Engineering Defense-in-Depth Cybersecurity for the Modern Substation

Chris Ewing
Schweitzer Engineering Laboratories, Inc.

Presented at the
12th Annual Western Power Delivery Automation Conference
Spokane, Washington
April 13–15, 2010

Engineering Defense-in-Depth Cybersecurity for the Modern Substation

Chris Ewing, *Schweitzer Engineering Laboratories, Inc.*

***Abstract*—Applying defense-in-depth cybersecurity measures to modern substation design provides a holistic and robust security posture for the substation environment. This paper discusses a multilayer security approach that applies to existing substation environments and can be integrated into the planning and design phases of new substation projects. Layer 1 focuses on perimeter security and the controls surrounding the protection of the ingress/egress point of the substation electronic security perimeter. Layer 2 focuses on the security controls for communication and devices that perform data aggregation. Layer 3 focuses on host-based cybersecurity controls used to provide security at the device level. Lastly, the paper suggests how this approach is a scalable solution that aligns with the U.S. Department of Energy "Roadmap to Secure Control Systems in the Energy Sector."**

I. Introduction

Traditionally, access to substation control system components was managed almost entirely through physical access control measures such as keys, locks, fences, and guards. The primary focus in design and development was availability and accessibility, including remote access, with cybersecurity as an afterthought.

Modern substation designs provide enhanced functionality through increased connectivity and communication. Increased communication, when implemented safely and securely, provides excellent opportunities for the automation and integration of systems, which can increase reliability and efficiency.

This same design, if implemented poorly and lacking cybersecurity, can result in security holes that may be exploited by attackers. The lack of cybersecurity can, and eventually will, lead to decreased reliability. Cybersecurity needs to be viewed as a facilitator for enhancing the reliability and operation of the power grid.

Applying defense-in-depth cybersecurity from the very beginning of the planning and design phases results in a robust and secure system that provides a reliable platform for future applications and improves the cybersecurity of existing implementations.

II. Multilayer Security

Multilayer security puts the critical assets at the most reliable and secure layer. With multiple layers, each layer can have unique yet complementary security controls. An attacker must then not only compromise security controls at the perimeter but must be able to compromise each layer behind the perimeter to reach the critical asset. Fig. 1 depicts a high-level architecture with multiple security layers. The dashed line indicates the perimeter, or logical boundary, of the substation network. Access to and from the substation network must pass through Layer 1, the access layer. If access is permitted, network traffic is allowed to enter or leave the substation network based on a set of rules. Layer 2, the data aggregation layer, is a concentration point on the substation network; it provides protocol conversion and also controls access to devices in Layer 3. Layer 3 is where the most protected devices reside. This layer hosts control and protection equipment that, if compromised by an attacker, could result in catastrophic failure of the safe and reliable operation of the substation primary function. Such an occurrence could have many impacts, including financial loss, equipment damage, or reduced customer and stakeholder confidence; in worst-case scenarios, diminished safety can result in human injury or death. These are just a few examples, but we can imagine the cascading effects and possible outcomes of a cyberattack.

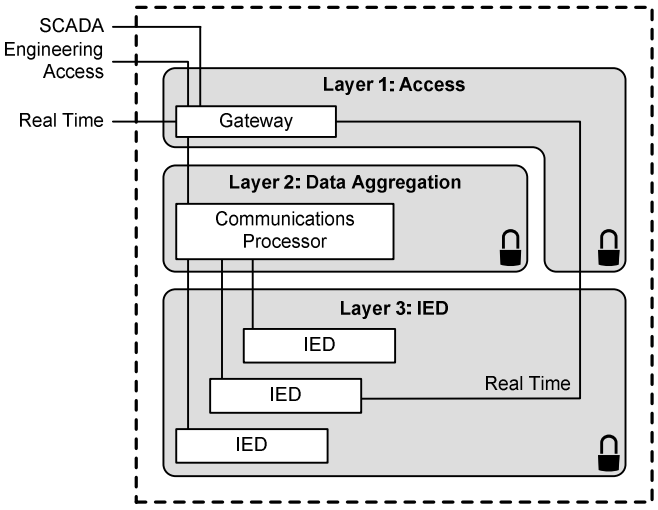

Fig. 1. Multilayer security architecture

A. Importance of a Security Program

While this paper focuses primarily on technical security measures, it is important to note that a good security program must include policies and procedures supported by upper-level management. Policies and procedures are the foundation of any security program and provide the "teeth" for implementing and enforcing technical security controls. For example, if an organization has a remote access policy for the substation stating that remote access must occur only from the control center network, this can be enforced through technical security controls.

