3

SEPARATION

CHAPTER OUTLINE

A limitation of firewalls is that they can only be as good as their access controls and filters. They might fail to detect subversive packets. In some situations, they might be bypassed altogether. For example, if a computer behind a firewall has a dial-up port, as is all too common, an intruder can get access by dialing the machine.

Dorothy Denning[1]

The separation of network assets from malicious intruders using a firewall is perhaps the most familiar protection approach in all of computer security. Today, you will find some sort of firewall deployed in or around virtually every computer, application, system, and network in the world. They serve as the centerpiece in most organizations' security functionality, including intrusion detection, antivirus filtering, and even identity management. An enormous firewall industry has emerged to support such massive deployment and use, and this industry has done nothing but continue to grow for years and years.

In spite of this widespread adoption, firewalls as separation mechanisms for large-scale infrastructure have worked to only a

[1] D. Denning, *Information Warfare and Security*, Addison–Wesley, New York, 1999, p. 354.

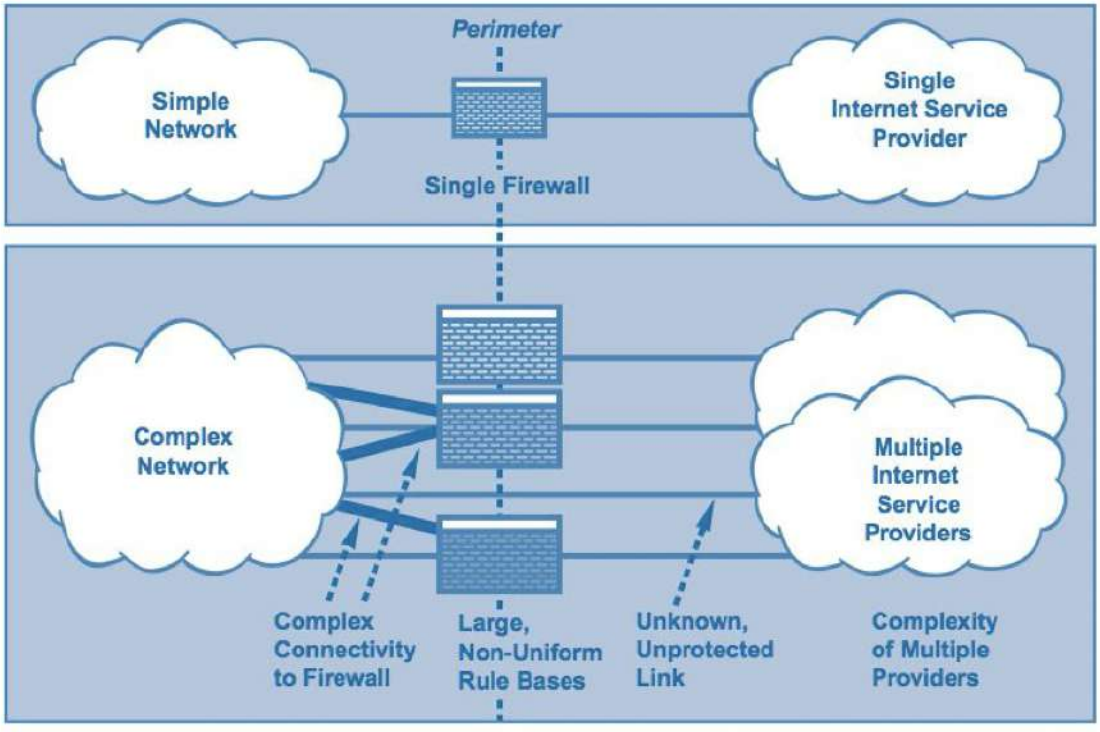


Figure 3.1 Firewalls in simple and complex networks.

Firewalls are valuable and frequently employed but may not provide enough protection to large-scale networks.

limited degree. The networks and systems associated with national infrastructure assets tend to be complex, with a multitude of different entry points for intruders through a variety of Internet service providers. In addition, the connectivity requirements for complex networks often result in large rule sets that permit access for many different types of services and source addresses. Worse, the complexity of large-scale networks often leads to unknown, unprotected entry points into and out of the enterprise (see Figure 3.1).

Certainly, the use of traditional perimeter firewalls will continue to play a role in the protection of national assets, as we will describe below. Egress filtering, for example, is often most efficiently performed at the perceived perimeter of an organization. Similarly, when two or more organizations share a private connection, the connection endpoints are often the most natural place to perform firewall filtering, especially if traditional circuit-switched connections are involved. To achieve optimal separation in the protection of large-scale national assets, however, three new firewall approaches will be required:

- *Network-based separation*—Because the perimeter of any complex national infrastructure component will be difficult to define accurately, the use of separation methods such as network-based firewalls is imperative. Such cloud-based functionality allows a broader, more accurate view of the egress and ingress activity for an organization. It also provides a

richer environment for filtering high-capacity attacks. The filtering of denial of service attacks aimed at infrastructure, for example, can only be stopped with special types of cloud-based filtering firewalls strategically placed in the network.

- *Internal separation*—National infrastructure protection will require a program of internal asset separation using firewalls strategically placed in infrastructure. This type of separation of internal assets using firewalls or other separation mechanisms (such as operating system access controls) is not generally present in most infrastructure environments. Instead, the idea persists that insiders should have unrestricted access to internal resources and that perimeter firewalls should protect resources from untrusted, external access. This model breaks down in complex infrastructure environments because it is so easy to plant insiders or penetrate complex network perimeters.
- *Tailored separation*—With the use of specialized protocols in national infrastructure management, especially supervisory control and data acquisition (SCADA), tailoring firewalls to handle unique protocols and services is a requirement. This is a challenge because commercial firewalls are generally designed for generic use in a wide market and tailoring will require a more focused effort. The result will be more accurate firewall operation without the need to open large numbers of service ports to enable SCADA applications.

Commercially available firewalls are not designed for the large-scale complexity of our national infrastructure networks.

The reader might be amused to consider the irony presented today by network connectivity and security separation. Twenty years ago, the central problem in computer networking involved the rampant interoperability that existed between systems. Making two computers connect over a network was a significant challenge, one that computer scientists worked hard to overcome. In some instances, large projects would be initiated with the goal of connecting systems together over networks. Amazingly, the challenge we deal with today is not one of connectivity, but rather one of separation. This comes from the ubiquity of the Internet Protocol (IP), which enables almost every system on the planet to be connected with trivial effort. Thus, where previously we did not know how to interconnect systems, today we don't know how to separate them!

Now that we are able to connect systems with ease, we must learn to separate them for protection!

What Is Separation?

In the context of national infrastructure protection, separation is viewed as a technique that accomplishes one of the following security objectives:

- *Adversary separation*—The first separation goal involves separating an asset from an adversary to reduce the risk of direct

attack. Whatever implementation is chosen should result in the intruder having no direct means for accessing national assets.

- *Component distribution*—The second separation goal involves architecturally separating components in an infrastructure to distribute the risk of compromise. The idea here is that a compromise in one area of infrastructure should not be allowed to propagate directly.

The access restrictions that result from either of these separation approaches can be achieved through functional or physical means. Functional means involve software, computers, and networks, whereas physical means include tangible separations such as locks, safes, and cabinets. In practice, most separation access restrictions must be designed to focus on either the insider or outsider threat. The relationship between these different separation options can be examined based on the three primary factors involved in the use of separation for protecting infrastructure (see box).

