quantify the risk of a successful attack to a system by combining all cyber vulnerabilities as branches" [12]. Similar to FTA, ATA allows designers to create and quantify metrics for cyber vulnerabilities to compare, design out, or mitigate. Designers use ATA to analyze the probability of success for each threat to the digital trip circuit to understand the vulnerabilities associated with private process bus communications or shared communications with and without secrecy protocols.

E. Cascading Impacts Can Lead to Dangerous Undetectable Faults

Failure of various international standards organizations to design for reliability over the years has had a dramatic impact on digital trip circuits. One example of the cascading impact of incomplete design choices is the IEC 61508 DU nature of a failed IEC 62439-3 (*Industrial Communication Networks – High Availability Automation Networks*) GOOSE or SV duplication based on the Parallel Redundancy Protocol (PRP). This proprietary method, described in Part 3 of [16] as repairable, puts the trip circuit design at risk; and worse, it does not self-announce failure, so operators remain unaware that the system is at increased risk [16]. High-availability seamless redundancy (HSR) is another duplication scheme, mistakenly named redundancy, that changes the GOOSE, SV, PTP, and other messages so that they are no longer Ethernet. Proprietary HSR creates DU failures and disables the use of Ethernet-based switches, monitoring devices, and cybersecurity methods.

Operators may continue to believe that their digital trip circuit system is operating as designed to have N-2 or N-1-1 availability, when in fact, it has been reduced to an N-1 single point of failure. With a failed and unsupervised copper wire interconnection or failed and undetected PRP duplication, the receiver relay continues to perform communications-assisted protection without raising an alarm, likely giving the end user a false sense of security. This DU failure remains undetected and unannounced until full system failure or someone detects the unannounced failure and makes a repair.

If a LAN A or LAN B PRP cable to an IED is broken, link status fails and is recorded as a self-announced failure in the IED. This cable failure creates but does not indicate a PRP subscription failure. Spanning tree networks do not self-announce data path failures between switches to the end devices, so it is possible to have a message path failure of one PRP data flow remain a DU failure. It is also true that methods can be easily developed that are not described within the standard to detect a failed GOOSE or SV PRP subscription, but they would be proprietary and not available from multiple suppliers. Numerous system level diagnostics exist to learn the

status of switch connections and port link status, but none of these represent the subscription status at each IED receiver.

F. The Impact of Digital Trip Circuits

The history of unintended consequences of using IEC 62439 PRP with DU failures prevents designs from preserving reliability (as explained in value engineering), performing risk mitigation (as described in LVD), and creating DU failure modes contrary to IEC 61508 methods. The impact on digital trip circuits is summarized as follows.

Initially, the utility communications architecture was harmonized into the IEC 61850 and IEEE 802.1Q method of virtual LAN (VLAN) tagging using the data link layer VLAN identifier (VID). This architecture was introduced to deliver GOOSE messaging at Layer 2 of the Open System Interface (OSI) network. The IEEE VLAN tag is used in Ethernet network engineering to appropriately configure the traffic navigation settings of publishers, subscribers, and switches.

Devices that implemented IEEE 802.1Q are fully capable of publishing separate and truly redundant GOOSE messages out to two relay ports, into two Ethernet LANs, and to two ports on the subscriber.

This IEEE 802.1Q method is compatible with automatically reconfigurable Ethernet LANs, based on Part 1 of [16], IEEE 802.1w, RSTP, and SDN. These both create high-availability automation networks via the automatic detection of Ethernet failures and automatic corrective action to restore communications using the SDN or spanning tree algorithm and RSTP in the switches. Ethernet path faults are detected and isolated, and network traffic is rerouted without human interaction.

Devices that do not fully implement IEEE 802.1Q VLANs are not useful for mission-critical applications requiring redundancy, but they are useful in single LAN designs for SCADA and demonstrations of GOOSE. However, without full VLAN tag support, these devices are not capable of increased availability based on redundant GOOSE and SV messages. They are prohibited from connecting two ports to two Ethernet LANs because of IEEE 802.1w RSTP restrictions.

The control system industry created the Ethernet frame duplication protocols (for a less mission-critical control system), rather than fully supporting IEEE 802.1Q. Over time, many industrial and some protective devices have begun to support these proprietary IED protocols to create identical duplicates of Ethernet messages using the same VLAN on each duplicate message. Therefore, proprietary PRP and HSR protocols perform replication and not actual redundancy.

