

Chapter 6

Defend the Enclave Boundary/ External Connections

An enclave is an environment under the control of a single authority with personnel and physical security measures. Enclaves typically contain multiple local area networks (LAN) with computing resource components such as user platforms; network, application, and communication servers; printers; and local switching/routing equipment. This collection of local computing devices is governed by a single security policy regardless of physical location. Because security policies are unique to the type, or level, of information being processed, a single physical facility may have more than one enclave present. Local and remote elements that access resources within an enclave must satisfy the policy of that enclave. A single enclave may span a number of geographically separate locations with connectivity via commercially purchased point-to-point communications (e.g., T-1, T-3, Integrated Services Digital Network [ISDN]) or using wide area network (WAN) connectivity such as the Internet.

The majority of enclaves have external connections to other networks. These external connections may be single-level connections, where the enclave and connected network are at the same privacy level, or the connection may be a High-to-Low/Low-to-High transfer, where the enclave is at a higher or lower level than the connected network. Enclaves may also have remote access connections to traveling users or users located in remote locations. The point at which the enclave's network service layer connects to another network's service layer is the enclave boundary. Figure 6-1 highlights the enclave boundary target environments within the high-level information infrastructure context. The placement of boundary protection mechanisms in Figure 6-1 is notional, representing only suggested, not necessarily actual, placement of information assurance (IA) components.

Defense of the enclave boundary is focused on effective control and monitoring of data flow into and out of the enclave. Effective control measures include firewalls, guards, virtual private networks (VPN), and identification and authentication (I&A)/access control for remote users. Effective monitoring mechanisms include network-based intrusion detection systems (IDS), vulnerability scanners, and virus detectors located on the LAN. These mechanisms work alone, and in concert with each other, to provide defenses for those systems within the enclave that cannot defend themselves or could be undermined by failures in systems operating at lower security levels or with less stringent security policies. Although the primary focus of the perimeter is on protecting the inside from the outside, enclave boundaries also provide some protection against malicious insiders who use the enclave to launch attacks or who facilitate outsider access through open doors or covert channels.

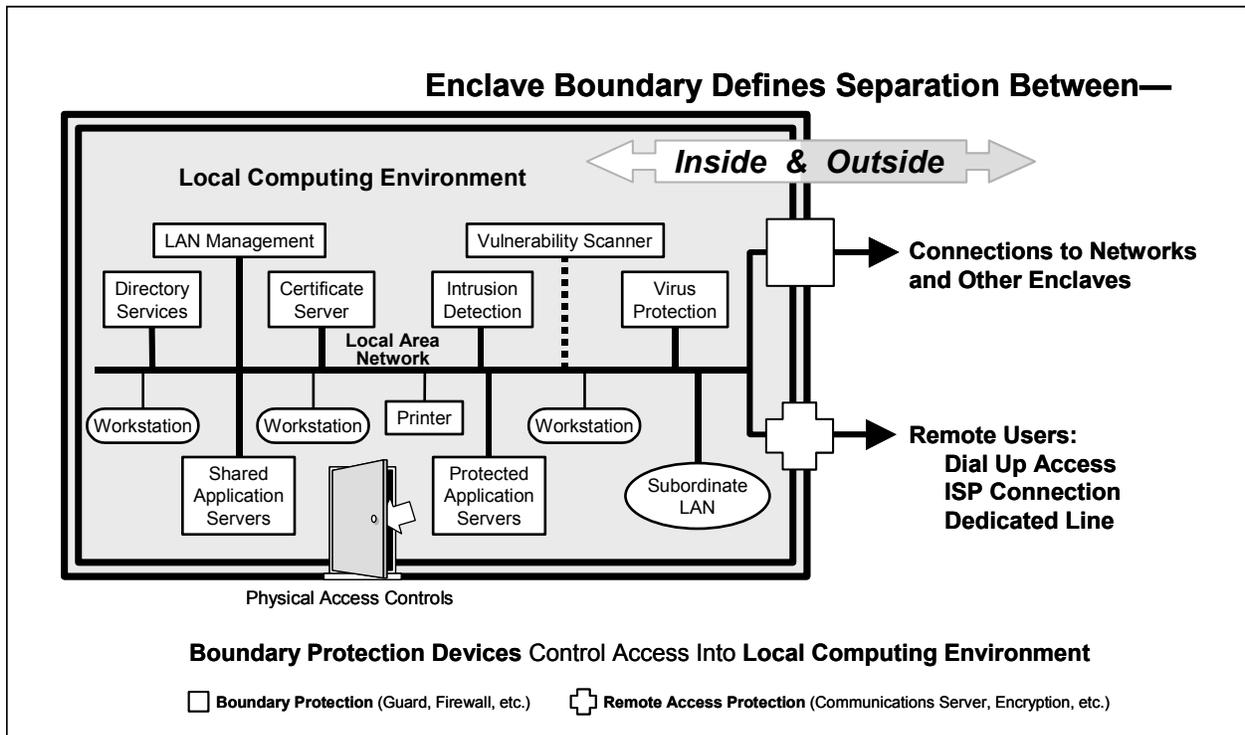


Figure 6-1. Defend the Enclave Boundary

The IA strategy for defending an enclave boundary includes a number of general defensive measures and specific capabilities that address remote access and interoperability across security levels. In general, the enclave perimeters must be established and must be equipped with professionally managed electronic access portals that enable effective control and monitoring. These portals should enable dynamic throttling of services in response to changing information conditions (INFOCON). They should establish mandatory Department of Defense (DoD) policy on the protocols that are allowed and disallowed between secure enclaves and external systems.

The strategy mandates the use of basic intrusion detection for all DoD enclaves, with additional detection mechanisms for mission-critical and mission-essential enclaves. VPNs, used to establish communities of interest (COI) (or intranets) will not be used between enclaves that provide different degrees of security, unless other adequate measures are used to protect the stronger enclave from the weaker one. An important strategy consideration is not losing detection capabilities when increasing the use of encryption. This requires that protection and detection capabilities be planned together. For VPNs, the DoD strategy is to install the VPNs in such a way that network-based monitors can be placed on their clear-text side.

Within the IA strategy, systems and enclaves that are provided with remote access to a secure enclave must comply with the security policy of the secure enclave. The remote enclave or system must comply with approved remote access protocols, be authenticated at the enclave perimeter, and ensure that the entire secure enclave is not jeopardized by overrun of remote access points. In all cases, remote access will require authentication using approved techniques.

UNCLASSIFIED

Defend the Enclave Boundary/External Connections
IATF Release 3.1—September 2002

At a minimum, this means using nonreusable passwords, preferably in encrypted form, or public key-based approaches.

Continuous authentication (versus authentication only at the beginning of a session) is preferred. For interoperability across security levels, the DoD infrastructures will be based on a multiple-security-level strategy in which separate system and network infrastructures are maintained at each security level. The use of devices that control data transfers across security levels will be minimized. When required by operational necessity, these shall be implemented by an official Secret and Below Interoperability (SABI) (or Top Secret and Below Interoperability [TSABI]) process. High-side servers that serve as gateways to receive Low-to-High transfers will use operating systems that are capable of enforcing user-level access controls, are properly configured and operated using the concept of least privilege, and include other appropriate layers of protection (including tripwires for protection against malicious software, preplaced forensics, reporting of incidents and anomalous activity, and host-based auditing).

The Defend the Enclave Boundary/External Connections chapter of the framework addresses the role of IA technologies in providing protection for the enclave. The Firewall section explores ways of protecting internal information systems from external attacks. While the Remote Access section reviews methods for users to securely access their LANs, the Guards section addresses technology used to enable users to exchange data between private and public networks. The Network Monitoring section considers ways to monitor the network infrastructure. The Network Scanners section has a slightly different focus, examining the system for vulnerabilities. Malicious code protection is covered along with multilevel security.

UNCLASSIFIED

Defend the Enclave Boundary/External Connections
IATF Release 3.1—September 2002

This page intentionally left blank.

6.1 Firewalls

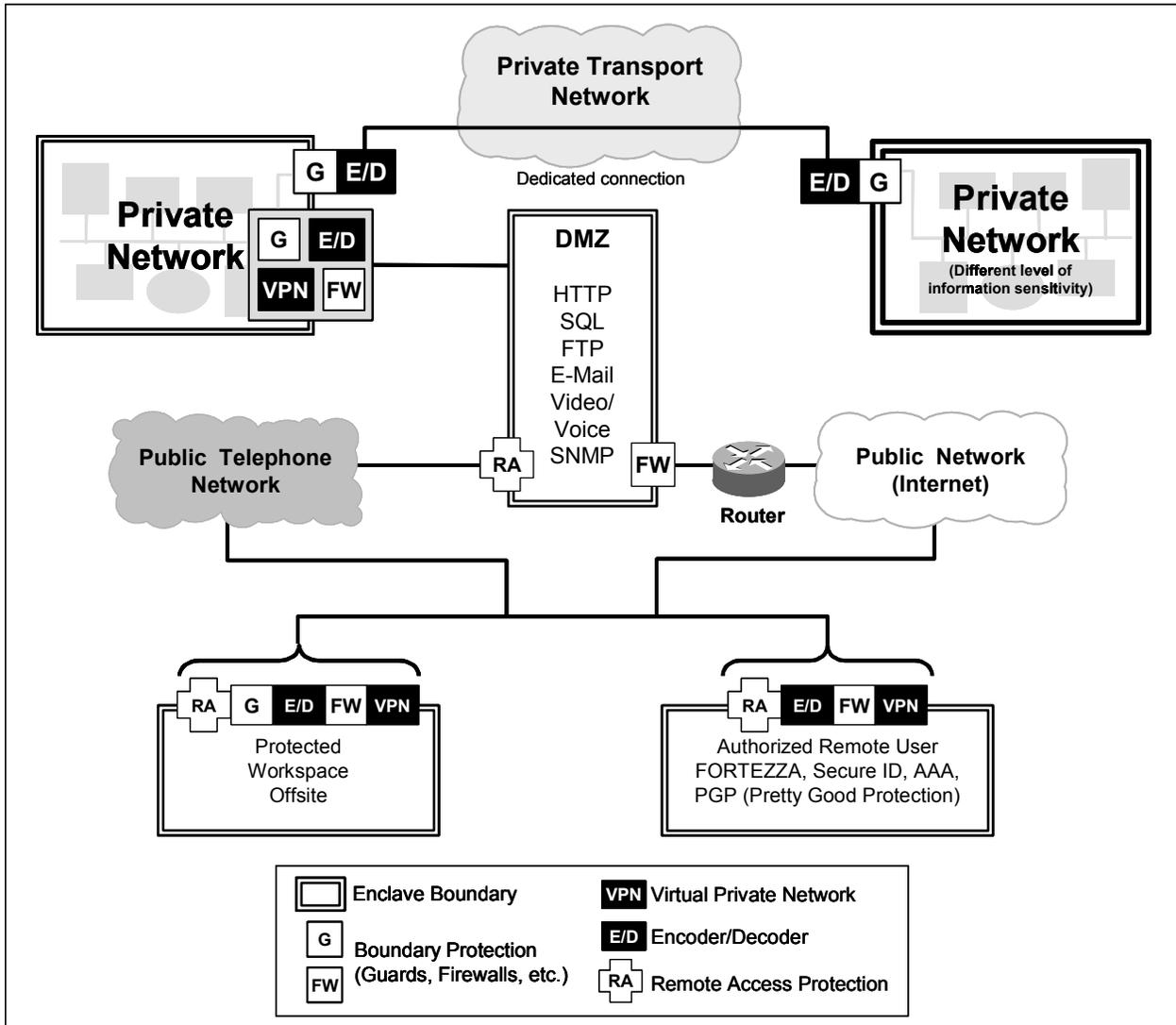
The purpose of a firewall is to protect internal information systems from external attacks. Firewalls address the requirement for authorized Local Area Network (LAN) users and administrators as well as individual workstation or personal-computer users, to safely access and be accessed-by untrusted (potentially hostile) external network connections. This means that all components inside the enclave boundary are protected against intrusion attacks: unauthorized extraction, modification, or deletion of data, denial-of-service, and theft of resources or services. This firewall section addresses all components used for protecting interconnected, digital-electronic processing, transmission, or storage of information.

The focus of this Firewall section is on external electronic intrusions through the enclave boundary into a LAN or workstation that may be possible due to electronic connections. Attacks such as those performed by insiders or passive intercepts of traffic traversing backbone networks are not directly addressed within this section of the Information Assurance Technical Framework (IATF). While the unique concerns of the other protection categories are primarily addressed elsewhere in the Framework, there are some fundamental protection countermeasures—common to most environments—addressed here. Clearly, the concerns and approaches relevant to external electronic intrusions are interdependent with those of other protection categories (such as remote access, system high interconnects, Multi-Level Security [MLS], or security for applications). Thus, the following firewall-focused sections are intended to be complementary and integrated rather than separate, distinct layers of protection. For further expansion of site security, refer to <http://www.ietf.org/rfc/rfc2196.txt?number=2196>, RFC 2196, Site Security Handbook.) [1]

6.1.1 Target Environment

Users within an enclave can access external information services via network connections, dedicated connections, or dial-up connections. The environment illustrated in Figure 6.1-1 includes various combinations of methods of access involving Internet Service Providers (ISP), Integrated Services Digital Networks (ISDN), Public Switched Telephone Networks (PSTN), X.25 Packet Exchange, wideband (cable-modems) and Internet and intranet networks/hosts that consist of both valid (trustworthy) agents and potentially hostile agents.

Included are those involving multiple access levels such as a private corporate LAN connecting to a public Wide Area Network (WAN), or a private corporate LAN connecting to a corporate intranet. The boundary protection approaches should be applied to many of the cases described in other categories (e.g., remote access, system high interconnections and virtual private networks [VPN]). Whenever networks (workstations) are interconnected, the Network Security Policy should require protection at the network access points; i.e., the enclave boundaries. Generally, the amount of protection needed increases as the sensitivity of the information increases, as differences in sensitivity levels increase, as the threat increases, and as the operational environment changes (likelihood for attack increases for high profile organizations).



iatf_6_1_1_0101

Figure 6.1-1. Enclave Boundary Environment

6.1.2 Firewall Requirements

6.1.2.1 Functional Requirements

The following have been identified as representative ideal requirements based on a customer's perspective of needs:

- The user, if authorized, should have maximum access to needed information and services available on the WANs using any of the existing and emerging networking technologies and applications.

- The user and user's system should be protected against the full range of network attacks, be able to locate the source and type of intrusions, be able to react to such intrusions, and be able to fully reconstitute the system following damage caused by intrusions.
- The approaches used to protect network access points should have minimal operational impact on the user.
- The approaches used to protect network access points should have minimal operational impact on performance of the associated components and networks.
- The approaches used to protect network access points should be a scalable solution to allow for future needs.

6.1.2.2 Boundary Protection Mechanism Requirements

Boundary protection mechanisms are used to limit access to the internal network and are provided through the use of some combination of routers, firewalls, and guards. Refer to Section 6.1.4.1, Technical Countermeasures, Boundary Protection via Firewalls, for further expansion of this subject. The following are typical requirements that boundary protection mechanisms should offer.

- Restrict sources, destinations, and services and block dangerous protocols such as certain Internet Control Message Protocol (ICMP) messages. Both incoming and outgoing communications should be restricted.
- Restrict executable services and download capabilities.
- Employ internal Access Control Lists (ACL) where appropriate.
- Use Identification and Authentication (I&A) mechanisms—to include the use of software or hardware tokens—to authenticate outsiders to the boundary point.
- Use encryption to prevent interception of data that could provide the attacker with access to the network and for access control. This should include the encryption of remote management data.
- Hide the internal network (addresses, topology) from potential attackers using a mechanism such as network address translation.
- Log and analyze source-routed and other packets and react to or restrict attacks.
- Scan for malicious software.
- Facilitate proper boundary protection configuration by operators, e.g., user-friendly graphical user interface (GUI).
- Be self-monitoring and capable of generating alarms.

Note that the intent of several of these countermeasures is to eliminate vulnerabilities of services that may not be needed by a particular user system. Current technologies do not permit complete user access to all desired services and destinations while simultaneously blocking all attacks. In addition, the use of encryption and certain identification and authentication mechanisms (such as hardware tokens) limits interoperability. Trade-offs must be made.

6.1.2.3 Interoperability Requirements

The boundary protection should not force users to employ any nonstandard protocols or modes of operation nor any procedures that would prohibit interoperability with those external users or systems with which users desire to communicate and are permitted by the organization's network security policy.