A Working Taxonomy of Separation Techniques

The three primary factors involved in the use of separation for protecting infrastructure include the source of the *threat* (insider or outsider), the *target* of the security control (adversary or asset), and the *approach* used in the security control (functional or physical). We can thus use these three factors to create a separation taxonomy that might help to compare and contrast the various options for separating infrastructure from adversaries (see Figure 3.2).

The first column in the taxonomy shows that separation controls are focused on keeping either insiders or outsiders away from some asset. The key difference here is that insiders would typically be more trusted and would have more opportunity to gain special types of access. The second column indicates that the separation controls are focused on either keeping an adversary away from some asset or inherently separating components of the actual asset, perhaps through distribution. The third column identifies whether the separation approach uses computing functionality or would rely instead on some tangible, physical control.

From the first two rows of the taxonomy, it should be clear that internal access controls demonstrate a functional means for separating insider adversaries from an asset, whereas Internet firewalls achieve roughly the same end for outside adversaries. These firewalls might be traditional devices, as one might find in an enterprise, or special filtering devices placed in the network to throttle volume attacks. The third and fourth rows show that logical separation of an application is a good way to complicate an insider attack; this is comparably done for outsiders by distributing the application across different Internet-facing hosts. The last four rows in Figure 3.2 demonstrate different ways to use physical means to protect infrastructure, ranging from keeping projects and people separate from an asset to maintaining diversity and distribution of infrastructure assets. The following sections provide more detail on these separation taxonomy elements.

Threat	Target	Approach	Example	
Insider	Adversary	Functional	Internal access control	Functional Adversary Techniques
Outsider	Adversary	Functional	Internet-facing firewall	
Insider	Asset	Functional	Application separation	Functional Asset Techniques
Outsider	Asset	Functional	Application distribution	
Insider	Adversary	Physical	Project compartmentalization	Physical Adversary and Asset Techniques
Outsider	Adversary	Physical	Information classification	
Insider	Asset	Physical	Internal network diversity	
Outsider	Asset	Physical	Physical host distribution	

Figure 3.2 Taxonomy of separation techniques.

Functional Separation

Functional separation of an adversary from any computing asset is most commonly achieved using an access control mechanism with the requisite authentication and identity management. Access controls define which users can perform which actions on which entities. The access rules should be predetermined in a security policy. They should specify, for example, which users can access a given application, and, obviously, the validation of user identity must be accurate. In some cases, security policy rules must be more dynamic, as in whether a new type of traffic stream is allowed to proceed to some Internet ingress point. This might be determined by real-time analysis of the network flow.

An access policy thus emerges for every organization that identifies desired allowances for users requesting to perform actions on system entities. Firewall policies are the most common example of this; for example, users trying to connect to a web server might be subjected to an access control policy that would determine if this was to be permitted. Similarly, the IP addresses of some organization might be keyed into a firewall rule to allow access to some designated system. A major problem that occurs in practice with firewalls is that the rule base can grow to an enormous size, with perhaps thousands of rules. The result is complexity and a high potential for error. National infrastructure initiatives must identify rewards and incentives for organizations to keep their firewall rule bases as small

as possible. Some organizations have used optimization tools for this purpose, and this practice should be encouraged for national assets.

Two broad categories of security can be followed when trying to achieve functional separation of adversaries from any type of national infrastructure assets. The first involves distributing the responsibility for access mediation to the owners of smaller asset components such as individual computers or small networks; the second involves deployment of a large, centralized mediation mechanism through which all access control decisions would be made (see Figure 3.3).

In large networks, firewall rules can become so numerous that they actually increase the margin for error.

The distributed approach has had considerable appeal for the global Internet community to date. It avoids the problem of having to trust a large entity with mediation decisions, it allows for commercial entities to market their security tools on a large scale to end users, and it places control of access policy close to the asset, which presumably should increase the likelihood that the policy is appropriate. The massive global distribution of computer security responsibility to every owner of a home personal computer is an example of this approach. End users must decide how to protect their assets, rather than relying on some centralized authority.

Unfortunately, in practice, the distributed approach has led to poor results. Most end users are unqualified to make good decisions about security, and even if a large percentage make excellent decisions, the ones who do not create a big enough vulnerability as to place the entire scheme at risk. Botnets, for example, prey on poorly managed end-user computers on broadband connections. When a home computer is infected with malware, there really is no centralized authority for performing a cleansing function. This lack of centralization on the Internet thus results in a huge security risk. Obviously, the Internet will never be redesigned to include centralized control; that would be impractical, if not impossible.

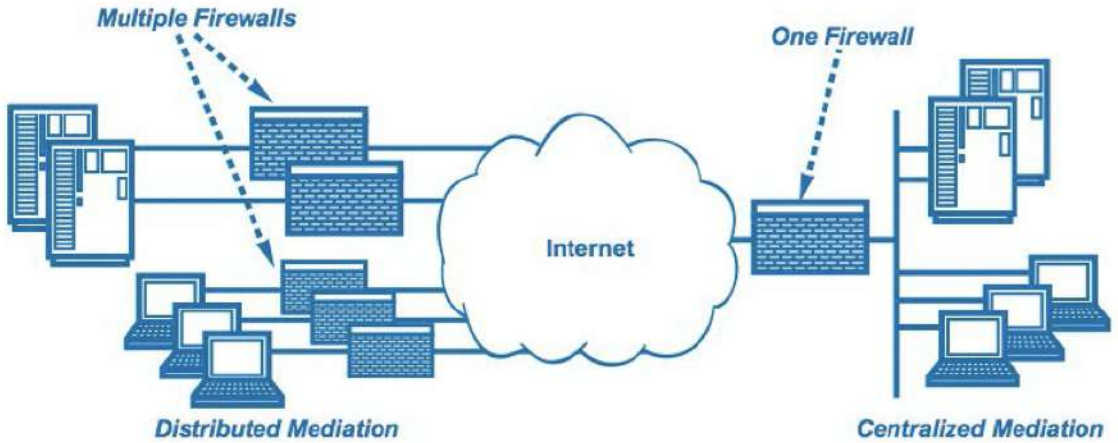


Figure 3.3 Distributed versus centralized mediation.

For national infrastructure, however, the possibility does exist for more centralized control. The belief here is that an increased reliance on centralized protection, especially in conjunction with the network service provider, will improve overall national asset protection methods. This does not imply, however, that distributed protection is not necessary. In fact, in most environments, skilled placement of both centralized and distributed security will be required to avoid national infrastructure attack.

Centralized control versus multiple, independent firewalls—both have their advantages, so which is best for national infrastructure?

National Infrastructure Firewalls

The most common application of a firewall involves its placement between a system or enterprise to be protected and some untrusted network such as the Internet. In such an arrangement for the protection of a national asset, the following two possibilities immediately arise:

- *Coverage*—The firewall might not cover all paths between the national asset to be protected and the untrusted network such as the Internet. This is a likely case given the general complexity associated with most national infrastructure.
- *Accuracy*—The firewall might be forced to allow access to the national asset in a manner that also provides inadvertent, unauthorized access to certain protected assets. This is common in large-scale settings, especially because specialized protocols such as those in SCADA systems are rarely supported by commercial firewalls. As a result, the firewall operator must compensate by leaving certain ports wide open for ingress traffic.

To address these challenges, the design of national security infrastructure requires a skillful placement of separation functionality to ensure that all relevant traffic is mediated and that no side effects occur when access is granted to a specific asset. The two most effective techniques include aggregation of protections in the wide area network and segregation of protections in the local area network (see Figure 3.4).

