

Mitigations for Security Vulnerabilities Found in Control System Networks

KEYWORDS

Control system, SCADA, cyber security, mitigation, firewall, IDS, encryption, DMZ

ABSTRACT

Industry is aware of the need for Control System (CS) security, but in on-site assessments, Idaho National Laboratory (INL) has observed that security procedures and devices are not consistently and effectively implemented. The Department of Homeland Security (DHS), National Cyber Security Division (NCSD), established the Control Systems Security Center (CSSC) at INL to help industry and government improve the security of the CSs used in the nation's critical infrastructures. One of the main CSSC objectives is to identify control system vulnerabilities and develop effective mitigations for them. This paper discusses common problems and vulnerabilities seen in on-site CS assessments and suggests mitigation strategies to provide asset owners with the information they need to better protect their systems from common security flaws.

INTRODUCTION

Events during recent years have increased awareness that the computer systems controlling our nation's critical infrastructures are vulnerable to cyber attack. The INL Critical Infrastructure Protection Division supports multiple programs sponsored by government and private sector clients to enhance critical infrastructure security. The CSSC is working to improve the cyber security of critical infrastructure CSs. A significant part of this effort is to assess existing cyber security vulnerabilities in various types of CSs at critical infrastructure sites. The National SCADA Test Bed, funded by the Department of Energy Office of Electricity Delivery and Energy Reliability (DOE-OE), is working to improve the cyber security of CSs that operate the nation's energy infrastructure. Under these two government programs, and in assessments for private clients, valuable insight has been gained into the security issues facing our nation's critical infrastructure CSs.

The intent of this paper is to provide general information regarding mitigation strategies for CSs. There are significant differences between various CS architectures, but CS will be used to refer to Supervisory Control and Data Acquisition (SCADA) systems, Process Control Systems (PCS) and Distributed Control Systems (DCS) which make up the various general configurations of CSs. Also, the use of "process" throughout this paper represents the physical system being monitored and controlled by the CS.

BACKGROUND

To better understand the reasoning behind the security recommendations that follow, the steps an attacker would take to compromise a CS are discussed first.

To remotely manipulate or attack a CS, an attacker must successfully accomplish the following:

1. Gain access to the CS Local Area Network (LAN).
2. Discover and understand the Process.
3. Control the Process.

The first step in gaining control of a CS is to get access to the CS LAN. Fortunately, most CS networks are no longer directly accessible from the Internet. It is now common industry practice to separate the business LAN from the CS LAN with a firewall. The firewall helps keep hackers out and isolates the CS LAN from worms and other maladies that may infect the corporate network. The firewall can also be used to separate the CS network into sub-networks known as demilitarized zones (DMZs), and control access between them. These sub-networks are used to safely share data between the corporate and CS LANs. DMZs also keep non-CS applications off the CS LAN. To access the CS LAN, the attacker must first bypass the perimeter defense provided by the firewall or find another avenue onto the CS LAN. The attacker can use a number of proven techniques, such as piggybacking on a connection allowed through the firewall, discovering an auto-answer modem or connection circumventing the firewall, or gaining access through a trusted peer site. The attacker needs to maintain access to the SCADA network in order to accomplish the rest of the attack.

After gaining access to the CS LAN, an attacker must discover details about the process under control of the CS. If the attacker's goal is merely to shut down the process, very little discovery is needed. However, if the attacker intends a surgical attack or process manipulation, specific details are needed. The main sources of information about the process on the CS LAN are the points database and the operator's Human-Machine Interface (HMI) screens. The points database provides useful information such as description, setpoints, point data type, etc. An attacker planning a surgical strike on the process needs the point information because, at the protocol level, each device is referred to by number only. The HMI is generally the easiest way to understand the process and assign meaning to the points (numbers) because it links to the database points that describe the interaction between the operator and the physical equipment. It also provides a graphical representation of the process and additional information related to HMI navigation points, additional process logic, etc.

After the intruder has discovered enough information regarding the process, the next step is to perform the attack directed at manipulation of the process. In general, the easiest way for an intruder to control the CS is to send commands directly to the front-end equipment which is the part of the CS that converts the CS point data to the various protocols for transmission to and from the field devices and controllers. Most front-end equipment, such as protocol converters or front end processors (FEPs) lack even basic authentication. To control such equipment, in most cases, an attacker need only establish a connection and issue a properly formatted command. The operator's screen could possibly be exported back to the attacker as well, giving him operator level awareness and control of the process.

