

Control Systems Cyber Security: Defense in Depth Strategies

David Kuipers
Mark Fabro

May 2006



The INL is a U.S. Department of Energy National Laboratory
operated by Battelle Energy Alliance

Control Systems Cyber Security: Defense in Depth Strategies

**David Kuipers
Mark Fabro**

May 2006

**Idaho National Laboratory
Idaho Falls, Idaho 83415**

**Prepared for the
U.S. Department of Homeland Security
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517**

Table of Contents

<i>Keywords.....</i>	<i>3</i>
<i>Introduction.....</i>	<i>3</i>
<i>Background.....</i>	<i>3</i>
<i>Overview of Contemporary Control System Architectures.....</i>	<i>5</i>
<i>Security Challenges in Control Systems</i>	<i>7</i>
<i>Security Profiles and Attack Methodologies.....</i>	<i>8</i>
<i>Isolating and Protecting Assets: Defense-in-Depth Strategies.....</i>	<i>16</i>
<i>Specific Recommendations and Countermeasures</i>	<i>25</i>
<i>Suggested Reading.....</i>	<i>27</i>
<i>Glossary.....</i>	<i>28</i>

Keywords

Defense in depth, industrial control system, SCADA, PCS, cyber security, mitigation, firewall, IDS, intrusion detection, encryption, DMZ

Introduction

Information infrastructures across many public and private domains share several common attributes regarding IT deployments and data communications. This is particularly true in the control systems domain. A majority of the systems use robust architectures to enhance business and reduce costs by increasing the integration of external, business, and control system networks. However, multi-network integration strategies often lead to vulnerabilities that greatly reduce the security of an organization, and can expose mission-critical control systems to cyber threats.

This document provides guidance and direction for developing ‘defense-in-depth’ strategies for organizations that use control system networks while maintaining a multi-tier information architecture that requires:

- Maintenance of various field devices, telemetry collection, and/or industrial-level process systems
- Access to facilities via remote data link or modem
- Public facing services for customer or corporate operations
- A robust business environment that requires connections among the control system domain, the external Internet, and other peer organizations.

Background

The critical infrastructure systems that support major industries, such as manufacturing, transportation, and energy, are highly dependent on information systems for their command and control. While there is still a high dependence on legacy control systems, critical infrastructure/key resource (CI/KR) systems are migrating to new communication technologies. As a result, the diverse and disparate proprietary mechanics of control systems are being replaced with common communications protocols and open architecture standards, which can have both positive and negative impacts.

On one hand, the migration empowers control system users and manufacturers to offer new and more efficient methods of communication, as well as more robust data, quicker time to market, and interoperability. On the other hand, empowering control system users introduces new risks. Cyber-related vulnerabilities and risks are being created that could not exist when the CI/KR information infrastructures that involve control systems were isolated. The interdependence of CI/KR systems, such as in the power sector, has been illustrated in a number of instances, including the 2003 North American blackout.

The new protocols and communication standards that are providing increased interoperability and control in the control system community are the same technologies that have been exploited and compromised in the Internet and networking domains. Historically, control system security meant locating and identifying problems in a closed-loop system; now unauthorized intrusion or attacks are evolving issues to be addressed.

Figure 1 illustrates the traditional separation of corporate architectures and control domains. This architecture provided means for data sharing, data acquisition, peer-to-peer data exchange, and other business operations. However, the security of any given system was based on the fact that few, if any, understood the intricate architecture or the operational mechanics of the resources on the controls system LAN. This ‘security by obscurity’ works well for environments that have no external communication connections which allow an organization to focus on physical security.

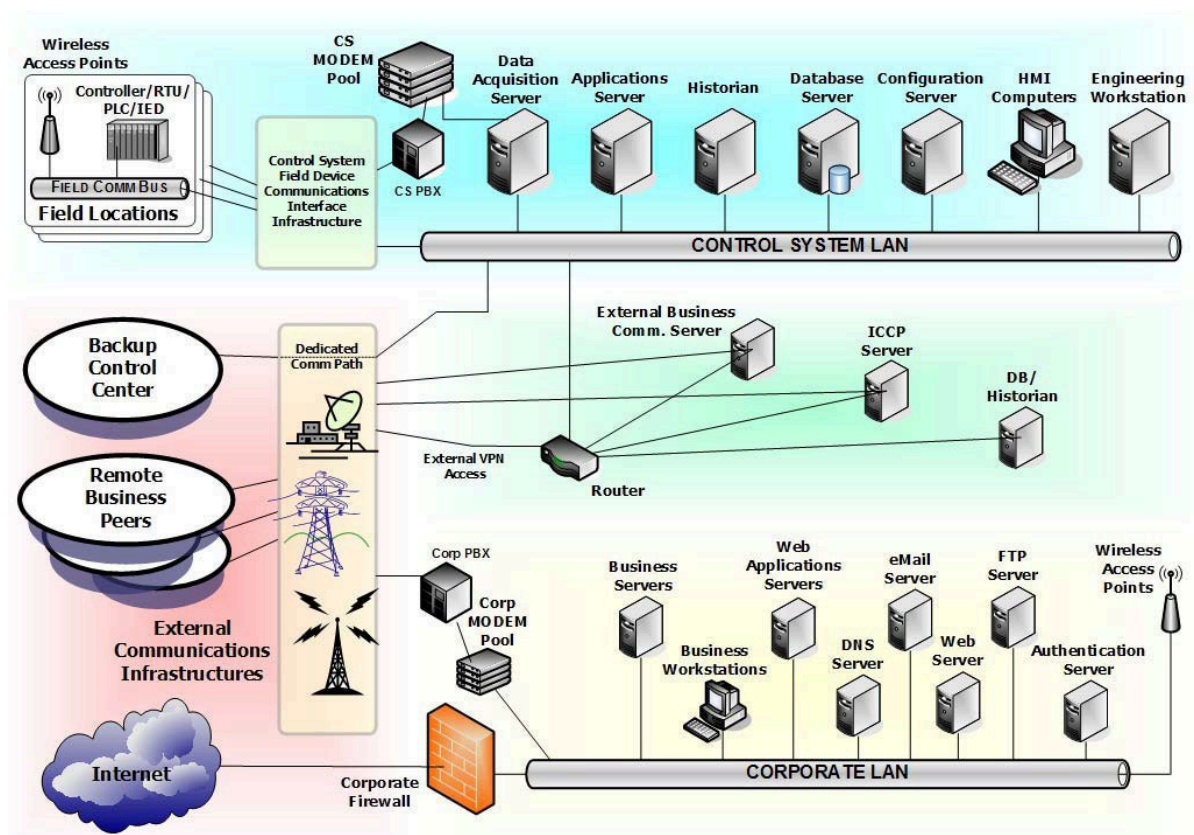


Figure 1 – Traditional isolation of corporate and control domains

Overview of Contemporary Control System Architectures

In today's competitive markets, isolated control system networks are being inter-connected. In connecting these networks, and introducing IT components into the control system domain, security problems arise due to:

- Increasing dependency on automation and control systems
- Insecure connectivity to external networks
- Usage of technologies with known vulnerabilities
- No business case for cyber security in control system environments
- Control system technologies have limited security, and if they do the vendor-supplied security capabilities are generally only enabled if the administrator is aware of the capability
- Control system communications protocols are absent of security functionality
- Considerable amount of open source information is available regarding control system configuration and operations.

Control system operational security has historically been defined by industry as the level of reliability of the system to operate. The total isolation from the external (and untrusted) network allowed the organization to reduce the level of communications security—threats to operations resided with physical access to a facility or plant floor. Thus, most data communications in the information infrastructure required limited authorization or security oversight. Operational commands, instructions, and data acquisition occurred in a closed environment where all communications were trusted. In general, if a command or instruction was sent via the network it was anticipated to arrive and perform the authorized function, as only authorized operators had access to the system.

Obviously, this arrangement is very different from effective network and IT cyber security systems. Merging a modern IT architecture with an isolated network that may not have any effective security countermeasures is challenging. Although simple connectivity using routers and switching is the most obvious means to provide interconnectivity, unauthorized access by an individual will provide unlimited access to the systems. Figure 2 shows an integrated architecture.