Part 3 of [16], which defines repairable replication methods that require manual intervention to detect and repair faults, defines PRP and its repairable behavior to duplicate messages but not support redundant paths. PRP supports generic supervision frame monitoring but does not have enough detail to support message subscription error detection. It includes no recovery or corrective action intelligence and requires human intervention to recognize and repair failures in data flow. Availability is improved by using IEC 62439-3 PRP in concert with LANs and switches using the IEC 62439-1 RSTP protocol or SDN; however, the DU failures remain part of the PRP link redundancy entity (LRE) design. As mentioned, proprietary methods may be available to overcome this DU failure.

This DU data path failure behavior is a critical cyber-defense risk because it allows undetected malicious or accidental disconnection of one of the duplicate data paths. This allows an accidental or intentional capture of Ethernet traffic and undetected installation of communications equipment.

Unique VLAN tags for each SV and GOOSE message enable the use of VLAN segregation in the local- and wide-area network to restrict messages to only the appropriate network segments. It is not standard practice to unplug Ethernet cables in a commissioned system and isolate messages to prevent delivery, replace with test signals, or deliver them to a test device. The best practice is done by using unique VLAN tags for ingress and egress filtering settings in the LAN switches to allow messages where they belong, restrict them from everywhere else, and modify these rules in real time for testing and diagnostics without physically changing the cabling. Through this method, VLAN filtering acts similarly to physical test switches used to isolate and open circuit physical signals in field panels to improve test and safety procedures. These safety practices are not possible with devices that do not support the full use of IEEE 802.1Q VLAN tags. Media access control (MAC) address filtering is possible but is awkward and only useful as egress filtering in most Ethernet switches.

Unique VLAN tags among multiple SV and GOOSE messages enable the use of VLAN segregation in the network to prevent messages from reaching IEDs on the network that are not subscribing to them. This noise, or unnecessary burden on the IED Ethernet channels may interfere with the communications, protection, and automation performance of the IEDs receiving unwanted messages. Perhaps the most dangerous part of this scenario is that a problem may be unclear throughout a demonstration or test of communications. It may not even be a problem in service until the system grows larger and the burden eventually affects IED performance. If VLAN segregation is not designed from the beginning, and if it becomes an issue in performance or troubleshooting, it is not reversible, and the data flow design has to be done over again.

Reference [27] describes that for PRP, “each node has two ports that operate in parallel and that are attached to the same upper layers of the communication stack through the LRE.” Reference [27] further states, “the LRE presents toward its upper layers the same interface as a non-redundant network adapter, so the upper layers are unaware of redundancy.” The unintended consequence of this is that the upper layers of applications within each IED are unaware of the failure of redundancy.

The lack of awareness of DU failures at the application level prohibits alarming, prevents corrective action, and puts IEC 62439-3 in conflict with IEC 61508 principles, since it reduces the availability of trip circuit designs using this method.

Reference [27] has been updated to describe that the LRE tracks the health of each LAN by “keeping a counter of received messages and of messages received with an error on each port.” However, this is not suitable for digital trip circuits that need

detail from many separate GOOSE messages for status, tripping, and possibly SV messages carrying raw analog samples on a single port.

The domino effect of inadequate design choices followed by compensation methods, rather than correcting design flaws, has created a challenge for modern trip circuit design. However, even this can be corrected as encouraged by value engineering, which identifies why the problem exists, but when necessary, takes stock of the present design flaws and moves forward with a new design to mitigate the risk. For example, IEC 62439 describes tracking the health of each port [28]. Devices can stay compliant with this task and add the capability to keep counters for each subscription, which eliminates DU errors, permits immediate alarming and corrective action, and consequently, dramatically improves availability.

VII. Performance Requirements

Designs of numerous in-service process bus MU solutions based on IEC 61850 include the necessary list of international standards used to define requirements for message delivery performance, message delivery quality, and device quality criteria [21]. The performance requirements are as follows.

A. International Standards That Define Device Performance Criteria

The standards that define signal dependability and security are as follows.

Signal dependability and security requirements:

- IEC 61850
- IEC 60834

Device availability requirements:

- IEC 61850
- IEC 60834
- IEEE 802.1

System reliability metrics:

- IEC 61850
- IEEE 1613
- IEC 60870

B. International Standards That Define Signal Transfer Criteria

Ethernet packet dependability and security requirements:

- IEC 61850
- IEC 60834
- IEC 15802
- IEEE 802.1

Ethernet packet latency specifications:

- IEC 61850
- IEC 60834
- IEC 15802
- IEEE 802.1

Protection signal packet speed:

- IEC 61850
- IEEE 1646
- IEC 61869

Review of these standards relative to the design of Ethernet communications provides the following acceptance criteria for mission-critical digital trip circuits as part of process bus systems.