Aggregating firewall functionality at a defined gateway is not unfamiliar to enterprise security managers. It helps ensure coverage of untrusted connections in more complex environments. It also provides a means for focusing the best resources, tools, and staff to one aggregated security complex. Segregation in a local area network is also familiar, albeit perhaps less practiced. It is effective in reducing the likelihood that external access to System A has the side effect of providing external access to System B. It requires management of more devices and does generally imply higher cost. Nevertheless, both of these techniques will be important in national infrastructure firewall placement.

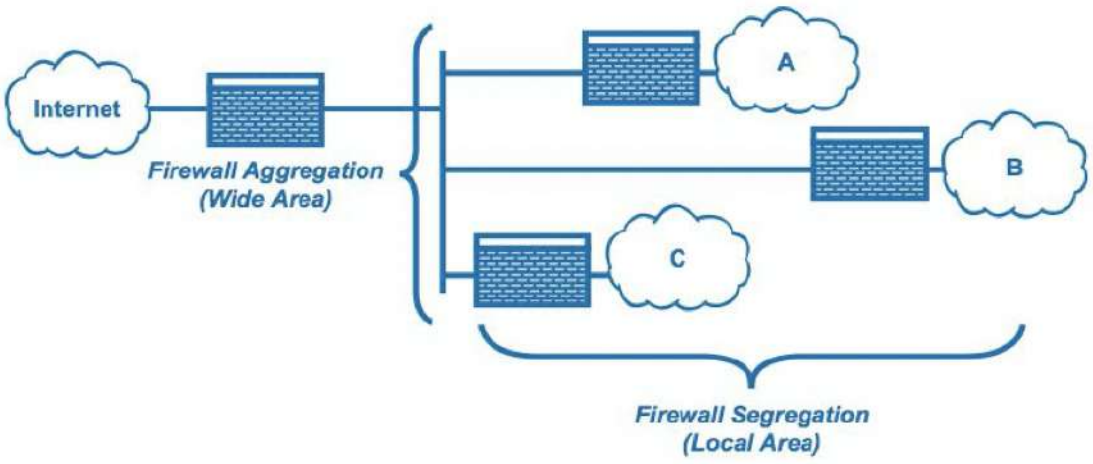


Figure 3.4 Wide area firewall aggregation and local area firewall segregation.

Effective protection of national infrastructure will undoubtedly be expensive due to the increased management of devices.

Smart devices have added another layer of complexity to network protection.

A major challenge to national infrastructure comes with the massive increase in wireless connectivity that must be presumed for all national assets in the coming years. Most enterprise workers now carry around some sort of smart device that is ubiquitously connected to the Internet. Such smart devices have begun to resemble computers in that they can support browsing, e-mail access, and even virtual private network (VPN) access to applications that might reside behind a firewall. As such, the ease with which components of infrastructure can easily bypass defined firewall gateways will increase substantially. The result of this increased wireless connectivity, perhaps via 4G deployment, will be that all components of infrastructure will require some sort of common means for ensuring security.

Massive distribution of security to smart wireless endpoint devices may not be the best option, for all the reasons previously cited. It would require massive distribution, again, of the security responsibility to all owners of smart devices. It also requires vigilance on the part of every smart device owner, and this is not a reasonable expectation. An alternative approach involves identifying a common transport infrastructure to enforce desired policy. This might best be accomplished via the network transport carrier. Network service providers offer several advantages with regard to centralized security:

- *Vantage point*—The network service provider has a wide vantage point that includes all customers, peering points, and gateways. Thus, if some incident is occurring on the Internet, the service provider will observe its effects.
- *Operations*—Network service providers possess the operational capability to ensure up-to-date coverage of signatures, updates, and new security methods, in contrast to the inability of most end users to keep their security software current.

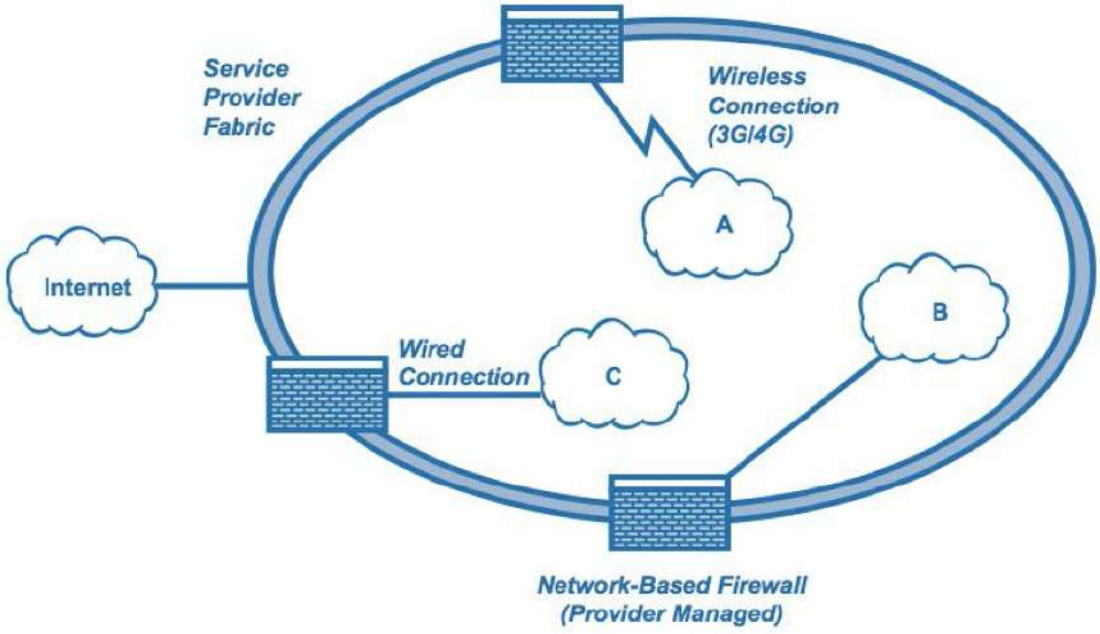


Figure 3.5 Carrier-centric network-based firewall.

- *Investment*—Where most end users, including enterprise groups, are unlikely to have funds sufficient to install multiple types of diverse or even redundant security tools, service providers can often support a business case for such investment.

For these reasons, a future view of firewall functionality for national infrastructure will probably include a new aggregation point—namely, the concept of implementing a network-based firewall in the cloud (see Figure 3.5).

A firewall in the cloud may be the future of firewall functionality.

In the protection of national infrastructure, the use of network-based firewalls that are embedded in service provider fabric will require a new partnership between carriers and end-user groups. Unfortunately, most current telecommunications service level agreements (SLAs) are not compatible with this notion, focusing instead on packet loss and latency issues, rather than policy enforcement. This results in too many current cases of a national infrastructure provider being attacked, with the service provider offering little or no support during the incident. Obviously, this situation must change for the protection of national assets.

DDOS Filtering

A major application of the network-based firewall concept includes a special type of mediation device embedded in the wide area network for the purpose of throttling distributed denial of service (DDOS) attacks. This device, which can be crudely

referred to as a *DDOS filter*, is essential in modern networking, given the magnified risk of DDOS attacks from botnets. Trying to filter DDOS attacks at the enterprise edge does not make sense given the physics of network ingress capacity. If, for example, an enterprise has a 1-Gbps ingress connection from the Internet, then a botnet directing an inbound volume of anything greater than 1 Gbps will overwhelm the connection.

