NISCC
Good Practice Guide
on
Firewall Deployment
for SCADA and Process Control Networks

Prepared for:
National Infrastructure Security Co-ordination Centre (NISCC)
By the:
British Columbia Institute of Technology (BCIT)

Revision Number: 1.4
Document Date: 15th February 2005

Internet Engineering Lab (IEL)
Group for Advanced Information Technology (GAIT)

Revision History

Revision	Date	Author(s)	Description
0.1	May 9 2004	Eric Byres, BCIT Internet Engineering Lab Ken Savage, BCIT Internet Engineering Lab	Draft
1.0	May 22, 2004	Eric Byres, BCIT Internet Engineering Lab John Karsch, BCIT Internet Engineering Lab Joel Carter, BCIT Internet Engineering Lab	Preliminary Release
1.1	July 8, 2004	Eric Byres, BCIT Internet Engineering Lab John Karsch, BCIT Internet Engineering Lab Joel Carter, BCIT Internet Engineering Lab	Grammatical Changes
1.2	September 27, 2004	Eric Byres, BCIT Internet Engineering Lab John Karsch, BCIT Internet Engineering Lab Joel Carter, BCIT Internet Engineering Lab	Grammatical Changes
1.3	January 15, 2005	Eric Byres, BCIT Internet Engineering Lab John Karsch, BCIT Internet Engineering Lab	Added Scoring System and 4.2 Architecture
1.4	February 15, 2005	Eric Byres, BCIT Internet Engineering Lab John Karsch, BCIT Internet Engineering Lab Joel Carter, BCIT Internet Engineering Lab	Public Release

Acknowledgements

The Group for Advanced Information Technology (GAIT) at the British Columbia Institute of Technology (BCIT) would like to thank all the vendors and end users that generously supported our efforts through numerous interviews and by providing us with documents that could only be described as extremely sensitive. Unfortunately we can not name you for obvious security reasons, but we appreciate your time, trust and encouragement.

Four people stood out in their contributions and advice for this document that we can acknowledge. These people are Darrin Miller of Cisco System's Critical Infrastructure Assurance Group (CIAG), Andy Cobbett and Ian Henderson of BP International and Justin Lowe of PA Consulting. Thank you for all your help.

Finally we would like to thank Karl Williams of the National Infrastructure Security Coordination Centre (NISCC) for his vision and support. Without him, this project never would have been possible.

NISCC
National Infrastructure Security Coordination Centre

PO Box 832
London SW1P 1BG

Tel:0870 487 0748

This page is deliberately left blank.

Table of Contents

EXECUTIVE SUMMARY **1**

1 INTRODUCTION **3**

2 WHAT IS A FIREWALL? **5**

2.1 TYPES OF FIREWALLS 5
2.2 CLASSES OF FIREWALLS 6
2.2.1 Packet Filter Firewalls *6*
2.2.2 Stateful Firewalls *7*
2.2.3 Application Proxy Firewalls *7*
2.2.4 Deep Packet Inspection Firewalls *8*
2.3 OTHER FIREWALL SERVICES 8

3 OVERALL SECURITY GOALS OF PCN/SCADA FIREWALLS **9**

4 COMMON SCADA/PCN SEGREGATION ARCHITECTURES **11**

4.1 DUAL-HOMED COMPUTERS 11
4.2 DUAL-HOMED SERVER WITH PERSONAL FIREWALL SOFTWARE 12
4.3 PACKET FILTERING ROUTER/LAYER-3 SWITCH BETWEEN PCN AND EN 13
4.4 TWO-PORT FIREWALL BETWEEN PCN AND EN 13
4.5 ROUTER/FIREWALL COMBINATION BETWEEN PCN AND EN 15
4.6 FIREWALL WITH DEMILITARIZED ZONES BETWEEN PCN AND EN 15
4.7 PAIRED FIREWALLS BETWEEN PCN AND EN 16
4.8 FIREWALL AND VLAN-BASED PROCESS NETWORK COMBINATIONS 17
4.9 SUMMARY OF FIREWALL ARCHITECTURES 18

5 FIREWALL IMPLEMENTATION AND CONFIGURATION **21**

5.1 GENERAL FIREWALL POLICIES 21
5.2 RULES FOR SPECIFIC SERVICES 23
5.2.1 Domain Name Service (DNS) *24*
5.2.2 Hyper Text Transfer Protocol (HTTP) *24*
5.2.3 File Transfer Protocol (FTP) and Trivial File Transfer Protocol (TFTP) *24*
5.2.4 Telnet *24*
5.2.5 Simple Mail Transfer Protocol (SMTP) *25*
5.2.6 Simple Network Management Protocol (SNMP) *25*
5.2.7 Distributed Component Object Model (DCOM) *25*
5.2.8 SCADA and Industrial Protocols *25*
5.3 NETWORK ADDRESS TRANSLATION (NAT) 25
5.4 SPECIFIC PCN FIREWALLS ISSUES 27
5.4.1 Data Historians *27*
5.4.2 Remote Support Access *27*
5.4.3 Multicast Traffic *28*

6 MANAGEMENT OF PCN/SCADA FIREWALLS **29**

7 SPECIAL OR FUTURE TECHNOLOGIES **31**

7.1 SCADA PROTOCOL AWARE FIREWALLS 31
7.2 DISTRIBUTED MICRO-FIREWALLS 31
7.3 QUALITY OF SERVICE (QOS) 31
7.4 ONE WAY COMMUNICATION PATHS 32

ACRONYMS **33**

REFERENCES **35**

Figures

FIGURE 1: A SIMPLIFIED EXAMPLE OF AN INTERNET-FACING FIREWALL PROTECTING DEVICES ON A NETWORK 5
FIGURE 2: NETWORK SEPARATION USING DUAL-HOMED COMPUTERS.....11
FIGURE 3: NETWORK SEPARATION USING DUAL-HOMED SERVER WITH PERSONAL FIREWALL INSTALLED 12
FIGURE 4: NETWORK SEPARATION ROUTERS OR LAYER-3 SWITCHES WITH ACL FILTERS 13
FIGURE 5: NETWORK SEPARATION WITH SINGLE FIREWALL 14
FIGURE 6: NETWORK SEPARATION WITH FIREWALL/ROUTER COMBINATION (INTERNET EXAMPLE).....15
FIGURE 7: FIREWALL WITH DEMILITARIZED ZONE FOR SHARED ENTERPRISE/PCN DEVICES 16
FIGURE 8: PAIRED FIREWALLS WITH DEMILITARIZED ZONE FOR SHARED ENTERPRISE/PCN ASSETS.....17
FIGURE 9: FIREWALL WITH DEMILITARIZED ZONE AND SCADA/PCN VLANS.....18
FIGURE 10: COMPARISON CHART FOR PCN/SCADA SEGREGATION ARCHITECTURES 19

Executive Summary

In recent years, Supervisory Controls and Data Acquisition (SCADA), process control and industrial manufacturing systems have increasingly relied on commercial information technologies such as Ethernet®, TCP/IP and Windows® for both critical and non-critical communications. While beneficial in other areas, use of these common protocols and operating systems has resulted in significantly less isolation from the outside world for vital SCADA and Process Control Networks (PCNs). These systems are now under risk of attack from a variety of threats, ranging from teenage script-kiddies to skilled and determined cyber-criminals.

Unfortunately, there are few proven methods available to the asset owner or engineer to protect these vital systems. One commonly suggested security solution is to isolate the SCADA and PCN systems from the Internet and corporate enterprise network (EN) through the use of firewalls, but there is little information available on exactly how these firewalls should be deployed in terms of architectures, configuration and management. Firewalls can be complex devices to design and deploy correctly, so guidance on how best to deploy them in the industrial setting would be very useful.

To address this need, the UK's National Infrastructure Security Coordination Centre (NISCC) commissioned the Group for Advanced Information Technology (GAIT) at the British Columbia Institute of Technology (BCIT) to investigate and compile the current practices in SCADA/PCN firewall deployment. The intent was to examine the "state of the art" in firewall architectures, deployment and management used to protect industrial control environments.

In March 2004, the research team sent out requests for information regarding the use of firewalls in industrial settings to approximately 60 organizations and industry news lists. A total of 10 vendors, including firewall manufacturers, IT security firms and control systems manufacturers, responded in some form. Approximately 15 industrial users from the petroleum, chemical, food, and electrical sectors also responded. The vendor and end-user organizations were a mix of North American and European-based firms. This information provided was in the form of personal interviews, white papers, policy manuals, network audit reports and security product literature. In addition, draft documents from standards organizations involved in industrial control security were obtained. These included documents from the American Petroleum Institute (API), the Industrial Automation Open Networking Association (IAONA), the International Electrotechnical Commission (IEC), the Institute of Electrical and Electronics Engineers (IEEE) and the Instrumentation, Systems and Automation Society (ISA).

All collected information was summarized by the research team in terms of firewall architecture, design, deployment and management to determine current security practises. These practices were then analysed and scored for their likely effectiveness in the industrial control environment. The results of this analysis indicate that there are a significant number of different solutions used by the industry and the security effectiveness of these can vary widely. The commonly dual-homed workstations, bridges and routers are unlikely to provide suitable isolation, while two zone architectures are marginally secure, but should be only be deployed with extreme care. In general, architectures that allow the establishment of a Demilitarised Zones (DMZ) between the enterprise network and SCADA/PCN network will provide the most effective security solution.

This page is deliberately left blank.

1 Introduction

In recent years, Supervisory Controls and Data Acquisition (SCADA), process control and industrial manufacturing systems have increasingly relied on commercial information technologies such as Ethernet, TCP/IP and Windows for both critical and non-critical communications. The use of these common protocols and operating systems has made the interfacing of industrial control equipment much easier, but there is now significantly less isolation from the outside world. Network security problems from the enterprise network (EN) and the world at large can be passed onto the SCADA and process control network (PCN), putting industrial production and human safety at risk.

At the same time, the core of our information infrastructure – the Internet – has come under increasing attack from a wide variety of sources, ranging from teenagers on a cyber joyride to professional hackers. People in almost every country on the globe can access a network that, in turn, is ultimately connected to networks that run critical functions throughout the world. Cyber attacks on information networks occur regularly and can have serious consequences if they impact connected PCN or SCADA systems. For example, in 2000 a hacker attacked a sewage plant in Queensland, Australia, causing millions of litres of sewage to pollute a river and park in that region[1]. Three years later, the Slammer Worm caused documented impacts on at least two power generation/distribution systems[2], a safety monitoring system in a nuclear reactor[3] and emergency services phone system. Future attacks could have serious consequences such as extended disruption of critical services and even loss of life.

While the importance of protecting these critical systems is widely acknowledged, there are few proven protection methods available to the asset owner or engineer. Certainly there are general information security practices and standards that can offer significant guidance towards securing the PCN, but there are also important differences that need to be considered. For example, the goals of information technology (IT) departments can be fundamentally different from those of a process control department. The IT world typically sees performance and data integrity as paramount, while the industrial world sees human and plant safety as its primary responsibility. Other documented distinctions include differences in reliability requirements, event impacts, performance expectations, operating systems, communications protocols and system architectures[4] [5]. This can mean significant differences in acceptable security practice.