An attacker could also perform man-in-the-middle attacks on CS protocols which provide the data communications between the CS LAN devices. Once the attacker knows the protocol, he can modify the packets in transit. By inserting packets into the network, he can issue arbitrary commands. By modifying replies, he can give the operator a false picture of the process. Thus, the attacker could both spoof the operator HMI and control the system process.

Security vulnerabilities can allow an attacker to carry out the steps necessary to access the CS LAN, discover the CS and process configurations, and control the process.

ON-SITE VULNERABILITIES AND MITIGATIONS

On-site assessment work provides an opportunity for cyber security professionals to meet with industry personnel to work toward increasing the security posture of a specific CS installation. Security issues are unique to each CS implementation, yet there are commonalities among a number of these installations. These commonalities are discussed in this paper, which outlines some general problems and mitigations that can be applied industry-wide.

Planning efforts need to be implemented for prioritization of the tasks necessary to enhance CS security. Important considerations in this process are cost, probability, and consequence. Decisions concerning methods of mitigating cyber vulnerabilities include balancing the risk of system compromise by an intruder with the risk of potentially degrading system operability. Above all, the CS must be reliable and perform its required mission. Therefore, the suggested approach is to build security into a system before it is put into production or add security into an existing system in small increments. When adding security to a production system, test on a backup system first to allow quick recovery to the previous configuration in the event any security measure affects system operation.

In the discussion that follows, security problems are categorized as relating to policies and procedures, operating system (OS) security, or network level security. The vulnerabilities and suggested mitigations are based on observations made during CS assessments. Each vulnerability section will conclude with an actual CS example.

SECURITY POLICIES AND PROCEDURES

Effective security policies and procedures are the first step to a secure CS network. Many of the same policies used for Information Technology (IT) security for corporate systems can be applied directly to CS networks. The SANS Institute provides free templates for many types of security policies, and can be a valuable resource for CS network administrators in developing their own policies [1]. CS-specific requirements can then be added to it, such as the North American Electric Reliability Council (NERC) cyber security requirements for electric systems [2].

Security policies in facilities where CSs are deployed are often non-existent or poorly enforced. 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. Network and system administrators have the technical knowledge, but also need authorization and support from management to implement the policy.

SECURITY TRAINING

In many cases, the individuals in charge of the CS network do not have adequate security training. This situation is generally due to a lack of funding or appreciation for the importance of this training. Training provides an understanding of the security implications of a network architecture and how to design a more secure network.

Network administrators require a constant retraining program to keep them up to date with the fast-paced changes and advancements of the network security field. This includes the latest network architecture designs, and firewall and Intrusion Detection System (IDS) configurations. New techniques are constantly being developed to attack, and therefore defend, computer networks. It is very important to have comprehensive computer security training, not only for system administrators, but each individual should be trained to protect against problems such as email phishing attacks. If formal training is cost prohibitive, some of this information can be gleaned from books, papers, and web sites on cyber and CS security.

On-site system assessments have identified cases in which CS network administrators are very competent at designing and maintaining reliable networks, but do not understand the security implications of their designs. For example, networks have been seen with redundant firewalls, IDSs, backup system networks, and DMZs for non-CS critical computers. At first glance, these appeared to be secure configurations, but a closer look revealed direct connections from the DMZ to the CS network circumventing the firewall, incomplete firewall rules, and non-validated IDS signatures. A greater understanding of security issues would have brought these problems to light.

An example that demonstrates the need for security training for all users is the classic phishing attack which can draw users normally not susceptible to exploitation into a vulnerable situation. One experience utilized a well crafted email announcing potential layoffs at a major corporation just after a large merger. Many of the recipients took the bait and visited the attacker's malicious web site, and even forwarded it on to others.

PASSWORD POLICY

In many CS operations, most user IDs and passwords are shared among the different operators of the system. This sharing must exist, in many cases, because of the criticality of the system operation. Unacceptable consequences might occur because of a locked user ID or a forgotten password. Typical continual manning of operating consoles provides additional physical security that reduces the need for distinct operator user IDs and passwords. If user-level authentication is therefore not an option, using different user IDs and passwords for the DMZ, as well as different user IDs and passwords for the business LAN, can help increase security. This prevents an attacker from using a user ID and password obtained from the business LAN to gain access to the CS DMZ and/or the CS LAN.

CS and networking equipment should not be left with the default password from the manufacturer. Default passwords can give an attacker easy access to the equipment that controls the process. Unless

required by the CS software, default passwords should always be changed to robust, unpublished passwords. Implement a password policy that enforces strong passwords to greatly impede password cracking and guessing. The SANS Institute's password policy provides guidance on creating, protecting, and changing passwords [3].