The risk of DDOS attacks must be effectively addressed.

The solution to this volume problem is to move the filtering upstream into the network. Carrier infrastructure generally provides the best available option here. The way the filtering would work is that volumetric increases in ingress traffic would cause a real-time redirection of traffic to a DDOS filtering complex charged with removing botnet-originating traffic from valid traffic. Algorithms for performing such filtering generally key on the type of traffic being sent, the relative size of the traffic, and any other hint that might point to the traffic being of an attack nature. Once the traffic has been filtered, it is then funneled to the proper ingress point. The result is like a large safety valve or shock absorber in the wide area network that turns on when an attack is under way toward some target enterprise (see Figure 3.6).

Moving the filtering functionality into the network will allow legitimate traffic to pass through and the discovery of potential DDOS attacks.

Quantitative analysis associated with DDOS protection of national infrastructure is troubling. If, for example, we assume that bots can easily steal 500 Kbps of broadband egress from the unknowing infected computer owner, then it would only require three bots to overwhelm a T1 (1.5-Mbps) connection. If one carries out this argument, then botnets with 16,000 bots are sufficient to overwhelm a 10-Gbps connection. Given the existence of prominent botnets such as Storm and Conficker, which some experts

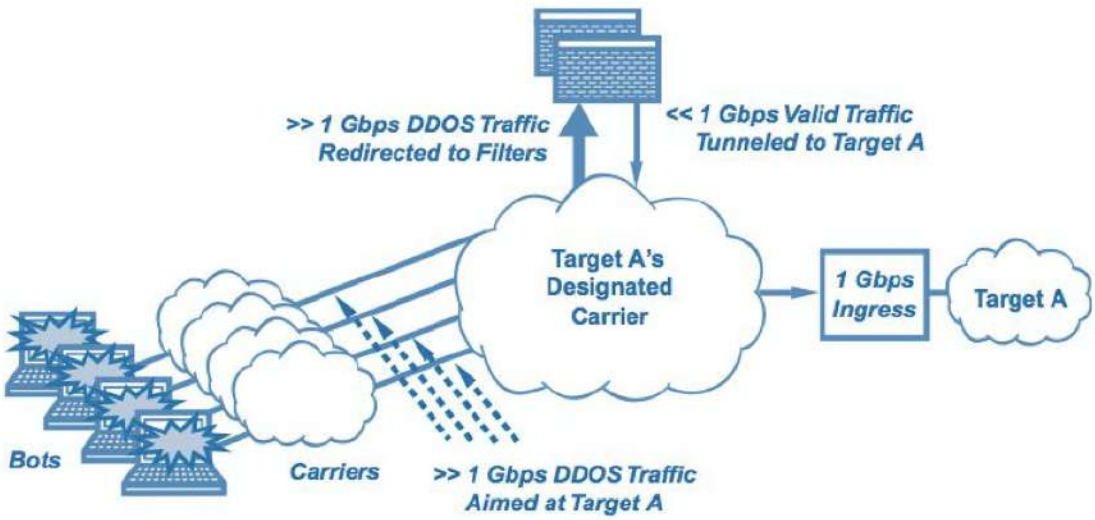


Figure 3.6 DDOS filtering of inbound attacks on target assets.

suggest could have as many as 2 or 3 million bots, the urgency associated with putting DDOS filtering in place cannot be understated. An implication is that national infrastructure protection initiatives must include some measure of DDOS filtering to reduce the risk of DDOS attacks on national assets.

A serious problem that must be addressed, however, in current DDOS attacks on infrastructure involves a so-called *amplification* approach. Modern DDOS attacks are generally designed in recognition of the fact that DDOS filters exist to detect large inbound streams of unusual traffic. Thus, to avoid inbound filtering in carrier infrastructure, adversaries have begun to follow two design heuristics. First, they design DDOS traffic to mimic normal system behavior, often creating transactions that look perfectly valid. Second, they design their attack to include small inbound traffic that utilizes some unique aspect of the target software to create larger outbound responses. The result is a smaller, less obvious inbound stream which then produces much larger outbound response traffic that can cause the DDOS condition.

Modern DDOS attacks take into account a more advanced filtering system and thus design the DDOS traffic accordingly.

The Great Challenge of Filtering Out DDOS Attacks

The great challenge regarding current DDOS attacks is that the only way to avoid the sort of problem mentioned in the text is through nontrivial changes in target infrastructure. Two of these nontrivial changes are important to mention here:

1. Stronger authentication of inbound inquiries and transactions from users is imperative. This is not desirable for e-commerce sites designed to attract users from the Internet and also designed to minimize any procedures that might scare away customers.
2. To minimize the amplification effects of some target system, great care must go into analyzing the behavior of Internet-visible applications to determine if small inquiries can produce much larger responses. This is particularly important for public shared services such as the domain name system, which is quite vulnerable to amplification attacks.

These types of technical considerations *must* be included in modern national infrastructure protection initiatives.

SCADA Separation Architecture

Many critical national infrastructure systems include supervisory control and data acquisition (SCADA) functionality. These systems can be viewed as the set of software, computers, and networks that provide remote coordination of controls systems for tangible infrastructures such as power generation systems, chemical plants, manufacturing equipment, and transportation

systems. The general structure of SCADA systems includes the following components:

- *Human-machine interface (HMI)*—The interface between the human operator and the commands relevant to the SCADA system
- *Master terminal unit (MTU)*—The client system that gathers data locally and transmits it to the remote terminal unit
- *Remote terminal unit (RTU)*—The server that gathers data remotely and sends control signals to field control systems
- *Field control systems*—Systems that have a direct interface to field data elements such as sensors, pumps, and switches

The primary security separation issue in a SCADA system architecture is that remote access from an MTU to a given RTU must be properly mediated according to a strong access control policy.[2] The use of firewalls between MTUs and RTUs is thus imperative in any SCADA system architecture. This separation must also enforce policy from any type of untrusted network, such as the Internet, into the RTUs. If this type of protection is not present, then the obvious risk emerges that an adversary can remotely access and change or influence the operation of a field control system.

Remote access from MTUs to RTUs opens the door for adversaries to take advantage of this separation.

As one might expect, all the drawbacks associated with large-scale firewall deployment are also present in SCADA systems. Coverage and accuracy issues must be considered, as well as the likelihood that individual components have direct or wireless connections to the Internet through unknown or unapproved channels. This implies that protection of RTUs from unauthorized access will require a combination of segregated local area firewalls, aggregated enterprise-wide firewalls, and carrier-hosted network-based firewalls (see Figure 3.7).

The biggest issue for SCADA separation security is that most of the associated electromechanical systems were designed and evolved in an environment largely separate from conventional computing and networking. Few computing texts explain the subtle details in SCADA system architecture; in fact, computer scientists can easily complete an advanced program of study without the slightest exposure to SCADA issues. Thus, in far too many SCADA environments, the computerized connections between tangible systems and their control networks have occurred in an *ad hoc* manner, often as a result of establishing local convenience such as remote access. For this reason, the likelihood is generally low that state-of-the-art protection mechanisms are in place to protect a given SCADA system from cyber attack.

Protection mechanisms must be updated to effectively protect a SCADA system from cyber attack.

[2] R. Krutz, *Securing SCADA Systems*, John Wiley & Sons, New York, 2006.