C. Performance Requirements for Protection Signal Transfer Via Ethernet

In accordance with these standards, the digital trip circuit must be designed to perform protection signal transfer via an Ethernet that meets the following criteria:

- Signal transfer success rate greater than 99.99 percent.
- Expected signal transfer time between devices of less than 3 milliseconds.
- Expected signal transit via LAN of less than 1 millisecond.

D. Station Bus Resilience Requirements

The LAN must be designed in accordance with these signal transfer performance criteria to avoid failure. However, the design must also anticipate failure and have built-in resilience that meets the following criteria:

- Boolean protection logic with fewer than 4 dropped GOOSE packets and momentary outages shorter than 16 milliseconds.
- Analog protection calculations with fewer than 4 dropped SV packets and momentary outages shorter than 417 microseconds.

E. Signal Transfer Operational Latency

In this paper, we focus on the time latency performance of communicating digital messaging among IEDs to perform communications-assisted logic and decision-making in support of digital trip circuits. As described in [1], each data publisher performs an analog-to-digital conversion of the analog signals to create a pool of process-level raw signal information. Then, with each microprocessor process interval, the IEDs create calculated values and binary logic values. “These local, raw, and calculated values are used to make local decisions about the health and performance of the primary equipment and to perform local control and protections” [1].

Each 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 IED (data consumer) 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, includes the eight steps defined previously as components of the information transfer operational latency.

The eight steps of the operational latency information transfer aggregate to the time latency associated with detecting, digitizing, and transferring a fault signal to a relay. Reference [1] also describes the consequential action of a protection decision in the relay and a transfer of the trip signal to the field breaker control device with the IEC 61850 transfer time class of TT6 and transfer time latency of 3 milliseconds.

Digital signaling transfer time requirements and associated applications are summarized in Part 90-4 in [28], as shown in Table 1.

TABLE 1
DIGITAL SIGNALING TRANSFER TIME REQUIREMENTS

Transfer Time Class	Transfer Time (ms)	Application Example
TT0	>1,000	Files, events, and log items
TT1	1,000	Events and alarms
TT2	500	Operator commands
TT3	100	Slow automatic interactions
TT4	20	Fast automatic interactions
TT5	10	Releases and status changes
TT6	3	Trips and blockings

VIII. CONSIDERATIONS FOR NEW DESIGNS

A. *Shorter Fault Clearing Time Via Direct Tripping in a Process Bus*

Reference [9] mentions that microprocessor-based relays generally have enough output contacts to provide direct tripping to multiple breakers and perform ancillary functions. "Direct breaker tripping also eliminates the delay incurred by an intermediate contact multiplying or lockout relay, thereby reducing the time to interrupt a fault" [9]. This direct tripping, previously illustrated as Scenario 5 and 6, benefits from fewer devices and associated latencies and failure modes, thereby, improving availability and reducing the time to interrupt a fault. According to [9], the shorter fault clearing time afforded by using an IMU to trip directly improves or preserves power system stability when the IMU has sufficient availability for installation in a harsh environment among the primary equipment.

B. *Difference Between Fault Tolerance and High Availability*

"In general terms, high availability means that a system will experience minimal outages, while fault tolerance means there will be no outages" [1]. Using true redundancy and devices that continue to function in the presence of a component failure or cyber attack adds resilience to the design. Systems that constantly monitor, detect, and self-announce faults improve system availability. Automatic corrective action may even operate quickly enough to avoid an outage and make the system fault-tolerant.

C. *Component Selection*

Even with competent designs, manufacturing, and quality MUs used in intelligent substation designs studied by the Chongqing CEPREI Industrial Technology Research Institute, China still experienced a high-field failure rate [29]. Reference [29] explains the challenges to the process bus based on unavailability of system components, which include:

- High-field failure rate.
- Unexpected maintenance and repair.
- Large economic losses.
- Serious accidents.

Analysis of the products and suppliers reveals the possible reasons for the poor reliability, which include:

- Insufficient reliability design and review.
- Material defects.
- Assembly process problems.
- Absence of reliability acceptance test.