One commonly suggested security solution is to isolate the SCADA and PCN system from the corporate and Internet systems through the use of firewalls. Unfortunately, while firewalls are widely used in the traditional IT sector, their effectiveness in SCADA/PCN environments is still in question. IT firewalls are generally unaware of SCADA/PCN protocols, may introduce unacceptable latency into time-critical systems and face operational constraints not typical in the IT world. To make matters worst, there is little information available on exactly how these firewalls should be deployed in terms of architectures, configuration and management. Typical guidance for the PCN or SCADA engineer rarely extends beyond:

> *"Never mix your office LAN with your industrial-control LAN. They should be separated by a firewall, or at minimum, a bridge or router."*[6]

This type of advice misses the fact that firewalls can be complex devices that need careful design, configuration and management if they are to be effective. For example, how should firewall rule sets be defined between the PCN and the corporate network? Should they be strictly based on address filtering or should protocol filtering be used as well? And if protocol filtering is used, which protocols should be blocked and which should be allowed? Certainly each SCADA or PCN system has unique characteristics that rule out a single definitive answer to these questions, but some guidance on how best to deploy firewalls in the industrial setting would be very useful.

To address this need, the UK's National Infrastructure Security Co-ordination Centre (NISCC) commissioned the Group for Advanced Information Technology (GAIT) at BCIT to investigate and compile the current practices in SCADA/PCN firewall deployment. The intent was to examine the "state of the art" in firewall architectures, deployment and management used to protect industrial control environments. The goals of the project were to:

1. Collect available SCADA/PCN firewall deployment recommendations from equipment vendors, users and industry standards bodies.

2. Summarise the collected firewall deployment recommendations and best practices.

3. Analyse and rate the compiled recommendations and best practices in terms of security effectiveness and possible weaknesses.

In March 2004, the research team sent out requests for information regarding the use of firewalls in industrial settings to approximately 60 organizations and industry news lists. A total of 10 vendors, including firewall manufacturers, IT security firms and control systems manufacturers, responded in some form. Approximately 15 industrial users from the petroleum, chemical, food, and electrical sectors also responded. The vendor and end-user organizations were a mix of North American and European-based firms. The information supplied was in the form of personal interviews, white papers, policy manuals, network audit reports and security product literature. In addition, documents from five standards organizations involved in industrial control security (API, IAONA, IEC, IEEE and ISA) were obtained.

This information was first summarized by the research team in terms of firewall architecture, deployment and management to determine current security practises. These practices were then analysed for their likely effectiveness in the industrial control environment. The results of this summary and analysis are contained in this report.

A follow-on project will test a number of the best practices discussed in this report in a lab setting. This will include creating a reference firewall architecture and rule set for SCADA/PCN security and then testing it, along with a number of commercial firewalls, in terms of security correctness and performance. Combined, these two studies hope to provide clear guidance to the SCADA/PCN engineer and asset owner on the design and implementation of firewalls in critical industrial environments.

2 What is a Firewall?

A firewall is a mechanism used to control and monitor traffic to and from a network for the purpose of protecting devices on the network. It compares the traffic passing through it to a predefined security criteria or policy, discarding messages that do not meet the policy's requirements. In effect, it is a filter blocking unwanted network traffic and placing limitations on the amount and type of communication that occurs between a protected network and other networks (such as the Internet, or another portion of a site's network). Figure 1 shows a simple firewall protecting a personal computer (PC) and programmable logic controller (PLC) from unwanted traffic from the Internet, while allowing requests coming into the corporate web server.

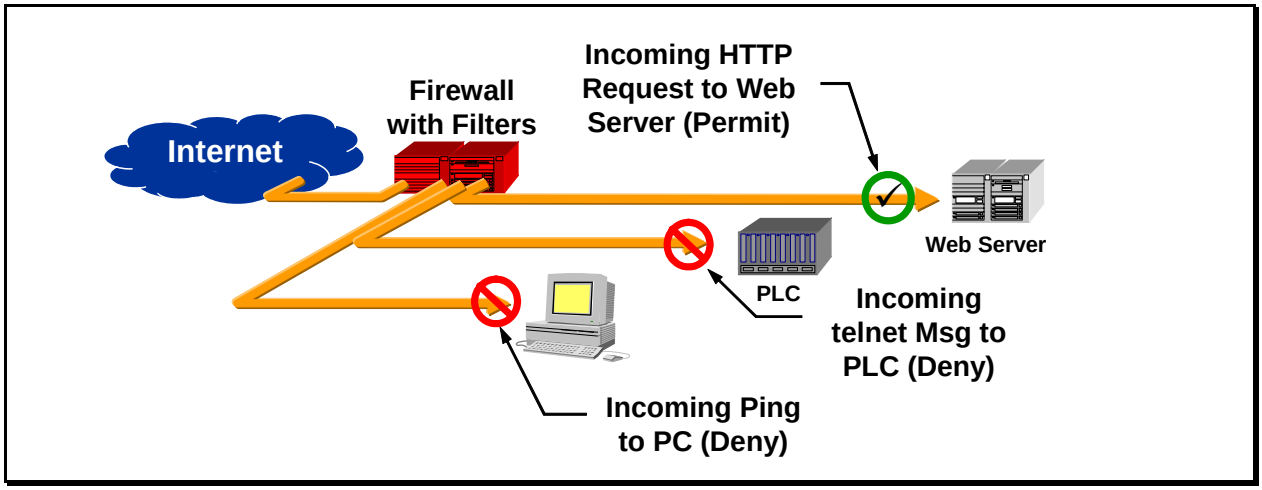

Figure 1: A Simplified Example of an Internet-Facing Firewall Protecting Devices on a Network

2.1 Types of Firewalls

A firewall can come in many different designs and configurations. It can be a separate hardware device physically connected to a network (such as the Cisco PIX® or the Symantec Security Gateway® firewalls), a hardware/software unit with OS-based firewall capabilities (such as 'iptables' running on a Linux® server), or even a completely host-based software solution installed directly on the workstation to be protected (such as Norton Personal Firewall® or Sygate Personal Firewall®).

Separate hardware or hardware/software units are often referred to as network firewalls and typically provide the most secure solution for the separation of the PCN from corporate network. They are dedicated-function units that can be hardened to resist all but the most ingenious assaults. In addition, network firewalls generally offer the best management options since they are typically designed to allow remote management.

Host-based firewalls must generally accept compromises since the primary function of the host device is not security, but workstation or server tasks such as database access or web services. As well, host-based firewall solutions are currently only available for Windows or Unix-based platforms and can do little to regulate traffic destined for embedded control devices, such as

PLCs, on the network. Host-based firewalls may have a place on the PCN/SCADA network, such as in the architecture described in section 4.2, but generally they are beyond the scope of this study. Thus with a few exceptions, this paper assumes that the PCN/SCADA firewall will be a dedicated hardware or hardware/software solution that, through a series of rules, allows or denies network traffic to pass on to the control network.

2.2 Classes of Firewalls

Network traffic is sent in discrete groups of bits, called a packet. Each packet typically contains a number of separate pieces of information, including (but not limited to) items such as the:

- Sender's identity (Source Address);
- Recipient's identity (Destination Address);
- Service to which the packet pertains (Port Number);
- Network operation and status flags;
- Actual payload of data to be delivered to the service.

A firewall, upon receiving a packet, analyses these characteristics, and determines what action to take with the packet. It may choose to drop the packet, allow it through immediately, buffer it momentarily to bandwidth limit a class of service or forward it to a different recipient than was initially intended - whatever the network security policy deems is an appropriate action to take. These decisions are based on a series of rules commonly referred to as Access Control Lists (ACLs). Different classes of firewalls exist, each with increasingly sophisticated analysis and action capabilities.

2.2.1 Packet Filter Firewalls

The simplest class of firewall is known as a packet filter firewall. It has a series of static rules and uses them to take action upon received packets on an individual basis. The following sample rules can easily be handled by a packet filter firewall:

- Accept Domain Name Service (DNS) response packets on User Datagram Protocol (UDP) port 53;
- Block any traffic to or from Internet Protocol (IP) address 24.116.25.21;
- Prevent web surfing by blocking outgoing packets from accessing Transmission Control Protocol (TCP) ports 80, 443, 3128, 8000 and 8080, unless they are directed to the corporate intranet web server's IP address;
- Drop source routed packets;
- Accept TCP port 23 (telnet) traffic to a specific Distributed Control System’s (DCS) IP address, from a specified range of IP addresses for engineering consoles.

Unfortunately, a packet filter firewall lacks the ability to understand the relationships between a series of packets. For example, the broad rule "Accept DNS response packets on UDP port 53" contains a serious flaw. What if no DNS query was ever issued, but a spoofed DNS "response" packet came in instead? This simplistic firewall would accept the packet and deliver it to the

"requesting" host, possibly causing it to become confused. A moderately intelligent hacker could use these firewall weaknesses to compromise internal systems.

The advantages of packet filtering firewalls include low cost and low impact on network performance, often because only the IP addresses and port numbers in the packet are examined. This method is also sometimes called Static Filtering[7]. It is often directly deployed in layer-3 switches or routers, rather than dedicated “firewalls”.

2.2.2 Stateful Firewalls

A more sophisticated firewall, known as a stateful firewall, intelligently tracks the interrelationships between packets allowed to flow through it. By keeping a history of accepted packets and the state of current connections, it can accept only “anticipated” traffic. Stateful firewall rule sets can be made conditional due to the firewall's intelligence. For example:

- Accept DNS response packets on UDP port 53 only if they match an outgoing DNS query with the same DNS request identification.
- Allow File Transfer Protocol (FTP) data channel traffic through only after successful control channel negotiation.
- Block all incoming web traffic originating from TCP port 80 unless it is specifically in response to a previous outgoing Hyper-Text Transfer Protocol (HTTP) request.

These types of firewalls offer a high level of security, good performance, and transparency to end users, but are more expensive. Due to their complex nature, they can be less secure than simpler types of firewalls if not administered by competent personnel. This method is also sometimes called Dynamic Packet Filtering[8].

2.2.3 Application Proxy Firewalls

Application Proxy firewalls open packets at the application layer, process them based on specific application rules, then reassemble them and forward them to the desired target device. Typically they are designed to concentrate on a variety of application protocols (such as Telnet, FTP, HTTP, etc.) through a single machine, but then forward the traffic to individual host computers for each service. Instead of connecting directly to an external server, the client connects to the proxy firewall which in turn initiates a connection to the requested external server. Depending on the type of proxy firewall used, it is possible to configure internal clients to perform this redirection automatically, without the user’s knowledge. Others might require that the user connect directly to the proxy server and then initiate the connection through a specified format.[9]

Proxy firewalls offer some significant security features for controlling protocols that the firewall can recognize. For example, it is possible to apply access control lists against the application protocols, requiring users or systems to provide additional levels of authentication before access is granted. In addition, rules sets can be created which understand specific protocols and can be configured to block only subsections of the protocol. For example, a firewall rule for HTTP could be created to block all inbound HTTP traffic containing scripts. By contrast, a filtering router could either block all HTTP traffic or none, but not a subset.

These types of firewalls can offer a high level of security, but can also have a significant impact on network performance. In addition, most application proxy firewalls products only support common Internet protocols such as HTTP, FTP or SMTP (Simple Mail Transfer Protocol). As a result, messages containing industrial application layer protocols such as Common Industrial Protocol® (CIP) or MODBUS/TCP® will require the firewall to process the traffic in stateful or packet-filter mode.[10]

In the past few years, most of firewalls on the market have utilized a combination of Stateful and Application-Proxy techniques and are often referred to as hybrid firewalls.