- The firewall command and control channel must be secure to prevent eavesdroppers from learning the rules, Media Access Control (MAC) secrets, and other controlling data communicated over the firewall command and control channel (e.g., Simple Network Management Protocol (SNMP), Remote Monitor (RMON), Application Program Interface (API), and Telnet).
- An authentication mechanism is needed to prevent unauthorized entities from changing the rules. In the simplest case, IP-address-based authentication may be satisfactory. If end-devices are allowed to modify the rules (as they are with SOCKS), secure user-based authentication would have to be deployed along with an administration policy. For example, the policy may permit authenticated user A to open pinholes from his host at high port numbers and deny anything else. (SOCKS is out of the scope of this chapter; for more information refer to <http://www.socks.nec.com> and <ftp://ftp.nec.com/pub/socks>). [2, 3]

6.1.2.4 Anticipated Future Requirements

The approach employed to protect network access should allow for the evolution and reconfiguration of the network and associated components. The chosen approach should be scalable to allow for future evolutions.

6.1.3 Potential Attacks

As previously stated, the focus of this firewall section is on external attacks into a LAN or workstation that may be implemented by virtue of its electronic connections through the enclave boundary. The types of attacks are discussed below: active-based attacks, distribution attacks, and insider attacks. Other attack categories (passive attacks and close-in attacks) are not directly addressed within the remainder of this chapter, but relate to this category and the technologies discussed. Refer to Section 4.2, Adversaries, Threats (Motivations/Capabilities), and Attacks, and for additional details refer to Section 5.3, System-High Interconnections and VPNs, regarding virtual private networking capabilities regarding security and protecting enclave assets from attacks.

6.1.3.1 Active Attacks

Attacks at the network access points generally fall within the active attacks category as defined in Section 4.2.1.4, Categories of Attacks. This type of attack also has been referred to as an active attack. Any attempt to gain unauthorized access to a network or break network security features is an active attack. For more description, refer to Section 4.2.1.4.2, Table 4-2, Examples of Specific Active Attacks. Listed below are various examples of active attacks.

- Trick the Victim (Social Engineering).
- Masquerade as Authorized User/Server.
- Exploit System-Application and Operating System Software.
- Exploit Host or Network Trust.
- Exploit Data Execution.
- Exploit Protocols or Infrastructure Bugs.
- Denial of Service.

6.1.3.2 Distribution Attacks

Distribution attacks are the hostile modification of hardware or software. Such attacks can occur anytime hardware or software is transferred. For additional information, refer to Section 4.2.1.4.4, Hardware/Software Distribution Vulnerabilities and Attacks and Table 4-3, Examples of Specific Modification Attacks. The following are examples of distribution attacks.

- Via software distribution computer disks that are transferred among firewalls.
- Software that is downloaded from the Internet, e-mail, or an internal LAN system.
- Modifications made to hardware or software at the factory before distribution or during distribution. Malicious changes to software code or malicious modification of hardware can occur between the time it is produced in the factory and the time it is installed and used.
- During firewall configuration, especially from remote locations.

6.1.3.3 Insider Attacks

Although the emphasis of protecting network access points is on protecting the inside from a potentially hostile outside world, mechanisms are needed for protection against outside and inside intruders. Thus, some of the technologies identified in this section apply to both insider and outsider threats. Further, once an outsider has successfully attacked a system to obtain access, the outsider, in effect, maneuvers within the system as an insider would. Technologies such as those designed to detect attacks by an insider may be used in a similar manner to detect outsider attacks.

Insider attacks can occur when an authorized user (i.e., a person who has authorization to access the system) remotely connects to the system and unintentionally causes damage to the information or to the information processing system. This nonmalicious attack can occur either from the user not having the proper knowledge or by carelessness. Malicious insider attacks are those in which an authorized user causes damage to the system or enters areas where the user is not authorized. Malicious attacks can also be caused by an unauthorized individual employing an authorized user's personal computer (PC) to maneuver within the system and cause damage. An example would be when an authorized user's laptop computer is stolen and then used to gain access into the system. For more information, refer to Section 4.2.1.4.3, Insider Vulnerabilities and Attacks.

6.1.4 Potential Countermeasures

Fundamentally, protecting network access points from potential attacks can be addressed by limiting access to and from the LAN or workstation. In the protection of a network, important issues that need to be addressed include detecting and identifying malicious or non-malicious insider attacks, identifying potential vulnerabilities, and attacks that may occur given the current configuration and responding to, deterring, and recovering from detected attacks. The following subsections describe security requirements applicable to addressing attacks through an enclave boundary. Several of the countermeasures are covered in detail within other IATF focus areas and are listed as applicable. The countermeasure requirements are grouped under the two primary headings of Technical Countermeasures and Administrative Countermeasures.

6.1.4.1 Technical Countermeasures

Boundary Protection via Firewalls

Connecting through the enclave boundary to external resources such as the Internet introduces a number of security risks to an organization's information and resources. The first step in minimizing those risks consists of developing a comprehensive network security policy. This network security policy framework should include firewalls as boundary protection mechanisms. Boundary protection mechanisms can provide a measure of protection for a network or an individual workstation within the enclave boundary. The boundary protection device is intended to operate primarily as an access control device, limiting the traffic that can pass through the enclave boundary into the network. In general, boundary protection is provided through the use of some combination of routers, firewalls, and guards. Refer to Section 6.1.1.2, Firewall Requirements, Boundary Protection Mechanism Requirements for additional information.

Although the main focus of this section is firewalls, a definition of routers and guards follows. A router that is configured to act as a firewall is a packet-filtering device that operates at multiple layers and permits or denies traffic through the enclave boundary into the internal network based on a set of filters established by the administrator. A guard is generally a highly assured device that negotiates the transfer of data between enclaves operating at different security levels. Refer

to Section 6.3, Guards, for more information. In contrast, a firewall is a boundary protection device between networks communicating at the same security level.

A firewall is a collection of components placed between two networks (or an individual workstation and a network) with the following properties.

- All traffic from inside to outside and vice versa must pass through this mechanism.
- Only authorized traffic, as defined by the local network security policy, will be allowed to pass.
- The mechanism itself is immune to penetration.

Thus the firewall is a tool for enforcing the network security policy at the enclave boundary and has several distinct advantages as a protected network access device. First, the firewall allows for centralized network security management, as it becomes the focal point for network security decisions. In addition, as the only directly accessible component of the enclave network, the firewall limits the exposure of the network to attack. By implementing and following a well-defined network security policy, maintaining cognizance of current vulnerabilities, reviewing audit data, and using available scanning tools, the security of the enclave is greatly enhanced.

However, there are disadvantages to using firewalls. They can be the single points of attack to the enclave. Firewalls do not protect the network and workstations within the enclave against most data-driven attacks, some denial-of-service attacks, social engineering attacks, and malicious insiders. Firewalls can thus potentially provide a false sense of security. Firewalls must be looked at as being only one part of a larger network security approach.

Access Constraint

Measures that should be taken to constrain access to facilitate defense of enclave boundaries include the following.

- Provide data separation. For data that is allowed access to the protected network or workstation, steps should be taken to constrain as much as possible the amount of the system that can be affected. Steps that could be taken include allowing executables to run only in a particular domain or only on a server reserved for such purposes as discussed in Section 6.3, Guards.
- Employ application-level access control. Access restrictions may also be implemented within the enclave—within workstations or at various points within a LAN—to provide additional layers and granularity of protection. See Access Control List under Section 6.3.5.3, Processing, Filtering, and Blocking Technologies.
- Provide authenticated access control and (as appropriate) encryption for network management. See a previous subheading in this category, Boundary Protection via Firewall and Section 6.3.5.1, Authenticated Parties Technologies.

6.1.4.2 Administrative Countermeasures

While defending the enclave boundary, administrative countermeasures should be implemented with the boundary protection mechanisms and throughout the enclave. Quality network management and network security administration are imperative in maximizing the security of the network's configuration and protection mechanisms and increasing the likelihood of detecting vulnerabilities and attacks. The following administrative mechanisms act as countermeasures to the various attacks mentioned in Section 6.1.3, Potential Attacks.

- Be prepared for severe denial-of-service attacks; i.e., institute and practice contingency plans for alternate services.
- Routinely inspect the firewall for physical penetrations.
- Educate users and staff on correct procedures when dealing with firewalls.
- Institute and exercise well-publicized firewall procedures for problem reporting and handling.
- Institute and exercise suspicious behavior-reporting channels.
- Institute and monitor critical access controls, e.g., restrict changeable passwords, require dial-back modems.
- Minimize use of the Internet for mission or time-critical connectivity.
- Require security-critical transactions to be conducted in-person; e.g., establishing identity when registering.
- Use trusted software where available and practical.
- Use subversion-constraining software and techniques wherever possible; e.g., avoid software that uses pointers that could be employed by a software developer to access unauthorized memory locations.
- Carefully map relationships between hosts and networks, constraining transitive trust wherever possible.
- Minimize cross-sharing between users and file systems, particularly for high-sensitivity or high-threat applications, allowing only essential functions that have compelling justifications for sharing.
- Where possible, do not rely on Domain Name Server (DNS) for security sensitive transactions where spoofing an Internet Protocol (IP) address could cause problems.
- Institute, exercise, and monitor a strict computer emergency response team alert and bulletin awareness and patch program.
- Institute and practice procedures for recovery from attack when the firewall is penetrated.

Countermeasure Effectiveness

The following is a list of attacks and the most successful countermeasures against them. More detailed information about the types of attacks is also provided in Section 4.2, Adversaries, Threats (Motivations/Capabilities), and Attacks.

Trick the Victim (Social Engineering). The best defense against this type of attack is to educate system/network users. The users must be aware that attempts may be made to obtain their passwords to enable access to the network or to secure areas of the network that the attacker may not be authorized to access.

Masquerade. The best technical countermeasure against this type of defense is to identify and authenticate outsiders and to use access constraints to authenticate and encrypt data. Administrative countermeasures that have high levels of effectiveness include using and monitoring access controls and minimizing the use of the Internet for critical communications.

Exploit Software Vulnerabilities. The highest defenses against attacks made by exploiting vulnerabilities of software include subverting constrained software, monitoring the Computer Emergency Response Team (CERT), obtaining patches, and minimizing the use of the Internet for critical communications.

Exploit Host or Network Trust. Minimizing use of the Internet for critical communications and subverting constrained software provides the highest level of defense against attacks exploiting the host or trust in the network.

Exploit via Executables. Attacks against the enclave boundary through executable applications can be fought through technical and administrative countermeasures. Overall technical measures that can be implemented include boundary protection, access constraints, and detection mechanisms. Boundary protection offers the best technical defense by restricting sources and services, by restricting the ability to download, and by restricting executables. Administrative measures to counteract attacks via executables are minimizing the use of the Internet for critical communications and using subversion-constraining software.

Exploit Protocol Bugs. To protect against protocol bugs, the two countermeasures providing the best defense are—once again—minimizing the use of the Internet for critical communications and using subversion-constraining software.

Denial of Service. The best technical defense for a denial-of-service attack against a system is to have a detection and response system in place. Administrative countermeasures include advance planning to be able to offer service alternatives, minimize Internet usage for critical communications, and to have documented and rehearsed recovery procedures in place to help reconstitute the system.

6.1.5 Firewall Technology Assessment

Access Control/Filtering

Access control/filtering is the main function of every firewall. This function can be accomplished in several ways ranging from a proxy at the application layer of the Open Systems Interconnection (OSI) model to stateful inspection at the IP layer. By its nature, the firewall implements a specific network security policy that corresponds to the level of sensitivity of the boundary the firewall is protecting. The main fundamental purpose of the security policy is to limit access to the network and systems inside the enclave boundary from external sources. Only necessary in-bound connections and services should be allowed. The firewall also restricts the connectivity of internal users to external destinations. Although internal users are generally trusted, they should be limited in what services they can use through the firewall to prevent them from unintentionally opening security vulnerabilities. The different firewall technologies offer different granularities of access control. Some firewalls are now capable of what were traditionally guard-like filtering functions. For example, firewalls incorporate software that filters access to either specific Universal Resource Locators (URL) or categories of URLs. Certain File Transfer Protocol (FTP) commands can be blocked while other commands are allowed through the firewall. Technology will continue to develop in this area. Very sophisticated and highly refined access control capabilities are likely to become standard firewall features.

Identification and Authentication

Identification and authentication is one of the major functions provided by the different firewall products. While users on the inside of a firewall, inside the enclave boundary, are often considered trusted, external users who require access to the internal network must be authenticated. Most security experts agree that passwords are not a strong method of authentication. In fact, cracking user passwords is one of the most common system attacks. Other authentication methods for screening access through a firewall include one-time passwords, time-based passwords, and challenge-response schemes. The most common one-time password system in use is S\key, a software-based authentication mechanism using Message Digest 4 (MD4) or Message Digest 5 (MD5). S\key works by starting with a seed and applying MD4 or MD5 to generate a sequence of keys. S\key encodes the keys into a series of short words and prompts the user for the previous key, n-1, then S\key applies the MD4 or MD5 to the user's answer and checks to see if the result is the key n that it knows. Time-based passwords are a special form of one-time password. In these systems, the password varies at a specified time interval based on an internal algorithm, thus adding the additional complication of maintaining clock synchronization. Challenge-response systems are more complex and involve something the user has (a smart card or PC card) and something the user knows (password). Although it is possible to implement these systems in software, using hardware tokens has numerous advantages. Commercial firewall products support a wide range of authentication mechanisms.

Mobile Code Blocking

In addition to more basic blocks of mobile code (Java, *Script, ActiveX, etc.), firewall systems are beginning to offer containment for the execution of mobile code. This includes sandbox machines isolated from the rest of the network and restricted environments to run the Java Virtual Machine (VM) within. Refer to RFC 1918—Address Allocation for Private Internets for more information: <http://www.ietf.org/rfc/rfc1918.txt?number=1918>. [4]

Encryption

Firewalls become a focal point for the enforcement of security policy. Some firewalls take advantage of this to provide additional security services, including traffic encryption and decryption. To communicate in encryption mode, the sending and receiving firewalls must use compatible encrypting systems. Current standards efforts in encryption and key management have begun to allow different manufacturers' firewalls to communicate securely. To address this situation, vendors have been working on a network-level encryption interoperability approach through the Internet Protocol Security (IPSec) standard, set forth by the Internet Engineering Task Force (IETF). However, these efforts require further development before the customer can assume compatibility. Firewall-to-firewall encryption is thus used for secure communication over the Internet between known entities with prior arrangement, rather than for any-to-any connections. Verifying the authenticity of system users is another important part of network security. Firewalls can perform sophisticated authentication, using smart cards, tokens, and other methods.

Auditing

Auditing refers to the tracking of activity by users and administrators. As opposed to accounting—where the purpose is to track consumption of resources—the purpose of auditing is to determine the nature of a user's network activity. Examples of auditing information include the identity of the user, the nature of the services used, when hosts were accessed, protocols used, and others.

Network Address Translation

Network Address Translation (NAT) is a method by which IP addresses are mapped from one realm to another to provide transparent routing to hosts. NAT enables a LAN to use one set of IP addresses for internal traffic and a second set of addresses for external traffic. Traditionally, NAT devices are used to connect an isolated address realm with private unregistered addresses to an external realm with globally unique registered addresses (Internet). That is, a NAT device sits at the enclave boundary between the LAN and the Internet and makes all necessary IP address translations.

Resist Penetration

Another important aspect of a firewall is how well it protects itself against attack. The firewall itself should resist penetration, because breaking into the firewall will give a hacker access to the entire network. Most firewalls run on stripped-down versions of the operating system; unnecessary executables, compilers, and other dangerous files are removed. In addition, some firewalls employ technology that makes penetrating the firewall operating system extremely difficult. These firewalls are built on trusted operating systems or use mechanisms such as type enforcement (i.e., controls based on factors that can only be changed by the system security administrator) to provide this extra protection against penetration. Although these types of additional safeguards are traditionally found on guard devices, firewalls are also beginning to offer this type of extra protection against enclave boundary penetration.

Configuration and Third Party Monitoring

Properly configuring the firewall components is critical to the security of the enclave boundary. Most vulnerabilities in firewalls arise from the improper configuration or maintenance of the firewall. For this reason, it is important to examine the administrative interface provided by the firewall. A GUI alone will not make the firewall any more secure. However, a well-designed operator interface can ease the administrative burden and more effectively illustrate how well the firewall has implemented the security policy. Firewalls also make use of various self-monitoring tools. These tools can provide additional access controls, can increase the auditing capability of the firewall, and can provide for an integrity check on the file system of the firewall. Some of these tools are proprietary and are provided with the firewall; other tools are available from the third parties and can be used to enhance the security of the firewall.