B. Layer 1: Access

The perimeter is defined in this paper as the logical boundary of the substation network, where the ingress/egress point connects to an external or untrusted network. In North American Electric Reliability Corporation (NERC) Critical Infrastructure Protection (CIP) terms, this is the electronic security perimeter [1]. Securing the perimeter provides controlled communication, access control, data integrity, data confidentiality, and a level of assurance that the assets behind the perimeter are protected from external attackers. The following security measures and principles help secure the access layer.

1) Limiting Ports and Services

An attacker must find a vulnerability to exploit in the access layer in order to gain entry. One of the first actions an attacker will take is to enumerate ports and services to determine which ports are open and accessible from the network perspective of the attacker. Once this list of open ports is compiled, the attacker looks for vulnerabilities to exploit in each of these services.

Firewalls are commonly used on the perimeter to filter ingress/egress traffic. This filtering limits which ports and services within and behind the access layer are visible to an attacker. A best practice for deploying firewalls is to identify all communications paths to and from the substation, determine which services the security policy permits over each communications path, and limit access to these services using a deny-by-default approach. A deny-by-default approach denies access to all ports and services unless specifically allowed through policy.

2) Virtual Private Networks

Virtual private networks (VPNs) are commonly used on the perimeter to create an encrypted and authenticated communications path over an untrusted network. As an example, a VPN can be set up by placing an Ethernet security gateway on the perimeter and establishing an Internet Protocol Security (IPsec) tunnel between the substation and the control center. The control center has another Ethernet security gateway that authenticates with the substation gateway. Upon successful authentication, a secure and encrypted communications path between the substation and control center is established. VPN technology complements firewall technology; the firewall determines what traffic is allowed, and the VPN encrypts the allowed traffic over the communications path.

3) Securing Serial Communication

Serial communication typically uses nonroutable protocols, which must still be protected. If an attacker has physical access to an unsecured serial communications channel, the attacker can view or modify the data, gain unauthorized access to the endpoint devices within the substation, and proceed with an attack. Unsecured dial-up connections are even more susceptible to unauthorized access.

Serial cryptographic modules offer security for serial communication at the access layer. These devices provide session authentication to prevent unauthorized access, message authentication to ensure that the message sent to the destination device has not been modified in transit, and encryption for data to be transmitted over an untrusted network.

C. Layer 2: Data Aggregation

Layer 2 hosts devices located inside the security perimeter that perform data aggregation functions. The data aggregation layer is where data are concentrated within the substation network. Typical devices include communications processors, Ethernet switches, or port servers. By using a multilayer architecture, the resources of these devices can be dedicated to their primary objectives, such as protocol conversion, substation automation, and port concentration. These devices do not need to provide perimeter security controls because such controls are provided by the access layer devices.

The devices in the data aggregation layer contain critical information and should not be placed on the perimeter network. Doing so would provide only one layer of security controls that an attacker would need to compromise and would also open the devices up to denial-of-service (DOS) attacks on the untrusted network. DOS attacks could impact device performance, even to the point of a device becoming unresponsive. A recommended approach is to place data aggregation devices behind the security perimeter, where they are protected by devices such as firewalls, VPN gateways, and serial cryptographic modules. This approach adds a layer of security, or buffer, between an attacker and a device in the data aggregation layer. While a level of security is provided by the access layer, the data aggregation layer still needs to have security controls of its own. The following are a few examples of security controls that can be found in the data aggregation layer.