2.2.4 Deep Packet Inspection Firewalls

The firewall market is currently moving to a fourth technique that is referred to as "deep packet inspection" (DPI) or "application firewalling". This typically offers filtering deeper into the application layer than the traditional application proxy and yet does not perform a full proxy on the TCP connection. For example, DPI firewalls will be able to inspect Simple Object Access Protocol (SOAP) objects in Extended Markup Language (XML) on web connections and enforce policy on what objects are allowed/disallowed into the network. This market is still in development.

2.3 Other Firewall Services

In addition to the core service of traffic filtering, most modern firewalls offer other network-based security services. These may include:

1. Acting as an intrusion detection system (IDS) by either logging packets that are denied access, recognizing network packets specifically designed to cause problems, or reporting unusual traffic patterns.

2. Deploying “frontline” anti-virus software on a firewall. If the characteristics of infected traffic can be formulated, such traffic can be blocked before entering the network.

3. Authentication services requiring users wishing to connect to devices on the other side of the firewall to authenticate to the firewall using either passwords or strong two-factor authentication methods such as public key encryption.

4. Virtual Private Network (VPN) gateway services where an encrypted tunnel is set up between firewall and remote host device.

5. Network Address Translation (NAT) where a set of IP addresses used on one side of the firewall are mapped to a different set on the other side.

Additional services offered will depend on the particular make and model of the firewall purchased. Obviously these additional functions can mean additional costs, increased complexity and reduced performance. However by utilizing them, one can often significantly improve the overall security of the PCN/SCADA network. An analysis of these types of additional security services is beyond the scope of this document, but the reader is advised to consider them when making any design or purchasing decisions.

3 Overall Security Goals of PCN/SCADA Firewalls

Ideally, a process control or SCADA network would be a closed system, accessible only by trusted internal components such as the Human Machine Interface (HMI) stations and data historians. Realistically, the need for external access from both corporate users and selected 3rd parties exists. For example, production and maintenance management information needs to be relayed to computers and users outside of the plant floor for management purposes, while vendors may need to access controllers for support purposes. Implicitly this means that some network paths exist from the outside, untrusted world to internal control components.

The goal of the firewall, simply stated, is to minimize the risk of unauthorized access (or network traffic) to internal components on the PCN or SCADA systems. Such a risk minimization strategy will typically include the following general objectives:

1. **No direct connections from the Internet to the PCN/SCADA network and vice-versa**. The reasons for this are fairly obvious and include:

   (a) Unsolicited inbound traffic could congest or corrupt the control network, preventing critical messages from reaching the control devices.

   (b) Invalid control messages or denial of service (DoS) attacks could be directed to control devices causing production upsets.

   (c) Outbound Internet traffic (FTP, HTTP, SMTP, etc.), even if driven from the PCN, could congest critical control traffic. Furthermore, embedded objects (attachments, Java® applets, ActiveX® components) received from the Internet could compromise control workstations, allowing the control network to fall victim to remote attacks.

   (d) Proprietary corporate production data could be intercepted.

2. **Restricted access from the enterprise network to the control network.** The enterprise network should not be able to directly query or control a PLC or other control device at the plant floor/SCADA level. Both physically and logically, the enterprise and process control networks should be isolated.

3. **Unrestricted (but only authorized) access from the enterprise network to shared PCN/enterprise servers.** Typically shared servers like a data historian or maintenance database has query access to the control network and remits relevant information back to the querying business services.

4. **Secure methods for authorized remote support of control systems.** Many companies need to allow third party access to PLC/DCS/SCADA devices by the system vendors for support purposes. Others allow for control network access to appropriate plant staff from a remote location for emergency maintenance. Both need to be explicitly addressed in the firewall design.

5. **Secure connectivity for wireless devices (if used).** Without the proper security in place, a wireless device on a plant floor is equivalent to an open connection with the outside world. As with an open Internet connection, any invalid or disruptive communication messages could cause serious operational failures.

6. **Well-defined rules outlining the type of traffic permitted on a network**. These are enforced via mechanisms such as access control lists (ACLs) and virtual local area networks (VLANs). If the traffic is not expected or wanted on the PCN, it should not be allowed to pass.

7. **Monitoring of traffic attempting to enter and on the PCN**. It is important to be aware of the nature and patterns of the communications traffic attempting to enter, or on, the PCN or SCADA network as this will often be the only indication of either an attack on the network or a failure in the firewall strategy. Typically this is achieved through an Intrusion Detection System (IDS) and can be either embedded in the firewall or running on an external device[i].

8. **Secure connectivity for management of the firewall.** Traffic for the monitoring and management of the firewall should be over a secure communication system and from a highly restricted set of management devices.

[i] It is important to note that there are other monitoring solutions that can be used to baseline PCN traffic in addition to, or instead of, the firewall. For example, netflow functions in routers and switches can collect traffic statistics and can be integrated into independent anomaly detection IDS systems.

4 Common SCADA/PCN Segregation Architectures

A survey of documentation from security/firewall/control system vendors, standards bodies and end-users resulted in eight overall architectures used to separate the PCN/SCADA network from the enterprise network. These range from hosts with dual network interface cards to multi-tiered combinations using firewalls, switches and routers. Each is described below, analysed in terms of advantages and disadvantages and then rated on a scale from one (poorest) to five (best) for the following criteria:

- **Security** – the likely effectiveness of the architecture to prevent possible attacks;
- **Manageability –**ability of the architecture to be easily and effectively managed (both locally or remotely);
- **Scalability –** ability of the architecture to be effectively deployed in both large and small systems or in large numbers.

4.1 Dual-Homed Computers

One commonly proposed security solution is the installation of dual network interface cards (NICs) in workstation or control devices requiring information access to both the enterprise and process control networks. This technique is often referred to as dual-homing and a typical schematic of this design is shown in Figure 2 below.

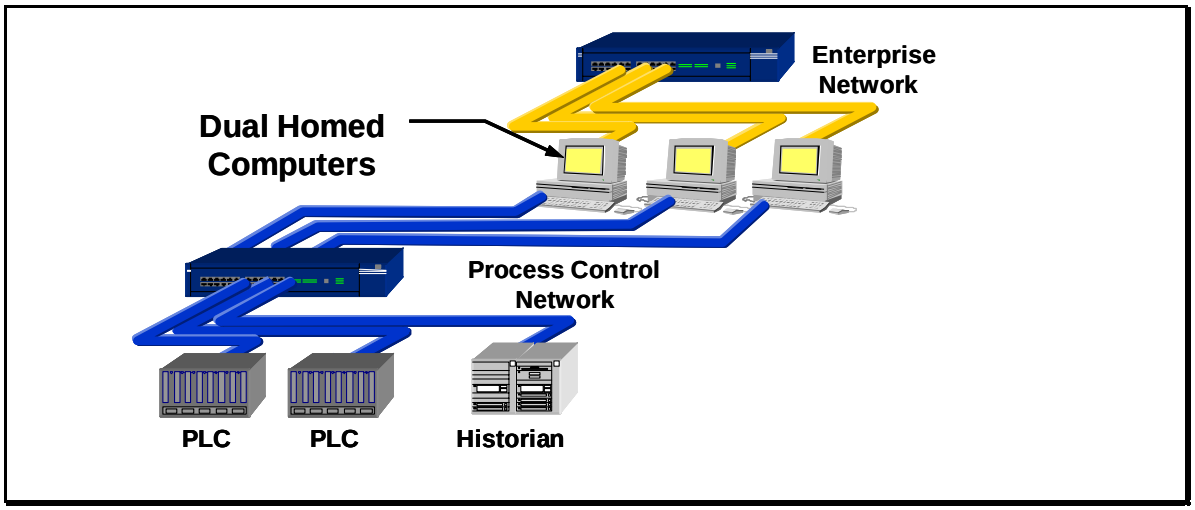

Figure 2: Network Separation Using Dual-Homed Computers

As noted in the appendix B of API Standard 1164:

> *"Dual homing a computer is certainly a convenient way of allowing the computer to connect to two different networks. However, in situations where dual homing is used to communicate between two networks that are being segregated for security reasons it makes the computer a significant security risk."*[11]

In an ideally configured system, this method allows for some minimal network separation. However, in most devices it is trivial to adjust network configuration parameters so that the

device automatically forwards packets arriving on one network over to the other network, thereby circumventing any premise of network isolation. The fact is that dual-homed servers are widely seen as easy targets by the hacking community. For example, the following discussion can be found on the web by searching for the phrase "Dual-Homed Computers":

> *How to get around the firewall? The juiciest targets are dual-homed machines -- that is, boxes with two NICs connected both to the DMZ and the internal net. In theory there should be none; in practice, users (well, power users and developers) frequently do this so they can get their work done more quickly.*[12]

In addition, this architecture can potentially violate the security objective of "No Direct Internet Connections from the Process Control Network" since the dual-NIC device is a part of the PCN, may be attached to a network with Internet access and can run typically insecure software such as web browsers. Furthermore, if the dual-homed computer is compromised it will likely compromise both networks. A number of process control incidents involving the Slammer worm in January of 2003 were a result of this type of architecture. Finally, while remote management software does exist for personal computers, the effort involved in securely and effectively managing more than a small number of dual-homed machines can be significant.

In summary, while dual-homed computers are useful for redundancy on the same network or on two networks within the same PCN system, their use as an enterprise/PCN segregation measure is less than ideal. **[Security = 1 Manageability = 2 Scalability = 1]**

4.2 Dual-Homed Server with Personal Firewall Software

One variation on the previous architecture is the installation of host-based personal firewall software on a single dual-homed machine (typically a data historian server), as shown in Figure 3 below. The idea is the only traffic between the PCN and enterprise network is the shared history data and this terminates on the server anyway. Thus if you use the personal firewall to allow only data requests from the server to the business users and then protect the server with the inexpensive personal firewall, you will have a very inexpensive security mechanism.

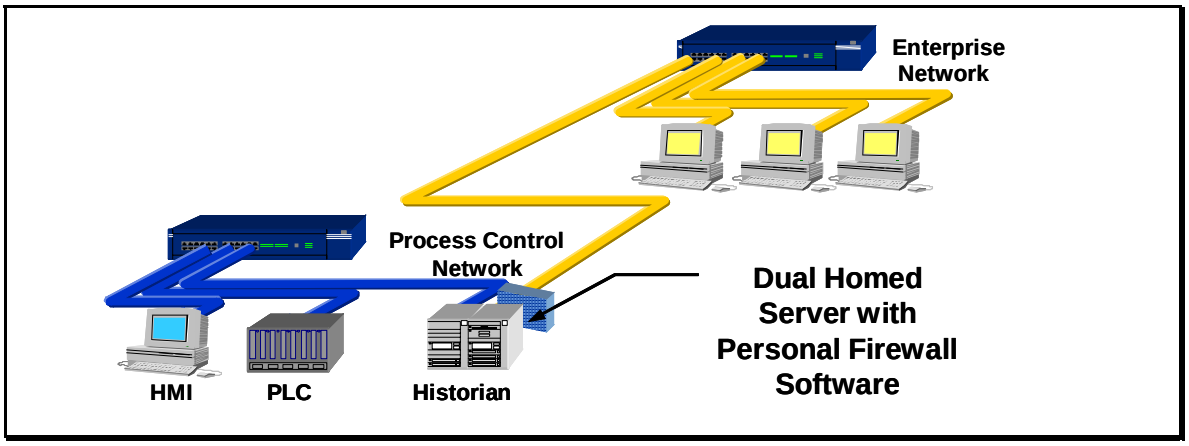

Figure 3: Network Separation Using Dual-Homed Server with Personal Firewall Installed

The first issue with this solution is that it will only provide a mechanism to allow the sharing of server data. If there is any other traffic that needs to traverse the PCN to EN boundary (such as remote maintenance access to a controller) then this architecture will either completely block that traffic or leave the PCN poorly secured. There is little opportunity for granularity in this design.