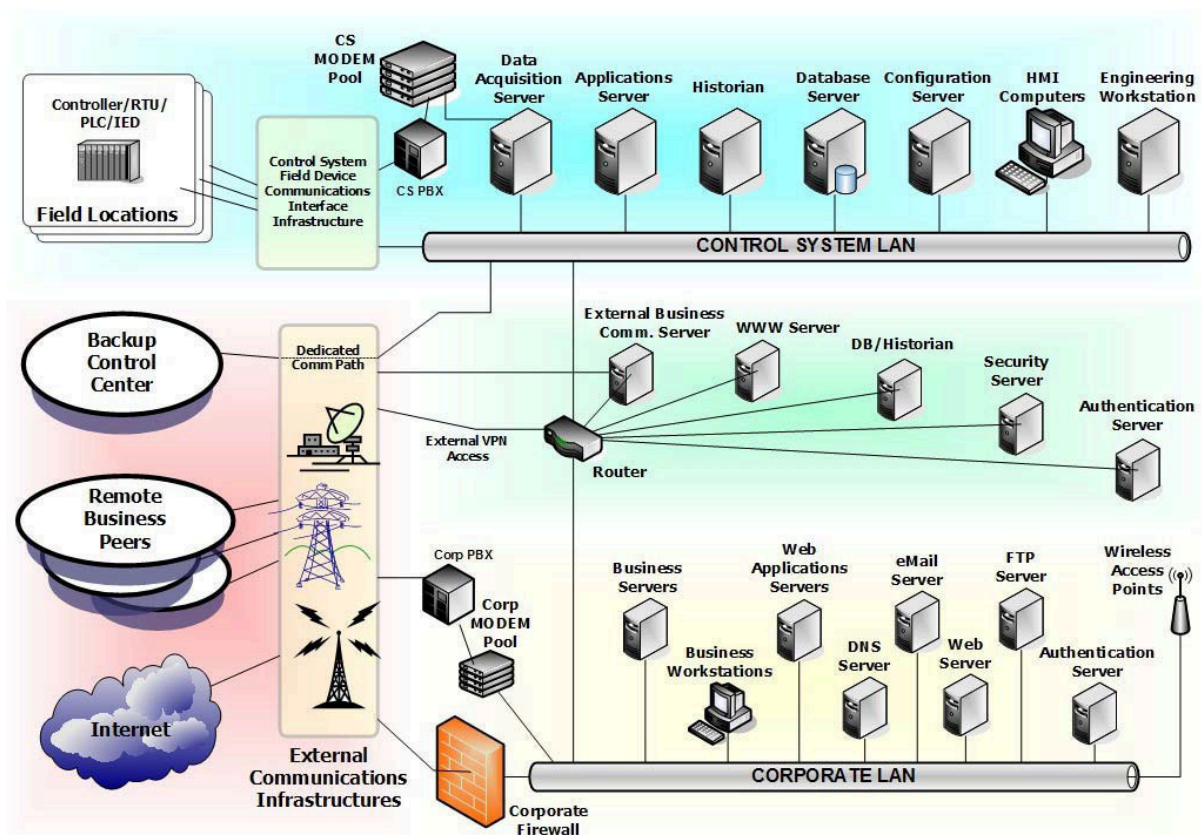


Figure 2 – Integrated networks

From Figure 2, it is clear that such architectures, if compromised, could provide an attacker with various avenues for accessing critical systems, either on the corporate LAN, the control LAN, or even the communications LAN. In addition, the very nature of such architectures demands the exchange of data from disparate information sources, a factor that could clearly be taken advantage of by an attacker.¹

¹ This type of architecture, and the back end control system, is vulnerable to both external attackers and internal attackers. Insider attacks have always been a major threat to IT systems, but architectures like that in Figure 2 exacerbate the issue by providing for access to a large, highly connected, and unprotected trusted information infrastructure. The insider has historically been a threat to control systems, but new connectivity creates opportunity for the external attacker as well.

Security Challenges in Control Systems

Within modern TCP/IP based environments, such as the corporate infrastructure for managing the business that drives operations in a control system, there are technology-related vulnerabilities that need to be addressed. Historically, these issues have been the responsibility of the corporate IT security organization, usually governed by security policies and operating plans that protect vital information assets. Clearly, the main concern as control systems become part of these large architectures is providing security procedures that cover the control system domain as well. Contemporary network-based communications have security issues that must be addressed in the control system domain, as unique vendor-specific protocols and assumed legacy system security is not adequate to protect mission critical systems.

Examples of threats in open systems architectures that can (and most likely will) migrate to control system domains include hostile mobile code (if applicable to the system), escalations of privileges through code manipulation, network reconnaissance and data gathering, covert traffic analysis, and unauthorized intrusions into networks either through or around perimeter defenses. With successful intrusion into control systems networks come new issues, such as reverse engineering of control system protocols, attacks on operator consoles, and unauthorized access into conjoined peer networks and remote facilities. To fully translate information security and information assurance into the control system realm, one must understand the key differences between traditional IT architectures and control systems technology.

From a mitigation perspective, simply deploying IT security technologies into a control system may not be a viable solution. Although modern control systems use the same underlying protocols that are used in IT and business networks, the very nature of control system functionality may make even proven security technologies inappropriate. Some sectors, such as energy, transportation, and chemical, have time sensitive requirements, so the latency and ‘throughput’ issues associated with security strategies may introduce unacceptable delays and degrade or prevent acceptable system performance.

There are several key differences between traditional IT environments and control system environments insofar as security is concerned. Table 1 shows some of the more common security elements an organization could leverage, and how they are addressed in IT domains, as opposed to architectures that run control systems.²

² NIST SP 800-82 will have a concise section discussing these differences.

SECURITY TOPIC	INFORMATION TECHNOLOGY	CONTROL SYSTEMS
Anti-virus/Mobile Code	Common Widely used	Uncommon/Impossible to deploy effectively
Support Technology Lifetime	2-3 Years Diversified vendors	Up to 20 years Single vendor
Outsourcing	Common Widely Used	Operations are often outsourced, but not diverse to various providers
Application of Patches	Regular Scheduled	Rare, Unscheduled Vendor specific
Change Management	Regular Scheduled	Highly managed and complex
Time Critical Content	Generally delays accepted	Delays are unacceptable
Availability	Generally delays accepted	24x7x365 (continuous)
Security Awareness	Moderate in both private and public sector	Poor except for physical
Security Testing/Audit	Part of a good security program	Occasional testing for outages
Physical Security	Secure (server rooms, etc.)	Remote/Unmanned Secure

Table 1 – Security focus in IT vs. Control Systems

Security Profiles and Attack Methodologies

Control networks have evolved from stand-alone islands to interconnected networks that co-exist with corporate IT environments, introducing security threats. For example, mobile code, in the form of viruses, worms, and parasitic code can manifest itself in network-enabled control system environments just as easily as in non-control system domains. For devices with embedded firmware, such as controllers and relays, hostile mobile code cannot generally have an impact through network propagation. However, should the compiled code these devices download on a regular basis be corrupted with hostile malware, the effects could be very damaging.³

³ Although the occurrence of this type of compromise is currently unlikely, such attack vectors should not be ignored when considering future attack scenarios.

Critical cyber security issues that need to be addressed include those related to:

- Backdoors and holes in network perimeter
- Vulnerabilities in common protocols
- Attacks on Field Devices
- Database Attacks
- Communications hijacking and ‘Man-in-the-middle’ attacks

Understanding attack vectors is essential to building effective security mitigation strategies. The level of knowledge in the control system community regarding these vectors may need to increase in order to mitigate these vulnerabilities. Effective security depends on how well the community of control system operators and vendors understand the ways that architectures can be compromised.⁴

Several in-depth technical discussions are provided by DHS in the Control Systems Security program via the DHS Computer Emergency Readiness Team (US-CERT). For this recommendations document, a discussion of various attack vectors may provide some insight into how a defense-in-depth strategy can be effective:

*Backdoor Attacks via Network Perimeter*⁵

As in common networking environments, control system domains are subject to myriad vulnerabilities and holes that can provide an attacker a ‘backdoor’ to gain unauthorized access. Often, backdoors are simple shortcomings in the architecture perimeter, or embedded capabilities that are forgotten, unnoticed, or simply disregarded. Adversaries (threats) often do not require physical access to a domain to gain access to it, and will use any and all discovered access functionality. Modern networks and especially those in the control system arena, often have inherent capabilities that are deployed without sufficient security analysis, and can provide access to attackers once they are discovered. These ‘backdoors’ can be accidentally created in various places on the network, but it is the network perimeter that is of greatest concern.

When looking at network perimeter components, the modern architecture will have technologies to provide for robust access. These technologies often include firewalls, public-facing services, and wireless access. Each of these will allow enhanced communications in and amongst affiliated networks, and will often be a sub-system of a much larger and more complex information infrastructure. However, each of these components can (and often does) have associated security vulnerabilities that an attacker will try to detect and leverage.