1) Port-Level Security

Port-level security can be found in communications processors and Ethernet switches. Many communications processors can limit the maximum access level associated with a specific port. For example, Port 1 may be limited to read-only access for all users, while Port 2 may be used by personnel at the appropriate privilege level to read and write settings. This allows one communications channel to connect to the read-only port and a separate communications channel to connect to a port that allows settings changes. Ports can also be configured to allow or deny remote connections. For example, a common procedure is to disable a port that can be enabled through a supervisory control and data acquisition (SCADA) command. The engineer calls the control center, and the control center operator verifies the identity of the engineer. The control center operator can then issue the SCADA command to enable the port for remote engineering access. Once the engineer completes the task, the port is then disabled.

Port security on Ethernet switches is slightly different. Ethernet ports in many switches can be associated with a media access control (MAC) address, which is used to identify an Ethernet device on the network segment. If another device is plugged into a port that was assigned only to a particular

MAC address, the port will be disabled, preventing access for what may be a rogue device on the network.

Ports not being used should be disabled. This applies both to physical ports (e.g., serial ports and Ethernet ports) and logical ports. Logical ports are used to access a service through an Internet protocol (IP) address and protocol. For example, most web servers listen for Hypertext Transfer Protocol (HTTP) requests on Transmission Control Protocol (TCP) Port 80. If a device runs a web server that is not being used, the web server should be disabled as long as doing so does not impact the operation of the device. This results in TCP Port 80 not being accessible over the network for this particular device and reduces the attack surface of the device.

2) Secure Protocols

Cleartext protocols are susceptible to eavesdroppers. For example, if a user logs in to a device using Telnet over an unsecured communications channel, the username and password will be passed over that communications channel in cleartext. If the channel is being "sniffed" by an attacker, the attacker will have the username and password of the user who just logged in to the device.

A best practice guideline is to use secure protocols whenever possible. The following are some examples:

- Use Secure Shell (SSH) instead of Telnet when possible. SSH encrypts all traffic between communicating end devices.
- Use HTTP Secure (HTTPS) instead of HTTP when possible. Similar to SSH, HTTPS encrypts all traffic between communicating end devices.
- Use Secure File Transfer Protocol (SFTP) or Secure Copy (SCP) instead of File Transfer Protocol (FTP) when possible. SFTP and SCP encrypt all traffic between communicating end devices.
- Encapsulate unsecure protocols. Some devices do not support secure protocols such as SSH and HTTPS. In that case, Telnet, HTTP, and FTP data can be encapsulated into a packet and encrypted. An example of this is sending Telnet, HTTP, and FTP traffic through an IPsec VPN tunnel. The Telnet, HTTP, and FTP packets are encapsulated into encrypted packets that are sent securely over the untrusted communications channel. Once the encapsulated packets reach the destination network, they are decrypted and sent on to their final destinations.

D. Layer 3: IED

Layer 3 focuses on security principles that apply to individual devices. Individual devices may include computers used as human-machine interfaces (HMIs) or intelligent electronic devices (IEDs) such as relays and meters. Access to a device in this layer through an electronic communications channel implies that the user has passed through both the access layer and the data aggregation layer successfully. Adding another layer of security to the end devices themselves further protects the most critical assets within the substation. The security controls and measures within the IED or host layer can be applied directly to a device; this hardens the device to thwart potential attacks over the electronic communications channel and provides additional security controls for users with physical access to the device. Some guidelines for security controls to be used at the IED or host layer are listed below.

1) Disabling Default Ports and Services

Similar to limiting ports and services within the access layer, turning off unused default ports and services decreases the attack surface of the device. For instance, if a firewall rule was configured improperly at the access layer and allowed communication with a device that would not normally be permitted, the device would still not be accessible if this service were disabled. For example, assume an organization implements a policy stating that all remote access must use secure protocols (e.g., SSH instead of Telnet). This could be enforced by filtering Telnet traffic through the firewall at the access layer. If a firewall is improperly configured and does not block Telnet, the device in the IED or host layer responds to the request and allows communication. Disabling Telnet on the device itself would prevent Telnet access to the device, regardless of whether or not the firewall is configured properly.

2) Firmware Integrity Verification

Ensure that all firmware updates to devices are installed from a trusted source. Some manufacturers digitally sign their firmware so that even if a compromised firmware installation were attempted, it would be rejected during the firmware update process.