Passwords have been found in control rooms on small pieces of paper on the bottom of the keyboard, in a drawer, etc. If a password is too complicated and difficult to remember, or changes too often, users will undermine their security in order to remember them. Complex passwords do protect against some of the advanced password cracking attacks, but they create a physical and social engineering vulnerability that could be exploited by an attacker. Passwords should, therefore, not be auto-generated, but instead created from passphrases or other memorable means.

INCIDENT RESPONSE PROCEDURE

An incident response procedure that instructs employees in the steps to take if a computer on the network has been compromised should be in place. 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) has developed a Computer Security Incident Handling Guide that provides guidance to security personnel in developing an incident response procedure [4].

OS LEVEL SECURITY

PATCHES

OS patches repair vulnerabilities in the OS that could allow an attacker to exploit the computer. The importance to system security of keeping OS patches up-to-date cannot be over emphasized. However, patching CS machines can present unique challenges. Among the factors to consider are system functionality, security benefit, and timeliness.

For security, patches can be downloaded to a trusted server off of the control network, and burned to a CD. The CD can then be used to patch the machines on the CS network. Other methods of patching could include the same process, but instead of loading each computer separately with the patch, the administrator could feed the new patch into a patch management server on a secure DMZ.

Patches must be tested for adverse affects on system functionality. The system vendor should test OS patches for compatibility with their system and supply the testing results to users. These results should be made available as soon as possible after the patch release, to limit the length of time the user's system is vulnerable to the OS exploit. Patches should always be tested on a backup system first, before being implemented on a production system. This testing period should be long enough and include full operational evolutions to make any side effects apparent. There have been cases in which the patch was tested and approved by the vendor, but when it was installed, it rendered the CS inoperable. Therefore, even if the vendor tests patches and updates, they should be tested on the backup and/or test system as well.

CSs have been seen that are still vulnerable to exploits that have had patches available for a long time. For example, a system that hasn't been patched for the RPC DCOM overflow vulnerability could be exploited with an off-the-shelf available tool like Metasploit, compromising the system. This particular exploit would allow complete graphical remote control of the system because the attacker has control of the HMI. This would complete the discovery phase of the attack because the graphical interface usually provides the details needed to understand the process.

APPLICATIONS AND SERVICES

Services or applications running on a system open up different network ports to be able to communicate to the outside world. Each open port provides a possible access path for an attacker that can be used to send exploits and receive data. An attacker can only gain access to, and receive information from, the CS through an open port. The more ports and services that are accessible, the greater the risk of successful exploits due to existing vulnerabilities in the services.

New vulnerabilities are found every day in the applications and services that run on computers. Some of these vulnerabilities are published shortly after their discovery, and some are kept a close secret, allowing a few hackers to exploit computers at will, with no patches available to stop them. Decreasing the number of installed applications and services decreases the likelihood of an attacker finding a vulnerability on the computer. All unneeded applications and services should, therefore, be removed. Also, adequate resources must be allocated to ensure that all services and applications are completely patched and up-to-date using the process described in the preceding patches section. The patching process should be worked closely with vendor support to ensure CS application integrity is maintained. Before stopping any services or programs, the vendor should confirm that the service is not needed for system functionality. This can be tested on a backup or development system first, to isolate the primary system from any potential damage. For example, a standard security measure is to shut off the auxiliary services such as echo, chargen, daytime, discard and finger. However, if the echo port is being used as the system pulse to confirm that the system is up and running, shutting off these services would disable the entire system.

The development system can be isolated for increased security. Development servers with system source code and system information warrant extra protection to ensure the information is not compromised. System code and configuration information can be used in the discovery phase of an attack to discover the brand of the CS and how it was implemented. This may involve moving the

development servers from the CS LAN onto its own protected LAN. If the development system is available on the CS LAN, an attacker may be able to retrieve all of the critical system information and have direct access to application development tools.

Applications that do not require network service should also be evaluated for whether they are required for the system operation. As an example, attackers can use compilers to exploit a system. A simple compiler like QBasic, which was a standard application on the older Windows builds, can be used to reassemble custom root kits on victim computers. An attacker could upload shell code in ASCII notation and use a simple QBasic program to decode the ASCII code into a binary file. This allows attackers to install root kits, gather sensitive data, and clean their tracks before anyone can find them. Applications that are not used should be kept off of the CS network, especially compilers.