Therefore, in addition to considering the use of field installed relays and controllers, it is necessary to require the same manufacturing processes and type tests for MU, IMU, and standalone MU (SAMU) design and manufacturing, as used for protective relays. IEC 61869-13 defines the required type tests, insulation, electromagnetic compatibility (EMC), and safety requirements for SAMU devices [28]. "Recognizing the fact that new devices are typically mounted in the immediate vicinity of the high voltage breakers, IEC TC 38 based their recommendations for the standard on the wealth of information available from substation yard-based relay installations" [28]. These end users conclude that new SAMU devices are exposed to similar conditions and must meet or exceed the general capabilities defined in the IEC 60255 series of standards. SAMU EMC requirements defined in IEC 61869-13 match IEC 60255-26 with safety requirements based on [30].

IX. CONCLUSION

Trip circuits have evolved from only being hardwired analog signals to include signals transferred via serial-based and Ethernet-based devices and protocols. The benefits of this transition include a reduction in physical wiring and the ability to supervise signal transfer, which improves reliability and safety. However, this evolution requires a paradigm shift in the way engineers design and evaluate digital trip circuit performance, speed, and reliability. Our tests show that the use of digital trip circuits may come with the cost of slight reduction of tripping times when compared to a relay directly tripping through high-speed output contacts. Our tests also show that digital trip circuits can be very deterministic when designed and implemented correctly. These results demonstrate the importance of performing tests, as illustrated in this paper, to make quality assessments of a system with digital trip signals before its installation.

To help design engineers assess their implementations of digital trip circuits, the paper defines latency and jitter and explains how to measure these on digital trip circuit designs while accounting for the accuracy of the timing source.

Latencies for several trip circuit technologies are measured in parallel. The paper demonstrates the use of an automation controller for the verification of total operational latency of each parallel trip path. Average times for each method are presented. Test Case 1 setup uses a discrete binary input contact as an event trigger, which is then propagated through parallel trip paths of various technologies. Test Case 2 setup injects a simulated fault using secondary current into MUs and IMUs. The IED is then programmed to trip at a current threshold, and the trip signal is transferred over various parallel methods.

The paper provides various resources and methodologies to help engineers design for reliability, including:

- NERC reliability fundamentals.
- Value engineering and how it can be used to not only identify what the problem is but why the problem exists. When necessary, it can take stock of the present design flaws and move forward with a new design to mitigate the risk.
- The U.S. Department of the Army's LVD approach to identify, recognize, investigate, and limit vulnerabilities and design gaps.
- FTA and ATA to evaluate design fault tolerance and cyber-attack tolerance.

The paper also provides performance and resilience requirements and considerations for new digital trip circuit designs, as follows:

- Have a signal transfer success rate greater than 99.99 percent.
- Achieve an expected signal transfer time of less than 3 milliseconds between devices.
- Achieve an expected signal transit via LAN of less than 1 millisecond.
- Have Boolean protection logic with fewer than 4 dropped GOOSE packets and momentary outages shorter than 16 milliseconds.
- Have analog protection calculations with fewer than 4 dropped SV packets and momentary outages shorter than 417 microseconds.

Lastly, the paper provides a cautionary tale of high-field failure rates of MUs used in intelligent substations and the challenges imposed due to the unavailability of critical components.

X. Acknowledgments

We gratefully acknowledge the contributions of Tim Grigg, Gale Nelms, Isaac West, and Edson Hernandez.

XI. References

[1] A. McDonald, A. Dolezilek, and D. Dolezilek, "Hidden in Plain Sight: Anticipating and Avoiding Hidden Failures in Communications-Assisted Protection," proceedings of the 47th Annual Western Protective Relay Conference, Virtual Format, October 2020.

[2] W. Allen and T. Lee, "Flexible High-Speed Load Shedding Using a Crosspoint Switch," proceedings of the 32nd Annual Western Protective Relay Conference, Spokane, WA, October 2005.

[3] D. Dawson, J. Gosalia, D. Blackburn, S. Conrad, C. Gromen, S. Grier, R. Haas, R. Hart, I. Hasenwinkle, J. Huddleston, M. McDonald, G. Moskos, R. Onate, J. Stephens, and D. Tziouvaras, "Summary of Relay Trip Circuit Design. A Summary of an IEEE Special Publication by the PSRC Relay Trip Circuit Design WG," proceedings of the 2000 Power Engineering Society Summer Meeting, Seattle, WA, July 2000.