Another method for checking firmware integrity is the use of hashing algorithms. Some manufacturers post Message-Digest Algorithm 5 (MD5) or Secure Hash Algorithm (SHA) hash values for firmware files. After receiving the firmware installation file from the manufacturer, a user can run a hashing algorithm against the received file and compare the results with the value posted by the manufacturer. Matching values assure the user that the firmware file has not been modified, either accidentally or maliciously. There are many freely available tools that hash files.

3) Malware Detection

Malware is very prevalent in operating systems that typically run on laptops, desktops, and server hardware platforms. Antivirus software and updated virus definitions mitigate the malware risk on these devices. While this approach has worked reasonably well for these platforms, it is not feasible to install antivirus software on embedded devices such as protective relays. One method of detecting malware on embedded devices is to continuously compare the code executing in runtime memory with the code stored on a read-only chip that is loaded into memory when the device boots. If the code does not match exactly, then the device treats the code running in memory as malware, and preventive or corrective actions should occur.

III. Universal Security Principles

This section describes security principles that apply to each of the three layers discussed in this paper.

A. Authentication, Authorization, Accountability (AAA)

Authentication is the process of verifying an identity being presented to the system for the purpose of gaining access. An identity is typically a user account that is authenticated through the use of a password. Other ways to authenticate an identity exist, such as tokens, certificates, and biometrics, but passwords are most commonly used. It is critical that passwords are strong and are changed on a regular basis.

User-based accounts are individual accounts created and assigned to individual users. Each user creates a password that is known to no one else, and the user is responsible for actions associated with the account. These accounts can be assigned different permissions, creating different authorization levels. User-based accounts also establish individual accountability.

Authorization determines which functions an authenticated user is able to perform on a particular device. Some devices have predefined authorization levels (e.g., administrator, operator, security officer) that are associated with certain permissions or functionalities within the device. Other devices have customizable authorization levels; an administrator can create granular permission levels within the system and associate users with administrator-defined authorization levels.

Accountability within a system provides an audit trail of *who* did *what* and *when* they performed the action. User-based accounts and a form of event logging provide individual traceability, which results in individual accountability. Actions such as successful and failed login attempts, configuration changes, and firmware updates are examples of events that should be logged. Some products log events through local event logs, and more recently, others have begun providing support for syslog-compatible event logging. This feature allows a device to report events to a remote syslog server for event correlation and analysis.

B. Time Synchronization

Time synchronization is important for event correlation so that event logs may be accurately analyzed to determine when a system event occurred. This is critical when following an audit trail of events from multiple devices. If the time is not synchronized between the devices in question, it is very difficult to correlate these events and produce a meaningful analysis.

C. Physical Security

Physical threats to critical assets remain present, and risk must be mitigated using physical security controls.

D. Security Awareness Training and Education

Informing personnel of their responsibilities when it comes to cybersecurity is an important step in implementing and enforcing policies and procedures. Promoting general security awareness throughout an organization can help send the message that everyone has a part in overall security.

E. Network Documentation

Documentation of system components and network communications paths is critical when developing a cybersecurity strategy that aligns with security policies and procedures. It is important that this documentation be updated when changes occur in order to remain current and to present an accurate depiction of the system. This documentation becomes very useful when planning system upgrades and improvements and is an essential tool in troubleshooting. Some key components that should be included in this documentation are the following:

- Communications paths to and from the substation network.
- Ports and services allowed to and from the substation network.
- Critical assets.
- Risk assessment methodology and the security measures applied to critical assets.

This documentation must be stored in a secure location, yet be accessible to personnel in a timely fashion.

IV. Scalable Solution

According to the U.S. Department of Energy (DOE) "Roadmap to Secure Control Systems in the Energy Sector," "In 10 years, control systems for critical applications will be designed, installed, operated, and maintained to survive an intentional cyber assault with no loss of critical function" [2]. Applying multilayer security to provide defense in depth, as recommended in this paper, is a fundamental step in accomplishing the DOE vision. The DOE roadmap outlines a strategic framework with four main goals, as discussed in this section.