PRIVILEGES

A common problem with applications and services is that they are run with system or root level privileges. If this is the case, and an attacker is able to cause the application to crash, the exploit code will run with those same privileges giving him full access to that device. A number of software products run with these super user permissions by default, and yet, will function running as a less privileged user. Permission levels of applications and services should therefore be lowered to that necessary for their required functions.

Another common problem is allowing users to operate a computer system (consoles, servers, etc.) with more permissions than necessary. User accounts used for interactive logon should be carefully evaluated for the lowest set of permissions necessary.

A related issue is file permissions. Share files to only the computers and accounts that require them. Restrict the read and write permissions to these shared files and directories to the minimum required for each user. Give each user and process the minimal privileges necessary for system operation. This is an emerging practice in the IT world that can be brought over to the CS domain, as well.

There have been many cases in which file access permissions needed to be restricted to prevent information gathering. Assessments have discovered systems with all system information and vendor code shared to all computers on the CS LAN. In one example, even though the CS network was segmented, an intruder gaining access to the CS network would have access to all CS-specific information. In another case, a corporation housed the critical CS diagrams and documentation on an anonymous ftp server for easy use by the engineers. This ftp server was viewable from the Internet, making all the data accessible to anyone seeking the information.

NETWORK LEVEL SECURITY

PERIMETER DEFENSES

Any connection into the CS LAN is considered part of the perimeter. Often these perimeters are not well documented and some connections are neglected. All entry points into the CS LAN should be known and strictly managed by a security policy. Most common entry points include:

- Vendor access – Vendor access provides a direct link to the CS LAN and, as such, is a potential way for an attacker to access the system without having to penetrate the firewall. Vendor access is typically through a VPN connection or by dialup modem. Know where the vendor access points are, and have a written, enforced procedure in place to connect/disconnect them. Disconnecting can be done by physically unplugging the connection, or by employing a time-out mechanism.
- Corporate LAN connection – Since the corporate LAN is connected to the Internet, this connection has the most potential for giving an attacker access to the CS LAN. Corporate connections exist to provide replicated CS data to corporate users and to allow remote VPN access for system operators and administrators. Again, know what and where the connections are, and have policies in place to manage these connections. Routinely check for unauthorized connections such as modems and direct cables to give operators, etc., remote access to the CS.
- Communication lines – Know where the communication lines for the CS go. For example, do any communication lines connecting the Remote Terminal Units (RTUs) to the CS pass through the networking equipment on the business network? Keep communication lines for the business and CS networks separate, or bring CS communications through the CS firewall. An attacker can compromise the business PBX and use this system to send data into the CS LAN, potentially taking control of different communication links or systems on the CS LAN.

Vendor connections are one of the biggest breaches in CS network security. Most CS LANs allow unsupervised vendor access directly into the heart of the CS network. Therefore, the vendor VPN should only be accessible when needed, and then immediately unplugged. This has been demonstrated at sites where the vendor VPN has a large tag on an unplugged cable labeled “Vendor VPN.”

NETWORK SEGMENTATION

Network segmentation has traditionally been accomplished by using multiple routers. Recently, firewalls have replaced these routers, providing the capability to add much tighter and more complex rules for communication between the different network segments. Firewalls should be used to create DMZs to protect the CS network. Multiple DMZs should be created for separate functionalities and access privileges, such as peer connections, the data historian, the InterControl Center Communications Protocol (ICCP) server in SCADA systems, the security servers, replicated servers, and development servers. FIG. 1 shows a multiple DMZ architecture. All connections to the CS LAN should be routed through the firewall, with no connections circumventing it. Network administrators need to keep an accurate network diagram of their CS LAN and its connections to other protected subnets, DMZs, the corporate network, and the outside.

In an on-site assessment, while scanning for vulnerabilities on the CS network, the assessment team discovered IP addresses belonging to a data sharing partner. The network at this connected utility was obviously not really isolated or protected from its partners. Because of situations like this, separate

network segments need to be created and secured, isolating data flows. Otherwise, an attacker on a shared link could have access to any other computer on the same LAN.