[4] IEC 61850-90-4, *Communication Networks and Systems For Power Utility Automation – Part 90-4: Network Engineering Guidelines*, 2013.

[5] L. Carpini, A. Kalra, D. Dolezilek, G. Vielmini, and T. Grigg, "Using Real-Time Testing Tools to Baseline the Performance of OT Networks for High-Speed Communications," proceedings of the 7th Annual PAC World Americas Conference, September 2020.

[6] IEEE Std 1646, *Communication Delivery Time Performance Requirements for Electric Power Substation Automation*, 2004.

[7] M. Rensburg, D. Dolezilek, and J. Dearien, "Case Study: Lessons Learned Using IEC 61850 Network Engineering Guideline Test Procedures to Troubleshoot Faulty Ethernet Network Installations," proceedings of the Power and Energy Automation Conference, Spokane, WA, March 2017.

[8] D. Dolezilek and J. Dearien, "Lessons Learned Through Commissioning and Analyzing Data from Ethernet Network Installations," proceedings of the 5th International Scientific and Technical Conference, Sochi, Russia, June 2015.

[9] Power System Relay Committee, "Relay Scheme Design Using Microprocessor Relays," proceedings of the 68th Annual Conference for Protective Relay Engineers, College Station, TX, April 2015.

[10] V. Skendzic, I. Ender, and G. Zweigle, "IEC 61850-9-2 Process Bus and Its Impact on Power System Protection and Control Reliability," proceedings of the 9th Annual Western Power Delivery Automation Conference, Spokane, WA, April 2007.

[11] A. Shrestha, M. Silveira, J. Yellajosula, and S. K. Mutha, "Understanding the Impacts of Time Synchronization and Network Issues on Protection in Digital Secondary Systems," proceedings of the PAC World Global Conference, Virtual Format, August–September 2021.

[12] M. Silveira, D. Dolezilek, S. Wenke, and J. Yellajosula, "Cyber Vulnerability Assessment of a Digital Secondary System in an Electrical Substation," proceedings of the 74th Annual Conference for Protective Relay Engineers, Virtual Format, March 2021.

[13] IEC 61850-9-2, *Communication Networks and Systems for Power Utility Automation – Part 9-2: Specific Communication Service Mapping (SCSM) – Sampled Values Over ISO/IEC 802-3*, 2011.

[14] D. Bekker, T. Tibbals, and D. Dolezilek, "Defining and Designing Communications Determinism for Substation Applications," proceedings of the 40th Annual Western Protective Relay Conference, Spokane, WA, October 2013.

[15] D. Holstein, "Development of an IEEE Standard for Integrated Substation Automation Communication (P1525): Specifications for Substation Integrated Protection, Control and Data Acquisition Communication," proceedings of the 2000 Power Engineering Society Summer Meeting, Seattle, WA, July 2000.

[16] IEC 62439, Industrial Communication Networks – High Availability Automation, March 2016.

[17] North American Electric Reliability Corporation, "Reliability Fundamentals of System Protection," *System Protection and Control Subcommittee Reference Documents*, January 2011.

[18] ANSI ASTM E1699-14, *Standard Practice for Performing Value Engineering (VE)/Value Analysis (VA) of Projects, Products and Processes*, 2020.

[19] Department of the Army, TM 5-602-1, *Utility Systems Terrorism Countermeasures for Command, Control, Communications, Computer, Intelligence, Surveillance, and Reconnaissance (C4ISR) Facilities*, February 2006.

[20] IEC 61508, *Functional Safety of Electrical/Electronic/ Programmable Electronic Safety-Related Systems*, 2010.

[21] D. Dolezilek, P. Lima, G. Rocha, A. Rufino, and W. Fernandes, "Cost and Performance Comparison of Numerous In-Service Process Bus Merging Unit Solutions Based on IEC 61850," proceedings of the 10th Annual Protection, Automation, and Control World Conference, June 2019.

[22] G. Scheer, "Answering Substation Automation Questions Through Fault Tree Analysis," proceedings of the 4th Annual Substation Automation Conference, College Station, TX, April 1998.

[23] D. Dolezilek, "Case Study of a Large Transmission and Distribution Substation Automation Project," August 1999. Available: selinc.com.

[24] R. Sandoval and J. L. Eternod, "Evaluation of Methods for Breaker-Flashover Protection," proceedings of the 31st Annual Western Protective Relay Conference, Spokane, WA, October 2004.