Unsecured wireless access is a reoccurring element in many organizations, and such deployments are common due to the ease-of-use of wireless communications as well as a low level of understanding regarding security implications of wireless deployments. Moreover, in the

⁴ The technical mechanics of attacks are beyond the scope of this paper.

⁵ http://www.us-cert.gov/control_systems/pdf/backdoors_holes0805.pdf

plant floor environment, wireless technology is easier to deploy than traditional wired infrastructures, which can require drilling through walls and laying cable.

Common security issues with wireless communications often include the residual effects of default installations. Attackers, once having discovered wireless communications points, can leverage the inherent functionality of wireless networks to their advantage, and take advantage of service set identifier (SSID) broadcasting, limited access controls, lack of encryption, and limited network segmentation.

Although much of the complexity in maintaining secure systems can be avoided by proper patch-management programs, there is a major problem for control system units when both geography and accessibility are a concern. Remotely located control system elements that can be accessed via remotely connected communications require special consideration. Often, if systems are based on commercial operating systems, the attacks can be via denial of service, escalated privilege exploits, or clandestine tools such as a Trojan horse or logic bomb.

With remote connectivity being commonplace, the security perimeter facilitating access to the control system has been moved back out to the corporate level or even the remote operator level. Clearly, compromising a computing resource that has access to a control system is the same as compromising the control system itself. This concern relates to the interception, modification, and re-injection of control data into a network, or the possibility of an attacker escalating privileges within the control domain to execute engineering level instructions across the control signal communications loop.

When considering the historical characteristics of controls system networks, especially those that impact security due to the presence of plaintext traffic and inherent trust relationships, unauthorized access via a wireless access point into the control domain provide an attacker with a very effective backdoor, often bypassing security perimeters.

To allow robust information transfer, organizations in many CI/KR sectors provide data to customers, providers, and affiliates through publicly accessible services. These services are critical to business operations in many sectors, such as electrical and water, as they provide data for calculating load expectations, billing futures, and associated operational information. As these services are in the public domain, they are often accessible from the Internet with little or no user access limitations. The data on these servers is usually sourced from the business domain (after it is collected from the control domain) as well as collected from the public domain.

This interconnected capability, as effective as it is, is also a vector for attackers to gain access into the protected business networks and perhaps the control systems networks. Attackers can often collect important information from these public servers, including data regarding operations, customers, and file transfers. Moreover, if the servers are compromised, the attackers can escalate their privileges and leverage the communication channel to the back-end business networks or even the control networks.

Networks with firewalls to separate public servers from internal networks often find it hard to defend against these types of attacks. To allow robust information to be provided via external services, such as a web or ftp server, communication must be made from the web server to the internal databases or historians, and this connection is made via the firewall. If deployed without effective security countermeasures, the trust relationship between the firewall and the web server allows data to flow from the external side to the internal domain. If this data is unauthorized, and is the product of an attack that has compromised the trusted web server, the attacker has a channel to access internal services on the business (or control systems) LAN.

In general, there is a delicate balance between business functionality and security. This balance has to be evaluated properly and revisited often. The deployment of modern technology to increase productivity and access requires special attention so prevent backdoors into the business or control system networks.

Attacks Using Common Protocols, i.e. OPC/DCOM attacks⁶

The impact of modern operating systems on control systems has been significant. Over the last several years, more and more organizations have started to use underlying services in these environments, some of them being the Object Link and Embedding (OLE), Distributed Component Object Model (DCOM), and Remote Procedure Call (RPC). OLE for Process Control (OPC) is a real-time data communications standard based in these services. Although many installations are moving away from the Microsoft-based OPC model, OPC is commonly used for efficient connectivity with diverse control systems equipment.

Historically, the data access standards provided by OLE, COM, and DCOM are found extensively in common computing platforms, and continue to be a significant target for attackers. The convergence of traditionally isolated control networks with business environments provides a new environment for attackers to exploit. What makes this very interesting, and also a concern, is that the traditional mitigation strategies for common networks are not always effective or practical in control systems architectures.

Applying security patches to operating systems and applications that run control systems is not a trivial endeavor. Prior to modification, rigorous testing must be completed to ensure that modifications do not impact operations. This, in itself, makes application of security patches, and thus the mitigation of security vulnerabilities, a challenge. In a simple example, SP2 for Microsoft XP, a platform commonly used in control systems, mitigates the security issues associated with some mobile code attacks by disabling DCOM. If this patch is deployed in a production environment where OPC is used for interoperability, OPC over DCOM will not work. There have been several reports of this patch bringing production facilities to a complete stop, or creating unexpected and irrational behavior in the control systems.

⁶ See US-CERT “Security Implications of OPC, OLE, DCOM, and RPC in Control Systems.”

With the convergence of control systems and modern networking technologies comes some inherited security vulnerabilities. Even though many of these vulnerabilities have solutions and available workarounds, the deployment of these mitigations in control systems architectures is not always feasible.

Attack into control system via field devices

Control systems architectures usually have a capability for remote access to terminal end points and telemetry devices. In some cases, the field equipment itself has the capability to be accessed a number of ways, including by telephonic or dedicated means. To provide for the collection of operational and maintenance data, some modern equipment has embedded file servers and web servers to facilitate robust communications. Engineers and administrators often have a secondary means of communicating with these field devices using this access capability in addition to other dedicated communications channels.

However, as has been previously discussed, these devices are part of an internal and trusted domain, and thus access into these devices can provide an attacker with an unauthorized vector into the control system architecture. By gaining access into a field device, the attacker can become part of the sensor network and ‘tunnel’ back into the control system network. Recognizing that field devices, such as RTUs, are an extension of the control domain, attackers can add these field devices to their list of viable targets to be investigated during reconnaissance and scanning phases of the attack. Although such attacks are typically not possible across serial connections, the security related to the convergence of modern networking protocols and traditional control protocols in remote devices requires attention.

If a device is compromised, and the attacker can leverage control over the device and escalate privileges, the attacker can begin to execute a number of procedures, including scanning back into the internal control network, altering the data that will be sent to the control master, or changing the behavior of the device itself. If the attacker decides to scan back into the control network, which is probable considering the assumed trust between resources, it may be possible to do so by using the communications protocols for the entire control system domain. This is of particular advantage to the attacker, as it is likely that the connections are not monitored for malicious or suspect traffic.⁷

*Database and SQL data injection attacks*⁸

Database applications have become core application components of control systems and their associated record keeping utilities. Traditional security models attempt to secure systems by isolating core control system components and concentrating security efforts against threats

⁷ Some Intrusion Detection Systems (IDS) can be updated with control systems signatures to help defend control domains. Usually, these systems are signature-based and will trigger on seeing recognized malicious traffic. In lieu of viable signature, IDS can be deployed to trigger on non-specific traffic, or upon seeing traffic that is not expected or unusual. See below for the discussion on IDS.

⁸ See US-CERT “Attack Methodology Analysis: SQL Injection Attacks.”

specific to those computers or software components. Database security within control systems follows these models by using generally independent systems that rely on one another for proper functionality. The high level of reliance between the two systems creates an expanded threat surface.

Databases used by control systems are often connected to databases or computers with web-enabled applications located on the business network. Virtually every data-driven application has transitioned to some form of database, and most use Structured Query Language (SQL).

The information contained in databases makes them high-value targets for any attacker. When control system databases are connected to business or financial databases or to computers with applications used to access the data, attackers can exploit the communications channel between the two networks and bypass the security mechanisms used to protect the control system environment.