6.1.5.1 Firewall Types

Packet Filtering

Because routers are commonly deployed where networks with differing security requirements and policy meet, it makes sense to employ packet filtering on routers to allow only authorized network traffic, to the extent possible. The use of packet filtering in those routers can be a cost-effective mechanism to add firewall capability to an existing routing infrastructure.

As the name implies, packet filters select packets to filter (discard) during the routing process. These filtering decisions are usually based on comparing the contents of the individual packet headers (e.g., source address, destination address, protocol, and port) against preset rule sets. Some packet filter implementations offer filtering capabilities based on other information beyond the header. These are discussed below in Stateful Pack Filtering. Packet filtering routers offer the highest performance firewall mechanism. However, they are harder to configure because they are configured at a lower level, requiring a detailed understanding of protocols.

Stateful Packet Filtering

Stateful packet filtering technology, also referred to as *stateful inspection*, provides an enhanced level of network security compared to the static packet filtering described above. The stateful packet filter—working at layer 3 of the OSI model to examine the state of active network connections—looks at the same header information as packet filters do, but can also look into the data of the packet where the application protocol appears. Based on the information gathered, stateful packet filtering determines what packets to accept or reject. More importantly this technology allows the firewall to dynamically maintain state and context information about *previous* packets. Thus, the stateful packet filter compares the first packet in a connection to the rule set. If the first packet is permitted through, the stateful packet filter adds the information to an internal database called a state table. This stored information allows subsequent packets in that connection to pass quickly through the firewall.

Network security decisions can then be based on this state information. For example, the firewall can respond to an FTP port command by dynamically allowing a connection back to a particular port. Because they have the capability of retaining state information, stateful packet filters permit User Datagram Protocol (UDP)-based services (not commonly supported by firewalls) to pass through the firewall. Thus stateful packet filters are advertised to offer greater flexibility and scalability. Stateful packet filtering technology also allows for logging and auditing and can provide strong authentication for certain services. Logging, or authentication as required by the rule set, occurs at the application layer (OSI layer 7). A typical stateful packet filtering firewall -may log only the source and destination IP addresses and ports, similar to logging with a router.

Unlike application-level gateways, stateful inspection uses business rules defined by the administrator and therefore does not rely on predefined application information. Stateful inspection also takes less processing power than application-level analysis. However, stateful inspection firewalls do not recognize specific applications and thus are unable to apply different rules to different applications.

Proxy Service, Application Gateways and Circuit Gateways

Figure 6.1-2, shows how proxy services prevent traffic from directly passing between networks. Rather, Proxy Services are software applications that allow for connections of only those application sessions (e.g., Telnet, FTP, DNS, Simple Mail Transfer Protocol (SMTP) for which there is a proxy. Thus, proxy services are application-level firewalls. The host running the proxy service is referred to as an application gateway. Since an application-level gateway is a system set up specifically to counter attacks from the external network, it is also referred to as a bastion host. If the application gateway contains proxies for only Telnet or DNS, only these sessions will be allowed into the subnetwork. If a proxy does not exist on the application gateway for a particular session (Telnet, DNS, FTP, SMTP), those sessions will be completely blocked. Therefore, only essential services should be installed on the bastion host, for if a service is not installed, it cannot be attacked. Proxy services can also filter connections through the enclave boundary by denying the use of particular commands within the protocol session

(e.g., the FTP put command) and by determining which internal hosts can be accessed by that service.

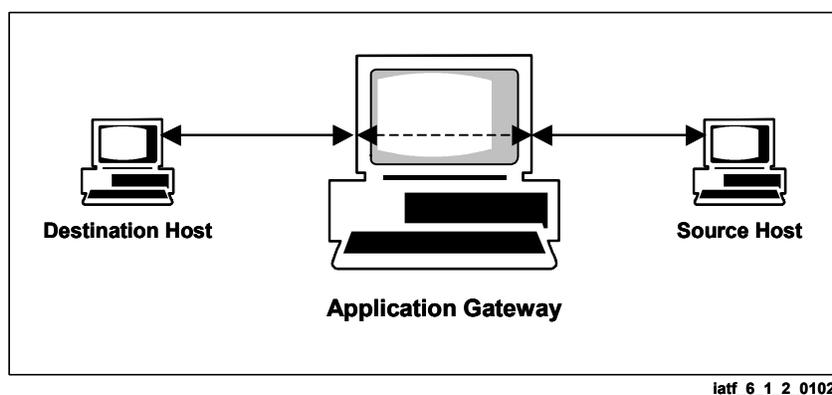


Figure 6.1-2. Application Gateway

By using an application gateway through which access to the subnetwork is permitted, internal information can be hidden from systems outside the enclave boundary. The application gateway can provide a means for strong authentication by requiring additional authentication such as an additional password or the use of a smart card. Each proxy contained within the bastion host can also be set up to require yet another password before permitting access. The bastion host and each proxy service can maintain detailed information by logging all traffic and the details of the connections. Logging helps in the discovery of, and response to, attacks. Each proxy is independent of all other proxies that may be running on the bastion host, so any operational malfunction of one proxy will not affect the operation of the other proxies. This also allows for ease of installation and removal of proxies from the system.

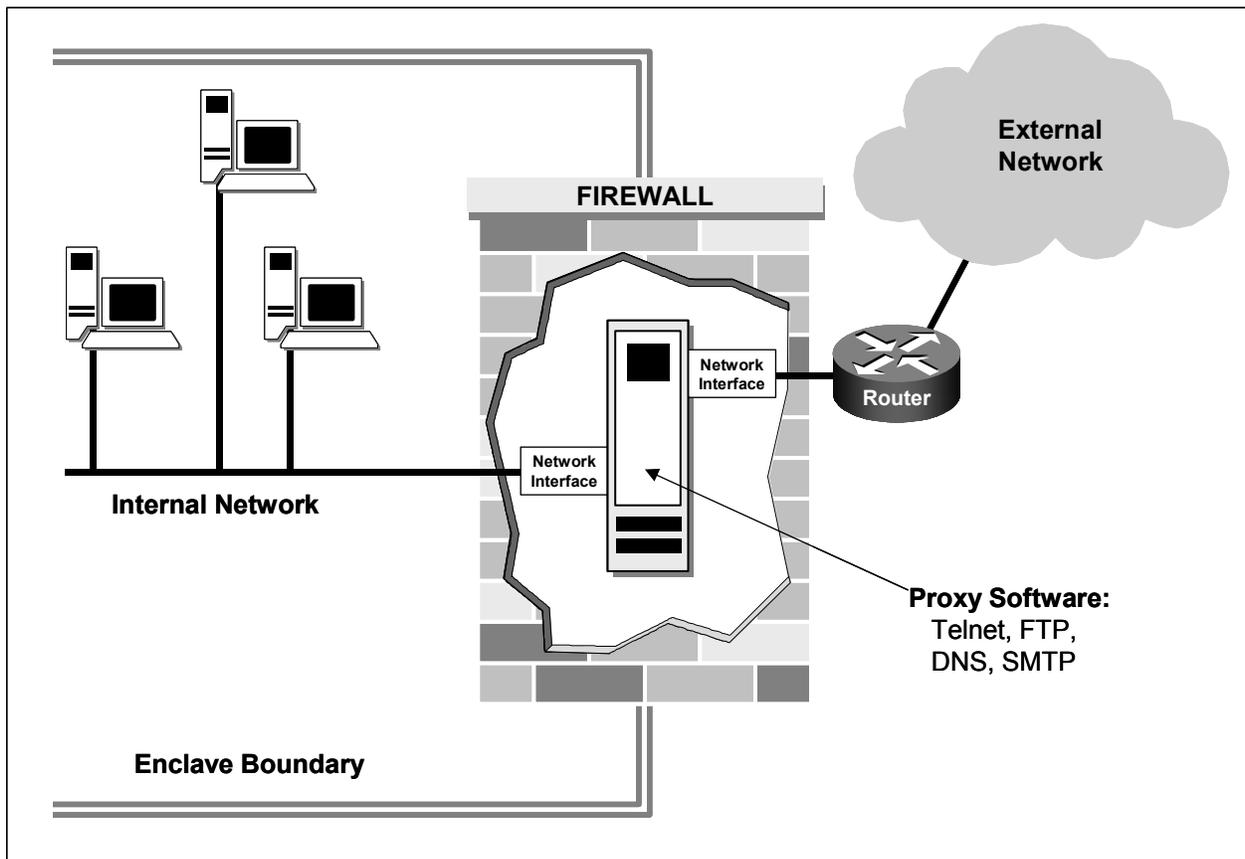
Circuit-level gateways are another type of firewall. A circuit-level gateway relays Transmission Control Protocol (TCP) connections without performing any additional packet processing or filtering. Circuit-level gateways are often used for outgoing connections where internal users are trusted. Outbound connections are passed through the enclave boundary based on policy and inbound connections are blocked. Permission is granted by port address, upon which management control is primarily based. Although a circuit-level gateway is a function that can be performed by an application-level gateway, it is not as secure as an application-level gateway. When completing a connection, checking is not conducted to verify if application protocols (proxies) exist on the application gateway. Therefore, a circuit relay will not detect the violation if approved port numbers are used to run unapproved applications. A circuit-level proxy, acting as a wire, can be used across several application protocols. A bastion host can be configured as a hybrid gateway supporting application-level or proxy services for in-bound connections and circuit-level functions for outbound connections. Circuit-level firewalls are less common than application-level firewalls due to the high probability that client modifications will be necessary to allow use of the circuit-level protocol.

Application gateways are generally dual-homed, which means that they are connected to both the protected network and the public network; however, they can be used in other configurations as discussed below. Packet filtering firewalls can also be dual-homed.

6.1.5.2 Firewall Architectures

Dual-Homed

A dual-homed gateway architecture has two network interfaces, one on each network, and blocks all traffic passing through it, as shown in Figure 6.1-3. That is, the host cannot directly forward traffic between the two interfaces. Bypassing the proxy services is not allowed. The physical topology forces all traffic destined for the private network through the bastion host and provides additional security when outside users are granted direct access to the information server.



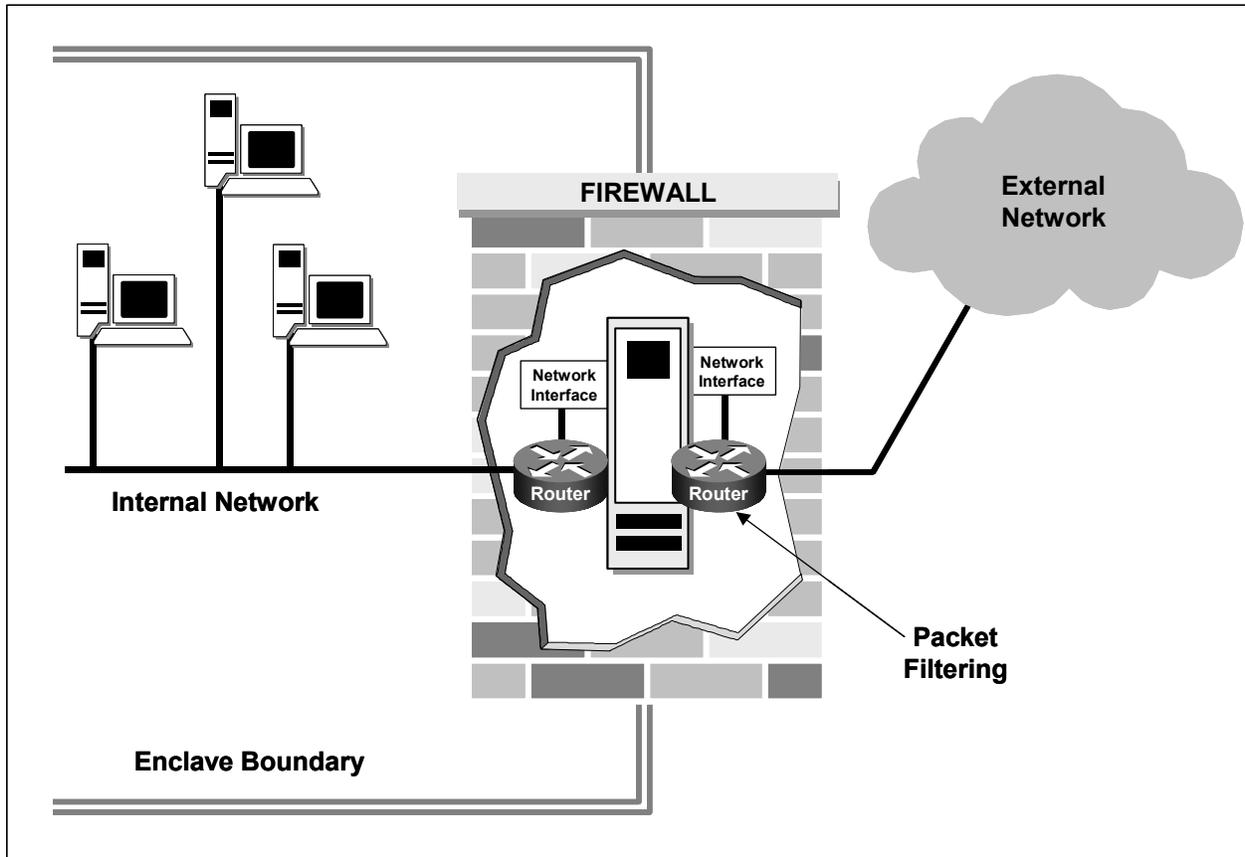
latf_6.1_3_0103

Figure 6.1-3. Dual-Homed Firewall Architecture

Screened Host (Hybrid)

A screened host is a type of firewall that implements both network-layer and application-layer security by using both a packet-filtering router and a bastion host. A screened host architecture is also known as a hybrid architecture. This type of firewall architecture provides a higher level of network security, requiring an attacker to penetrate two separate systems. The system is set up with a packet filtering router sitting between an untrusted (external) network and the bastion host on the protected network so that only allowable traffic from untrusted networks pass to or

from the internal bastion host. (See Figure 6.1-4.) The packet filtering router is configured in such a manner that outside traffic has access only to the bastion host. An additional router may be set up between the Bastion Host and the internal network for a greater level of security.



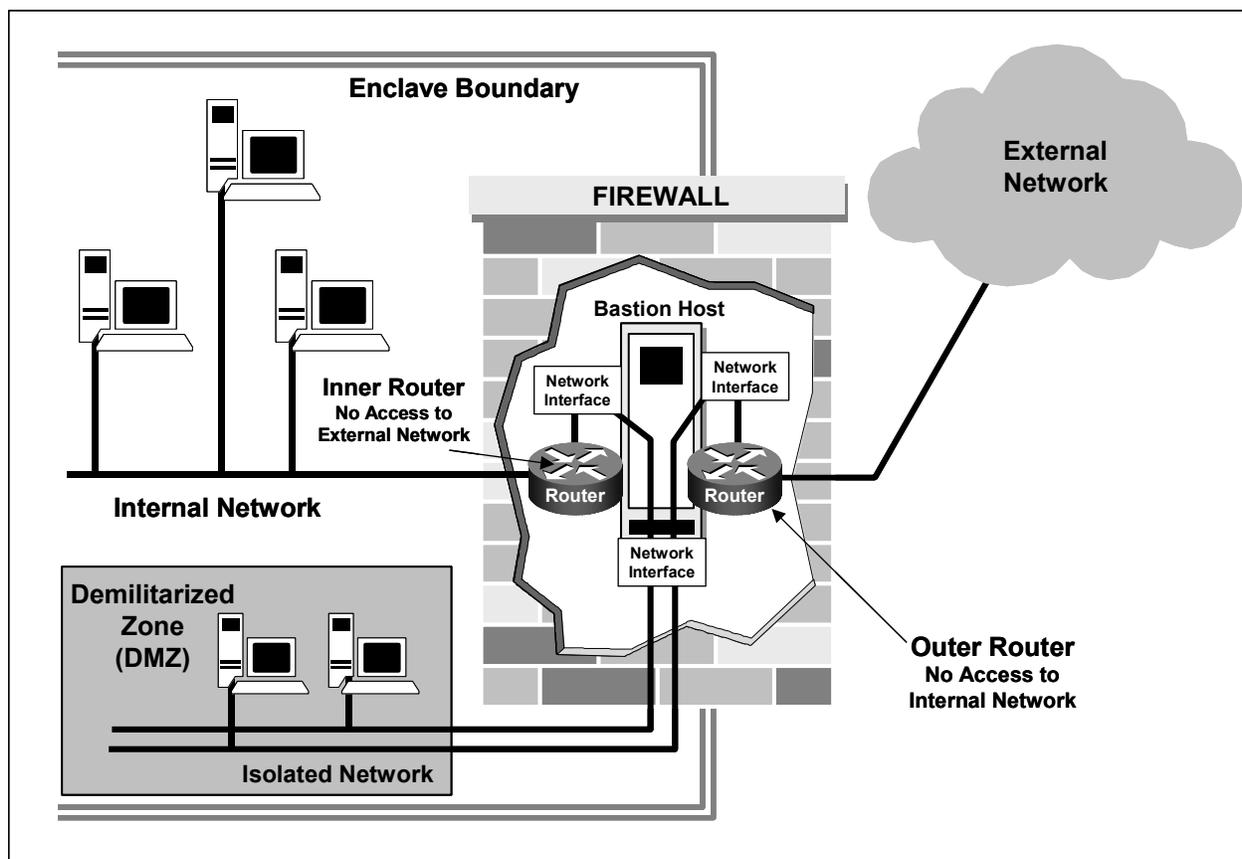
iatf_6_1_4_0104

Figure 6.1-4. Screened Host Firewall Architecture

Screened Subnet

In the Screened Subnet firewall architecture, see Figure 6.1-5, a host is set up as a gateway with three NIC's, one connected to the external network through a router, one to the internal network, and one to the Demilitarized Zone (DMZ). Packet forwarding is disabled on the gateway and information is passed at the application level or the network layer depending on the type of firewall used. The gateway can be reached from all sides, but traffic cannot directly flow across it unless that particular traffic is allowed to pass to the destination it is requesting.