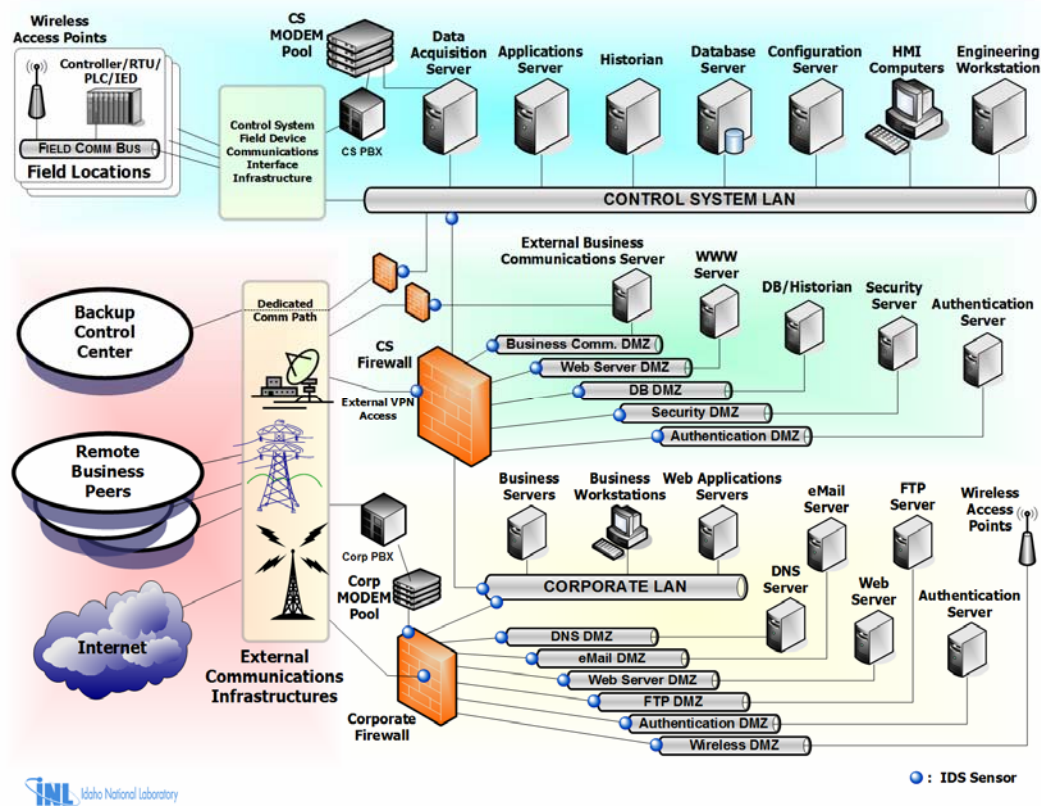


FIG. 1 - RECOMMENDED NETWORK ARCHITECTURE - COMPARTMENTALIZING COMMUNICATION AND DEFENSE-IN-DEPTH

FIREWALL RULE-SET

Well configured firewalls are critical to CS LAN security. Communications should be restricted to only that which is necessary for system functionality. CS 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 the implicit outbound firewall rules. Firewall rules should consider both directions through the firewall. Most administrators effectively block traffic into the CS network, but do not filter traffic out of the network. Outbound rules should be created starting with no exceptions; working toward a rule set that excludes all unnecessary traffic. Once the necessary

outbound traffic has been determined, a safer configuration can then be created that blocks all traffic with exceptions for necessary communication.

An attacker needs to obtain information from and send files and commands to the CS network. To remotely control exploit code running on a CS computer, a return connection must be established from the CS network. Because of the nature of most vulnerabilities, exploit code must be small and contain just enough code to get an attacker onto the computer; there is not enough space to add expensive logic 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.

NETWORK SECURITY DETECTION

The old network security adage is “prevention is ideal but detection is essential [5].” There are several common methods for monitoring a network for unusual or unauthorized activity. This includes host and network based IDSs, intrusion prevention systems (IPSs), Address Resolution Protocol (ARP) monitoring (for man-in-the-middle attacks), anti-virus software, and even simple honey pots i.e. canary computers that should never be used and generate alerts when scanned.

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 utilized. Implementing a variety of these tools will help create a defense-in-depth architecture that will be more effective in identifying attacker activities

One of the common problems observed in industry is that network monitoring tools 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.

During an IDS validation experience, common exploits were sent from the DMZ to different computers with harmless payloads. Two different IDSs were running on the CS network and the CS DMZ. Neither IDS alerted from the several different exploits launched, even though it was confirmed that the IDSs did have rules for the exploits. The IDSs did catch altered TCP options in the exploits, proving that the traffic was traveling past them. This is an example of why IDSs need to be tested to ensure they are working correctly.