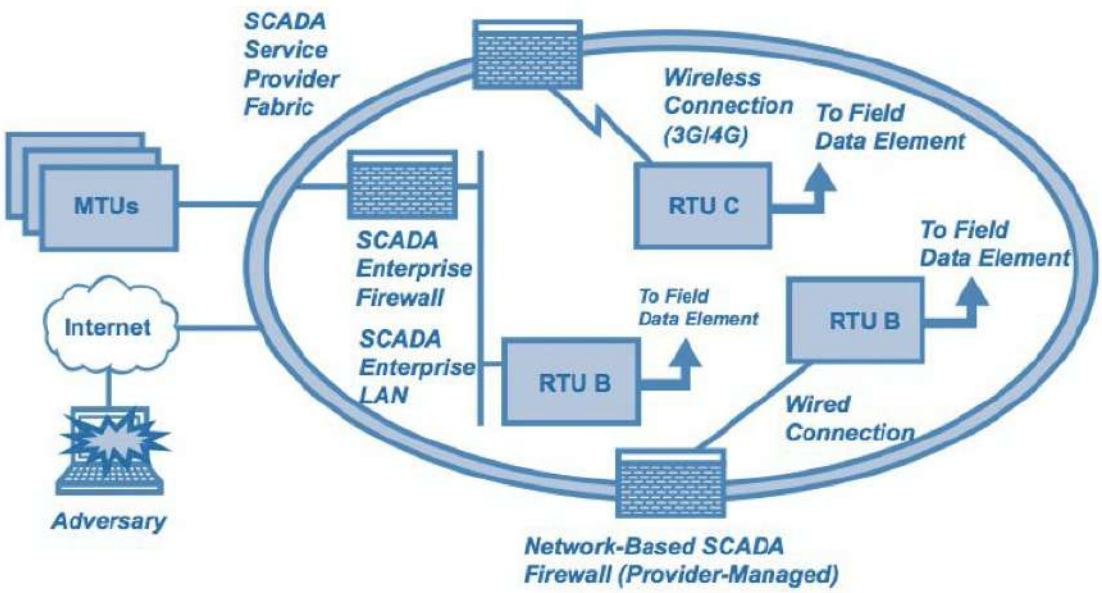


Figure 3.7 Recommended SCADA system firewall architecture.

An additional problem that emerges for SCADA firewall usage is that commercial firewalls do not generally support SCADA protocols. When this occurs, the firewall operator must examine which types of ports are required for usage of the protocol, and these would have to be opened. Security experts have long known that one of the great vulnerabilities in a network is the inadvertent opening of ports that can be attacked. Obviously, national infrastructure protection initiatives must be considered that would encourage and enable new types of firewall functionality such as special proxies that could be embedded in SCADA architecture to improve immediate functionality.

Opening ports, although necessary, is a risky endeavor, as it subjects the SCADA system to increased vulnerabilities.

Physical Separation

One separation technique that is seemingly obvious, but amazingly underrepresented in the computer security literature, is the physical isolation of one network from another. On the surface, one would expect that nothing could be simpler for separating one network from any untrusted environment than just unplugging all external connections. The process is known as *air gapping*, and it has the great advantage of not requiring any special equipment, software, or systems. It can be done to separate enterprise networks from the Internet or components of an enterprise network from each other.

Air gapping allows for physical separation of the network from untrusted environments.

As a company grows, physical separation as a protection feature becomes increasingly complex.

The problem with physical separation as a security technique is that as complexity increases in some system or network to be isolated, so does the likelihood that some unknown or unauthorized external connection will arise. For example, a small company with a modest local area network can generally enjoy high confidence that external connections to the Internet are well known and properly protected. As the company grows, however, and establishes branch offices with diverse equipment, people, and needs, the likelihood that some generally unrecognized external connectivity will arise is high. Physical separation of network thus becomes more difficult.

So how does one go about creating a truly air-gapped network? The answer lies in the following basic principles:

- *Clear policy*—If a network is to be physically isolated, then clear policy must be established around what is and what is not considered an acceptable network connection. Organizations would thus need to establish policy checks as part of the network connection provision process.
- *Boundary scanning*—Isolated networks, by definition, must have some sort of identifiable boundary. Although this can certainly be complicated by firewalls embedded in the isolated network, a program of boundary scanning will help to identify leaks.
- *Violation consequences*—If violations occur, clear consequences should be established. Government networks in the U.S. military and intelligence communities, such as SIPRNet and Intelink, are protected by laws governing how individuals must use these classified networks. The consequences of violation are not pleasant.
- *Reasonable alternatives*—Leaks generally occur in an isolated network because someone needs to establish some sort of communication with an external environment. If a network connection is not a reasonable means to achieve this goal, then the organization must provide or support a reasonable work-around alternative.

Dual-homing creates another area of vulnerability for enterprise networks.

Perhaps the biggest threat to physical network isolation involves dual-homing a system to both an enterprise network and some external network such as the Internet. Such dual-homing can easily arise where an end user utilizes the same system to access both the isolated network and the Internet. As laptops have begun to include native 3 G wireless access, this likelihood of dual-homing increases. Regardless of the method, if any sort of connectivity is enabled simultaneously to both systems, then the end user creates an inadvertent bridge (see Figure 3.8).

It is worth mentioning that the bridge referenced above does not necessarily have to be established simultaneously. If a system connects to one network and is infected with some sort of

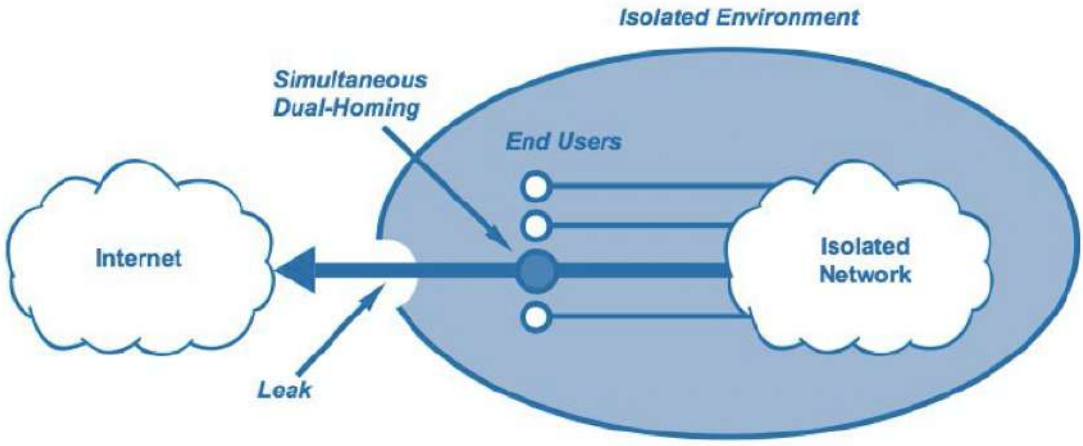


Figure 3.8 Bridging an isolated network via a dual-homing user.

malware, then this can be spread to another network upon subsequent connectivity. For this reason, laptops and other mobile computing devices need to include some sort of native protection to minimize this problem. Unfortunately, the current state of the art for preventing malware downloads is poor.

A familiar technique for avoiding bridges between networks involves imposing strict policy on end-user devices that can be used to access an isolated system. This might involve preventing certain laptops, PCs, and mobile devices from being connected to the Internet; instead, they would exist solely for isolated network usage. This certainly reduces risk, but is an expensive and cumbersome alternative. The advice here is that for critical systems, especially those involving safety and life-critical applications, if such segregation is feasible then it is probably worth the additional expense. In any event, additional research in multimode systems that ensure avoidance of dual-homing between networks is imperative and recommended for national infrastructure protection.