The router should also be setup with ACLs or IP filtering so connections are allowed between the router and the firewall only. The screened subnet provides external, untrusted networks restricted access to the DMZ for services such as World Wide Web (WWW) or (FTP). It allows the enclave to place its public servers in a secure network that requires external sources to traverse the firewall and its security policy to access the public servers, but will not compromise the operating environment of the internal networks if one of the networks is attacked by hackers.



iatf_6_1_5_0105

Figure 6.1-5. Screened Subnet Firewall Architecture

The screened subnet firewall may be more appropriate for sites with large traffic volume or high-speed traffic. A screened subnet can be made more flexible by permitting certain trusted services to pass from the external network to the protected network, but this may weaken the firewall by allowing exceptions. Greater throughput can be achieved when a router is used as the gateway to the protected subnet. Because routers can direct traffic to specific systems, the application gateway does not necessarily need to be dual-homed. However, a dual-homed gateway is less susceptible to weakening. With a dual-homed gateway, services cannot be passed for which there is no proxy. The screened subnet firewall could also be used to provide a location to house systems that need direct access to services.

6.1.5.3 Firewall Selection Criteria

When selecting a firewall system the following should be considered.

- The firewall should be able to support a “deny all services except those specifically permitted” design policy, even if that is not the policy used.
- The firewall should support your network security policy, not impose one.

UNCLASSIFIED

Firewalls

IATF Release 3.1—September 2002

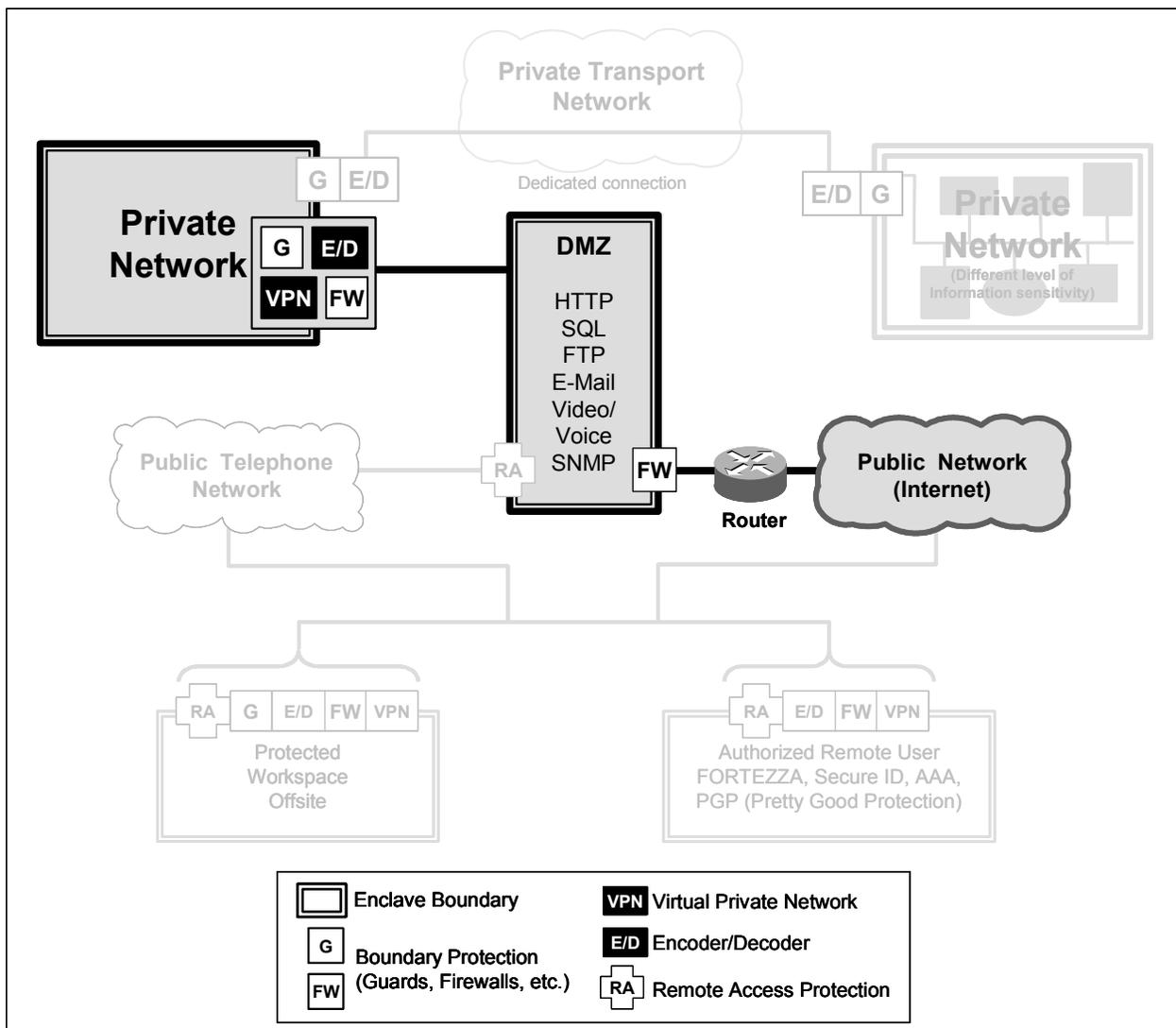
- The firewall should be flexible; it should be able to accommodate new services and needs if the network security policy of the organization changes.
- The firewall should contain advanced authentication measures or should contain the hooks for installing advanced authentication measures.
- The firewall should employ filtering techniques to permit or deny services to specified host systems as needed.
- The IP filtering language should be flexible, user-friendly to program, and should filter on as many attributes as possible, including source and destination IP address, protocol type, source and destination TCP/UDP port, and inbound and outbound interface.
- The firewall should use proxy services for services such as FTP and Telnet, so that advanced authentication measures can be employed and centralized at the firewall. If services such as Network News Transfer Protocol (NNTP), X Window System (X), Hypertext Transfer Protocol (HTTP), or gopher are required, the firewall should contain the corresponding proxy services.
- The firewall should have the ability to centralize SMTP access to reduce direct SMTP connections between site and remote systems. This results in centralized handling of site e-mail.
- The firewall should accommodate public access to the site in such a way that public information servers can be protected by the firewall, but can be segregated from site systems that do not require public access.
- The firewall should have the ability to concentrate and filter dial-in access.
- The firewall should have mechanisms for logging traffic and suspicious activity and should contain mechanisms for log reduction to ensure logs are readable and understandable.
- If the firewall requires an operating system such as UNIX, a secured version of the operating system should be part of the firewall, with other network security tools as necessary to ensure firewall host integrity. The operating system at start up should have all current and approved patches installed.
- The firewall should be designed and implemented in such a manner that its strength and correctness is verifiable. It should be simple in design so it can be understood and maintained.
- The firewall, and any corresponding operating system, should be maintained with current and approved patches and other bug fixes in a timely manner.

6.1.6 Cases

Case 1

A user communicating from a protected network to a public network. The information that is being sent is unclassified but private.

This is a case of the typical user connecting and passing information across the Internet. In Figure 6.1-6, a workstation within the protected network is communicating with the Internet. When connecting to a network of a lower protection level, mechanisms should be in place at the enclave boundary to provide protection for the users' workstation and the protected network.



iatf_6_1_6_0106

Figure 6.1-6. Case 1—Private to Public Network Communication

UNCLASSIFIED

Firewalls

IATF Release 3.1—September 2002

A firewall can be deployed as part of an effective boundary protection function. Other components of boundary protection that can be implemented are through e-mail, browsers, operating system configuration; and router configuration. Once mechanisms are in place to protect the enclave boundary, vulnerability checking and scanning procedures need to be implemented and exercised on the network and on the firewall.

As part of the boundary protection plan a site survey should be performed to ensure that the network operations and configuration is well understood. To assist with the site survey, a mapping tool can be used to construct the networks' topology and to examine the physical security of the network. The network map should detail which systems connect to public networks, and which addresses occur on each subnetwork. The network map should also identify which systems need to be protected from public access and identify which servers need to be visible on the outside and perimeter networks and what type of authentication and authorization is required before users can access the servers. The site survey should also examine which applications are used by authorized users of the network, what the anticipated growth of the network is, and what a users' privileges are including system administrators and firewall administrators. In general, the site survey that should be attempted is directly related to the following.

- Technical expertise of the individual conducting the scanning.
- Level of threat.
- Sensitivity of potentially vulnerable information.
- Integrity of the source of the scanning software.

The placement of the firewall is of critical importance to the security of the network. The network needs to be configured to ensure that if an intruder accesses one part of the system, the intruder does not automatically have access to the rest of the system. A firewall should be placed at egress points to the network.

The recommended procedures that should be implemented relative to the firewall for protecting the enclave boundary include:

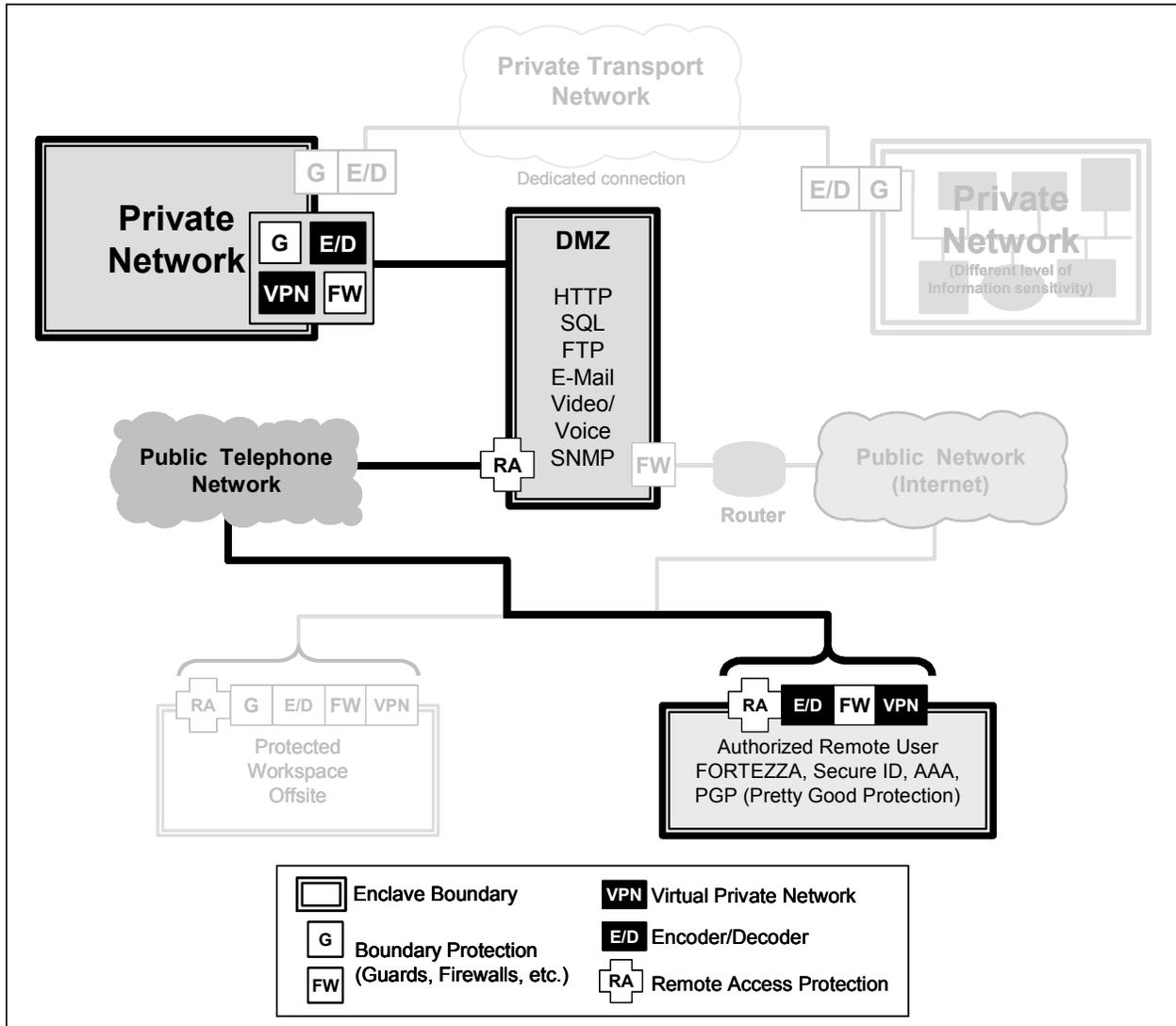
- Ensure that the virus-scanning application is no more than a few weeks old. Viruses may infect the firewall itself as well as resources behind the firewall.
- Ensure that passwords and logins are not in clear text. Clear text passwords and logins are unencrypted and unscrambled and therefore vulnerable to sniffers on the Internet, allowing hackers to obtain passwords.
- Ensure that passwords and Secure Sockets Layers (SSL) are not cached by proxy agents on the firewall.
- Train personnel on firewall operations and administration.
- Audit for intrusive or anomalous behavior employing operating system, browser, and e-mail built-in audit capabilities.

- Routers can be configured as a firewall and for port mappings. With routers, anti-spoofing can be implemented, especially at the enclave boundaries or between domains of network administration. Source address spoofing and denial-of-service protection can also be provided with access lists. The goal of creating an access list at the firewall level to prevent spoofing is to deny traffic that arrives on interfaces on nonviable paths from the supposed source address. For example, if traffic arrives on an interface sitting on the corporate side, yet the source address states that the traffic originated from the Internet, the traffic should be denied, as the source address has been falsified, or “spoofed.” Antispoofing access lists should always reject broadcast or multicast traffic.
- Routers could also be configured to hide the real network identity of internal systems from the outside network through port address translation. Port address translation minimizes the number of globally valid IP addresses required to support private or invalid internal addressing schemes.
- Configure operating system, browser, and applications for firewall functions and to permit specific access (make use of a proxy-based/application gateway). All traffic passing through the firewall should be proxied and/or filtered by the firewall. Proxies reduce the probability that flaws in the service can be exploited. Filtering limits the services that can be used and the user communities that have permission to use a service. The fewer services allowed through the firewall, the fewer opportunities there are to attack the protected network/system.
- Develop and exercise plans to handle any security incidents that may occur. These plans need to cover such things as:
 - How to handle detected port scans or more malicious attacks.
 - Recovery from any incident that degrades the performance of the network.
 - The procedure for adding new services to the firewall.

Case 2

A privileged user remotely connecting to a private network from dedicated workstations situated within a DMZ of a different protected network.

This case is an example of remotely accessing a company’s network from an off-site location. This off-site location is a protected network and has dedicated workstations connecting through that corporation’s DMZ. Multiple connections through the DMZ can be established. Figure 6.1-7 illustrates a valid remote user connecting through the DMZ to the protected network. A DMZ allows authenticated authorized users to tunnel through the firewall. A DMZ also allows access to a Web or FTP server inside the firewall without exposing the rest of the network to unauthorized users. Otherwise, intruders could gain control over the FTP or Web server and attack other hosts in the network. Therefore, servers should be placed so they can be accessed from any address in a separate subnetwork. Organizations can design, deploy, and proactively update and monitor a multi-zoned security network through a single firewall strategy. Administrators can create multiple DMZs within the network by simply adding rules to the existing firewall.



latf_6_1_7_0107

Figure 6.1-7. Case 2—Remotely Accessing a Private Network

Modem banks should be established as part of the firewall protection approach so that users can dial out and remote users can dial in via a modem bank. Modems should not be allowed on networked computers within the protected enclave boundary. By bypassing the implemented firewall and using a modem to connect to the Internet, all control over network security is lost. By using modem pools (a single dial-in point), all users are authenticated in the same manner. In addition, anti-spoofing controls can be applied at dial-up pools and other end-use connection points (also refer to <http://www.ietf.org/rfc/rfc2267.txt?number=2267>, RFC 2267). [5]

Before a user can access anything on the network, a username and password check should be completed. A stringent password policy is beneficial. One-time password schemes can also be used to further enhance the password security policy when establishing remote connections.