[25] W. Denson, G. Chandler, W. Crowell, A. Clark, and P. Jaworski, "Nonelectronic Parts Reliability Data 1995," Reliability Analysis Center, 1995.

[26] 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, United Kingdom, March 2020.

[27] IEC 62439, *Industrial Communication Networks – High Availability Automation Networks*.

[28] V. Skendzic and D. Dolezilek, "News from the Editor – Updates and Enhancements to Process Bus IEC 61850-9-2 and Related Standards," proceedings of the 5th Annual PAC World Americas Conference, Raleigh, NC, August 2018.

[29] Z. Jia-yong, H. Sheng-zong, C. Tie-zhu, and Y. Sheng-jun, "Study of Highly Accelerated Life Test for Merging Unit of Intelligent Substation," proceedings of the 18th International Conference on Electronic Packaging Technology, Harbin, China, August 2017, pp. 980–983.

[30] IEC 60255-27, *Measuring Relays and Protection Equipment – Part 27: Product Safety Requirements*, 2013.

XII. Biographies

Matt Ross has a Bachelor of Science in Electrical Engineering from the Illinois Institute of Technology and a Bachelor of Science in Physics from DePaul University. He is a general engineer at Commonwealth Edison Company (ComEd). Matthew joined ComEd's testing and commissioning engineering (TCE) department in 2016. ComEd TCE is responsible for the testing, commissioning, and troubleshooting of all transmission and distribution substation equipment on the ComEd system. He is a PE in Illinois.

John Bettler has a BSEE from Iowa State and an MSEE from Illinois Institute of Technology (IIT). John has worked at Commonwealth Edison Company (ComEd), a power company in the Chicago area, for 29 years. He has experience as a field engineer and protection engineer. Currently, he is the principal engineer for ComEd's relay section. His team's purview includes 4 kV and 12 kV feeders up to 765 kV transmission lines and all transmission and distribution equipment in between (e.g., TR, buses, cap, and inductors). John's team also reviews interconnections, independent power producers, and distribution generation projects. John is also adjunct faculty at IIT and University of Wisconsin–Madison, teaching power and protection classes. He is a PE in Illinois.

Andrew Sprenger received his Bachelor of Science in Electrical Engineering from Seattle University. He joined Puget Sound Energy (PSE) in 2014 and is currently an engineer in the substation controls and automation group. He started in this group designing traditional control systems for PSE's substations, but has since shifted focus to operational technology and automation work. He is currently midway through design on PSE's first IEC 61850-based, fully-digital secondary system.

Jesse Silva received his Bachelor of Science in Electrical Engineering from the University of California, Irvine. He is a senior engineer at Southern California Edison (SCE). In 2006, Jesse joined SCE in the substation automation department as an automation engineer, where he supported the commissioning of substation automation systems across the SCE territory. He currently works in the Grid Technology Innovation organization, where he evaluates and demonstrates new substation automation technologies for SCE.

Austin Wade received his BS in electronic and electrical engineering, summa cum laude, from California State University, Sacramento, in 2013. He joined Schweitzer Engineering Laboratories, Inc. (SEL) in 2012. He is presently a senior application engineer in research and development. He has experience in protection and control system design, communications-based protection schemes, substation automation, testing, and maintenance. Prior to receiving his BS, he worked for an electrical testing company as a testing technician and certified electrician. He is a registered professional engineer in the state of California and is a senior member of the IEEE Power and Energy Society. He holds 14 patents and has authored several papers related to power system protection.

David Dolezilek is a principal engineer at Schweitzer Engineering Laboratories, Inc. (SEL) 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.

Mauricio Silveira is an electrical engineer with a BS earned from São Paulo State University in 2013. Since 2014, he has been with Schweitzer Engineering Laboratories, Inc. (SEL), where he has held positions in SEL Engineering Services, Inc. (SEL ES), sales and customer service, and research and development. He is currently a lead integration and automation engineer. His work includes development of protective relay protocols and communications, network design for critical infrastructures, power system modeling, and cybersecurity assessment.

Rodrigo Abboud received his BS in electrical engineering from Washington State University in 2021, with a minor in computer science. He started working at Schweitzer Engineering Laboratories, Inc. (SEL) as an integration and automation engineering intern in 2020 and was hired as an associate engineer in 2021, while completing his studies. Rodrigo has participated in a variety of projects during his two years at SEL, becoming acquainted with multiple aspects of power systems and developing efficient programs for communication with relays and networks that assist work in research and development.