The second problem comes about if there is more than one server involved. Remote management of most personal firewall packages is fairly limited and as a result, maintaining consistent rule sets and management services across multiple servers becomes very time consuming. Finally, personal firewalls can not compete with proper network firewalls in terms of system hardening, stateful inspection, throughput or additional features like IDS. Thus, the result of this architecture is a very inexpensive, but very limited segregation solution.
[Security = 2 Manageability = 1 Scalability = 1]

4.3 Packet Filtering Router/Layer-3 Switch between PCN and EN

A number of vendor design documents and early industry papers[13] recommended the use of a bridge, router or layer-3 switch with basic filters in place to control traffic. A typical schematic of this design is shown in Figure 4 below. Many of these devices are effectively acting as packet-filter firewalls and, as discussed earlier, offer limited protection against any sophisticated assault. They are effective in enforcing basic device to device rule sets, but because they lack stateful inspection, will not prevent attacks that take advantage of techniques such as packet fragmentation to obscure the packet contents.

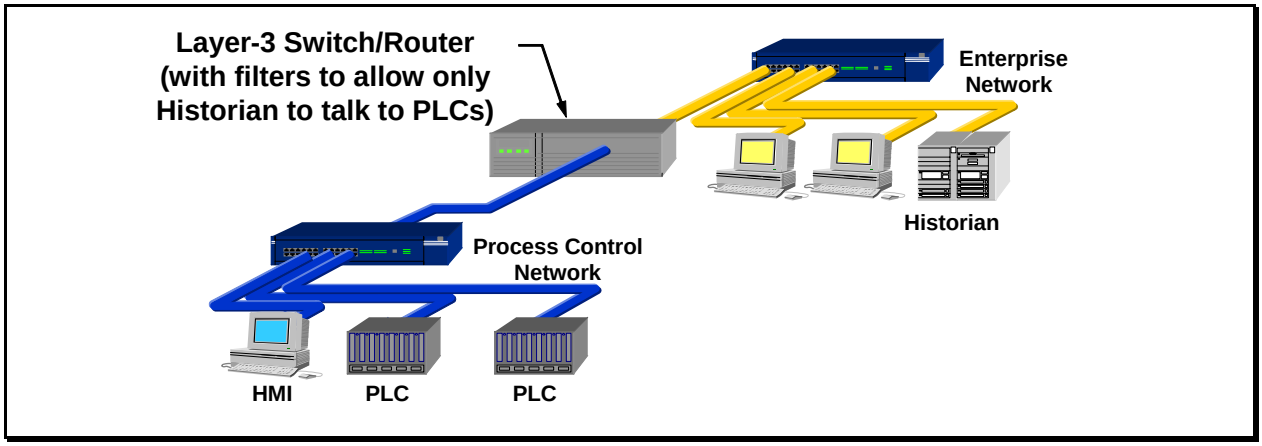

Figure 4: Network Separation Routers or Layer-3 Switches with ACL Filters

This type of packet filter design is only secure if the enterprise network is known to be highly secure in its own right and is not generally subject to attacks. On the other hand, a number of higher end routers on the market today can be upgraded to a stateful firewall relatively inexpensively. If this is done, then the architecture could be considered equivalent to those discussed in Sections 4.4 or 4.6. **[Security = 2 Manageability = 2 Scalability = 4]**.

4.4 Two-Port Firewall between PCN and EN

By introducing a simple two-port firewall between the enterprise and process control networks, a significant security improvement can be achieved. Most firewalls on the market today offer

stateful inspection for all TCP packets and application proxy services for common Internet application layer protocols such as FTP, HTTP and SMTP. Aggressively configured, the chance of a successful external attack on the PCN is significantly reduced. Many of the companies interviewed for this study used this as their standard design for PCN/SCADA security.

Unfortunately two issues still remain with this design. First, on which network are the servers shared enterprise/PCN (such as the data historian) located? If the data historian resides on the enterprise network, a rule must exist within the firewall allowing the data historian to communicate with the control devices on the PCN. A packet originating from a malicious or incorrectly configured host on the enterprise network (appearing to be the data historian) would be forwarded to individual PLCs.

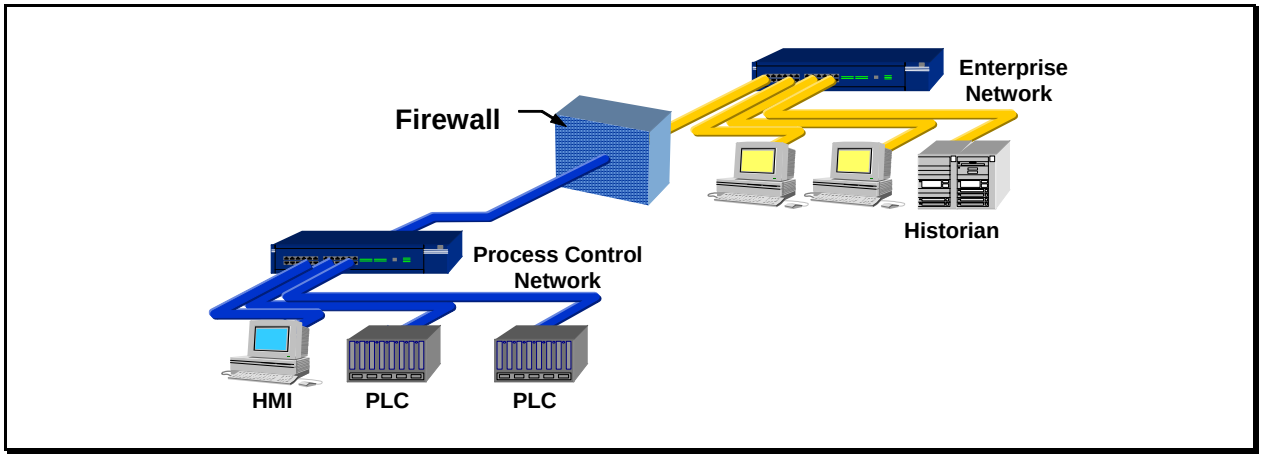

Figure 5: Network Separation with Single Firewall

If the data historian resides on the process control network, a firewall rule must exist allowing for all hosts from the enterprise to communicate with the historian. Typically, this communication occurs as Structured Query Language (SQL) or HTTP requests and flaws in the historian's application layer code could result in a compromised historian. Once the historian is compromised, the remaining nodes on the process control network are vulnerable to a worm/virus propagating or an interactive attack.

Secondly, spoofed packets can be constructed that can affect the PCN and covert data may be able to be tunnelled in allowed protocols. For example, if HTTP packets are allowed through the firewall then Trojan software accidentally introduced on a HMI or PCN laptop could be both controlled by a remote entity and send data (such as captured passwords) to that entity, disguised as legitimate traffic.

In summary, while this architecture is a significant improvement over the previous two, it requires the opening of rule sets that allow direct communications between enterprise and PCN/SCADA devices. This can result in possible security breaches if not very carefully designed and monitored. **[Security = 3 Manageability = 5 Scalability = 4]**

4.5 Router/Firewall Combination between PCN and EN

A slightly more sophisticated design is the use of a router/firewall combination, where the router sits in front of the firewall and offers basic packet filtering services, while the firewall handles the more complex issues using either stateful inspection or proxy techniques. This type of design is very popular in Internet-facing firewalls because it allows the faster router to handle the bulk of the incoming packets, especially in the case of DoS attacks, and reduces the load on the firewall. It also offers improved defence in depth since there are two very different devices an attacker must bypass. Finally, simple rules in the router can prevent packets from circumventing the firewall in complex architectures.

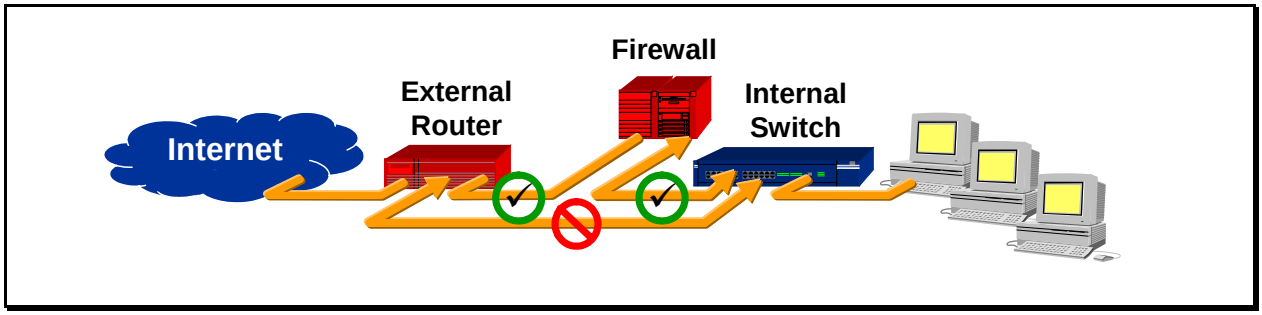

Figure 6: Network Separation with Firewall/Router Combination (Internet Example)

While mentioned in the occasional industrial document (and very popular in the corporate IT world), this design appears to be rarely used in the PCN/SCADA environment. In the few examples the study team did uncover, the routers placed in front of PCN/SCADA firewalls did not appear to be used for security functions. **[Security = 3.5 Manageability = 3 Scalability = 4]**

4.6 Firewall with Demilitarized Zones between PCN and EN

A significant improvement is the use of firewalls with the ability to establish a number of Demilitarised Zones (DMZ) between the enterprise and process control networks. Each DMZ holds a separate "critical" component, such as the data historian, the wireless access point or remote and third party access systems. In effect, the use of a DMZ-capable firewall allows the creation of an intermediate network often referred to as a Process Information Network (PIN).

To create a DMZ requires that the firewall offer three or more interfaces, rather than the typical public and private interfaces. One of the interfaces is connected to the enterprise; the second is connected to the PCN/SCADA network and the remaining interfaces to the shared or insecure devices such as the data historian server or wireless access points. Figure 7 illustrates a typical DMZ firewall design in a PCN/SCADA setting.