Remote access connections use standard authentication techniques (refer to Section 6.1.5, Firewall Technology Assessment, for more information regarding authentication).

Authentication, Authorization, and Accounting (AAA) for network access provides an additional level of security. AAA is the act of verifying a claimed identity, determining if the user has permission to access the requested resource, and collecting resource usage information for analyzing trends, auditing, billing or allocating costs. Message authentication plays a role when handling encrypted information. This verifies that the purported message sender is the person who really sent the message and that the message contents have not been altered. Although data can be authenticated at any hop on the way to the end destination, only the final destination may decrypt the data.

Refer to www.ietf.org/rfc/rfc2989.txt. [6] When remotely connecting to a company system, an alternative that also provides security is to establish a VPN. (See Section 5.3, System High Interconnections and Virtual Private Networks.)

Encryption of data is another common security measure. Encryption may be co-located with the firewall to provide secure tunnels to remote authorized users. Encoder/decoder products can be hardware- or software-based. Hardware-based solutions include PC cards (i.e., FORTEZZA), smart cards, or separate boxes attached to a network (for example, TACLANE, FASTLANE). For more information about FORTEZZA®, refer to <http://www.fortezza-support.com>. [7] There are also encryption software packages for encrypting e-mail such as Pretty Good Privacy (available free on the Internet, the site address is <http://www.wtvi.com/teks/pgp/>). [8] Software-based encoders/decoders also offer the capability of remote authentication, remote control, auto-answer secure data, and operation in both attended and unattended environments, therefore providing protection for facsimiles, e-mail, and computer communications. For further information on the FASTLANE and TACLANE refer to the FASTLANE category under Products & Services on General Dynamics' Web page, www.gd-cs.com. [9]

Users can also connect to their company's intranet via the Internet from a remote location. If a company's intranet is not configured properly, with some modification to the Internet site's URL, a hacker can gain access to the private intranet site. When setting up an intranet, access should be restricted to internally managed IP addresses only. Subnetting and access lists should also be implemented to allow only those permissible users within a company access to the Internet or certain intranet sites. Also, when establishing a virtual web or naming Web pages, make the names cryptic so the content is not obvious and make all pages that contain private information password protected. This will prevent unauthorized people—from outside and inside the organization—from gaining unauthorized access to information.

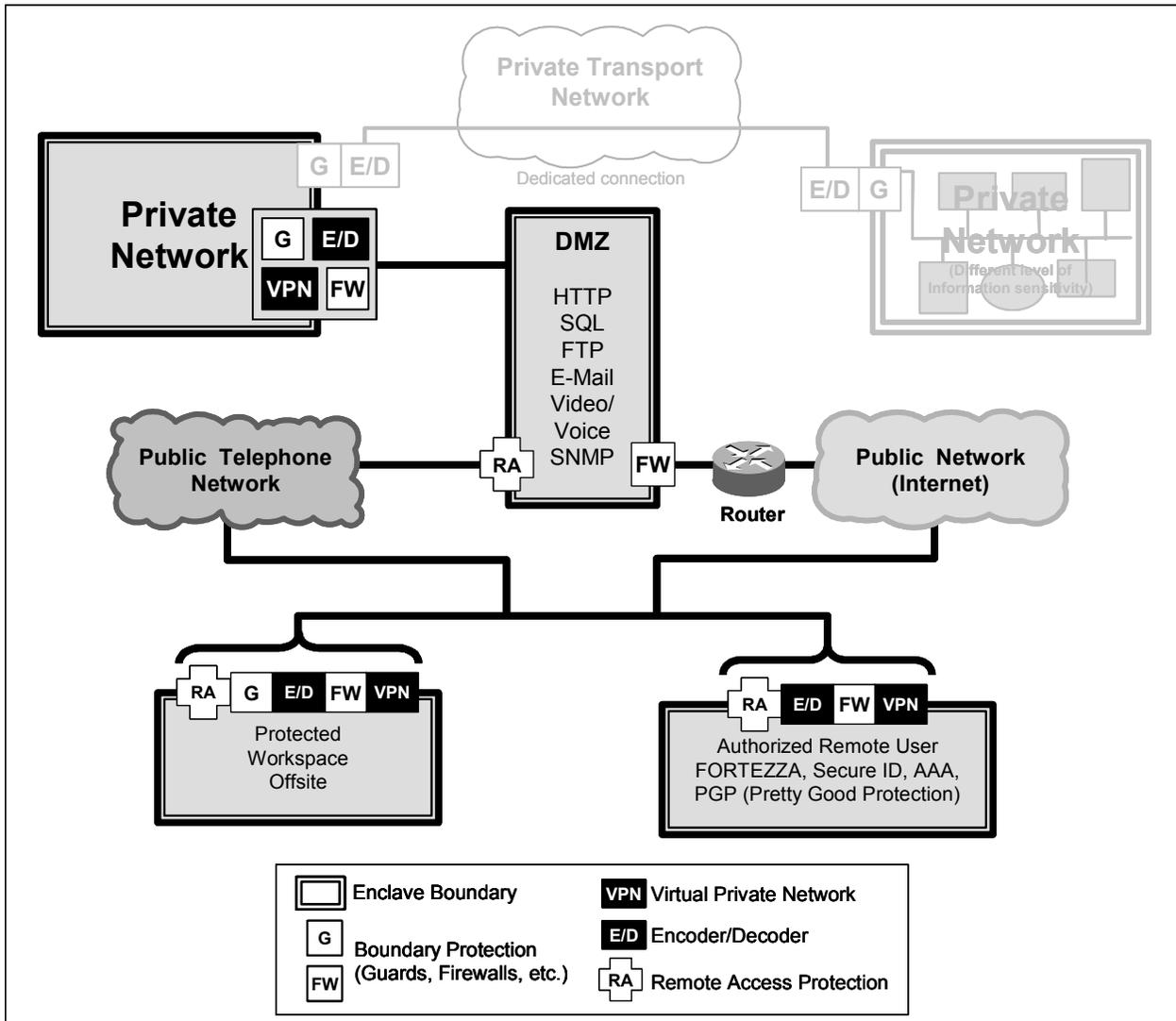
Case 3

Sensitive private network containing valuable information communicated through a lower level network to another network of equal classification/value (system high interconnects).

This case involves networks that are interconnected at essentially the same information sensitivity level, using a lower sensitivity level unprotected, public transmission media (Internet,

wireless). Referring to Figure 6.1-8, this scenario begins with the protected network containing proprietary data connecting via a public network to remote protected workspaces or valid remote users. At a minimum, this case requires:

- A boundary protection device (Firewall).
- A secure data connection device, i.e., encoder/decoder (KG, FASTLANE, TACLANE, FORTEZZA or other commercial-off-the-shelf [COTS]/government-off-the-shelf [GOTS]).
- A proactive audit capability to include COTS/GOTS intrusion detection products.



iatf_6_1_8_0108

Figure 6.1-8. Case 3—Private Network Connectivity via a Lower-Level Network

Medium assurance levels are required for the enclave boundary protection implementations. For this case, the recommended boundary protection procedures that should be implemented in priority order are:

- Institutionalize border security awareness and procedures as outlined in Chapters 3 and 4.
- Configure the local computing environment (home network) with built-in features and services for enclave boundary protection. Installation of firewall and/or comparable firewall feature set technology.
- Enable available audit capabilities to include firewall ingress and egress points and auditing of attempted resource connections.
- Scan for viruses using current virus definitions and profiles. Ensure that definition file databases are no more than a couple of weeks old.
- Perform a non-hostile vulnerability scan. Non-hostile scans include scans of: HTTP, FTP, Post Office Protocol (POP), SMTP, SNMP, ICMP, Telnet, Netbios, ensuring no deviations from initial network baseline scan.
- Perform comprehensive vulnerability scans to include: scans for non-standard UDP/TCP ports, unauthorized protocols, shares, unencrypted passwords, potential operating system related vulnerabilities.
- Add intrusion detection. Intrusion detection methods should include the ability to proactively monitor packets, log and alert appropriate personnel based on level of threat/probe, identify and record addresses of threat initiator(s).
- Couple scanning, monitoring, and testing with intrusion detection. A network is only as strong as its weakest link. By coupling scanning, monitoring, and testing—with intrusion detection—weaknesses and potential threats can be proactively identified upon first appearance or during the manifestation stage.

In addition, it is recommended that at least one staff person with an understanding of boundary protection be employed to configure and monitor the security parameters, perform virus and vulnerability scanning, and continually update the boundary protection and other security measures as vulnerabilities are detected and new intrusion detection capabilities become available.

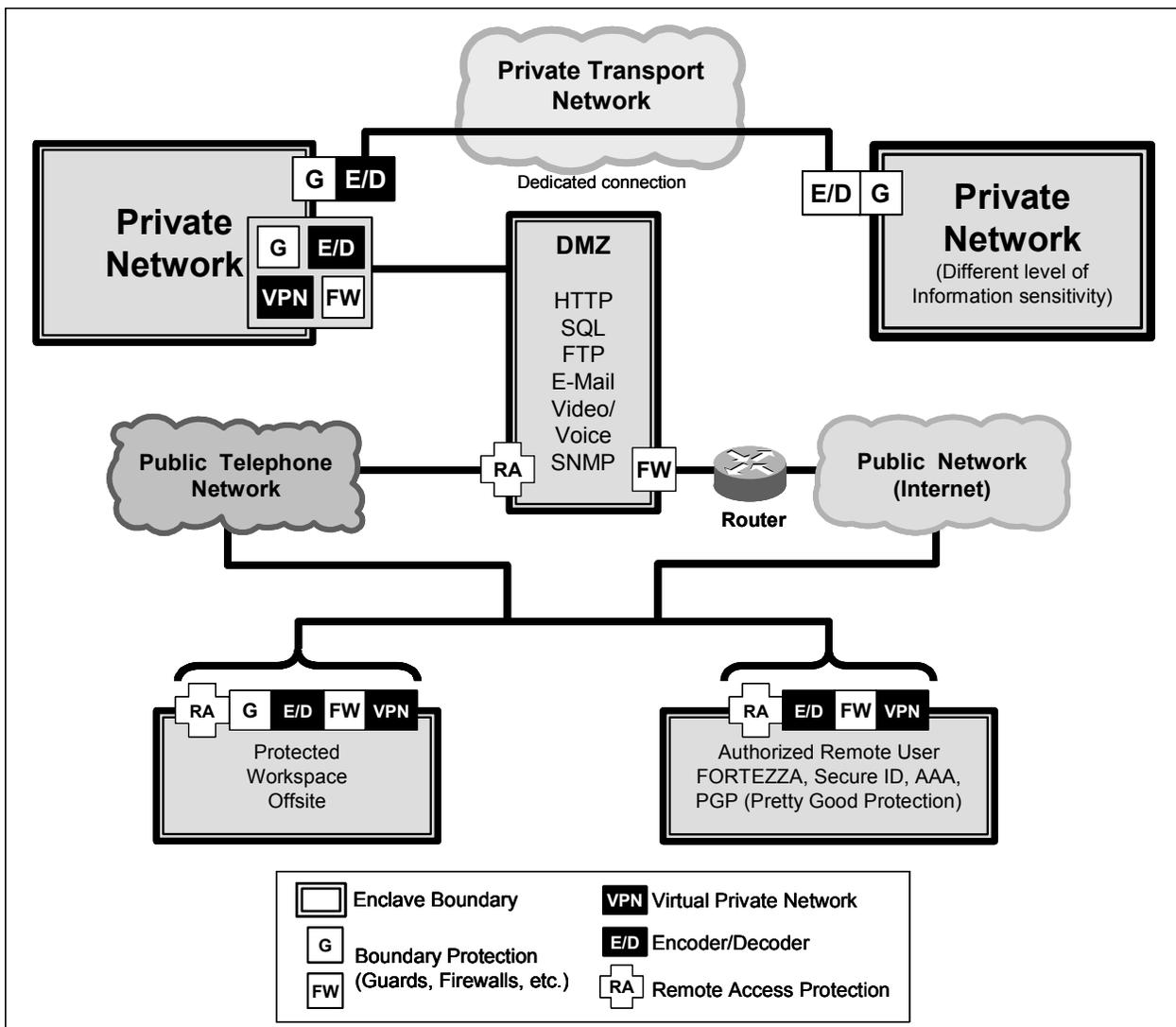
Software associated with the operating system, firewalls, and routers should be updated as the software continues to evolve with respect to built-in security features, especially as they relate to authentication and intrusion detection.

Case 4

Collaborating organizational LAN connecting to the main backbone network of the same classification, with public WAN connections to remote protected networks; e.g., North Atlantic

Treaty Organization (NATO) or foreign trusted network connected to main backbone network which is also connected to remote protected LAN(s) via a public WAN (Internet).

This case involves connections that may jeopardize interconnected high-level systems if users and administrators are not aware of the public-level WAN connection. As Figure 6.1-9 depicts, the unprotected network with proprietary data connects across a dedicated connection to the protected network with proprietary data, which is also connected to the public network/Internet and to remote users. The most basic level of protection for an enclave boundary includes employing the best available boundary protection technology (e.g., high assurance guards and intrusion detectors). Frequent virus and vulnerability scanning should also be performed by highly skilled personnel. An extensive security awareness program with institutionalized procedures for reporting and tracking is mandatory.



latf_6_1_9_0109

Figure 6.1-9. Case 4—Collaborative LAN's with Public Network Connections

The following scenarios require comprehensive protection from enclave boundary or network access point penetrations, employing the best available technology.

Collaborating LAN connecting to main LAN via dedicated connection.

The collaborating LAN (foreign company, NATO agency, etc.) is of the same information sensitivity level, and the anticipated threat level is at a minimum. Because the collaborating agency is accessing peripheral data, limited network resource access is required. Full access to all enclave contained information assets is not needed. Initiating an internal proxy server with a strict access security list is recommended (protected Solaris, local/global user access list via Microsoft's NT File System (NTFS) with auditing enabled). The collaborating LAN should be connected via a secure means, either through a data encoder/decoder (KG) or similarly approved security device. Intrusion detection monitoring products should include real-time auditing and tracking capabilities.

Protected off-site LAN with same security level connecting to main LAN via public WAN (Internet) with main site having a directly connected collaborating site.

All previously outlined security precautions need to be met (as defined by case studies 1, 2, and 3). The main LAN needs to have a strict access list in place (protected Solaris, local/global user access list via Microsoft's NTFS with auditing enabled). This precaution is to ensure that the connected collaborating LAN is able to access only predetermined enclave information assets, including resources at the main LAN as well as the off-site protected resources. To further ensure that only approved data is exchanged from the off-site LAN to the collaborating agency, it is recommended that guards be installed at both the ingress and egress location on the enclave boundary of the home enclave LAN.

The guards are present to ensure that only approved filtered data is exchanged between trusting and trusted networks/domains. Implemented intrusion detection monitoring products need to include real-time auditing and tracking capabilities.

Collaborating LAN connecting to protected remote site using main LAN's backbone.

All previously outlined security precautions need to be met (as defined by case studies 1, 2, and 3). If the *collaborating* LAN needs to connect directly to the off-site LAN without accessing any main LAN resources the following need to be addressed:

- A router or layer 3 switch is needed at the point of presence of the main LAN.
- A static route needs to be configured to route traffic directly to the off-site LAN via the main LAN's backbone.
- Data traffic needs to travel over the main LAN's encoders/decoders and through its DMZ.
- A guard needs to be installed at the boundary of the off-site LAN.

UNCLASSIFIED

Firewalls

IATF Release 3.1—September 2002

The purpose of this type of configuration is to prevent a direct association between an off-site and collaborative LAN (i.e., a foreign organization/agency that is communicating with a local company or agency, the main LAN, acts as a go-between).