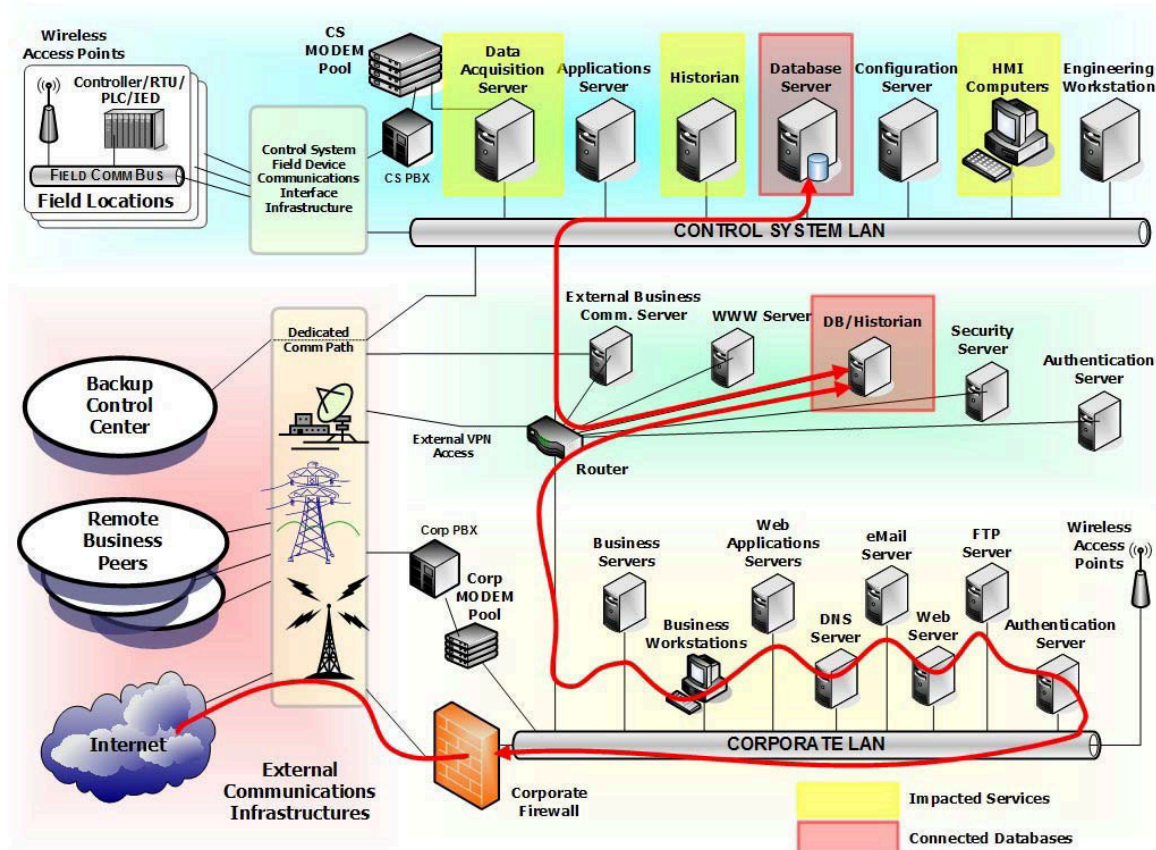


Figure 3 – Attacking via databases

Figure 3 shows an example of the open connectivity between databases. This example illustrates a communication path between the servers that an attacker would be able to leverage to gain access to the control network. Injection into a database with valuable data can have far-reaching

effects, especially in a control system environment where data accuracy and integrity are critical for both business and operational decision-making. The cascading effect of corrupted database content can impact data acquisition servers, historians, and even the operator HMI console. Control systems are more adversely affected by SQL injection than are many general IT databases because they are so reliant on data accuracy and integrity. Moreover, compromise of key trusted assets, such as a database, creates additional resources the attacker can use for both reconnaissance and code execution.

Given the reliance of control systems on the storage, accuracy, and accessibility of command and control data, as well as the prevalence of SQL databases on these types of networks, standard SQL injection techniques against control system components pose a major threat to control system security.

Man-in-the-middle attacks⁹

Control system environments have traditionally been (or been intended to be) protected from non-authorized persons by air gapping. In these networks, data that flows between servers, resources, and devices is often less secured. Three of the key security issues that arise from assumed trust are (1) the ability for an attacker to re-route data that is in transit on a network, (2) the ability to capture and analyze critical traffic that is in plaintext format, and (3) the ability to reverse engineer any unique protocols to gain command over control communications. By combining all of these, an attacker can assume exceptionally high control over the data flowing in a network, and ultimately direct both real and ‘spoofed’ traffic to network resources in support of the desired outcome. To do this, a ‘man-in-the-middle’, or MITM, attack is executed.

Management of addresses in a network, be it a control system or a business LAN, is critical to effective operations. Address Resolution Protocol (ARP) helps maintain routing by helping map network addresses to physical machine addresses. Using ARP tables in each of the network devices ensures that computers and other devices know how to route their traffic when requesting communication. Thus manipulation (or poisoning) of the ARP tables is a key goal of the attacker, as poisoning the ARP tables can force all network traffic (including control traffic) to be routed through the computer the attacker has compromised. In this manner, all resources on the network will have to talk to the attacker without knowing they are not communicating with the desired host. Moreover, the attacker can see, capture, replay, and inject data into the network and have it interpreted as if it were authorized and coming from a valid source.

Data analysis is a significant problem in the control system realm. For control system installations that are governed by common (open) network protocol and technologies, the threat of data analysis while the data is in transit is a major concern. The data that often traverses today’s local control environment is in plaintext. Historically, this vulnerability has allowed attackers to observe and re-use username/password combinations to increase access within a compromised network.

⁹ http://www.us-cert.gov/control_systems/pdf/csvul1105.pdf

Assuming an attacker has gained access onto the controls systems network, perhaps using any of the aforementioned attacks, he will use network reconnaissance to determine resources that are available on that network. As the attack is on the control domain, this plaintext traffic can be harvested (sniffed) and taken offline for analysis and review. This allows the attacker to review and re-engineer packet and payload content, modify the instruction set to accommodate the goal of the attack, and re-inject the new (and perhaps malicious) packet into the network. Control traffic, regardless of its unique nature, is not complex insofar as the nomenclature used for instruction in data payloads, and the data contained in the packets is used to control the action of the field devices and to provide input as to what is seen by the operator at the Human Machine Interface (HMI) station.

Using ARP poisoning and collecting traffic, the attacker can establish and maintain complete control over the communications in the network. If the attacker needs to acquire and analyze unique control system protocols, control data can be seen, captured, and manipulated. The time to reverse engineer key control data, and manipulate that data for nefarious purposes, is now available to the attacker.

In any environment, this MITM attack is exceptionally dangerous. However, in the control systems networks this mode of attack becomes even more critical. By assuming control of a key information resource, and performing a MITM attack, the attacker can continue to attack the system by:

- Stopping operations
- Capturing, modifying, and replaying control data
- Injecting inaccurate data to falsify information in key databases, timing clocks, and historians
- Replaying normal operational data to the operator HMI while executing a malicious attack on the field device (while preventing the HMI from issuing alarms).

Isolating and Protecting Assets: Defense-in-Depth Strategies

Modern IT architectures that involve both business and control network components share many common characteristics, regardless of how diversified their applications may be. In general, there are four main domains (zones) that provide:

- External connectivity to the Internet, peer locations, and back-up facilities (Zone 1)
- External connectivity for corporate communications (Zone 2)
- Control systems communications from external services (Zone 3)
- Control systems operations, be they process based or SCADA (Zone 4).

Figure 4 illustrates a common modern architecture that contains all of these zones.

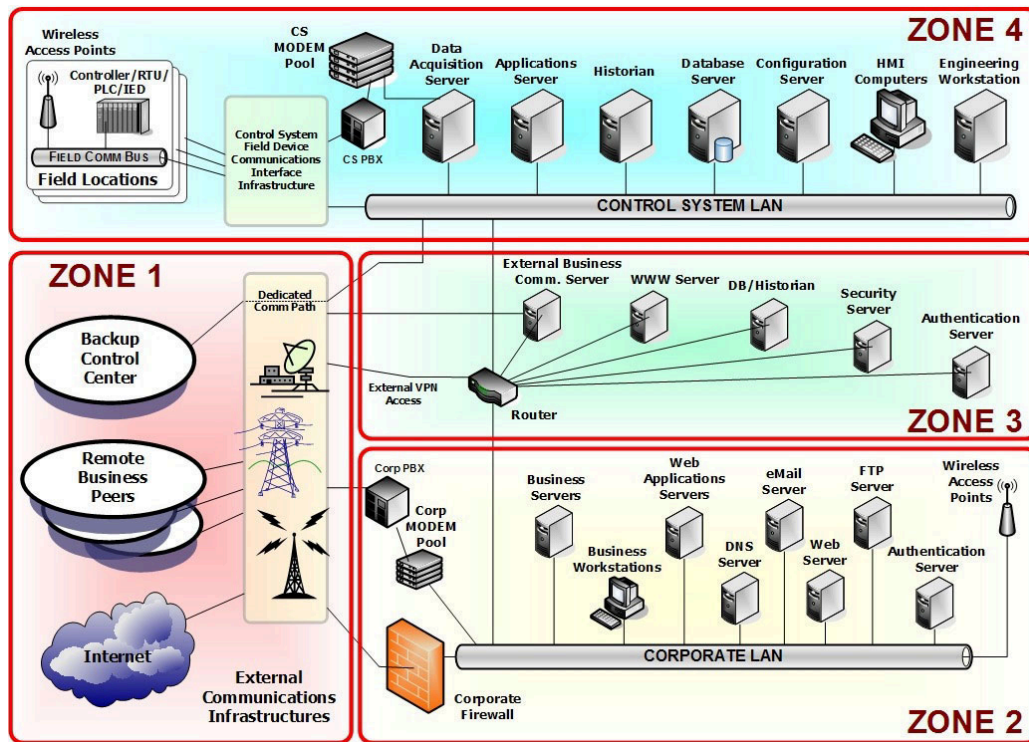


Figure 4 – Common architecture zones

Each of these zones requires a unique security focus. A ‘peel-the-onion’ analysis shows that an attacker trying to affect a critical infrastructure system would most likely be after the core control domain.¹⁰ Manipulation of the control systems information resources can be devastating if this critical zone is compromised. In many sectors the malicious attack on the control system will have real-world, physical results.