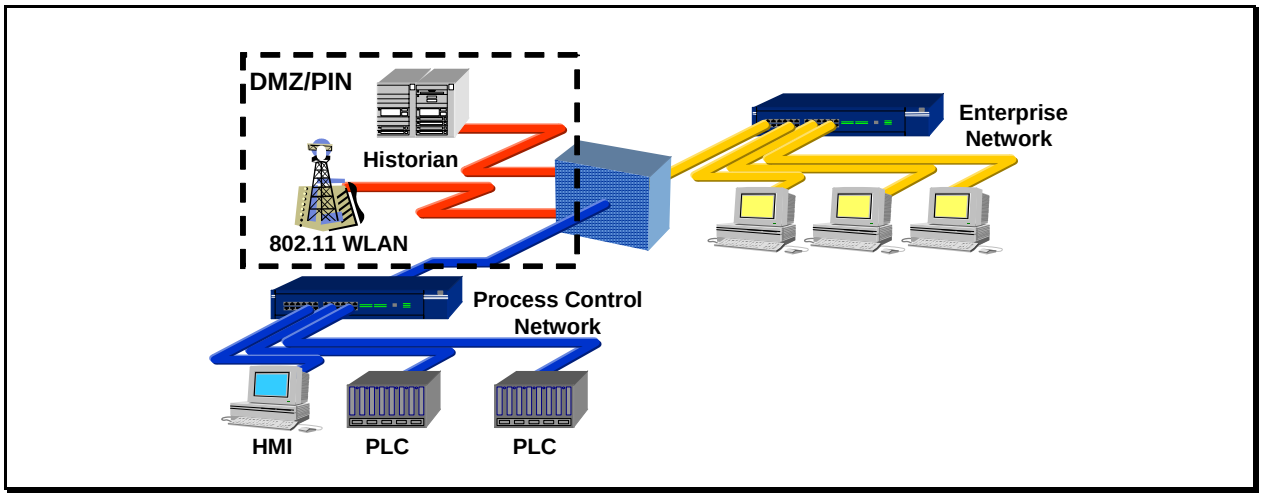

Figure 7: Firewall with Demilitarized Zone for Shared Enterprise/PCN Devices

By placing enterprise-accessible items in the DMZ, no direct communication paths are required from the enterprise network to the plant floor and each network effectively ends in the DMZ. Most sophisticated firewalls can allow for multiple DMZs, and specify what type of traffic is forwardable between zones. As above, the firewall blocks arbitrary packets from the enterprise network from entering the PCN, but also regulates traffic from the other network zones. By judicious use of access control lists a clear separation can be maintained between the PCN and other networks, with very little to no traffic directly between enterprise and PCN/SCADA networks.

The primary security risk in this type of architecture is that if a computer in the DMZ is compromised, then it can be used to launch an attack against the PCN via application traffic permitted from the DMZ to the PCN. This risk can be greatly reduced if a concerted effort is made to harden and actively patch the servers in the DMZ and if the firewall rules set permit only connections between the PCN and DMZ are initiated by PCN devices.

The concerns expressed by interviewed end users with this design were the added complexity (one company pointed out that the multiple interfaces can increase ACL rule complexity and increase the likelihood for errors) and the increased cost of firewalls with three or more ports. However, the study team believes that the improved security more than offsets these disadvantages. Several standards such as API-1164 and a number of the more sophisticated vendor[14 15] and end user documents propose this type of design.
[Security = 4 Manageability = 4.5 Scalability = 4]

4.7 Paired Firewalls between PCN and EN

A variation on the firewall with DMZ solution is to use a pair of firewalls positioned between the enterprise and process control networks. Common servers (such as the data historian) are situated between the firewalls in a DMZ-like network zone called either a PIN or Manufacturing Execution System (MES) layer[16]. As in the architectures above, the first firewall blocks arbitrary packets from proceeding to the process control network or the shared historians. The second firewall can prevent unwanted traffic from a compromised server from entering the PCN or PCN traffic from impacting the shared servers.

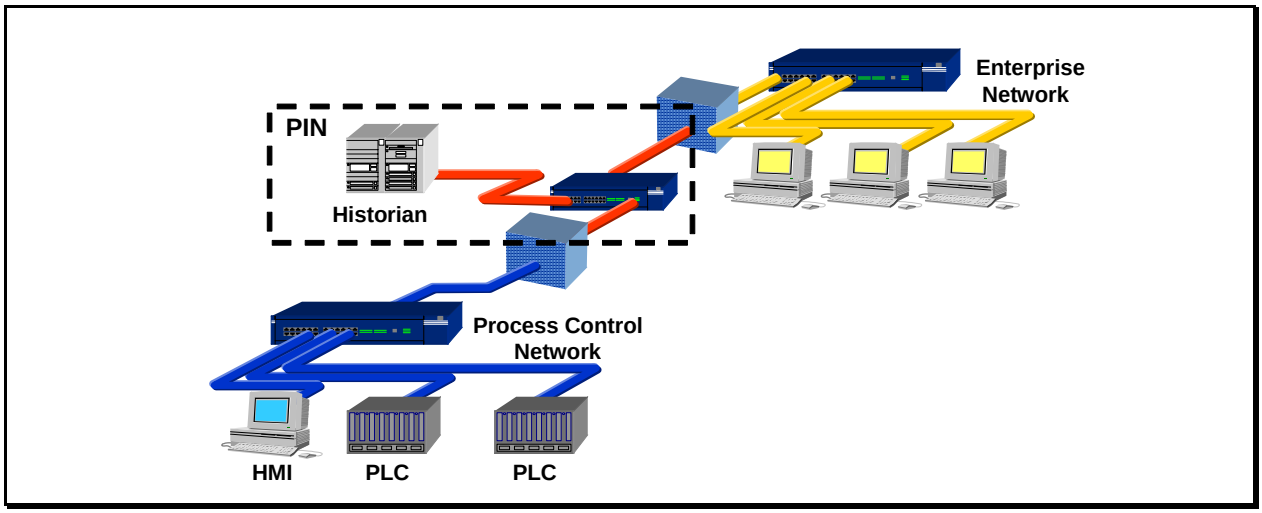

Figure 8: Paired Firewalls with Demilitarized Zone for Shared Enterprise/PCN Assets

If firewalls from two different manufacturers are used, then this solution may offer a "defence in depth" advantage[ii]. It also allows process control groups and the IT groups to have clearly separated device responsibility since each can manage a firewall on its own. In fact it is the study team's understanding that this design is recommended in the Federal Energy Regulatory Commission (FERC) Proposal for Security Standards for this reason.

A variation on this architecture is the basis for the IEC/SC 65C/WG 13 draft for "*Enterprise Network – Control Network Interconnection Profile (ECI)*[17]". It, however, proposed a router for the EN to DMZ firewall, possibly weakening the overall security. The packet filtering services offered by the router could leave the servers inside the DMZ exposed to more sophisticated attacks from the EN and they, if compromised, could be used as a staging point for attacks into the PCN.

The primary disadvantage with all two firewall architectures is the increased cost and management complexity. For environments with severe security requirements or the need for clear management separation, this architecture has some strong advantages.
[Security = 5 Manageability = 3 Scalability = 3.5]

4.8 Firewall and VLAN-based Process Network Combinations

Thus far, the designs have treated the PCN/SCADA network as a single entity. However in many cases there are functional areas or cells in the PCN or SCADA system where inter-area communication is not necessary (or perhaps not wanted). Extending the above architectures by further dividing the SCADA or PCN into a number of VLANs allows for inter-VLAN communication that can be controlled with simple packet filters in Layer-3 switches. Below the Layer-3 switch is a number of VLAN-capable Layer-2 switches that allow direct communication between devices on the same VLAN, but force all inter-VLAN traffic back to the Layer-3 switch for filtering.

[ii] There is considerable debate regarding whether there is a significant advantage in using two different firewalls, since rarely is the firewall itself compromised. Most currently observed infiltrations of PCNs are the result of either viruses/worms through allowed services or mis-configuration of the firewall, independent of vendor.

VLANs prevent the propagation of unwanted traffic across the entire PCN from either accidental access by PCN staff or from introduced viruses. This design also helps mitigate the "virus introduced by a laptop on a PCN" issue. Assuming that the PCN/SCADA environment is highly secure as a result of the other firewall(s) between it and the enterprise networks, then more sophisticated firewalls are not required at this level.

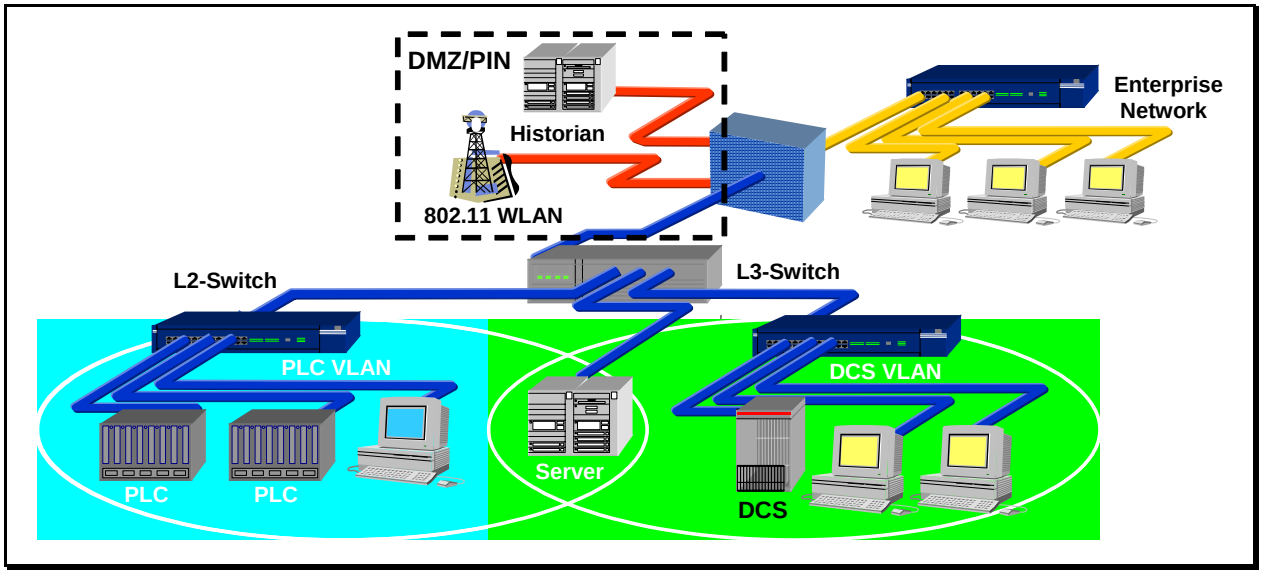

Figure 9: Firewall with Demilitarized Zone and SCADA/PCN VLANs

Several major auto manufacturers utilize variations on this design. Provided that DMZs are also used, then this architecture is very secure. Its primary disadvantages are added management complexity and cost. **[Security = 4.5 Manageability = 3 Scalability = 5]**

4.9 Summary of Firewall Architectures

The eight PCN/SCADA to EN segregation architectures recorded during this survey can be broken down into three general classifications:

1. Separation using non-firewall devices such as dual-homed workstations, bridges and routers (Architectures 1,2 and 3)

2. Two zone firewall-based designs without a DMZ (Architectures 4 and 5)

3. Three zone firewall-based designs with a DMZ (Architectures 6, 7 and 8)

In Table 1 we have summarized our ratings of security, manageability and scalability for each of the eight segregation architectures, while Figure 10 shows the same information in a graphical format. It is important to note that these ratings are approximate only and are based on the assumptions that only technologies that are widely available at the time of the report are used and that they are deployed in typical configurations. Unusual environments, technologies or deployments will require the reader to adjust these ratings accordingly.