- For this case and the associated scenarios, the recommended boundary protection procedures are similar to the previous recommendations, but require higher-assurance boundary protection technology implementations. The following recommendations should be implemented as a comprehensive package with reference to which scenario the network most resembles.
- Institutionalize boundary security awareness and procedures. As outlined in Chapters 3 and 4.
- Configure the home enclave network using built-in features and services for boundary protection. Installation of firewall and or comparable firewall feature set technology.
- Enable available audit capabilities to include firewalls, ingress and egress points and auditing of attempted resource connections.
- Scan for viruses using current virus definitions and profiles. Ensure that definition file databases are no more than a couple of weeks old.
- Perform a non-hostile vulnerability scan. Non-hostile scans include scans of HTTP, FTP, POP, SMTP, SNMP, ICMP, Telnet, Netbios, ensuring no deviations from initial network baseline scan.
- Frequently perform comprehensive vulnerability scans including scans for non-standard UDP/TCP ports, unauthorized protocols, shares, unencrypted passwords, potential operating system-related vulnerabilities.
- Incorporate enterprise-wide intrusion detection. Intrusion detection methods should include the ability to proactively monitor packets, log and alert appropriate personnel based on level of threat/probe, identify and record routing addresses of threat initiator(s).
- Incorporate infrastructure attack “early warning.”
- Employ supplementary boundary protection between off-site locations. (firewall/guard services).
- Couple scanning, monitoring, testing, and intrusion detection. A network is only as strong as its weakest link. By coupling scanning, monitoring, testing, and intrusion detection, weaknesses and potential threats can be identified upon first appearance or during the manifestation stage.

6.1.7 Enclave Boundary Protection Framework Guidance

The technologies discussed in this section and the types of techniques they employ should typically be composed to form a solution set to defend the enclave boundary. Although the technologies overlap, each focuses on a different subset of security countermeasures. Additional access control mechanisms should also be used in forming mitigation approach sets. These include encryption or application-layer discretionary access controls to permit or deny access to specific data within an enclave. Given these countermeasures, it must be determined how, where, in how many places, and how many times they should be applied. Places to which the countermeasures can be applied include at the enclave boundary, workstation/LAN interface, individual workstations, servers, operating systems, or at the application level. A layered security approach can be used, determining how many places a countermeasure should be applied. How many times a countermeasure should be applied is the choice between per session authentication and per packet authentication. It must also be determined how strong the security measures must be.

A number of factors generally influence the selection of firewall approaches. The mission needs and services desired by the users are primary factors in shaping mitigation approach sets. The risks to a given system must be assessed in terms of:

- The differences in information value and threat between the protected enclave information assets and the external networks to which it is connected.
- The environments and architecture.
- The impacts of potential attacks.

In addition, cost, policy mandates, scalability, maintainability, and overhead (including performance degradation and manpower) must be considered. Clearly, the specific protection approaches and products selected also must be those that can address the specific services, protocols, operating systems, applications, and components employed in the user's environment. Ideally, the technologies that incorporate all prescribed countermeasures, at the appropriate levels, and addressing all aspects of the specific user environment should be implemented. As indicated in Section 6.1.5, Firewall Technology Assessment, and below, there are gaps in successful achievement of countermeasures, performance, and other areas.

Potential negative impacts are associated with any of the technology solutions. Desired performance of a firewall must be determined when implementing a firewall to defend the enclave boundary. There is a trade-off between speed and security. A network can be more secure when the firewall performs more checking on the packets. However, the amount of checking that a firewall performs has an effect on the volume and the speed at which traffic can transverse the enclave boundary protection.

UNCLASSIFIED

Firewalls

IATF Release 3.1—September 2002

In addition, while greater restrictions to operations do yield greater protection of the enclave assets, the restriction of dangerous operations also restricts useful operations. There comes a point at which the tradeoff for greater security becomes more than the users want to pay in lost capability or hampered performance. For example, some antiviral and disinfectant (subversion-constrained) software may actually do as much damage to operational performance as viruses themselves might. Some systems may fail to prevent infections but prevent the user from eliminating the virus. Some antiviral systems may actually delete files without alerting the user or offering alternative approaches. Disinfecting has been known to leave workstations in a worse state than the infection did. The primary approach to selection of security protection should be to maximize benefits while minimizing harm. Only through a comprehensive risk analysis, with knowledge of the characteristics and trade-offs of different technologies and specific products including cost and resource constraints, can effective enclave boundary protection be implemented and maintained.

The first step in any effort to implement an enclave boundary protection mechanism and additional technology to protect the enclave information assets is to develop a security policy. The boundary protection mechanisms will then serve to implement this security policy. An in-depth requirement analysis forms the basis for the development of the policy and subsequent selection of protection devices.

Clearly, the environment in question will dictate the level of security robustness. For example, in connecting enclaves of different classifications, whether through a direct connection or through another network, additional security precautions must be taken. Remote access to the enclave through the boundary protection mechanism will require security mechanisms designed specifically for this situation. Firewalls, for example, generally have the capability to form an encrypted link to the remote user. Boundary protection mechanisms, which are used inside the enclave to limit access to restricted information, on the other hand, tend to be cheaper and less complex than those devices located at the boundary of the entire enterprise. Firewall technology has evolved so that firewalls are now developed and marketed specifically for intranet firewall applications.

In addition to the specific environment in question, there are a number of general trade-offs, which should be addressed when implementing firewall technology. One important trade-off with regard to firewall technology is between security and ease-of-use. The more rigorous the checks for user identity and user activity, the more inconvenience the user must endure. On the other hand, if the firewall simply passes everything through to the internal network, security is inadequate, even for the least sensitive data. In choosing a firewall, both the needs of the users for services and the security requirements must be balanced; otherwise, the users will find ways to bypass the firewall, weakening the protection of the enclave boundary.

Packet filters and stateful packet inspection technologies focus on flexibility. In general, these firewalls are able to support many services, and additional services can be easily added. However, this flexibility comes with a price. It is quite easy to configure these types of firewalls to permit dangerous access to services through the firewall. The ease-of-use administrative interfaces and preconfigured support for many services lend themselves to configuration errors. Application gateways, on the other hand, provide better auditing and finer grained control. For

example, application gateways can be used to allow certain activities, such as sending a file to an untrusted network, while blocking a user from copying a file from an untrusted network. In general, router-based firewalls are best for a dynamic environment where lots of things change in a short time frame. Application-level firewalls are better if a more deliberate approach to security is necessary.

Other considerations in selecting a firewall include the skill level available for maintaining the firewall. As noted above, proper configuration and maintenance of the firewall is a critical security element. If an organization does not have the staffing to assign qualified personnel to operate and maintain the firewall, there are options to purchase firewall maintenance services, from either the firewall company or the ISP. These costs of staffing or services should be considered, as well as the corporate credentials of the firewall vendor, and the quality of the documentation available with the firewall.

UNCLASSIFIED

Firewalls
IATF Release 3.1—September 2002

References

1. B. Frasier. Site Security Handbook RFC 2196. September 1997
<http://www.ietf.org/rfc/rfc2196.txt?number=2196>.
2. SOCKS. 1 May 2000 <http://www.socks.nec.com>.
FTP Directory. 1 May 2000 <ftp://ftp.nec.com/pub/socks>.
3. Rekhter Y., et al. "Address Allocation for Private Internets. RFC 1918." February 1996
<http://www.ietf.org/rfc/rfc1918.txt?number=1918>.
4. Ferguson P. and D. Senie. "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing." 18 May 2000
<http://www.ietf.org/rfc/rfc2267.txt?number=2267>.
5. AAA Working Group. "Criteria for Evaluating AAA Protocols for Network Access." 26 April 2000. On line posting. 11 May 2000
<http://www.ietf.org/rfc/rfc2989.txt>.
6. FORTEZZA Cryptography of the 21st Century. 12 May 2000.
<http://www.fortezza-support.com>.
7. Pretty Good Privacy Software. 12 May 2000 <http://www.wtvi.com/teks/pggp/>.
8. General Dynamics Communications System. 12 May 2000 www.gd-cs.com.

Additional References

- a. Cisco Systems, Inc. "How Data Moves Through The Firewall." 19 May 2000
http://www.cisco.com/univercd/cc/td/doc/product/iaabu/pix/pix_v41/pixcfg41/pix41int.htm#xtocid297201.
- b. Computer Security Resource Center. 1 May 2000 <http://csrc.nist.gov/>.
- c. Internet/Network Security. 1 May 2000
<http://www.netsecurity.about.com/compute/netsecurity>.
- d. Defense Information Systems Agency. Firewall Configuration Guide, 12 June 1998.
- e. Internet/Network Security site. "The Secure Telecommuters FAQ" Page 10 May 2000
<http://netsecurity.about.com/compute/netsecurity/library/weekly/aa020200c.htm>.
- f. National Security Agency/Network Boundary IA. Department of Defense Firewall Guidance. Version 1.0 Draft, 31 March 2000.
- g. Network Vulnerability Analysis and Penetration Testing. 8 May 2000
<http://www.blackmagic.com/assessment.html>.
- h. The Source of JAVA™ Technology. "Applets." 8 May 2000
<http://www.java.sun.com/applets/index.html>.

UNCLASSIFIED

Firewalls
IATF Release 3.1—September 2002

- i. United States Navy Web Information Service. 12 May 2000
<http://infosec.navy.mil/products/securevoice/stu3.html>.
Enter at <<http://infosec.navy.mil>>, then, navigate to:
<http://infosec.navy.mil/products/securevoice/stu3.html>.

UNCLASSIFIED

Firewalls
IATF Release 3.1—September 2002

This page intentionally left blank.

6.2 Remote Access

Remote access enables traveling or telecommuting users to securely access their Local Area Networks (LAN), local enclaves, or local enterprise-computing environments via telephone or commercial data networks. Remote access capability draws on both the virtual private networks (VPN) and the Defending the Enclave Boundary sections of this document. The remote access user connects by a shared commercial path, and can maintain the privacy of his or her connection using encrypting modems, technologies applicable to VPN needs (as discussed in Section 5.3, System-High Interconnections and Virtual Private Networks), or other technologies suitable to this requirement. Because the user entry point into the enterprise-computing environment could be used by a hostile connection, the enterprise must implement enclave boundary protection (as discussed in Section 6.1, Firewalls). The remote user's computing assets are also physically vulnerable, requiring additional protection. This section draws on the preceding two and explores protection for information storage to address the specific problem of remote access.

Note that although section 5.3, System High Interconnections and Virtual Private Networks, discusses VPNs, the discussion in that section focuses more on 'tunneling' data between enclaves over public networks or private networks of equal or lesser classifications. The discussion also covers what is termed 'bulk-encryption,' where it is an all or nothing protection paradigm. In the context of remote access, a more up-to-date definition of a VPN is a protected communications channel that protects data-in-transit between two points concurrently with unprotected data over a common, untrusted communications infrastructure. Therefore, this section will also discuss the importance of VPNs for the remote access user.

6.2.1 Target Environment

Within this section, traveling users and telecommuters are both treated as remote users. However, the environment of these two groups differs in the degree of physical exposure of the remote computer. The traveler's computer is vulnerable to theft and tampering while the user is in transit and while their computer is in storage. These risks are particularly great overseas. The telecommuter's computer is also vulnerable to theft and tampering, but to a much lesser extent if the physical location of the hardware is within Continental United States (CONUS). In addition, because the telecommuter's remote location is relatively fixed, additional steps can be taken for physical protection that are not feasible for traveling users. Conversely, the telecommuter's fixed remote location makes targeting by an adversary easier than in the case of mobile traveling users.

As depicted in Figure 6.2-1, remote users access their enterprise-computing environments by communication paths shared with others. Many remote users employ the Public Switched Telephone Network (PSTN) to access their home enclave directly or use the PSTN to connect to a data network such as an Internet Service Provider (ISP) that connects users to their enterprise-computing environment. Other remote users employ broadband communications technologies, including digital wireless service, cable modems, Integrated Services Digital Network (ISDN), and other high-data-rate media. Remote access via these networks increases the level of threat and imposes architectural constraints to the security solution. This section of the Information

Assurance Technical Framework (IATF) treats remote access, via these networks, separately from direct dial-in to an enterprise-computing environment via PSTN.

Note that for this section, remote access is limited to the capability of providing access to the information contained in users' local system-high LANs, enclaves, or enterprise-computing environments from remote locations, which, during the period of connectivity, are assumed to be controlled at the same system-high level as the local system. In other words, remote users with authorized access to unclassified information that is either sensitive or not will be given access to the unclassified information contained in their local unclassified system-high enclaves and remote users authorized access to secret information will be given access to secret information contained in their local secret system-high enclaves.

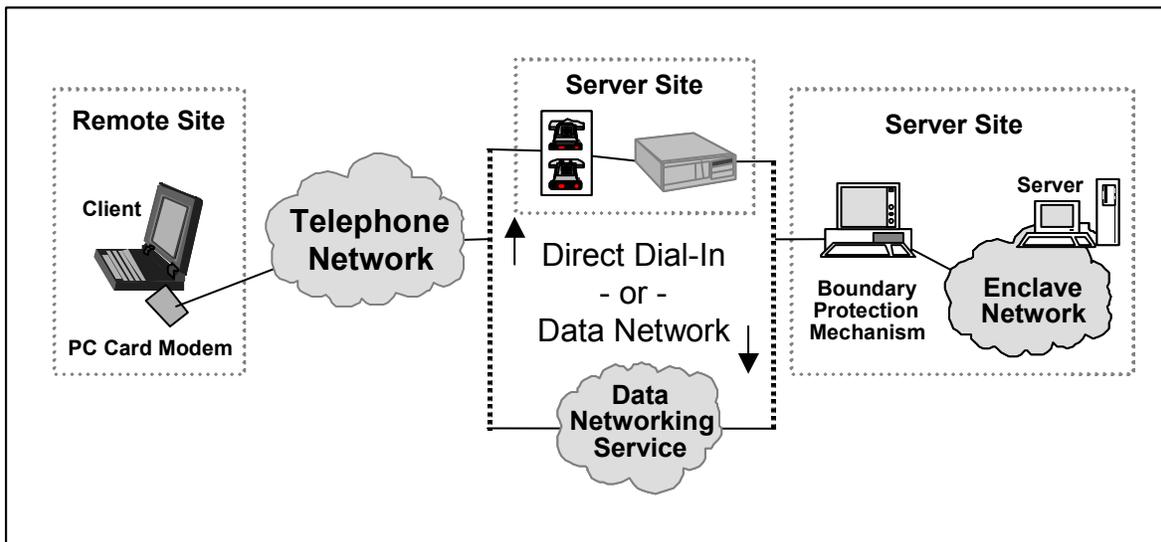


Figure 6.2-1. Typical Remote Access Environment

In the case of secret remote connectivity, the proposed remote connectivity approach will give the remote user the ability to store information on the remote terminal (typically a notebook computer) hard drive in an encrypted format, thereby declassifying the terminal when it is not in operation. However, during the period of connectivity to the home system, the remote user must provide sufficient physical protection and safeguarding of the secret information being processed.

6.2.2 Consolidated Requirements

6.2.2.1 Functional Requirements

The following requirements are from the user's perspective.

- Remote users should have access to all information stored on their remote computers, stored on their home enclave workstation, or available within their home enclave

information infrastructure. Because remote users need to conduct their business using familiar tools while traveling to a remote location, cryptographic application interfaces on the remote user's terminal should be similar and have the "same look and feel" as those provided at their home enclave. Applications that may be launched from a system-high enclave as a result of a remote user request, shall continue to support all security services as required by the enclave system security policy and procedures.

- The user should know when security features are enabled. Indications should not be intrusive, but the user should be able to tell easily when security features are working, and more important, when they are not. Feedback to the user is very important in any security solution.
- The security solution should have minimal operational impact on the user. It should not impose a significant performance penalty, or require extensive training.
- The traveling user's security suite should not include any external devices. Some remote users simply do not have room for these devices in their computing packages. Solutions that are unobtrusive to the user (e.g., user tokens and software products) are preferred.
- The remote user's equipment should be unclassified when it is unattended. Both the data stored on the remote user's computer and the approved configuration of the remote user's computer must be protected from unauthorized disclosure, modification, or manipulation when out of the direct control of the authorized remote user. This protection must effectively protect the computer and stored data from compromise if the computer is lost, stolen, or used to communicate with lesser security level authorized hosts. Assuming the data stored on the remote user's equipment is appropriately protected, the user is required to safeguard the terminal as would be required of high-value items.
- The remote user should not have greater access than would be available if accessing the enclave information resources from within the enclave.

6.2.2.2 Interoperability

Remote access systems that implement interoperable solutions facilitate the movement of users between organizations and increase the likelihood that the system can be supported and upgraded in the future. Interoperability also provides for the maximum evolution of this security solution in the commercial marketplace. For these reasons, the following interoperability requirement is added.