¹⁰ This of course depends on the overall objective of the attacker. In general it is believed that complete control over core services and operational capability of the control system has high value.

In this document, and in the suggested supporting documentation provided by DHS through US-CERT, numerous categories of attacks and outcomes have been discussed. In each of those scenarios, the intrusion begins at some point outside the control zone and the attacker pries deeper and deeper into the architecture.

Thus, defensive strategies that secure each of the core zones can create a defensive strategy with depth, offering the administrators more opportunities for information and resources control, as well as introducing cascading countermeasures that will not necessarily impede business functionality.

Firewalls

Firewalls provide additional levels of defense that support the traditional routers, providing the capability to add much tighter and more complex rules for communication between the different network segments or zones. Of critical importance to control systems is how the firewall is implemented and, to a certain degree, how the core functionality of the firewall impacts the overall business functionality of the environment.

There are many types of firewalls, and some research is required to ascertain what type of firewall is right for a given control architecture. The concept of security zones, as discussed earlier, provides some insight as to how an organization can determine what risk and consequence is associated with a particular zone. This analysis can be used to select the type of firewall and attributes that are best suited for protecting the assets. In general, there are four main types of firewalls: packet filter, circuit level gateways, proxy gateways, and stateful inspection.

Packet filter firewalls – These firewalls analyze the packets going into and out of separated networks, and either permit or deny passage based on a pre-established set of rules. Packet filtering rules are based on port numbers, protocols, and other defined data that correlates to the type of data request being made. Although usually flexible in assigning rules, this type of firewall is well suited for environments where quick connections are required and rules can be developed based on device addresses. It is effective for environments, such as control systems, that need security based on unique applications and protocols.

Proxy Gateway Firewalls – These firewalls are critical in hiding the networks they are protecting, and are used as primary gateways to proxy the connection initiated by a protected resource. Often called *Application-level* gateways, they are similar to circuit-level gateways except that they address the application. They filter at the Application Layer of the OSI model, and do not allow any connections for which there is no proxy available. These firewalls are good for analyzing data inside the application (POST, GET, etc.), as well as collecting data about user activities (login, admin, etc.). They are gateways and require users to direct their connection to the firewall. They also have some impact on network performance due to the nature of the analysis. In control systems environments, this type of firewall is well suited to separating the

business and control LANs, as well as providing protection to a demilitarized zone (DMZ) and other assets that require application-specific defenses.

Stateful Inspection Firewalls – These firewalls include characteristics of all the other types of firewalls. They filter at the network layer, determine the legitimacy of the sessions, and evaluate contents of the packets at the application layer. They tend to use algorithms to process data rather than run proxies. These firewalls execute a considerable amount of inspection of packets that are arriving on the interfaces. These firewalls look at the ‘state’ of the packets and analyze against pre-observed activities, thus allowing for a higher level of trust when deciding what is allowed and what is not. These firewalls are capable of keeping track of valid sessions, and make a good choice for protecting key assets in the control domain. Because many of the vulnerabilities in control systems have their roots in trust shared amongst servers and devices, being able to track and react to valid and invalid sessions is advantageous’

Figure 5 illustrates the deployment of layered firewalls in a multi-zone architecture.

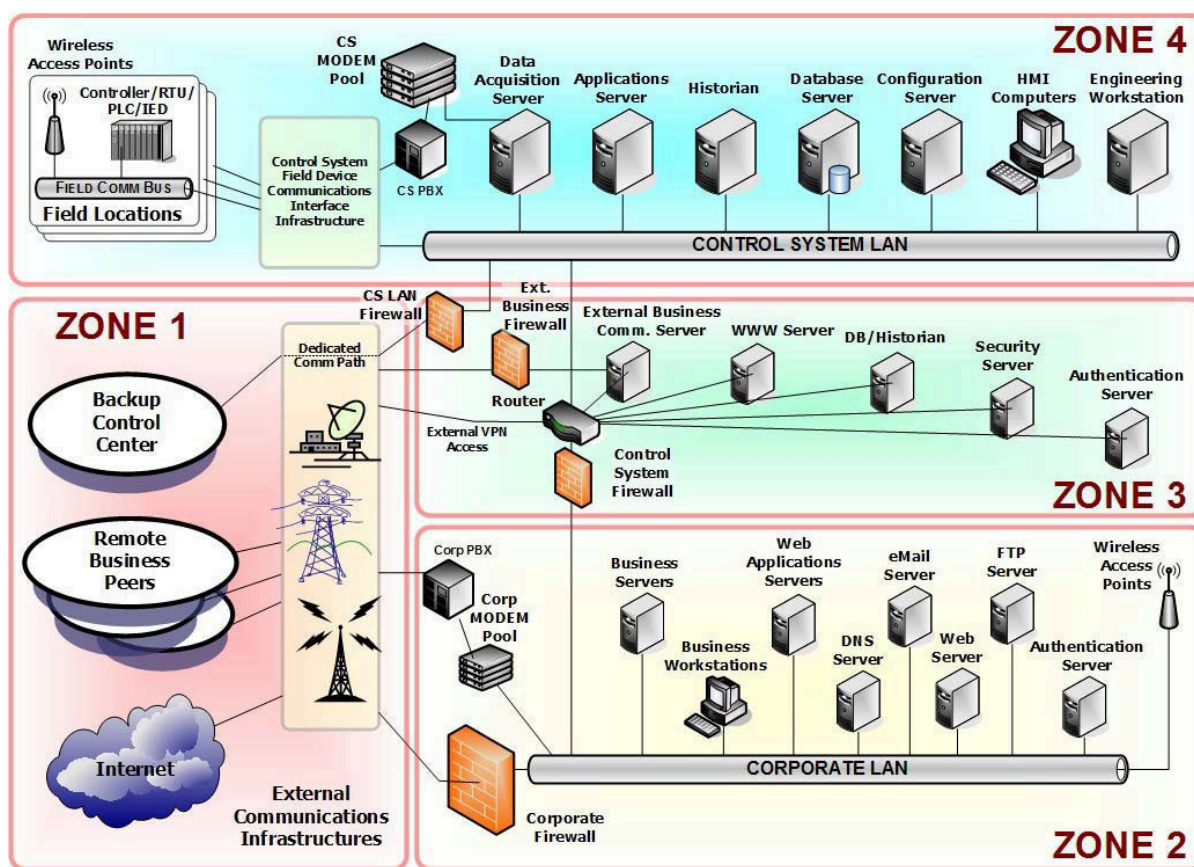


Figure 5 – Firewalls protecting architecture zones

To help understand how firewalls can be deployed to provide robust network separation, Figure 6 shows an example of firewall deployment in the energy management domain. In this case, the firewall isolates from the corporate domain and protects the Energy Management System (EMS) technologies found in the controls center. To provide defense-in-depth, an additional firewall has been deployed to separate the EMS domain from the SCADA system.