Imposing strict policies regarding connection of laptops, PCs, and mobile devices to a network is both cumbersome and expensive but necessary.

Insider Separation

The insider threat in national infrastructure protection is especially tough to address because it is relatively easy for determined adversaries to obtain trusted positions in groups with responsibility for national assets. This threat has become even more difficult to counter as companies continue to partner, purchase, and outsource across political boundaries. Thus, the ease with which an adversary in one country can gain access to the internal, trusted infrastructure systems of another country is both growing and troubling.

An adversarial threat may come from a trusted partner.

Traditionally, governments have dealt with this challenge through strict requirements on background checking of any

The commercially run components of our national infrastructure do not have the same stringent personnel requirements as the government-run components.

individuals who require access to sensitive government systems. This practice continues in many government procurement settings, especially ones involving military or intelligence information. The problem is that national infrastructure includes so much more than just sensitive government systems. It includes SCADA systems, telecommunications networks, transportation infrastructure, financial networks, and the like. Rarely, if ever, are requirements embedded in these commercial environments to ensure some sort of insider controls against unauthorized data collection, inappropriate access to customer records, or administrative access to critical applications. Instead, it is typical for employees to be granted access to the corporate Intranet, from which virtually anything can be obtained.

Techniques for reducing the risk of unauthorized insider access do exist that can be embedded in the design and operation of national infrastructure operation. These techniques include the following:

- *Internal firewalls*—Internal firewalls separating components of national assets can reduce the risk of insider access. Insiders with access to component A, for example, would have to successfully negotiate through a firewall to gain access to component B. Almost every method for separating insiders from assets will include some sort of internal firewall. They can be implemented as fully configured firewalls, or as packet filtering routers; but regardless, the method of separating insiders from assets using firewalls must become a pervasive control in national infrastructure.
- *Deceptive honey pots*—As we discussed in Chapter 2, internal honey pots can help identify malicious insiders. If the deception is openly advertised, then malicious insiders might be more uncertain in their sabotage activity; if the deception is stealth, however, then operators might observe malicious behavior and potentially identify the internal source.
- *Enforcement of data markings*—Many organizations with responsibility for national infrastructure do not properly mark their information. Every company and government agency must identify, define, and enforce clearly visible data markings on all information that could be mishandled. Without such markings, the likelihood of proprietary information being made available inadvertently to adversaries increases substantially. Some companies have recently begun to use new data markings for personally identifiable information (PII).
- *Data leakage protection (DLP) systems*—Techniques for sniffing gateway traffic for sensitive or inappropriate materials are becoming common. Tools called DLP systems are routinely deployed in companies and agencies. At best, they provide weak

protection against insider threats, but they do help identify erroneous leaks. Once deployed, they provide statistics on where and how insiders might be using corporate systems to spill information. In practice, however, no knowledgeable insider would ever be caught by a data leakage tool. Instead, the leak would be done using non-company-provided computers and networks.

One of the more effective controls against insider threats involves a procedural practice that can be embedded into virtually every operation of an organization. The technique is known as *segregation of duties*, and it should be familiar to anyone who has dealt with Sarbanes-Oxley requirements in the United States. Security researchers will recognize the related *separation of duties* notion introduced in the Clark-Wilson integrity model. In both cases, critical work functions are decomposed so that work completion requires multiple individuals to be involved. For example, if a financial task requires two different types of activities for completion, then a segregation of duties requirement would ensure that no one individual could ever perform both operations.

Segregation of duties offers another layer of protection.

The purpose of this should be obvious. By ensuring that multiple individuals are involved in some sensitive or critical task, the possibility of a single insider committing sabotage is greatly reduced. Of course, multiple individuals could still collude to create an internal attack, but this is more difficult and less likely in most cases. If desired, the risk of multiple individuals creating sabotage can be reduced by more complex segregation of duty policies, perhaps supported by the use of security architectural controls, probably based on internally positioned firewalls. In fact, for network-based segregation tasks, the use of internal firewalls is the most straightforward implementation.

Internal firewalls create a straightforward *de facto* separation of duties.

In general, the concept of segregation of duties can be represented via a work function ABC that is performed either by a single operator A or as a series of work segments by multiple operators. This general schema supports most instances of segregation of duties, regardless of the motivation or implementation details (see Figure 3.9).

The idea of breaking down work functions into components is certainly not new. Managers have decomposed functions into smaller tasks for many years; this is how assembly lines originated. Unfortunately, most efforts at work function decomposition result in increased bureaucracy and decreased worker (and end-user) satisfaction. The stereotyped image arises of the government bureau where customers must stand in line at this desk for this function and then stand in line at that desk for that function, and so on. The process is clearly infuriating but, ironically, is also difficult to sabotage by a malicious insider.

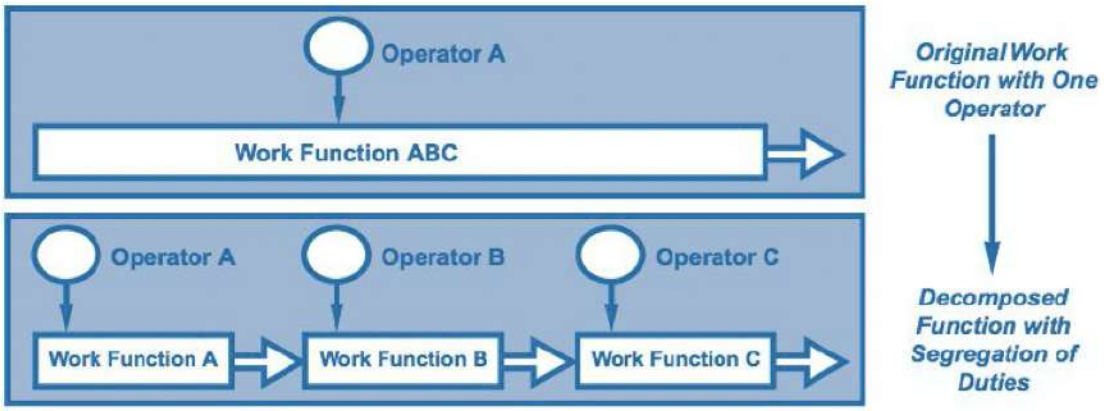


Figure 3.9 Decomposing work functions for segregation of duty.

How to effectively separate duties without increasing the unwieldy bureaucracy is a challenge that must be addressed.

The challenge for national infrastructure protection is to integrate segregation of duty policies into all aspects of critical asset management and operation, but to do so in a manner that minimizes the increased bureaucracy. This will be especially difficult in government organizations where the local culture always tends to nurture and embrace new bureaucratic processes.

Asset Separation

Asset separation involves the distribution, replication, decomposition, or segregation of national assets to reduce the risk of an isolated compromise. Each of these separation techniques can be described as follows:

- *Distribution* involves creating functionality using multiple cooperating components that work together as a distributed system. The security advantage is that if the distributed system is designed properly then one or more of the components can be compromised without breaking the overall system function.
- *Replication* involves copying assets across disparate components so that if one asset is broken then replicated versions will continue to be available. Database systems have been protected in this way for many years. Obviously, no national asset should exist without a degree of replication to reduce risk.
- *Decomposition* is the breaking down of complex assets into individual components so that isolated compromise of a component will be less likely to break the overall asset. A common implementation of a complex business process, for example, generally includes some degree of decomposition into smaller parts.
- *Segregation* is the logical separation of assets through special access controls, data markings, and policy enforcement. Operating systems, unfortunately, provide weak controls in this regard, largely because of the massive deployment of single-user

machines over the past couple of decades. Organizations thus implement logical separation of data by trying to keep it on different PCs and laptops. This is a weak implementation.