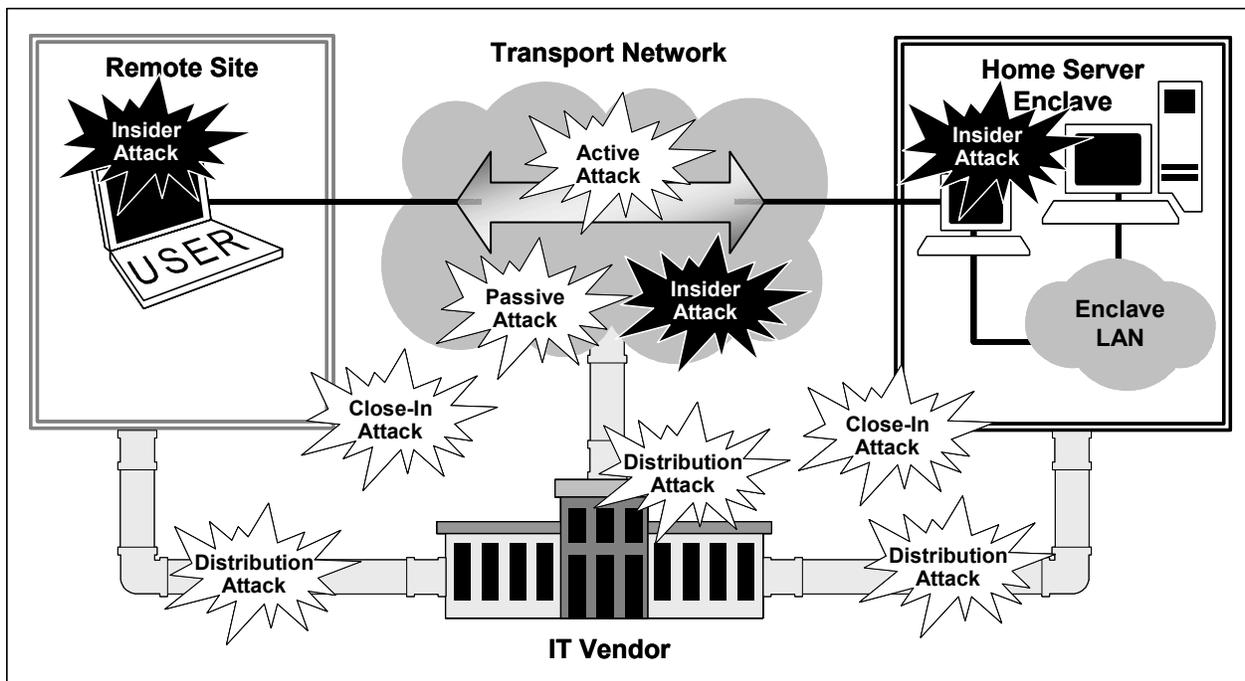
Security solutions should be based on open standards. The use of proprietary implementations creates significant issues related to interoperability and logistics support. To ensure an effective solution, the remote access mechanism should integrate easily into existing information systems and provide a path for upgrading to emerging technology (as discussed below).

6.2.2.3 Emerging Technology

It is desirable that the security solutions be capable of evolving to higher data rates and be adaptable to alternative means of communication, such as cellular telephony, wireless networks and ISDN.

6.2.3 Potential Attacks

All five classes of attacks introduced in Chapter 4, Technical Security Countermeasures are of concern in the remote access scenario. Section 6.1, the Firewalls section goes into detail on network attacks. The VPN's section's (Section 5.3) treatment of passive, network, and insider attacks is directly relevant to remote access. Since proper configuration and execution of software is critical to the proper functioning of security mechanisms, distribution attacks are also a concern. Remote access places the user's computer in public environments, adding the possibility of physical attack to the five generic attack classes. With reference to Figure 6.2-2, the following summarizes potential attacks against the remote access scenario.



iatf_6_2_2_0111

Figure 6.2-2. Attacks Against the Remote Access Scenario

6.2.3.1 Passive Attacks

An attacker monitoring the network could capture user or enclave data, resulting in compromise of information. Capture of authentication data could enable an attacker to launch a subsequent

network attack. Analysis of traffic captured by passive monitoring can give an adversary some indication of current or impending actions. Compromising emanations could also be intercepted.

6.2.3.2 Active Attacks

These attacks are most likely to originate from the Internet, but, with more effort, could also be mounted through the PSTN. Also attacks can target the remote user's computer, the user's enclave, or the user's connection to the enclave, potentially resulting in the loss of data integrity and confidentiality, and ultimately in the loss of use of the network by authorized users (e.g., a denial-of-service attack).

6.2.3.3 Insider Attacks

An insider is anyone having physical access to the remote user's computer or the network enclave from within the user organization's corporate boundaries. These attacks could be motivated by malice or could result from unintentional mistakes by the user. Deliberate attacks can be especially damaging to the organization's information system due to the attacker's access to the information, their advantage of knowing the network's configuration, and thus their capability to exploit the network's vulnerabilities.

6.2.3.4 Distribution Attacks

Distribution attacks could occur at the Information Technology (IT) provider's site while the product is developed, manufactured and shipped, while the remote user's computer is being configured or maintained, or when software is passed to the user's computer (including software passed over the network). This type of attack could result in a network's device (e.g., firewall, router, etc.) being used to perform a function for which it was not intended, thus making the remote access capability or the enclave vulnerable to attack.

6.2.3.5 Close-In Attacks

The remote user's computer is subject to theft and tampering. Physical attack also could result in the theft of the traveling user's computer, a denial-of-service attack. Typically, there are non-technical countermeasures (e.g., procedures) available for dealing with physical threats. The Framework addresses these since there are also technical countermeasures available that could help to mitigate those threats.

6.2.4 Potential Countermeasures

The following security services are required to counter the potential attacks against the enclave.

- Strong and continuous user authentication should be the basis for allowing access to the enclave. Strong continuous two-way authentication protects the enclave, the remote user, and the connection from network attacks. Cryptography-based authentication at the

enclave boundary ensures that only authorized users can gain access to the network. Use of a boundary protection mechanism is used in conjunction with cryptography-based authentication to provide a basis for controlling a user's access to individual network services. Continuous authentication prevents an unauthorized user from hijacking the remote user's session.

- Confidentiality may be invoked for all information flowing between the enclave and the remote user's computer. Confidentiality guards the enclave and the remote user from passive intercept attacks. Although encryption does little to guard against traffic analysis, the data and metadata (information about data) are protected against direct intercept and compromise. This security service is dependent, of course, on the level of required protection afforded the data.

- The information in the remote user's computer should be protected:

When the computer is not in use. This protects the information in case of theft of the workstation, or unauthorized physical access.

When the computer is connected to unclassified or untrusted networks. This guards against network attacks (e.g., session hijacking) from an unclassified and/or unauthorized network.

- The integrity of the remote user's hardware and software should be protected. Detection and protection mechanisms can guard against distribution attacks, tampering by an outsider, and physical access by an unauthorized user.
- The integrity of data flowing between the remote user's computer and his enterprise-networking environment should be protected. This protection is typically provided at the applications layer. See Section 7.1, Security for System Applications of the Framework for details.

6.2.5 Technology Assessment

The three technologies—media and file protection, workstation integrity, and enclave and connection protection—are included in this section and depicted in Figure 6.2-3 counters specific types of attacks. Some attacks, such as tampering, are only partially addressed by technical measures. Non-technical security measures, as discussed in Chapter 4, Technical Principles—physical protection of the laptop, prevention of casual “over-the-shoulder” observation of classified information—are critical to overall system security and should be considered a vital part of a remote access user policy. This section of the Framework only covers those technical measures that will counter attacks relevant to the remote access category.

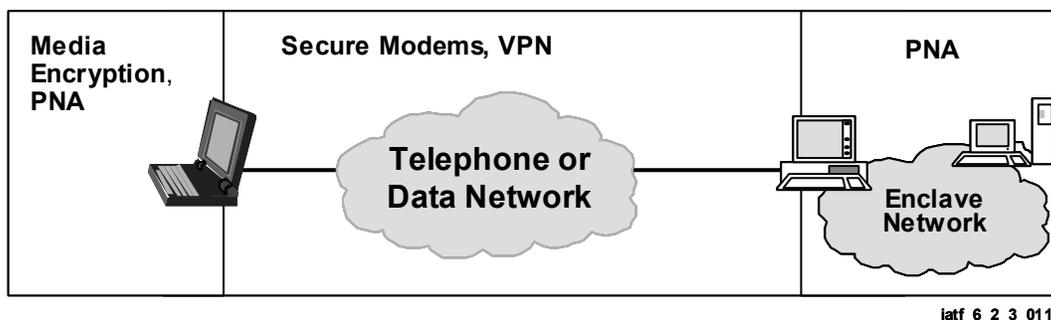


Figure 6.2-3. Security Technologies in the Remote Access Scenario

6.2.5.1 Media and File Encryptors

In some cases, physical removal of the remote computer storage media (typically a hard drive) between remote connection sessions is not acceptable. Encryption of the information on the storage media can provide confidentiality and integrity, alleviating the need for physical removal of the media. Media encryptors and file encryptors protect the information in the computer in the event of unauthorized physical access to the computer. File encryptors can protect the confidentiality and integrity of individual files, provide a means of authenticating a file's source, and allow the exchange of encrypted files between computers. Media encryptors protect the confidentiality and integrity of the contents of data storage media. For example, they can help maintain the integrity of the remote user's computer by verifying the Basic Input/Output System (BIOS) and ensuring that configuration and program files are not modified.

With the exception of some system files, media encryptors encrypt the entire contents of the drive. The media encryptors must leave some system files unencrypted so that the computer can boot from the hard drive. The integrity of most of these unencrypted system files can be protected by a cryptographic checksum; this protection will not prevent a tamper attack, but it will alert the user that that data has been altered. System files contain data that changes when the computer is booted and cannot be protected.

File encryptors typically implement a graphical users interface (GUI) that allows users to choose files to be encrypted or decrypted. This protects individual files, but it does not protect all files on the drive. Many applications generate temporary files that may contain user data. These files are normally closed (but not necessarily erased) when the application is terminated. However, the application does not terminate in an orderly fashion; these temporary files may remain open. Some operating systems do not actually erase data when files are closed or deleted. Instead, they alter the name of the file in the file allocation table or de-allocate the storage locations on the media. The user's data then remains on the hard drive until the space is allocated to another file and overwritten. Thus, unencrypted and potentially classified user data can remain on the hard drive after system shutdown, either because of the application's failure to erase temporary files or by the design of the operating system's file closure function. For these reasons, media encryptors provide better protection for the information on the disk drive—especially while the computer is not in use—than do file encryptors.

Media encryption's robustness is an advantage only when proper key management is used in protecting the information. There must be provisions to allow trusted key management to protect the key when encrypting the media and when the key is in storage. See Section 6.2.7, Framework Guidance of this chapter for further discussion of the secret dial-in case. Media encryption also supports workstation integrity, the topic of the next section.

6.2.5.2 Workstation Integrity

Workstation integrity components are necessary to protect the integrity of a remote computer's operation and data against active (network-based) and software-distribution threats. Active attacks include attempts to steal data by circumventing or breaking security features, or by introducing malicious code. The software distribution threat refers to the potential for malicious modification of software between the time it is produced by a developer and its installation and use on the remote user's computer.

Workstation integrity mechanisms to counter active attacks are addressed in the Firewalls section of the Framework. Products for detecting and removing computer viruses are available for both the workstation and boundary protection mechanism. Media encryption protects the configuration and software of the remote user's computer against malicious modification during the operational phase; it does not address this modification during the developmental or the distribution phases. Trusted operating systems can ensure the policy-enforced relationships between subjects and objects, thus limiting any effects the malicious code introduced into the machine might have on the system's integrity.

Software distribution attacks are discussed in Chapter 4, Technical Security Countermeasures. Most software distribution attacks can be thwarted by the use of digital signatures. Software can be signed at the manufacturer before distribution; these signatures are verified before the software is installed on the user's computer. Commercial file encryption packages containing this capability are available.

6.2.5.3 Enclave Boundary and Connection Protection

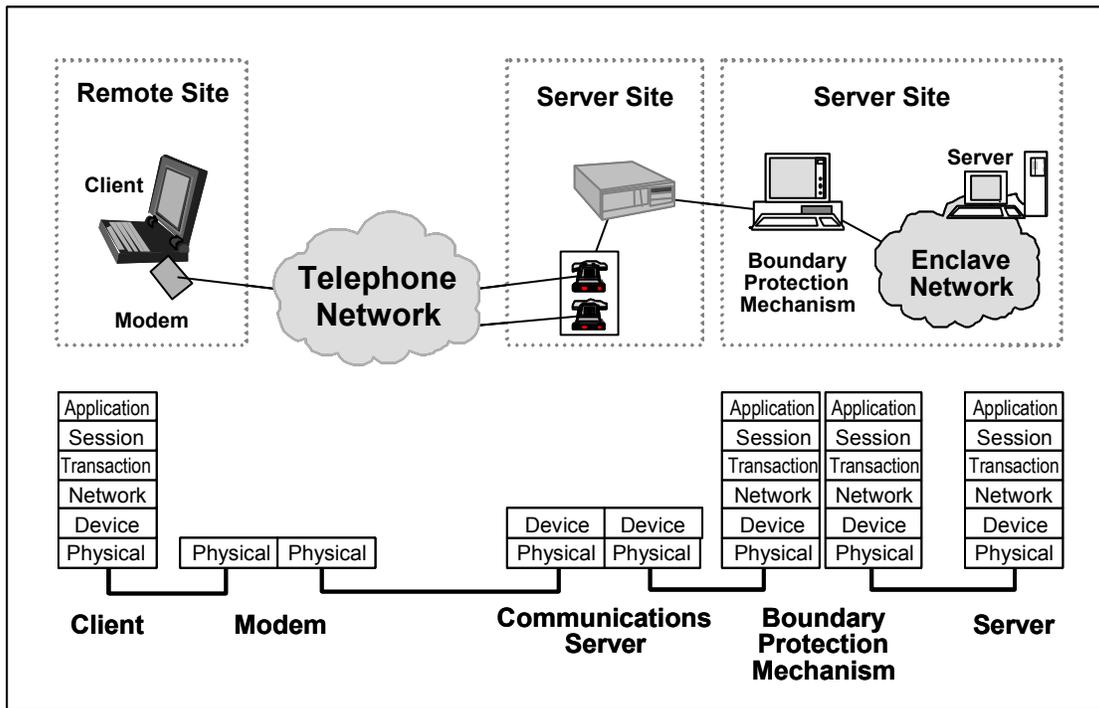
Components to implement authentication, confidentiality, and integrity mechanisms can operate at several layers in the protocol stack, with trade-offs in assurance, performance, and networks supported. Starting toward the bottom of the protocol stack, options include secure modems, data link layer technologies, network layer products, transport and session layer products, and application layer products. The protocol layer chosen does not necessarily imply a certain level of information assurance. There are mechanisms that can provide either at a high level of assurance, a low level of assurance, or something in-between at any protocol layer. Connection protection is dependent on an organization's risk management decision concerning the level of assurance placed on these mechanisms. All of these approaches, except application layer protocols are discussed in the VPN section (Section 5.3, System-High Interconnections and Virtual Private Networks). The authentication mechanism should provide mutual authentication of the remote user and the enclave's boundary protection mechanism, which is described in the Firewalls and Guards sections (Sections 6.1 and 6.3, respectively) and shown in Figure 6.2-1. It

also shows both options for connecting to the enclave—by direct dial-in to the enclave and by an ISP. Figure 6.2-4 shows the protocol layers associated with the remote access scenario.

Secure Modems (Physical Layer Mechanisms)

Secure modems offer an inherent means of boundary protection: the identity of the remote user’s modem is established by strong authentication before any network connections are initialized, preventing unauthorized modems from attempting an active attack. The invocation of encryption within a modem provides a high level of assurance provided that the encryption function is properly invoked and is protected from tampering. However, the implementation of additional features, such as plaintext bypass, can reduce some of that assurance. For instance, a secure modem needs a means of bypassing the encryption engine if it is also to interoperate with a nonsecure modem. Any bypass feature in a secure modem must be carefully implemented so it is not possible to bypass the cryptography accidentally or maliciously.