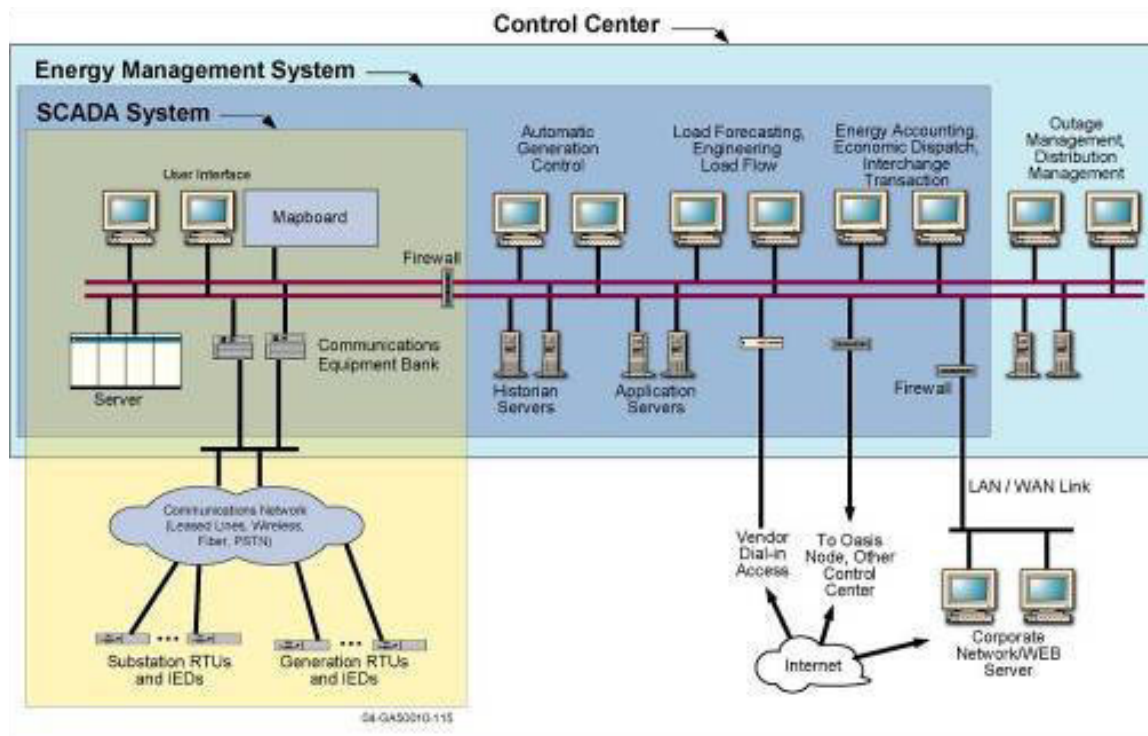


Figure 6 – Firewall deployment in an energy management network

Well-configured firewalls are critical to control system security. Communications should be restricted to that necessary for system functionality. Control system traffic should be monitored, and rules should be developed that allow only necessary access. Any exceptions created in the firewall rule set should be as specific as possible, including host, protocol, and port information.

A common oversight is not restricting outbound traffic. Firewall rules should consider both directions through the firewall. Most administrators effectively block traffic into the control network, but do not filter traffic out of the network. Outbound traffic rules should also be created, and such rules should initially have no exceptions. These rules should be fine-tuned so a rule set that excludes all unnecessary traffic is created. Once the necessary outbound traffic has been determined, a safer configuration can then be created that blocks all traffic with exceptions for necessary communication.

Traditionally, the role of the firewall in defending networks is straightforward. Like attacks on business networks, an attacker targeting a control system needs to obtain information from and send files and commands to the control system network. To remotely control any exploit code

running on a control system computer, a return connection must be established from the control network. With regards to attacking resources in the control system domain, exploit code must be small and contain just enough code to get an attacker onto the target computer, as there is generally not enough space to add logic onto the device for the attacker to get advanced functionality. Therefore, additional instructions are needed from the attacker to continue with the discovery portion of the attack. If outbound filtering is implemented correctly, the attacker will not receive this return connection and cannot discover and control the exploited machine.¹¹

Creating Demilitarized Zones (DMZs)

Network segmentation has traditionally been accomplished by using multiple routers. Firewalls should be used to create DMZs to protect the control network. Multiple DMZs could also be created for separate functionalities and access privileges, such as peer connections, the data historian, the Inter Control Center Communications Protocol (ICCP) server in SCADA systems, the security servers, replicated servers, and development servers. Figure 7 shows a robust architecture with multiple DMZ deployments.

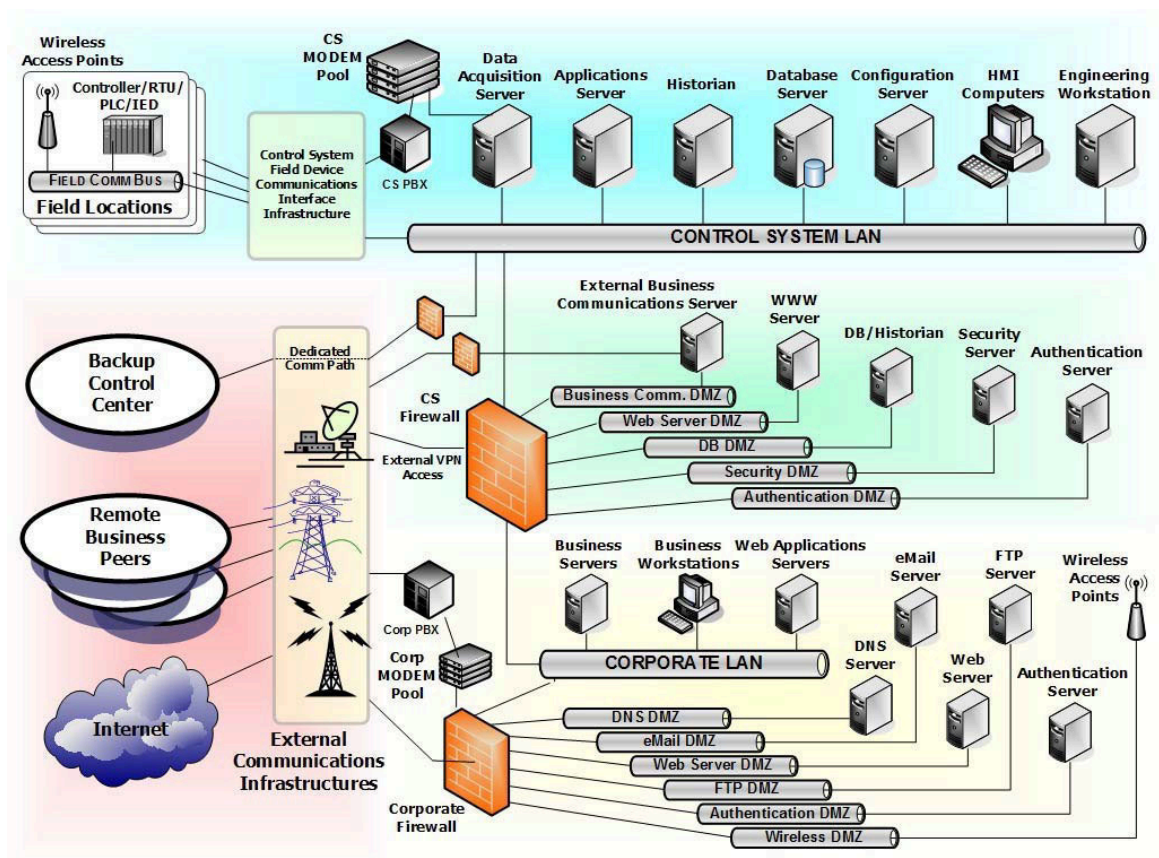


Figure 7 – Architecture with DMZ deployments

¹¹ www.niscc.gov.uk/niscc/docs/re-20050223-00157.pdf

All connections to the Control System LAN should be routed through the firewall, with no connections circumventing it. Network administrators need to keep an accurate network diagram of their control system LAN and its connections to other protected subnets, DMZs, the corporate network, and the outside.

Multiple DMZs have proved to be very effective in protecting large architectures comprised of networks with different operational mandates. A perfect example, illustrated in Figure 7, is the conjoined networks for control systems and business. In this example, the secure flow of data into and out of the different environments is critical to operations. Having multiple DMZs protects the information resources from attacks using Virtual-LAN (VLAN) hopping and trust exploitation, and is a very good way to enhance the security posture and add another layer to the defense-in-depth strategy.

Intrusion Detection Systems

When considering the most logical route an attacker will take in compromising a control network, it is easy to visualize an attack path that pries deeper and deeper into the architecture. Starting from the external environment, an attacker will move past perimeter devices and ultimately strive for access to both the network and hosts on that network. Bear in mind that this access may be via field devices where remote access requirements can introduce vulnerabilities into control system architectures. Once on the target network, the attacker must begin to collect intelligence through reconnaissance, followed by attempts at compromising more and more components. In each of these cases, unusual and unauthorized activity would be present in the network, and this activity can be monitored (and acted upon) to provide another level of defense.