Architecture	Security	Manageability	Scalability
1 Dual Homed Computers	1	2	1
2 Dual Homed Server with Personal Firewall	2	1	1
3 Packet Filtering Router/Layer-3 Switch	2	2	4
4 Two-Port Firewall	3	5	4
5 Router/Firewall Combination	3.5	3	4
6 Firewall with Demilitarized Zones	4	4.5	4
7 Paired Firewalls	5	3	3.5
8 Firewall/VLAN-based Combination	4.5	3	5

Table 1: Approximate Security, Manageability and Scalability Ratings for PCN/SCADA Segregation Architectures

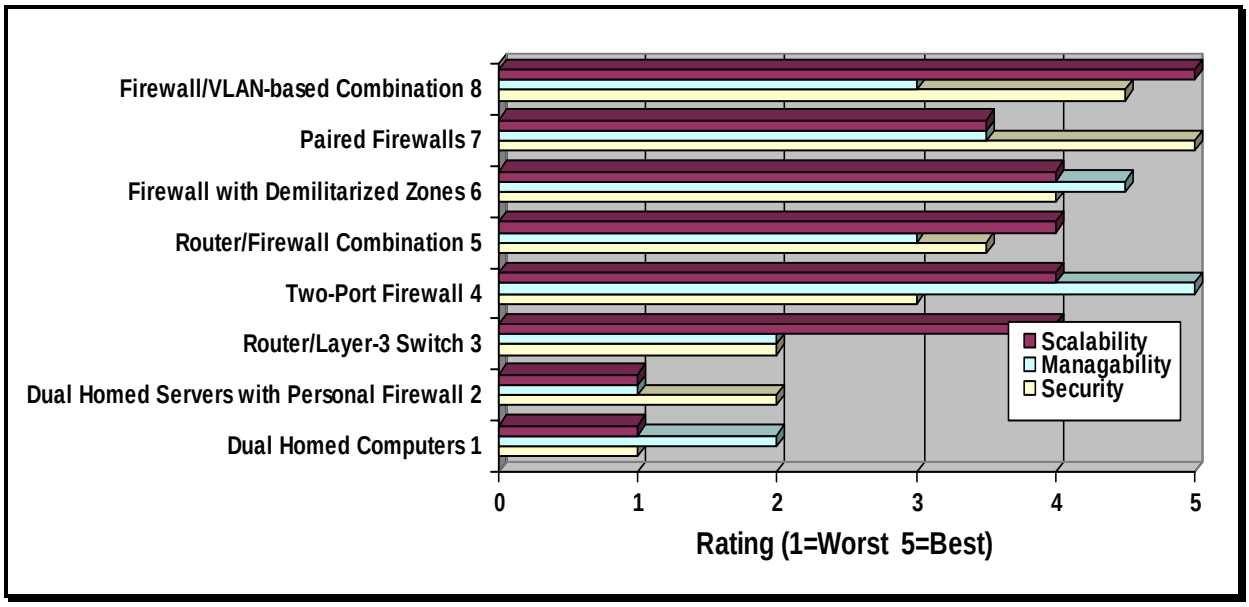

Figure 10: Comparison Chart for PCN/SCADA Segregation Architectures

In summary, it is the study team's opinion that the non-firewall based solutions will generally not provide suitable isolation between PCN/SCADA networks and enterprise networks. The two zone solutions are marginally acceptable but should be only be deployed with extreme care. The most secure, manageable and scalable PCN/EN segregation architectures should be based on a three zone system such as those described in Sections 4.6, 4.7 and 4.8.

This page is deliberately left blank.

5 Firewall Implementation and Configuration

5.1 General Firewall Policies

Once the firewall architecture is in place, the work of determining exactly what traffic you want to allow through the firewall begins. As Request for Comment (RFC) 2196 - Site Security Handbook puts it:

> *The difficult part is establishing the criteria by which the packets are allowed or denied access through the doors.*[18]

Configuring the firewall to "Deny All" except for pin-holes absolutely required for business needs is every company's basic premise, but the reality is much more difficult. Exactly what does "absolutely required for business" mean and what are the security impacts of allowing those "pin-holes" through? For example, many companies considered allowing SQL traffic through the firewall as required for business for many data historian servers. Unfortunately SQL was also the vector for the Slammer worm. The fact is, many important protocols used in the industrial world, such as HTTP, FTP, OPC®/DCOM®, EtherNet/IP™ and MODBUS/TCP, are significant security risks.

If one is installing a single two-port firewall without a DMZ for shared servers (i.e. the architecture described in Section 4.4) then particular care needs to be taken with the rule design. At a minimum, all rules should be stateful rules that are both IP address and port (application) specific. The address portion of the rules should restrict incoming traffic to a very small set of shared devices (e.g. the data historian) on the PCN from a controlled set of address on the enterprise network. Allowing any IP addresses on the enterprise network to access the server inside the PCN is not recommended. In addition, the allowed ports should be carefully restricted to relatively secure protocols such as HTTPS. Allowing HTTP, FTP or any unencrypted SCADA protocol to cross the firewall is a security risk due to the potential for traffic sniffing and modification.

On the other hand, if the DMZ architecture is being used, then it is possible to configure the system so that no traffic will go directly between enterprise and PCN/SCADA networks. With a few special exceptions (noted below) all traffic from either side can terminate at the servers in the DMZ. This allows more flexibility in the protocols allowed through the firewall. For example, MODBUS/TCP might be used to communicate from the PLCs to the Data Historian, while HTTP might be used for communication between the historian and enterprise clients. Both protocols are inherently insecure, yet in this case they can be used safely as neither actually crosses between enterprise and PCN.

An extension to this concept is the idea of using "disjoint" protocols in all PCN-enterprise communications. That is, if a protocol is allowed between the PCN and DMZ then it is explicitly NOT allowed between DMZ and enterprise networks. This design greatly reduces the chance of a worm such as Slammer actually making its way into the PCN/SCADA network since the worm would have to deploy two different exploits over two different protocols.

One area of considerable variation in practice is the control of outbound traffic from the PCN. While some organizations seem to have a very open attitude towards this type of traffic, the study team believes this traffic to represent a significant risk if unmanaged. One example, noted earlier in Section 4.3, is trojan software that uses HTTP tunnelling to exploit poorly defined outbound rules. Thus it is important that outbound rules be as stringent as inbound rules. Appendix-A of ISA's SP-99 Technical Report #2 contains some example guidelines that help clarify this. A summary of these follows:[19]

1. Outbound traffic through the PCN firewall should be limited to essential communications only;

2. All outbound traffic from the PCN to the enterprise network will be source and destination restricted by service and port using static firewall rules;

3. Mapped drives across the PCN firewall should be avoided.

In addition to these rules, the firewall should be configured with outbound filtering to stop forged IP packets from leaving the PCN/SCADA network or the DMZ. In practice this is achieved by checking the source IP addresses of outgoing packets against the firewall's respective network interface address. The intent is to prevent the PCN network from being the source of spoofed (i.e. forged) communications which are often used in DoS Attacks. Thus the firewalls should be configured to forward IP packets only if those packets have a correct source IP address for the PCN/SCADA or DMZ networks[20].

Finally, the question of PCN/SCADA connection and communication with the Internet elicits a wide variety of practices. For example, API -1164 states:

> *"Internet connections should not terminate directly into the SCADA network. A firewall should be used to isolate the SCADA network from the Internet."*[21]

In contrast, one of the end-user documents states:

> *"PCNs shall not be directly connected to the Internet, even if protected via a firewall"*.

While there will be special situations (such as access for remote support) that might support the first rule, we believe that the second is far more secure and should be the goal. Either way, Internet access by devices on the PCN should be strongly discouraged.

In summary, the study team believes that the following should be considered as recommended practice for general firewall rule sets:

1. The base rule set should be DENY ALL, PERMIT NONE.

2. Ports and services between the PCN environment and an external network should be enabled and permissions granted on a specific case by case basis. There should be a documented business justification with risk analysis and a responsible person for each permitted incoming or outgoing data flow[22].

3. All "permit" rules should be both IP address and TCP/UDP port specific, and stateful if appropriate;

4. All rules shall restrict traffic to specific IP address or range of addresses;

5. All traffic on the PCN/PIN is typically based only on routable IP protocols, either TCP/IP or UDP/IP. Thus any non-IP protocol should be dropped;

6. Prevent traffic from transiting directly from the PCN/SCADA network to the enterprise network. All traffic should terminate in the DMZ;

7. Any protocol allowed between the PCN and DMZ is explicitly NOT allowed between DMZ and enterprise networks (and vice-versa);

8. All outbound traffic from the PCN to the enterprise network should be source and destination restricted by service and port using static firewall rules;

9. Allow outbound packets from the PCN or DMZ only if those packets have a correct source IP address assigned to the PCN or DMZ devices;

10. PCN devices should not be allowed to access the Internet.

11. PCNs shall not be directly connected to the Internet, even if protected via a firewall.

12. All firewall management traffic be either via a separate, secured management network (e.g. out of band) or over an encrypted network with two-factor authentication. Traffic should also be restricted by IP address to specific management stations.

The reader is cautioned that these should only be considered as guidelines. A careful assessment of each control environment is required before implementing any firewall rule sets. Furthermore, there will always be exceptions to these rules. For example, certain direct database to database connections may be needed to synchronize a data historian on the PCN with a corporate-wide database, violating suggestion #4. Similarly, time service communication to a corporate time server may have to traverse the PCN firewall to synchronize device times on PCN. These exceptions and how they might be addressed will be discussed in more detail in later sections of the report.

5.2 Rules for Specific Services

Beside the general rules described above, it is difficult to outline all purpose rules for specific protocols. The needs and best practices vary significantly between industries for any given protocol and should be analysed on a company by company basis. The Industrial Automation Open Networking Association (IAONA) offers a template for conducting such an analysis[23], assessing each of the protocols commonly found in industrial environments in terms of function, security risk, worst case impact and suggested measures. Below we will summarize some of the key points from the IAONA document, suggested practices from the ISA TR2 Appendix A and several user policy documents. The reader is advised to consult these documents directly when developing rule sets.

5.2.1 Domain Name Service (DNS)

Domain Name Service (DNS) is primarily used to translate between domain names (such as control.com) and IP addresses (such as 192.168.1.1). Most Internet services rely heavily on DNS, but its use on the plant floor is relatively rare at this time. In most cases there is little reason to allow DNS requests out of the PCN to the enterprise network and no reason to allow DNS requests into the PCN. DNS requests from the PCN to DMZ should be addressed on a case by case basis. Local DNS or the use of host files is recommended.

5.2.2 Hyper Text Transfer Protocol (HTTP)

Hyper Text Transfer Protocol (HTTP) is the protocol underlying web browsing services on the Internet. Like DNS, it is critical to most Internet services. It is seeing increasing use on the plant floor as an all purpose query tool. Unfortunately it has little inherent security and has the ability to be a transport mechanism for a very large number of manual attacks and worms. In addition HTTP applications are renowned for having vulnerabilities that can be exploited.

In general, HTTP should not be allowed to cross from the enterprise to the PCN. If it is, then HTTP proxies should be configured on the firewall to block all inbound scripts and Java applications. Incoming HTTP connections should not be allowed into the PCN as they pose significant security risks. If HTTP services into the PCN are absolutely required, it is recommended that HTTPS be used instead and only to very specific devices.

5.2.3 File Transfer Protocol (FTP) and Trivial File Transfer Protocol (TFTP)

The File Transfer Protocol (FTP) and Trivial File Transfer Protocol (TFTP) are used for transferring files between devices. They are implemented on almost every platform including many DCS, PLCs and Remote Terminal Units (RTUs), since they are extremely well known and use minimum processing power. Unfortunately neither protocol was created with security in mind; in the case of FTP the login password is not encrypted and, in the case of TFTP, no login is required at all. Furthermore, some FTP implementations have a history of buffer overflow vulnerabilities. As a result all TFTP should be blocked while FTP should be allowed on outbound sessions only or if secured with additional token-based two-factor authentication and an encrypted tunnel.