PLAIN TEXT PROTOCOLS

One of the greatest security issues the assessment teams have identified is the widespread use of unencrypted plain-text network communications protocols. Many applications and services use protocols that include human-readable characters and strings. Network “sniffing” tools, many of which are freely downloadable, can be used to view this type of network traffic. As a result, the content of the CS communication packets can be intercepted, read, and manipulated. This includes usernames, passwords, and CS commands. Examples of these applications and services are proprietary CS protocols and remote access services, such as telnet, ftp, rsh, rexec, and rlogin, which don’t even encrypt the password or obfuscate it with a one-way hash function.

If an attacker is able to capture a username and password, he is able to legitimately log onto the system with that user's privileges. In addition, in order to strategically attack a CS, the attacker must perform discovery of the particular CS environment. This is done by monitoring the CS communication traffic to see which computers are performing specific functions and the protocols and commands used. These issues illustrate the security issue of data confidentiality. For this reason, plain-text protocols should be eliminated where possible; at a minimum, plain-text remote access services should be replaced with encrypted services such as secure shell (ssh). Encrypting other communications, such as proprietary CS protocols, is a complex task and should be carefully addressed by the system vendor. Another issue to consider prior to encrypting everything is that it prevents the ability to implement network monitoring tools on communication channels.

Encryption is used for confidentiality, but does not provide assurance of data integrity. An alternative to CS protocol encryption is to add authentication methods, such as passwords with each command, or update and data integrity checks, such as checksums, to ensure that the commands and updates have not been altered in transit. This will help protect against spoofing attacks, in which false information is sent to the operator's console in order to give them an altered view from reality. Authentication also protects against unauthorized commands being sent to the CS process devices.

The need for encryption was illustrated at one facility where a user ID and password for the system were captured as they came across a connection in plain text. While monitoring traffic from the communication equipment that controlled the communication links to all of the remote substations at a utility, the assessment team discovered a connection going to the communication server. The team performed a simple man-in-the-middle attack, which reset the TCP link, forcing the user to re-login. The team was then able to see the user ID and password in clear text. This enabled control of the communication line and led to direct access to the equipment in the field, bypassing the controls on the CS network. This could have been prevented by an encrypted authentication method.

SUMMARY

Cyber vulnerabilities in CSs can allow attackers to gain access to and exploit the systems that control our nation's critical infrastructures. Steps can be implemented to greatly enhance the security of these critical infrastructure CSs and reduce the opportunities for attack.

Policies and procedures should be developed that outline personnel security requirements and provide guidance on passwords and incident response. Valuable free resources are available from the SANS Institute [1]. Network administrators need to understand security and be kept up-to-date with on-going security training. If budgets cannot afford to support formal training, books and web research can help. Free advice is available for setting password policies [3], incident response procedures [4], etc.

OSs and application patches should be applied as they are made available, always testing for negative impacts on system functionality first. All unnecessary applications and services should be removed. The principle of "least privilege" should be applied in granting system access permissions to users and

applications, and in allowing access to files. Plain text protocols should be eliminated and data validation added, where appropriate.

Network segmentation can greatly increase security. All connections into the CS LAN must be known, managed and monitored. Firewalls can be used to segment the network into security zones and add tight, complex rules to allow only necessary communication between network segments. A variety of security monitoring tools can then be used to validate the effectiveness of security measures.

The suggested approach to increase security is to either build it in to a new installation or add improvements to an existing system in small increments. Security implementation plans should be worked in close coordination with the CS vendor to ensure the system maintains full operating functionality. Procedures must be in place to quickly reverse the system configuration in the event that any security measure conflicts with system operation. Always weigh the risks and add the appropriate amount of security measures for the specific situation.

REFERENCES

1. <http://www.sans.org/resources/policies/>
2. <http://www.nerc.com/~filez/standards/Cyber-Security-Permanent.html>
3. "Password Policy," http://www.sans.org/resources/policies/Password_Policy.pdf .
4. Grance, Tim, Kent, Karen, Kim, Brian, "Computer Security Incident Handling Guide," NIST Special Publication 800-61, National Institute of Standards and Technology: Gaithersburg, MD, January, 2004.
5. Todd Thompson, "Where Did All My Bandwidth Go?" June 18, 2001, SANS Institute Reading Room. http://www.giac.org/practical/gsec/Todd_Thompson_GSEC.pdf (p.1).

ACKNOWLEDGEMENTS

This work was supported by the U.S. Department of Energy under DOE Idaho Operations Office Contract No. DE-AC07-05ID14517. Many of the concepts in this paper are based on lessons learned working with Jason Larsen, head cyber security researcher at the INL.