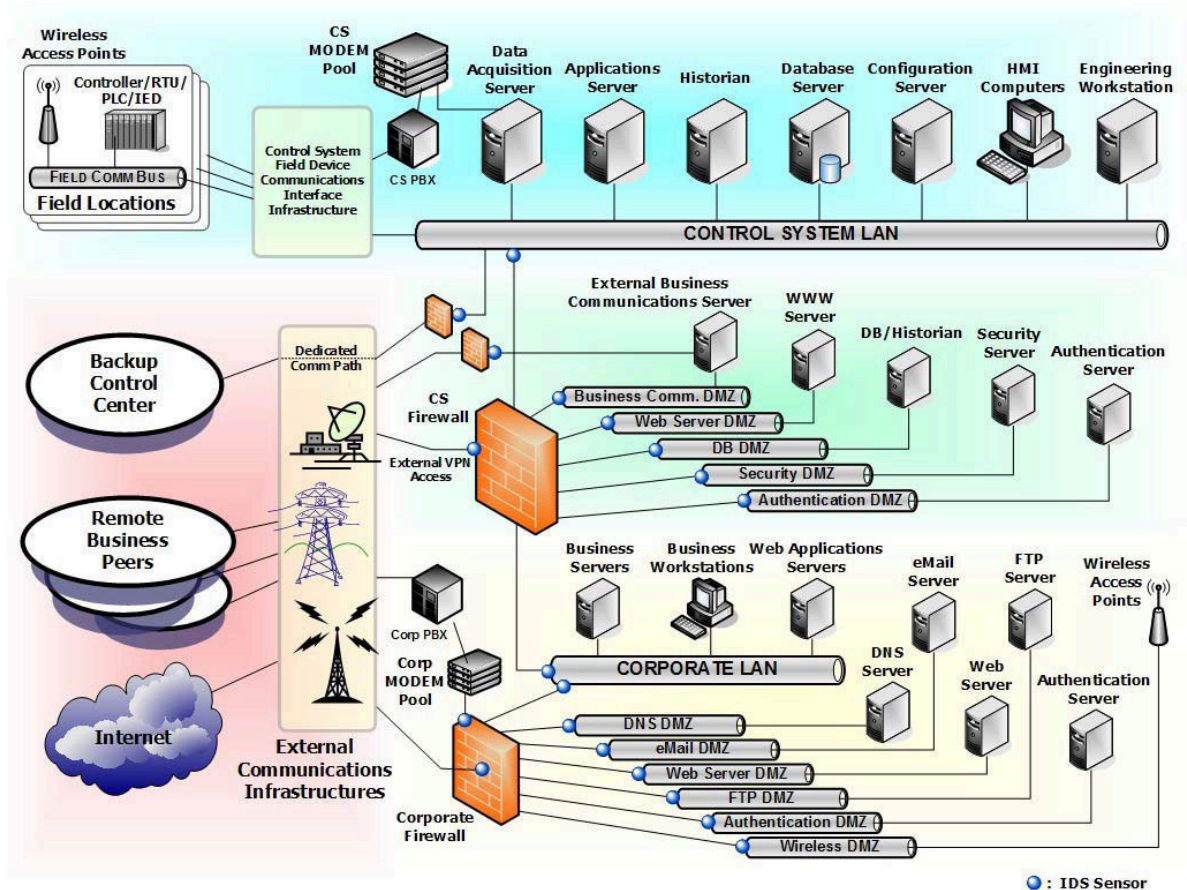
There are several common methods for monitoring a network for unusual or unauthorized activity, with one of the most effective being Intrusion Detection Systems (IDS). It is important to keep in mind that intrusion detection is not a single product or technology. It is a comprehensive set of tools providing network monitoring that can give an administrator a complete picture of how the network is being used. Implementing a variety of these tools helps to create a defense-in-depth architecture that can be more effective in identifying attacker activities.

An intrusion detection system, by its very nature, is passive. In a network deployment, the function of the IDS is to watch and assess the traffic or network activity without impacting that traffic. Having collected data, IDS compares it against a pre-defined rule set, as well as against a set of known attack ‘signatures’. The IDS will investigate port numbers and data payload to determine if any nefarious activity is occurring. Having recognized an attack pattern, or any deviation from what has been defined as normal/allowable traffic, the systems will carry out a set of instructions that can include alerting a systems administrator. Extensive logging is also a function of most IDSs available today.

Most IDS are signature based. In modern business environments, this is very acceptable, as there is an abundance of signatures for many network and host architectures using modern protocols and modern operating platforms. Security vulnerabilities in the contemporary business domain are also common, so it is often easy to fine-tune IDS for networks and hosts using ubiquitous technologies. Like the issues surrounding the deployment of patches and other security technologies in controls systems, the configuration and deployment of IDS is not straightforward. For example, even though many contemporary IDS signatures files are very robust and can detect a wide range of attacks, the signatures required to monitor for malicious traffic in control networks are not adequate. When looking at the unique communications protocols used in control systems, such as Modbus or DNP3, specific payload and port numbers have traditionally not been a part of the signatures seen in contemporary IDS. In short, modern IDS deployed on a control systems network may be blind to the types of attacks that a control system would experience.

When deploying IDS in a control system, the ability to add unique signatures must be used. It is also commonplace to remove some of the default signatures and response capability, as it may have no relevance to a control system network. However, analysis must be made to ensure some of the inherent capability of the IDS is leveraged, with some of the capability refined and augmented. Many security vendors, including those specializing in control systems security, have created signatures for the IDS that are deployed in control architectures. It is imperative, when deploying IDS on control system networks, that rules sets and signatures unique to that domain be used. It has been shown that developing security signatures and rules in a cooperative relationship with the control system vendor is very advantageous.

One of the common problems observed in industry is that tools deployed for network monitoring are implemented but improperly updated, monitored, or validated. Assigned individuals should be trained and given the responsibility of monitoring system data logs and keeping the various tool configurations current.



Complete defense-in-depth strategy with IDS

Deploying IDS at the host level is similar to deploying it at the network level, but rather than monitoring network activity, the IDS monitors with respect to rule sets. These rules can be very robust and extensive, and can include alerting on pre-defined signatures that are unique to the platform or operating systems the host is running. IDS placement at the host level provides yet another level of defense-in-depth, and can be used to augment the defense strategies deployed at the perimeter and network levels.

It is important to realize that due to the passive nature of IDS, security mitigation and attack realization is a function of how often (and how effective) the analysis of log files is done. Robust policies directing the timely analysis of IDS log is very important. If an attacker is able to gain access to a system and execute an attack prior to the log files being reviewed, IDS and the ability to counter an attack becomes a moot point.

The Security Policy

Effective security policies and procedures are the first step to a secure control systems network. Many of the same policies used for Information Technology (IT) security for corporate systems can be applied directly to control system networks. The SANS Institute provides free templates

for many types of security policies, and can be a valuable resource for control system network administrators in developing their own policies. Control system-specific requirements can then be added to it, such as the North American Electric Reliability Council (NERC) cyber security requirements for electric systems.¹²

To make the security policy effective, it must be practical and enforceable, and it must be possible to comply with the policy. The policy must not significantly impact productivity, be cost prohibitive, or lack support. This is best accomplished by including both management and system administrator personnel in policy development.

Network and control system administrators have technical knowledge, but they also need authorization and support from management to implement the policy. Management must support the appointment and development of appropriate personnel resources to implement and administer control system security.

Security Training

In many cases, the individuals administering a control system network may not have adequate security training. This situation is generally due to a lack of funding or appreciation for the importance of this training. Training is a core component of an overarching security awareness program, and is comprised of several key attributes used to support the protection of key information and information resources.

Security training and robust security awareness programs that are specific to the controls systems domain are critical to the security of the control systems, as well as the safety of those involved with any automated processes. Like the security awareness programs that are developed for the corporate domains, the programs that will support control systems domains have key components that can help drive a continuous and measurable security posture. Within common security awareness programs, such as those listed in NIST SP800-50 '*Building an Information Technology Security Awareness and Training Program*'¹³, organizations can create applicable security awareness and training curricula that can include:

- Purpose and Scope
- Materials Development
- Implementation Strategies
- Monitoring and Feedback
- Success Measurement

Network security administrators require continuous training to keep up to date with the fast-paced changes and advances in the network security field. This includes the latest network architecture designs, firewall and IDS configurations. New techniques are constantly being

¹² http://www.nerc.com/~filez/standards/Cyber_Sec_Renewal.html

¹³ <http://www.csrc.nist.gov/publications/nistpubs/800-50/NIST-SP800-50.pdf>

developed to attack, and to defend, computer networks. It is very important to have comprehensive computer security training, not only for system administrators, but also for each user.