5.2.4 Telnet

The Telnet protocol defines an interactive, text based communications session between a client and a host. It is mainly used for remote login and simple control services to systems with limited resources or to systems with limited needs for security. It is a severe security risk because all telnet traffic, including passwords, is unencrypted and it can allow a remote individual considerable control over a device. Thus inbound Telnet session commands from the enterprise to the PCN should be prohibited unless secured with token-based two-factor authentication and an encrypted tunnel. Outbound telnet sessions should be allowed only over encrypted tunnels to specific devices.

5.2.5 Simple Mail Transfer Protocol (SMTP)

The Simple Mail Transfer Protocol (SMTP) is the primary email transfer protocol on the Internet. Email messages are notorious for containing viruses so inbound e-mail should not be allowed to any PCN device. Outbound SMTP mail messages from the PCN to the enterprise are acceptable.

5.2.6 Simple Network Management Protocol (SNMP)

The Simple Network Management Protocol is used to provide network management services between a central management console and network devices such as routers, printers and PLCs. While SNMP is an extremely useful service for maintaining a network, its security weaknesses are infamous. Version 1 and 2 of SNMP uses unencrypted passwords to both read and configure devices (including devices such as PLCs) and in many cases the passwords are well known and cannot be changed. Version 3 is considerably more secure but is still limited in use. In addition, SNMP implementations in embedded devices have been shown to be seriously flawed in many instances. Thus SNMP commands both to and from the PCN should prohibited unless it is over a separate, secured management network.

5.2.7 Distributed Component Object Model (DCOM)

The Distributed Component Object Model (DCOM) is the underlying protocol for both the popular OLE for Process Control (OPC) and ProfiNet®. It utilizes Microsoft's Remote Procedure Call (RPC) service which has known vulnerabilities that were the basis for the Blaster Worm exploits. In addition, OPC (DCOM) dynamically opens a wide range of ephemeral ports (#1024 – #65535) that can be extremely difficult to filter at the firewall. This protocol should only be allowed between PCN and DMZ networks and explicitly blocked between the DMZ and enterprise network. Also, users are advised to restrict the port ranges used by making registry modifications on devices using DCOM.

5.2.8 SCADA and Industrial Protocols

SCADA and industrial protocols, such as MODBUS/TCP, EtherNet/IP and DNP3, are critical for communications to most control devices. Unfortunately, these protocols were designed without security in mind and do not typically require any authentication to remotely execute commands on a control device. Thus these protocols should only be allowed between the PCN and PIN and not allowed to cross into the enterprise network.

5.3 Network Address Translation (NAT)

One of the more contentious "best practices" is whether or not to use Network Address Translation (NAT) on the firewall. For example, from ISA SP-99 Technical Report #2:

> *"NAT will not be used on the PCN"*[24]

While in an end-user firewall practices the guideline reads:

> *"NAT is used on the interface connecting to the (Enterprise Network)"*

As noted in section 2.3, NAT is a service where IP addresses used on one side of the firewall can be mapped to a different set on the other side on an as-needed basis. It was originally designed for IP address reduction purposes so that a company with a large number of devices that occasionally needed Internet access could get by with a smaller set of assigned Internet addresses.

To do this, NAT relies on the premise that not every internal device is actively communicating with external hosts at a given moment. The firewall is configured to have a limited number of outwardly visible IP addresses. When an internal host seeks to communicate to an external host, the firewall remaps the internal IP address and port to one of the currently unused, more limited, public IP addresses -- effectively concentrating outgoing traffic into fewer IP addresses. The firewall must track the state of each connection, and how each private internal IP address and source port was remapped onto an outwardly visible IP address/port pair. When returning traffic reaches the firewall, the mapping is reversed and the packets forwarded to the proper internal host.

For example, a PCN-based device may need to establish a connection with an external, non-PCN host (for instance, to send a critical alert email). NAT allows the internal IP address of the initiating PCN host to be replaced by the firewall; subsequent return traffic packets are remapped back to the internal IP address and sent to the appropriate PCN device. More specifically, if the PCN is assigned the private subnet 192.168.1.xxx and the Internet network expects the device to use the corporate assigned addresses in the range 142.232.yyy.zzz, then a NAT firewall will substitute (and track) a 142.232.yyy.zzz source address into every outbound IP packet generated by a PCN device.

Today NAT is often promoted as a security feature since NAT’ed traffic ends up appearing (from an IP address perspective) as though the firewall is the initiator of the connection. NAT has effectively camouflaged the internal host's identity by acting as an intermediary between the internal and external hosts. If a device is not actively communicating to an external network then no NAT mapping for that device will exist in the firewall and thus no externally routable address is visible to the outside world. If internal devices are addressed with private addresses (as per RFC 1918), no router should allow direct traffic to reach them from the Internet; the only way to communicate with the internal devices is in fact through the NAT'ed connection. Of course any device with an active connection through the firewall will be visible to the outside world.

Whether NAT is really a true security technique or just “security through obscurity" is an open debate in the networking community. However, there is little doubt that using NAT can have negative consequences in some certain situations. First of all, debugging a firewall or validating a firewall's rule set as secure can be made more difficult using NAT, due to the dynamic nature of the currently open connections and the IP address/port combinations that must be allowed to pass back in.

Furthermore, certain protocols are specifically broken by NAT because of the lack of direct addressing. For example, OPC requires special 3rd party tunnel software to work with NAT. Producer-consumer protocols, such as EtherNet/IP and Foundation Fieldbus HSE®, are particularly troublesome as NAT does not support the multicast-based traffic that these protocols need to offer their full services.

In summary, while NAT offers some distinct advantages, its impact on the actual industrial protocols and configuration should be assessed carefully before it is deployed.

5.4 Specific PCN Firewalls Issues

5.4.1 Data Historians

As noted earlier in this report, the existence of shared PCN/enterprise servers such as data historians and asset management servers can have a significant impact on firewall design and configuration. In three-zone systems the placement of these servers in a DMZ is relatively straightforward, but in two-zone designs the issues become complex. Placing the historian on the enterprise side of the firewall means that a number of insecure protocols, such as MODBUS/TCP or DCOM, must be allowed through the firewall and that every control device reporting to the historian is exposed to the enterprise side of the network. On the other hand, putting the historian on the PCN side means other equally questionable protocols, such as HTTP or SQL, must be allowed through the firewall and there is now a server accessible to nearly everyone in the corporation sitting on the PCN.

In general, the best solution is to avoid two-zone systems and use a three-zone design, placing the data collector in the PCN and the historian component in the DMZ or PIN. Even this can prove problematic in some situations. Heavy access from the large numbers of users on the enterprise network to a historian in the DMZ may tax the firewall's throughput capabilities. One suggested solution is to install two servers; one on the PCN to collect data from the control devices and a second on the enterprise network mirroring the first server and supporting client queries. Of course this requires a special hole to be put through the firewall to allow direct server to server communications, but if done correctly, this poses only minor risk.

5.4.2 Remote Support Access

Another issue for PCN/SCADA firewall design is the use of 3rd party or remote access into the PCN. Obviously any users accessing the PCN from remote networks should be required to authenticate using an appropriately strong mechanism such as token-based authentication. While it is possible for the controls group to set up their own remote access system with two-factor authentication on the DMZ, in most companies it is typically more efficient to use existing systems set up by the IT department. In this case a connection through the firewall from the IT remote access server is needed.

Several documents suggest that remote support personnel connecting over the Internet or via dialup modems should run the corporate VPN connection client and authenticate using the token based two-factor authentication scheme in order to connect to the general corporate network. Once connected there they should be required to authenticate a second time at the PCN firewall (using two factor authentication) to gain access to the PCN network. For companies that don't allow any control traffic traversing the enterprise network in the clear, this will require a cascading or secondary tunnelling solution to gain access to the PCN, such as Secure Sockets Layer (SSL) VPNs inside an Internet Protocol Security (IPsec) VPN.

5.4.3 Multicast Traffic

Multicasting is a communication method in which a source device can send a single data message simultaneously to a group of destination devices. As opposed to sequentially sending the same message to each destination, the message can be sent once over a network and any host "tuned in" to the transmission will receive it.

Most industrial producer-consumer (or publisher-subscriber) protocols operating over Ethernet, such as EtherNet/IP and Foundation Fieldbus HSE, are IP multicast-based. The first advantage of IP multicasting is network efficiency; by not repeating the data transmission to the multiple destinations, a significant reduction in network load can occur. The second advantage is that the sending host need not be concerned with knowing every IP address of every destination host listening for the broadcast information. The third, and perhaps most important for industrial control purposes, is a single multicast message offers far better capabilities for time synchronization between multiple control devices than multiple unicast messages.

Multicasting in IP environments typically occurs through the use of multicast group ID's which are mapped directly onto the class D IP address range (from 224.0.0.0 to 239.255.255.255). Each address is considered a separate “transmission frequency”; a host listening to multicast packets must "tune in" to the group ID (IP address) of the transmission it wishes to receive.

If the source and destinations of a multicast packet are connected with no intervening routers or firewalls between them, the multicast transmission is relatively seamless[iii]. However, if the source and destinations are not on the same LAN, forwarding the multicast messages to a destination becomes more complicated. To solve the problem of multicast message routing, hosts need to join (or leave) a group by informing the multicast router on their network of the relevant group ID through the use of the Internet Group Management Protocol (IGMP). Multicast routers subsequently know of the members of multicast groups on their network and can decide whether or not to forward a received multicast message onto their network. A multicast routing protocol is also required. From a firewall administration perspective, monitoring and filtering IGMP traffic becomes another series of rule sets to manage, adding to the complexity of the firewall.

Another firewall issue related to multicasting is the use of NAT. A NAT'ing firewall receiving a multicast packet from an external host has no reverse mapping for which internal group ID to send the data to. It could, if IGMP-aware, broadcast it to every group ID it knows about (one of them will be correct!), but this could cause serious issues if an unintended control packet was broadcast to a critical node. The safest action for the firewall to take is to drop the packet. Thus multicasting is generally considered NAT-unfriendly.

[iii] Since it's easy to set up and requires little administration, multicasting is gaining in popularity among the control system suppliers. However, it can be extremely disruptive in some situations. At least one large end-user reported significant problems with legitimate multicast traffic on PCNs with large VLANs and many PLCs.

6 Management of PCN/SCADA Firewalls

The firewall is a focal point for PCN/SCADA security and requires considerable resources, not only for initial design and commissioning, but also for on-going monitoring, incident management, upgrading and technical support. The complexity of this task should not be underestimated, but unfortunately it often is. For example, a recent paper on firewall configuration errors by Avishai Wool showed that even core firewalls in major corporations can be enforcing poorly written rule sets and vulnerable to attack[25]. In the study the author defined 12 serious firewall configuration errors (each very general in nature) and then inspected the firewall configurations of 27 major corporations. He found an average 7 serious errors per firewall, with some having as many as 12 errors. The results clearly indicate the complex nature of firewall management.

The following are some of the management tasks that should be considered when setting up a firewall:

1. **Change Management and Documentation:** The configuration and rule sets in the SCADA/PCN can have a significant impact on production and safety in most facilities. Thus all firewall policies must be carefully documented and subject to the same change management requirements any PLC or DCS would be subjected to.