Segregation is one method of separation.

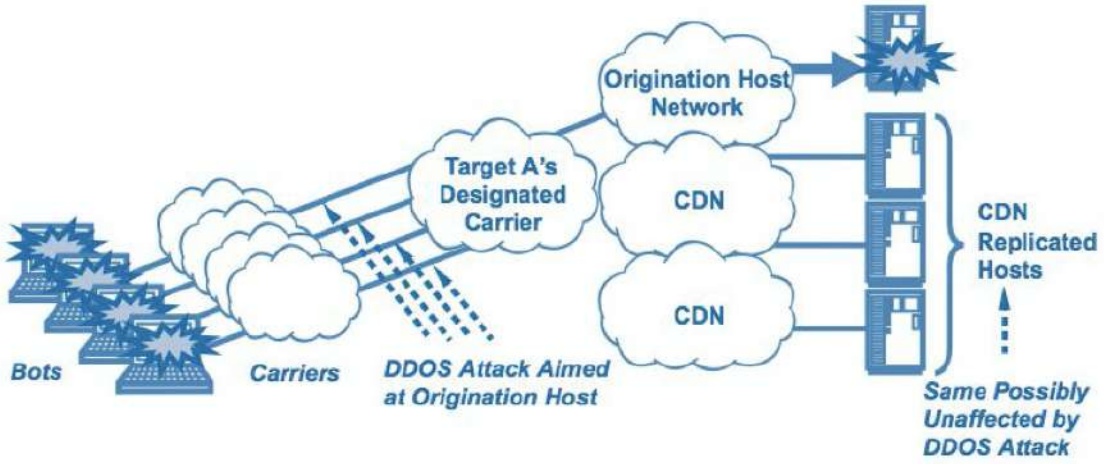
Each of these techniques is common in modern infrastructure management. For example, content distribution networks (CDNs) are rarely cited as having a positive impact on national infrastructure security, but the reality is that the distribution and replication inherent in CDNs for hosting are powerful techniques for reducing risk. DDOS attacks, for example, are more difficult to complete against CDN-hosted content than for content resident only on an origination host. Attackers have a more difficult time targeting a single point of failure in a CDN (see Figure 3.10).

It is important to emphasize that the use of a CDN certainly does not ensure protection against a DDOS attack, but the replication and distribution inherent in a CDN will make the attack more difficult. By having the domain name system (DNS) point to CDN-distributed assets, the content naturally becomes more robust. National infrastructure designers and operators are thus obliged to ensure that CDN hosting is at least considered for all critically important content, especially multimedia content (streaming and progressive download) and any type of critical software download.

This is becoming more important as multimedia provision becomes more commonly embedded into national assets. In the recent past, the idea of providing video over the Internet was nothing more than a trivial curiosity. Obviously, the massive proliferation of video content on sites such as YouTube.com has made these services more mainstream. National assets that rely on video should thus utilize CDN services to increase their robustness. Additional DDOS protection of content from the backbone service provider would also be recommended.

The increase in multimedia components within national infrastructure networks argues for increased reliance on CDN services.

Figure 3.10 Reducing DDOS risk through CDN-hosted content.



Multilevel Security (MLS)

A technique for logical separation of assets that was popular in the computer security community during the 1980s and 1990s is known as multilevel security (MLS). MLS operating systems and applications were marketed aggressively to the security community during that time period. A typical implementation involved embedding mandatory access controls and audit trail hooks into the underlying operating system kernel. Assurance methods would then be used to ensure that the trusted component of the kernel was correct, or at least as correct as could be reasonably verified. Today, for reasons largely economic, MLS systems are no longer available, except in the most esoteric classified government applications.

The familiar notion of "top-secret clearance" comes from MLS systems.

The idea behind MLS was that, by labeling the files and directories of a computer system with meaningful classifications and by also labeling the users of that system with meaningful clearances, a familiar security policy could be enforced. This scheme, which was motivated largely by paper methods used to protect information in government, produced a logical separation of certain assets from certain users, based on the existing policy. For example, files marked "secret" could only be read by users with sufficient clearances. Similarly, users not cleared to the level of "top secret" would not be allowed to read files that were so labeled. The result was an enforced policy on requesting users and protected assets (see Figure 3.11).

Several models of computer system behavior with such MLS functionality were developed in the early years of computer security. The Bell-La Padula disclosure and Biba integrity models are prominent examples. Each of these models stipulated policy rules that, if followed, would help to ensure certain desirable security

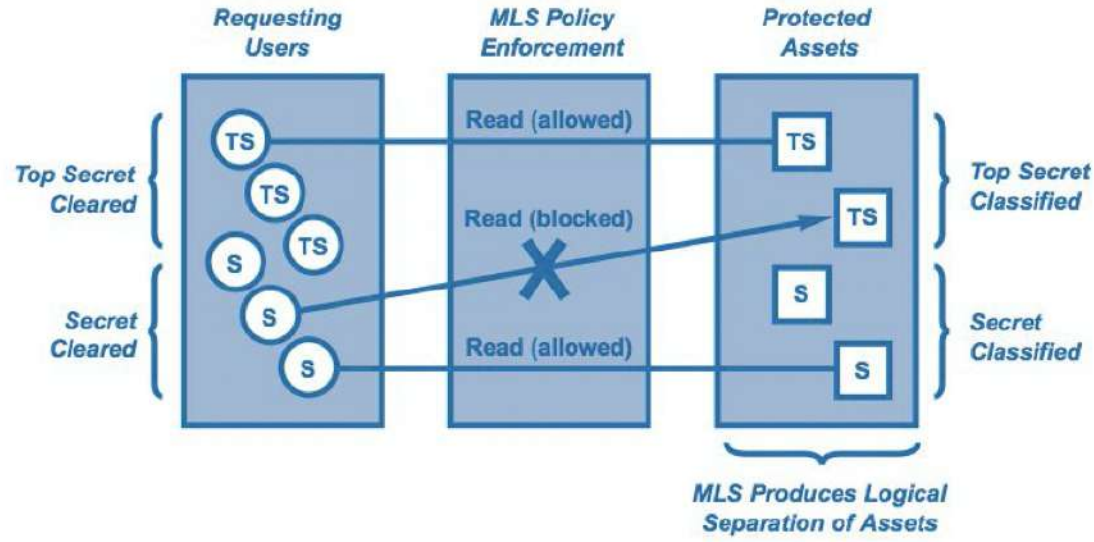


Figure 3.11 Using MLS logical separation to protect assets.

properties. Certainly, there were problems, especially as networking was added to isolated secure systems, but, unfortunately, most research and development in MLS dissolved mysteriously in the mid-1990s, perhaps as a result of the economic pull of the World Wide Web. This is unfortunate, because the functionality inherent in such MLS separation models would be valuable in today's national infrastructure landscape. A renewed interest in MLS systems is thus strongly encouraged to improve protection of any nation's assets.

MLS systems seem to have gone by the wayside but should be revived as another weapon in the national infrastructure protection arsenal.