If formal training is cost prohibitive, some of this information can be gleaned from books, papers, and web sites on cyber and control systems security. As an example, and as a resource to help build content for a control systems specific training curriculum, US-CERT maintains a website dedicated to control systems cyber security, at http://www.us-cert.gov/control_systems/

Incident Response

To fully support a defense-in-depth strategy, a robust incident response capability is required. In the event there is a security-related incident in the controls system domain, activities to recognize, respond, mitigate, and resume need to be established. An incident response procedure will instruct employees on the steps to take if a computer on the network has been compromised. All employees should be trained on, and have access to, the procedure before an incident occurs. Examples of questions to be answered in the incident response procedure include:

- What are the indications that an incident has occurred or is currently in progress?
- What immediate actions should be taken (e.g., should the computer be unplugged from the network)?
- Who should be notified, and in what order? Should law enforcement be consulted?
- How should forensic evidence be preserved (e.g., should the computer be left on to preserve the evidence in memory)?
- How can the affected computers be restored?

The National Institute of Standards and Technology (NIST) developed a Computer Security Incident Handling Guide, SP 800-53, which provides guidance to security personnel in developing an incident response procedure. In addition, US-CERT has extensive information and reporting capabilities available for any control system security incident. This reporting can be done at http://www.us-cert.gov/control_systems/.

Specific Recommendations and Countermeasures

When protecting any information infrastructure, good security starts with a proactive security model. This iterative model is comprised of several key security strategies that are illustrated in Figure 8.

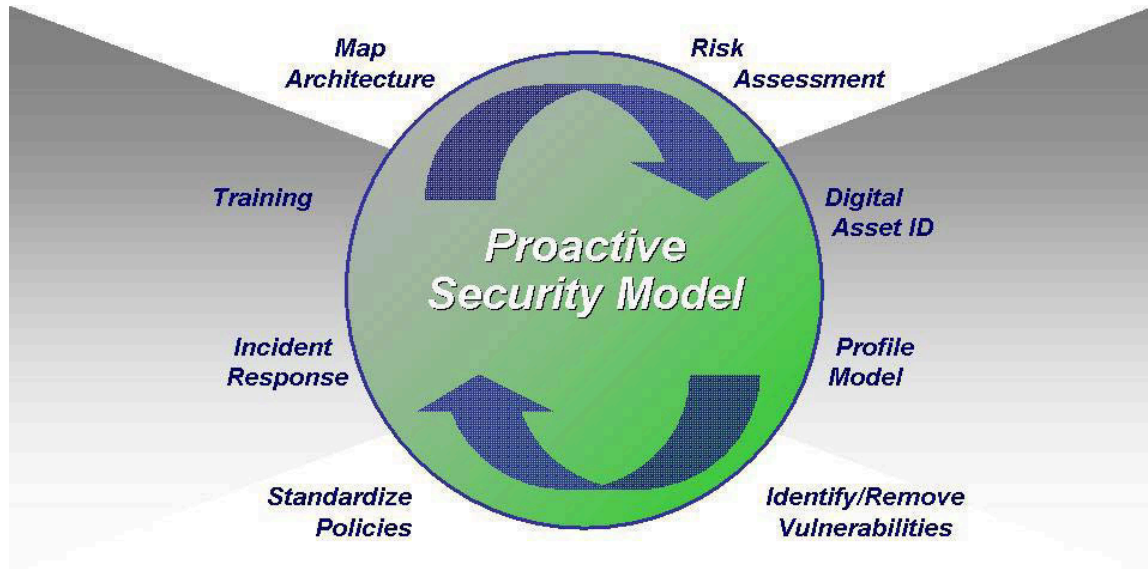


Figure 8 – Proactive security model

Traditionally, development of a defense-in-depth strategy starts with mapping the control systems architecture. Having an accurate and well-documented architecture can enable an organization to be very security-conscious, deploy effective security countermeasures, and be equipped to understand security incidents more readily. Having an understanding of the architecture will allow the administrators to know what it is they want to protect. A robust understanding of architecture also allows for effective risk assessments, as the development of the assessment parameters and processes can be easily aligned to the existing (and known) information assets in the control system environment.

Having been able to execute a security assessment, the organization can now assign asset ID's within the control domain, leading to definition of the overall profile of the command and control environment. Following the development of the profile, the defense-in-depth strategy can be deployed, as much of the final phases of the mitigation strategy involves the deployment of technology that is supported by recursive and on-going security training.

5 Key Security Countermeasures for Control Systems







Here are 5 key countermeasures that can be used to drive cyber-security activities in control systems environments.

1. Security policies. Security policies should be developed for the control system network and its individual components, but they should be reviewed periodically to incorporate the current threat environment, system functionality, and required level of security
2. Blocking access to resources and services. This technique is generally employed on the network through the use of perimeter devices with access control lists, such as firewalls

or proxy servers. It can be enabled on the host via host-based firewalls and anti-virus software.

3. Detecting malicious activity. Detection activities of malicious activity can be network or host based and usually requires regular monitoring of log files by experienced administrators. Intrusion detection systems are the common means of identifying problems on a network, but they can be deployed on individual hosts as well. Auditing and event logs should be enabled on individual hosts when possible.
4. Mitigating possible attacks. In many cases, a vulnerability may have to be present due to the fact that removal of the vulnerability may result in an inoperable or inefficient system. Mitigation allows administrators to control access to vulnerability in such a fashion that the vulnerability cannot be exploited. Enabling technical workarounds, establishing filters, or running services and applications with specific configurations can often do this.
5. Fixing core problems. The resolution of core security problems almost always requires updating, upgrading, or patching the software vulnerability or removing the vulnerable application. The software hole can reside in any of the three layers (networking, operating system, or application). When available, the mitigation should be provided by the vendor or developer for administrators to apply.

Suggested Reading

- INL Risk Doc (when available)
- INL SQL doc (when available)
- INL OPC/DCOM doc (when available)
- [Control Systems Cyber Security Awareness](#) 
- [An Undirected Attack Against Critical Infrastructure: A Case Study for Improving your Control System Security](#) 
- [Backdoors and Holes in Network Perimeters: A Case Study for Improving your Control System Security](#) 
- [Common Control System Vulnerability](#) 
- [A Comparison of Electrical Sector Cyber Security Standards and Guidelines](#) 
October 2004
- [Intruder Detection Checklist](#)
- [Personnel Security Guidelines](#) 

- [A Comparison of Oil and Gas Segment Cyber Security Standards](#) 

Glossary

ASN – Abstract Syntax Notation
CC – Common Criteria
CERT – Computer Emergency Response Team
COE – Common Operating Environment
COM – Common Object Model
DCE – Data Communications Equipment
DCOM – Distributed Common Object Model
DCS – Distributed Control System
DHS NCSD – DHS National Cyber Security Division
DOS – Denial of Service
DNP – Distributed Network Protocol
DOI – Domain of Interest
DTE – Data Terminal Equipment
EOP – Emergency Operating Procedures
EPA – Enhanced Performance Architecture
ES-ISAC – Energy Sector ISAC
FIPS – Federal Information Processing Standard
FTP – File Transfer Protocol
I&W – Indications and Warning
ICS – Industrial Control System
IEC – International Electrotechnical Commission
IEC TC – International Electrotechnical Commission technical Committee
IED – Intelligent Electronic Device
IP – Internet Protocol
ISAC – Information Sharing and Analysis Center
ISO – International Standards Organization
NIPC – National Infrastructure Protection Center
NIST – National Institute of Standards and Technology
OEM – Original Equipment Manufacturer
OLE – Object Linking and Embedding
OPC – OLE for Process Control
OSI – Open Systems Interconnectivity
PCS – Process Control System
PLC – Programmable Logic Controller
POTS – Plain Old Telephone Service
PSTN – Packet Switched Telephone Network
RPC – Remote Procedure Call
RTU – Remote Terminal Unit/Remote Telemetry Unit
SCADA – Supervisory Control and Data Acquisition
SIRC – Security Incident Response Capability
SMTP – Simple Mail Transfer Protocol
SNMP – Simple Network Message Protocol
SOP – Standard Operating Procedures
SSID – Service Set Identifier: A code attached to all packets on a wireless network to identify each packet as part of that network.
TCP – Transmission Control Protocol
TCSEC – Trusted Computer System Evaluation Criteria (a.k.a Orange Book)

TFTP – Trivial File Transfer Protocol

UDP – User Datagram Protocol

WARDIALING – Recursive dialing of phone numbers from a modem-enabled PC in an attempt to locate other unadvertised modems resulting in unauthorized access into a computing or PCS domain

WARDRIVING – Recursive searching for wireless access points in an attempt to access a communication network resulting in unauthorized access into a computing or control system domain

X86 – pronounced 'ex-eighty six': The standard abbreviation for the Intel 32-bit processor