   It is recommended that all access control lists be documented and that this documentation includes the purpose of each rule, interdependencies, and security considerations addressed. The rules on these devices need to be reviewed regularly to ensure that the business case for the connection is still valid and the security controls in place.

2. **Monitoring of Firewall Logs and IDS:** The team responsible for the firewall must perform monitoring and associated alerting of both logs and IDS events.

3. **Define an Incident Response Plan:** An incident response plan defines the emergency response process to a suspected or actual incident and defines the steps the management team uses to address the incident. This will include possible disconnection of the PCN from the enterprise network and the later re-engagement process. Where a full disconnect is not practical it should define the maximum possible restrictions that can be enabled when required.

   The re-engagement process can include details of so-called ‘pinhole’ configurations which allow temporary access to an extremely small set of trusted resources from which information and, for instance, anti-virus measures can be picked up so that the process control environment defences can be brought up to date prior to restoring full connectivity.

4. **Support, patches and updates:** The management team should monitor the appropriate vulnerability lists, vendor update lists and Computer Emergency Response Team (CERT) security alerts. Suitably qualified and authorised personnel should perform updates, upgrades, anti-virus upgrades, user/account management and capacity monitoring on a regular basis.

The above management services can either be provided by local staff such as the process control or IT departments or it can be provided from a central resource, either corporate or a 3rd party. Of the 15 companies interviewed, most of the large corporations managed their PCN firewalls centrally, while the smaller firms managed it locally with internal resources. A suggested rule of thumb was that central management is preferred for groups with more than 10 PCN/SCADA firewalls.

Regardless of whether central or local management is selected, nearly all end users pointed out the need to include control engineers as part of the team designing, implementing and maintaining the firewall systems. In general, the people who understand and set-up firewalls well do not know process control specifics. For example, most security experts are Internet focused and do not know that most control systems do not use DNS in the PCN. As a result they might assume that a DNS request is legitimate when in many cases it is an attempt to breach the firewall or an indication of improper activity on the PCN. Creating cross-functional controls/IT security teams is highly recommended for PCN/SCADA firewall management.

7 Special or Future Technologies

7.1 SCADA Protocol Aware Firewalls

At the present time, commercially-available firewalls are focused on traditional Internet and corporate application layer protocols and are unaware of the existence of industrial protocols such as MODBUS/TCP or EtherNet/IP. As a result, they can not examine or filter SCADA packets at the application layer or offer proxy services for these protocols. What would be ideal is a firewall that would understand the protocols well enough to allow rules that would block certain SCADA functions. For example, a rule might allow only MODBUS read commands to cross the firewall and drop all packets with invalid or unauthorized function codes.

While no commercial products are offered at the present time, open source MODBUS-aware firewall extensions to the Linux kernel have been developed by Matthew Franz and Venkat Pothamsetty of the Cisco Systems Critical Infrastructure Assurance Group (CIAG). These are available at no charge from http://modbusfw.sourceforge.net/

7.2 Distributed Micro-Firewalls

Several researchers have proposed the idea of a distributed micro-firewall appliance that could be installed in front of each PLC or RTU that is in need of critical protection. Each unit would be configured with security applications specific to its purpose, while tied back to a centralized management system. These firewall appliances would communicate over a secure channel, further protecting both the appliance and the PLC behind it from malicious attacks. The benefit of this approach is that it would offer a significant second layer of defence inside the PCN firewall and would protect critical devices from internal attack.

No products of this type are currently available for the SCADA market, but a research project to develop a prototype is currently underway at the British Columbia Institute of Technology. In addition, Siemens has announced plans to release in late 2004 a VPN gateway that can also act as a distributed micro-firewall.

7.3 Quality of Service (QoS)

One suggestion uncovered during this study was the use of Quality of Service (QoS) features in network equipment as a way to ensure that DoS attacks are of limited effectiveness in PCNs. QoS is a concept where packets are tagged with a special priority field that indicates how they should be handled by each router or switch they encounter. The IEEE 802.1q variation of QoS is getting some attention in the industrial setting for this reason, but is not yet widely deployed. However, should QoS gain wide usage in the industrial world then this could be a useful tool in the security arsenal.

It should be noted that QoS mitigation techniques need a detection mechanism (such as IDS/AV) to be effective. The mitigation effectiveness is also increased significantly if the QoS mechanism implemented on the device using Network-Based Application Recognition (NBAR). If the device can not inspect the traffic, then QoS is effective only if you can trust the markings

of the packets coming into the networking device. Unfortunately, a couple of recent of worms actually make the QoS tags high to get better propagation characteristics.

7.4 One Way Communication Paths

The concept of one-way communication paths that allow traffic out of the PCN only is discussed in Annex G of IEEE Std 7-4.3.2[26], "IEEE Standard for Digital Computers in Safety Systems of Nuclear Power Generating Stations". While this discussion appears to be designed with serial links in mind, it is possible to create these one-way paths over IP connections using a UDP, instead of the more well known TCP. At the present time we are unaware of any companies deploying this technique in the field, but it shows promise for sites requiring more stringent security than ACLs can provide.

Acronyms

ACL – Access Control List
AIChE – American Institute of Chemical Engineers
API – American Petroleum Institute
BCIT – British Columbia Institute of Technology
CERT – Computer Emergency Response Team
CIAG – Critical Infrastructure Assurance Group (Cisco Systems Inc.)
CIP – Common Industrial Protocol
DCOM – Distributed Component Object Model
DCS – Distributed Control System
AV – Anti-Virus
DH –Data Historian
DMZ – DeMilitarised Zone
DNS – Domain Name Service
DoS – Denial of Service
DPI – Deep Packet Inspection
EN – Enterprise Network
ESD – Emergency Shutdown System
FERC – Federal Energy Regulatory Commission
FTP – File Transfer Protocol
GAIT – Group for Advanced Information Technology
HMI – Human Machine Interface
HTTP – Hyper-Text Transfer Protocol
IAONA – Industrial Automation Open Networking Association
IDS – Intrusion Detection Systems
IEC – International Electrotechnical Commission
IEEE – Institute of Electrical and Electronics Engineers
IGMP – Internet Group Management Protocol
IP – Internet Protocol
IPsec – Internet Protocol Security
ISA – Instrumentation, Systems and Automation Society
ISO – International Organization for Standardization
IT – Information Technology
LAN – Local Area Network
MES – Manufacturing Execution System
NAT – Network Address Translation
NBAR – Network-Based Application Recognition
NIC – Network Interface Card
NISCC – National Infrastructure Security Coordination Centre
OPC – OLE for Process Control
OS – Operating System
PC – Personal Computer
PCN – Process Control Network
PIN – Process Information Network
PLC – Programmable Logic Controllers

QoS – Quality of Service
RFC – Request For Comment
RPC – Remote Procedure Call
RTU – Remote Terminal Unit
SCADA – Supervisory Control and Data Acquisition
SOAP – Simple Object Access Protocol
SQL – Structured Query Language
SMTP – Simple Mail Transfer Protocol
SNMP – Simple Network Management Protocol
SSL – Secure Sockets Layer
TCP – Transmission Control Protocol
TFTP – Trivial File Transfer Protocol
UDP – User Datagram Protocol
VLAN – Virtual Local Area Network
VPN - Virtual Private Network
WLAN – Wireless Local Area Network
XML – Extended Markup Language

References

[1] Smith, T.; “Hacker jailed for revenge sewage attacks,” *The Register*, October 31, 2001, http://www.theregister.co.uk/content/4/22579.html

[2] "SQL Slammer Worm Lessons Learned For Consideration By The Electricity Sector", *North American Electric Reliability Council*, Princeton NJ, June 20, 2003

[3] "NRC Information Notice 2003-14: Potential Vulnerability of Plant Computer Network to Worm Infection", United States Nuclear Regulatory Commission, Washington, DC, August 29, 2003

[4] E. Byres, J. Carter, A. Elramly and D. Hoffman; “Worlds in Collision: Ethernet on the Plant Floor”, ISA Emerging Technologies Conference, *Instrumentation Systems and Automation Society*, Chicago, October 2002

[5] E.J. Byres and D. Hoffman; “IT Security and the Plant Floor”, InTech Magazine, *Instrumentation Systems and Automation Society*, Research Triangle Park, NC, p. 76, December 2002

[6] The Ten Commandments of Industrial Ethernet, B&B Electronics Manufacturing Company, March 30, 2004

[7] Technical Report ISA-TR99.00.01-2004: Security Technologies for Manufacturing and Control Systems, *Instrumentation, Systems and Automation Society (ISA),* March 2004

[8] ibid - Technical Report ISA-TR99.00.01-2004

[9] B. Fraser, “RCF 2196 - Site Security Handbook”, *Internet Engineering Task Force*, September 1997, Pg. 22

[10] ibid - Technical Report ISA-TR99.00.01-2004

[11] “API Standard 1164 – SCADA Security (Draft) – Appendix B”, *American Petroleum Institute*, March 2004

[12] http://udell.roninhouse.com/bytecols/1999-10-13.html

[13] E.J. Byres; “Designing Secure Networks for Process Control”, IEEE Industry Applications Magazine, *Institute of Electrical and Electronics Engineers*, New York, Vol. 6, No. 5 p. 33 -39, September/October 2000

[14] “Process Control Network Reference Architecture v 1.0”, *Invensys Inc*., January 2004, pg. 2, 5

[15] “Experion PKS Network and Security Planning Guide EP-DSX173, Release 210”, Honeywell Limited Australia, October 2004

[16] “Presentation: Securing SIMATIC PCS7 and SIMATIC IT in Networks”, *Siemens,* 2003

[17] IEC/SC 65C/WG 13 Draft v1.04 “Enterprise Network – Control Network Interconnection Profile (ECI)”, *International Electrotechnical Commission*, December 2004

[18] B. Fraser, “RCF 2196 - Site Security Handbook”, *Internet Engineering Task Force*, September 1997, Pg. 21

[19] "Technical Report ISA-TR99.00.02-2004: Integrating Electronic Security into the Manufacturing and Control Systems Environment", *Instrumentation, Systems and Automation Society (ISA),* April 2004, Page 79

[20] Process Control Digital Security Practice: Firewall Practice Release 1.0, (Name withheld on request), 2003.

[21] "API Standard 1164 – SCADA Security (Draft)", *American Petroleum Institute*, March 2004, Page 20

[22] ibid - IEC/SC 65C/WG 13 Draft v1.04, Page 11

[23]"The IAONA Handbook for Network Security - Draft/RFC v0.4", *Industrial Automation Open Networking Association (IAONA)*, Magdeburg, Germany, 2003

[24] ibid, "Technical Report ISA-TR99.00.02-2004" Page 77

[25] Avishai Wool, "A quantitative study of firewall configuration errors" IEEE Computer Magazine, *IEEE Computer Society*, June 2004, Pages 62-67

[26] IEEE Standard 7-4.3.2, "IEEE Standard for Digital Computers in Safety Systems of Nuclear Power Generating Stations.", *Institute of Electrical and Electronic Engineers*

[22] John C. Netzel, *Network Security Across Wide Area Networks & the Internet*, IndComm 2003, Melbourne Australia, May 2003.

[23] Technical Architecture Whitepaper Network Zone Model v1.2E, The Dow Chemical Company, 2004.

[24] Process Network Security: Firewall Configuration and Policies, Invensys Inc., 2004.

[25] Montgomery Watson Harza, "Security for SCADA IP Networks", 2003.