Obviously, once a national program is in place, consideration of how one might separate assets between different cooperating nations would seem a logical extension. Certainly, this would seem a more distant goal given the complexity and difficulty of creating validated policy enforcement in one nation.

Implementing a National Separation Program

Implementation of a national separation program would involve verification and validation of certain design goals in government agencies and companies with responsibility for national infrastructure. These goals, related to policy enforcement between requesting users and the protected national assets, would include the following:

- *Internet separation*—Certain critical national assets simply should not be accessible from the Internet. One would imagine that the control systems for a nuclear power plant, for example, would be good candidates for separation from the Internet. Formal national programs validating such separation would be a good idea. If this requires changes in business practice, then assistance and guidance would be required to transition from open, Internet connectivity to something more private.
- *Network-based firewalls*—National infrastructure systems should be encouraged to utilize network-based firewalls, preferably ones managed by a centralized group. The likelihood is higher in such settings that signatures will be kept up to date and that security systems will be operated properly on a 24/7 basis. Procurement programs in government, in particular, must begin to routinely include the use of network-based security in any contract with an Internet service provider.
- *DDOS protection*—All networks associated with national assets should have a form of DDOS protection arranged before an attack occurs. This protection should be provided on a high-capacity backbone that will raise the bar for attackers contemplating a capacity-based cyber attack. If some organization, such as a government agency, does not have a suitable DDOS protection scheme, this should be likened to having no disaster recovery program.
- *Internal separation*—Critical national infrastructure settings must have some sort of incentive to implement an internal separation policy to prevent sabotage. The Sarbanes-Oxley requirements in the United States attempted to enforce such separation for financial systems. While the debate continues about whether this was a successful initiative, some sort of program for national infrastructure seems worth considering. Validation would be required that internal firewalls exist to create protection domains around critical assets.
- *Tailoring requirements*—Incentives must be put in place for vendors to consider building tailored systems such as firewalls for specialized SCADA environments. This would greatly reduce the need for security administrators in such settings to configure their networks in an open position.

Finally, let's briefly look at some practical ways to protect the critical national infrastructure through use of separation techniques. Current threats and vulnerabilities are also covered.

Protecting the Critical National Infrastructure Through Use of Separation

No single separation technique is sufficient enough to fully protect the critical national infrastructure networks. A combination of practical separation security measures, working together, is required to provide a strong defense-in-depth protection (see "An Agenda for Action in Using Separation to Protect the Critical National Infrastructure"). These practical separation security measures are as follows:

- Implement real-time threat protection.
- Segment and protect critical national infrastructure assets from interconnected networks.
- Control user access and network activities.
- Protect information about critical national infrastructure assets from data leakage.
- Implement strong security without jeopardizing availability, integrity, and reliability requirements.

An Agenda for Action in Using Separation to Protect the Critical National Infrastructure

When completing the Use of Separation to Protect the Critical National Infrastructure Checklist, the IT administrator should adhere to the provisional list of actions for preparing for contingencies in the event that separation fails. The order is not significant; however, these are the activities for which the research would want to provide a detailed description of procedures, review, and assessment for ease of use and admissibility. Current separation measures that must be adhered to, in order to protect the critical national infrastructure, include (check all tasks completed) the following:

1. Implement real-time threat protection.
2. Separate and protect critical assets from interconnected networks by taking the following actions:
   a. Control port access based on a positive security model (i.e., they deny all access except that which is explicitly allowed).
   b. Operate at gigabit speed and, therefore, do not interfere with control system availability and integrity standards.
   c. Include specific capabilities designed for control systems.
   d. Deliver a truly hardened operating system (not just a modified commercial one) that can defend itself from attacks, prevent or

eliminate root access, and restrict access escalation or arbitrary code execution by any outside party.

   e. Eliminate all unconstrained privileges and extraneous services, including network stack separation and control of super-user privileges, while providing triggers for intrusion detection.
   f. Provide easy-to-deploy and manage architecture with central policies, reporting, and strong forensics.
   g. Automatically filter out connections from locations that are suspicious or unnecessary to normal operations.
   h. Scan encrypted traffic (HTTPS, SSL, SSH, SFPT, SCP, etc.) to uncover and block hidden attacks
   i. Provide strong industry and government certifications and references (Common Criteria certification of EAL4+ is the minimum level suggested).
   j. Deliver a security architecture that has a long and proven history of never being breached or hacked.
3. Provide a suitable intrusion prevention security (IPS) solution by taking the following actions:
   a. Provide real-time protection from known, zero-day, DoS, DDoS, SYN flood, and encrypted attacks, as well as threats such as spyware, VoIP vulnerabilities, botnets, malware, worms, Trojans, phishing, and peer-to-peer tunneling.
   b. Maximize accuracy by using multiple advanced detection methods, including signature, application, and protocol anomaly; shell-code detection algorithms; and next-generation DoS and DDoS prevention.
   c. Parse over 200 protocols and review over 6,000 high-quality, multitoken, multitrigger signatures with stateful traffic inspection.
   d. Offer proactive, out-of-the-box blocking for hundreds of attacks by featuring preconfigured "recommended for blocking" policies.
   e. Receive continuous threat updates 24/7 from global research teams.
4. Control user access and network activities.
5. Protect information about critical assets from data leakage by taking the following actions:
   a. Drop
   b. Blind copy
   c. Replace
   d. Drop a portion or even the entire message
   e. Forward in line or as an attachment
   f. Quarantine
   g. Reroute
   h. Prepend
   i. Log
   j. Encrypt for secure delivery
   k. Rewrite the subject line
   l. Notify employees, managers, compliance officers, etc.
   m. Archive
   n. Educate users on rules

6. Implement strong security without jeopardizing availability, integrity, and reliability requirements by taking the following actions:
   a. Perform automatic updates that don't require critical assets be taken off line
   b. Support the long asset lifetimes of critical assets
   c. Minimize the need for extensive testing and downtime before patches can be applied
   d. Protect against threats that have yet to be identified
   e. Prevent privilege escalation vulnerabilities
   f. Support the custom and relevant signatures specific to critical networks
   g. Perform security at speeds that won't impact network performance
   h. Deploy a trusted security model based on reputation and an in-depth understanding of applications

Summary

This chapter focused on practical ways to use separation to protect the world's critical cyber national infrastructure and offered insights into current threats and vulnerabilities. It brought home the fact that critical asset security or critical national infrastructure protection are the vital networks and systems' practical measures that are relied on to control electricity, water, oil and gas, public transportation systems, and manufacturing. These critical assets have been separated from the rest of the computing world. This separation means that anyone in charge of critical assets has to worry about cyber attacks.

Furthermore, the rapid rise of the Internet and the spread of inexpensive bandwidth have put critical systems in jeopardy. The vast majority of these critical systems are interconnected with IT systems and accessed by remote users via wireless devices. These critical systems are also used by nontrusted operators to provide data mining opportunities for their corporations and tied in to independent systems operators and other third-party networks for multienterprise coordination.

As a result, the security threats that have dogged IT systems for decades can now be spread into the critical national infrastructure systems virtually undetected, which makes them vulnerable to hackers, saboteurs, and cyber criminals located anywhere in the world.

Finally, let's move on to the real interactive part of this chapter: review questions/exercises, hands-on projects, case projects, and optional team case project. The answers and/or solutions by chapter can be found online at http://www.elsevierdirect.com/companion.jsp?ISBN=9780123918550.