A. Measure and Assess Security Posture

Cybersecurity should be applied commensurately with the risk and the value of the asset requiring protection. A proper risk analysis identifies threats, vulnerabilities, and potential impacts to critical assets, which will dictate the level of cybersecurity controls that needs to be applied to mitigate risk to an acceptable level.

Risk assessment identifies critical assets, along with threats and vulnerabilities affecting these assets. Risk levels can then be assigned to each asset and to the system as a whole, providing the information required to implement proper security policies and procedures and technical cybersecurity controls. As automated security state monitoring tools become available, the assessment methodology remains the same and scales toward this trend.

B. Develop and Integrate Protective Measures

Identifying critical assets based on a risk assessment is a good starting point, but security must not end there. Once assets are identified, secure architectural design considerations and security controls should be analyzed, tested, and implemented. One of the major challenges the industry faces is securing legacy devices because of resource constraints within the devices. Implementing cybersecurity controls

within resource-constrained legacy devices could potentially have a negative impact on their primary function and result in decreased reliability. A well-implemented security architecture with multiple layers of defense provides compensating measures that help protect legacy devices from cyberattack. When a legacy device has reached the end of its life cycle, it can be replaced without requiring a restructure of the surrounding security architecture. The security architecture should scale independently of individual device life cycles, providing a means to integrate newly developed protective measures into the architecture.

C. Detect Intrusion and Implement Response Strategies

Incident response consists of policies, procedures, and technical measures that enable the identification of potential cyberintrusions and the structure to react to and remediate the event. As depicted in Fig. 2, this is a continual process that establishes a proactive stance related to cyberintrusion. Following this cycle improves the incident response capability as new threats emerge and before they are realized within the system.

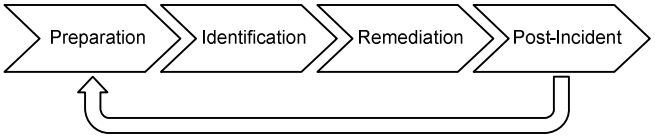

Fig. 2. Incident response cycle

A well-developed incident response program is the foundation for incorporating future tools that may automatically generate contingency measures and take remedial action in response to attempted intrusions. The incident response program should outline the remedial actions for certain events, which are executed today through manual procedures. Defining these manual procedures provides the essential framework for automating the same events when automated remediation tools become available.

As more sophisticated tools become available, data are needed for analysis. The capability to automate alerts for events flagged as cyberincidents exists even today. Structuring the system architecture to provide data yields a scalable solution for implementing new tools in the future that may take remedial action based on certain events.

D. Sustain Security Improvements

Sustaining security improvements requires identifying emerging threats, performing periodic vulnerability assessments, and determining the risk to critical assets. As new threats emerge, security improvements can be implemented within the existing security architecture and overall security program. New equipment and installations will be able to leverage the existing security measures and architecture and build upon the secure and reliable operation of the installed environment.

V. Conclusion

At first glance, multiple layers of security may appear to add complexity, especially for operators of the system. However, a properly designed and implemented multilayer architecture actually improves reliability, and most security controls are transparent to authorized end users of the system. Additionally, well-documented policies, procedures, and security approaches establish a security framework that can be used to provide the information necessary for a compliance program. In summary, the following steps provide a high-level approach for engineering defense in depth for the modern substation:

1. Define and implement management-approved policies and procedures.
2. Identify and document critical assets and communications paths.
3. Perform risk and impact analyses.
4. Test, implement, and document security controls.
5. Perform routine analysis and keep information updated.

VI. References

[1] NERC Standard CIP-005-1, June 2006. Available: http://www.nerc.com/page.php?cid=2|20.

[2] U.S. Department of Energy, "Roadmap to Secure Control Systems in the Energy Sector," January 2006. Available: http://www.oe.energy.gov/DocumentsandMedia/roadmap.pdf.

VII. Biography

Chris Ewing is a lead product engineer with the Schweitzer Engineering Laboratories, Inc. (SEL) security solutions division. Prior to joining SEL, he consulted as an information security engineer in both the private and public sectors. He received his B.S. in computer science and his M.S. in network security, and he holds the CISSP professional security certification. Chris has over ten years of experience in the cybersecurity field.