Strong authentication requires a significant cryptographic processing capability both in the calculations required to validate a signature and in the verification of the identity contained in a certificate (e.g., checking against a list of authorized users). The identity that is established by modem authentication may not necessarily be made available to the network. This requires the remote user to log into the network separately.



iatf_6_2_4_0113

Figure 6.2-4. Protocol Layers In Remote Access Scenario

Data Link Mechanisms

Data link layer protocols such as Point-to-Point Protocol (PPP) and Serial Line Internet Protocol (SLIP) encapsulate network layer packets for transmission via modems. Security services can be applied to these protocols to allow authentication and protect the connection between the remote user and the home enclave's communication server. Unlike the large bandwidth data links discussed in the VPN section, the remote user's data link is dedicated, so authentication of individual users is possible. This assumes, of course, that the remote machine is dedicated to one (and only one) user because authentication at the data link layer relies on lower level physical addresses versus those on higher layers that can distinguish among multiple users (e.g., with user Identifications [ID]).

Data link mechanisms allow users to choose their own modem hardware and upgrade or change it at their convenience, provided that the hardware can interoperate with the enclave's boundary communications hardware. A server implementing a data link mechanism could use the results of cryptographic authentication as a basis for access to the enclave. Data link security mechanisms are likely to be implemented in workstation software, where processing power and memory are more readily available than in the case of special-purpose security hardware. This makes implementation functions such as continuous authentication and certificate path validation more practical. However, it also makes these functions dependent on the integrity of the workstation on which they are running and more vulnerable to implementation errors and subversion.

At the data link layer, no information is available about the network resources or services the remote user is attempting to access. Any filtering mechanism would need to be implemented at a higher layer of the protocol stack.

Network Layer Mechanisms

Network layer protocols, such as Internet Protocol (IP), assign addresses to devices and pass data packets between them. ISPs assign an IP address to the remote user and pass IP packets for the remote user. For this reason, the network layer is the lowest layer at which security services can be applied in the ISP case. The VPN section addresses IP connections across public networks, and recommends the use of Internet Protocol Security (IPSec) with both Encapsulated Security Protocol (ESP) and Authentication Headers (AH). The VPN section also recommends the use of external encryptors. The current generation of external encryptors must be configured by a trained operator and are expensive and relatively bulky, so external encryptors are currently unfeasible for remote access. However, IPSec mechanisms are implemented in network card hardware, in modem cards, and in software on the user's computer (as before, the proper functioning of software mechanisms depends on the integrity of the user's computer).

Network layer mechanisms allow strong authentication directly from the remote user's computer to the boundary protection device, allowing the boundary protection device to base access control decisions on the user's identity. Network layer information allows the boundary protection mechanism to filter access to individual machines in the enclave. The downside is that they leave all of the enclave's dial-in equipment before the network device—specifically the

modems and the communications server—exposed to network attacks. Provided that the communications servers are properly configured and controlled, the potential for successful attacks against a communications server is relatively low (except for denial-of-service attacks). Remote control and administration of these devices can make the network vulnerable to attack by providing potential access to root level privileges. Please refer to Section 6.1 (Firewalls) for more information.

Transport and Session Layer Mechanisms

The transport layer forms a reliable channel between devices. The session layer establishes and synchronizes a communication session between two devices. The transport or socket layer is the lowest layer with information on the service being accessed so that security services can be called on a per application basis. The transport and session layers are discussed in the VPN section (Section 5.3). For the remote access scenario, these layers share many of the advantages and disadvantages of network layer mechanisms—they can allow continuous authentication directly to the boundary protection mechanism and allow further access control decisions based on the cryptographically authenticated identity. Transport and session layer mechanisms are not likely to be hardware-based, making them vulnerable to tampering and dependent on the integrity of the user's computer.

The Transport Layer Security (TLS) protocol, which sits at the top of the transport layer, is listed on the Internet Engineering Task Force (IETF) website www.ietf.org as RFC 2246. Product implementations of socket mechanisms should comply with the IETF standard, which is currently TSL.

The Remote Access Dial-in User Service (RADIUS) protocol (RFC 2138) was designed to authenticate remote users using a shared secret. The RADIUS protocol is currently an Internet Draft published by the IETF. Authentication requests are handled by a centrally located authentication server, which provides a method of supporting the management of remote users. The access requests made by RADIUS clients are capable of carrying attributes that include user name, user password, client identification, physical port identification, or other information. When passwords are present, they are protected by using RSA MD5. The ability of RADIUS to support a wide range of client attributes used in access control decisions makes this protocol very flexible. Access privileges can be varied for each user, as well as for the access method each user attempts. Maintaining a central RADIUS server, which controls the privileges for each user, makes RADIUS authentication scalable to handle large numbers of remote users.

Application Layer Mechanisms

Application layer security, invoked based on-site policy, supports the highest level of filtering. Individual commands within applications, as well as access to specific machines and services, can be permitted or denied. Application layer mechanisms are discussed in the opening part of the VPN Section 5.3. One of the major shortcomings of application layer mechanisms is that they rely on platforms with minimal trust mechanisms and that connections must be established at a lower level in the protocol stack (network and transport layer) before the application mechanisms are applied. This leaves the machine vulnerable to network attacks that are

unaffected by higher-layer security mechanisms. The other drawback of application layer security is the number of applications that need to be covered. As application protocols evolve, security is usually a secondary consideration. The number of application software packages offered in the commercial market (for example, e-mail packages) makes it difficult to add security services to every package as a retrofit. Efforts to standardize the interface to security services will help this problem, but are ineffective if the vendor is simply not interested in implementing security services in the product.

6.2.6 Cases

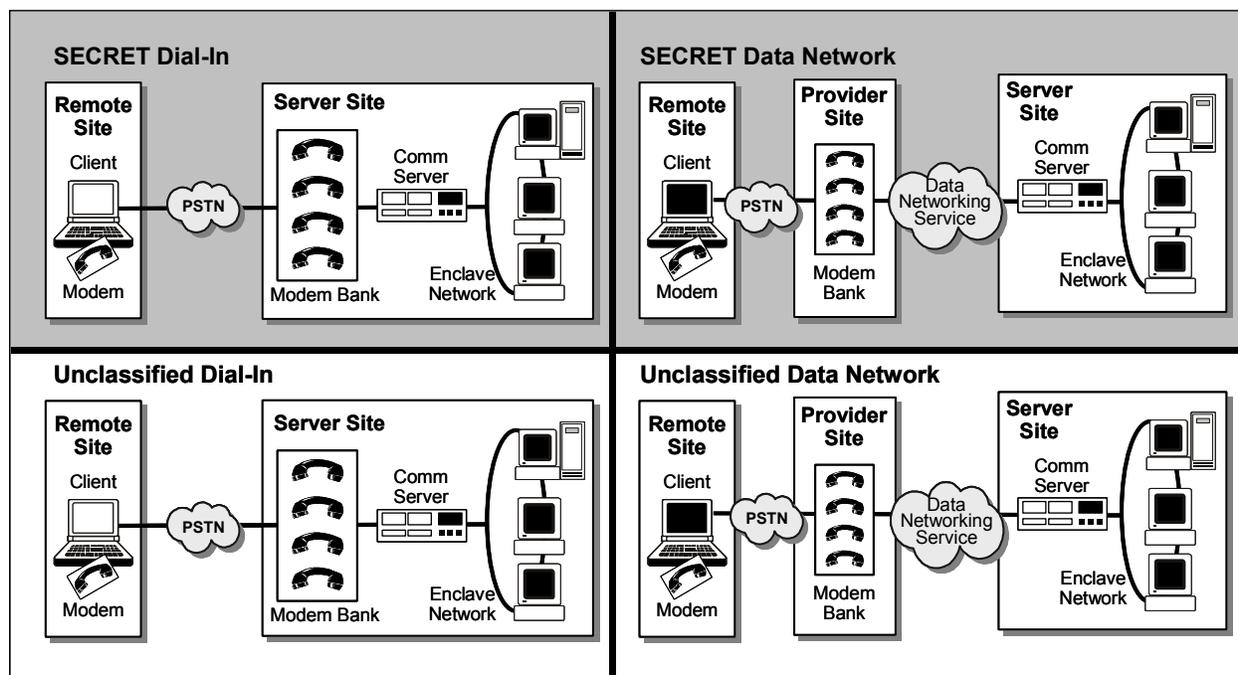
This version of the Framework does not address remote access of top secret or higher sensitivity level information. By definition, the disclosure of this information can cause exceptionally grave damage to national security. Remote access to top secret information presents extreme risk and should be handled on a case-by-case basis.

This section considers remote access to information at the unclassified level that is sensitive or not sensitive and the remote access to classified information up to the secret level as separate cases. Secure remote access to top secret information may be addressed in future versions of this document.

As depicted in Figure 6.2-5, the two different access paths combined with the two sensitivity levels produce four generic cases: secret dial-in access, secret ISP access, unclassified dial-in access, and unclassified ISP access. For each case, the underlying network options include PSTN, ISDN, and other digital and wireless services.

The specific requirement cases include the following.

- Remote access to secret enclave via direct connection through PSTN, ISDN, wireless connections, and other digital connections.
- Remote access to secret enclave via ISP connection through PSTN, ISDN, wireless connections, and other digital connections.
- Remote access to unclassified enclave via direct connection through PSTN, ISDN, wireless connections, and other digital connections.
- Remote access to unclassified enclave via ISP connection through PSTN, ISDN, wireless connections, and other digital connections.



iatf_6_2_5_0114

Figure 6.2-5. Remote Access Cases

6.2.7 Framework Guidance

The following guidance is based on the premise that the home site has properly followed an information systems security engineering process. This process will identify the organization's assets and vulnerabilities and provide a total system solution that mitigates the risk to the level decided by the organization. The discussion here is at a generic level. The level of risk acceptance and the availability of products and services will determine a site's remote access security solution.

6.2.7.1 Case 1: Remote Access to Secret Enclave via Direct Connection over PSTN

Guidance for this case is summarized in Tables 6.2-1a through 6.2-1d. Each of these tables is followed by a discussion of the rationale behind the recommendations.

Media Encryption

A media encryptor is recommended to protect the information stored in the remote user computer. The rationale for this is that media encryption provides confidentiality for data on the user's hard drive. It also performs a workstation integrity function by protecting the integrity of the computer's configuration; e.g., by verifying the BIOS and making sure that the user is notified of any modifications to applications and hardware configuration files.

**Table 6.2-1a. Summary Guidance for Remote Access
Direct Dial-up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
Media Encryptor	Role of this Component	To protect the confidentiality and integrity of all data stored on the hard disk in the event that the user's laptop is lost, stolen, or tampered with. To keep the laptop unclassified when not in use.	RASP	HARA
	Security Functions	Dynamically encrypt all data (but system boot files) stored on the hard disk. Protect the private key used to encrypt the data by storing it on a token that is physically removed when not in use. Require user PIN to unlock the token.	Hardware token-based, software media encryption for Windows platforms	WIN95 and WIN NT versions
	Cryptographic Strength (If applicable)	Cryptographic algorithm and key length should be of robustness level 2.	Type II algorithm (SKIPJACK) w/ 80 bit key	TBD
	Common Criteria Assurance Level	EAL 4	N/A	Three assurance levels
	SMI/PKI/KMI Services	Generation of file encryption keys Data recovery in event of lost token or user PIN		
	SMI Assurance	KMI level 2	TBD	TBD
	Interoperability Requirements	No requirement	No commercial standards exist. Current solutions are not compatible with each other.	Interoperability

The remote computer needs certain system files in order to boot, so these files should remain unencrypted on the storage media. However, the proper functioning of the media encryptor depends on the integrity of the boot process, so the integrity of these unencrypted system files must be verified. The media encryptor also should verify the integrity of the computer's BIOS configuration. All other space on the storage media should be encrypted. The media encryptor should verify the system's integrity upon boot-up and notify the operator if integrity checks fail.

The media encryptor should use algorithms approved for the protection of secret information. To help mitigate concerns about weak or compromised keys, the media encryptor should be capable of accepting keys from an outside source; e.g., FORTEZZA® card and its associated security management infrastructure. The implications of having a split-key are discussed in Chapter 8, Supporting Infrastructures of this Framework. The media encryptor should support both: user and system administrator roles. Only the system administrator should have the ability to change the configuration of the remote computer and the media encryptor. Depending upon the user’s environment and the organization’s security policy, the media encryptor also could be used to preclude the booting of the remote computer via an unencrypted floppy disk. If the remote user wants to access unclassified systems, it is recommended that a separate hard drive be used for this purpose, since the costs of implementing and maintaining a trusted operating system (to maintain data separation and integrity) typically would be prohibitive.

Remote Workstation Integrity

Recommendations concerning remote workstation integrity are contained in, Section 6.1, Firewalls, and are summarized here. Enclave boundary and protection components should be chosen in accordance with the site’s security policy. The user’s home enclave should choose a network boundary protection mechanism (e.g., guards, firewalls) paying close attention to the tradeoffs among security, performance, and cost. An intrusion detection system may be implemented. A virus scanning policy should be implemented, with scans occurring periodically or after certain events. Network vulnerability scanners should be run periodically, and identified deficiencies should be addressed.

**Table 6.2-1b. Summary Guidance for Remote Access
Direct Dial-up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
	Workstation Integrity	Role of this Component	Protect the remote user’s workstation against unauthorized modification	RASP
Security Functions		Digital signature and integrity hash function	Digital Signature Standard and Secure Hash Algorithm	
Cryptographic Strength (If applicable)				
Common Criteria Assurance Level		EAL4	N/A	Three Assurance Levels
SMI/PKI/KMI Services				
SMI Assurance				
Interoperability Requirements				

Remote user and enclave software should be kept up-to-date, since many discovered vulnerabilities are patched in later versions. In addition, software should be protected from tampering by cryptographic checksums applied by the manufacturer and should be checked when the software is installed (on the user’s workstation or the enclave components). New versions of software could also inject new vulnerabilities into the system and thus should be tested before operational use.

Other mechanisms used to protect the integrity of the remote user’s workstation include trusted operating systems, hardware tokens, user password authentication, and so on. At least in the case of a secret enclave, the remote user should be afforded the same protection mechanisms that are provided to the user’s workstation located in the user’s home enclave. In addition, the user’s environment will dictate extra security services, as required by the organization’s security policy. For instance, special policy and procedures are typically required in higher threat environments in which physical security is not at the same level as provided at the home enclave. Additional security mechanisms should give the user the tools to mitigate the loss of workstation integrity.

**Table 6.2-1c. Summary Guidance for Remote Access
Direct Dial-Up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
Secure Modem	Role of this Component	Authenticate and encrypt the connection between the remote user and the home enclave	RASP	HARA
	Security Functions	Mutual authentication Continuous authentication Full period encryption at the secure modem layer In-line encryption Hardware device Removable hardware token to store and protect private keys User PIN to unlock token	Encrypting modem supporting KEA and SKIPJACK	
	Cryptographic Strength (If applicable)	Secret	Secret w/ NAG-68 Interim Policy	Secret
	Common Criteria Assurance Level	EAL3	N/A	Three Assurance Levels
	SMI/PKI/KMI Services			
	SMI Assurance	KMI level 2	TBD	TBD
	Interoperability Requirements	Support for AT command set and communications protocol standards Software compression	56Kbps.X.90	Interoperability

Enclave Boundary and Connection Protection

A link-encrypting device should be used to protect the communications link between the remote user and its home classified enclave. To be used in a classified environment, the device must provide strong authentication and confidentiality services. Modems should meet the applicable commercial standards, such as V.nn and MNPnn. The modem should provide an AT commands interface. To authenticate the remote user to the modem, the modem should require the entry of a personal identification number (PIN) to enable the encrypted data mode. The modem must pass I&A information to the boundary protection mechanism for system access (See Section 6.2.5, Technology Assessment). GUI software should be provided to allow the entry of the PIN and it should display authenticated identities and security modes of operation. The modem may have a plaintext mode of operation (other than that required by the initial handshaking done before a secure session is established). Use of this mode should require overt action on the part of the user so this mode is not selected by accident or by default. Explicit requirements for secure modems will be provided in later releases of the Framework.

In addition to the encrypting modem, a boundary protection device should identify and authenticate the dial-in user at the point of presence of the classified network to the local PSTN. This is discussed in more detail in the next section.

**Table 6.2-1d. Summary Guidance for Remote Access
Direct Dial-up Access to Secret Enclave**

Primary Solution Components	Guidance Categories	Desired Solution	Best Commercially Available Solution	Gap Between Needed & Available Solution
Enclave Boundary Protection		Mutual and continuous authentication Full period encryption at the secure modem layer In-line encryption Hardware device User PIN to unlock token	Secure communications server supporting encrypting modem	
Solution Residual Risks		None	Acceptable	Difference

Authentication Mechanism

An additional authentication mechanism should be implemented that will provide strong authentication directly to the boundary protection mechanism to implement a “that which is not explicitly permitted is denied” policy. For example, many remote users only need e-mail while they are traveling; in addition, some may need access to a particular file server. Providing the minimum access needed to do the job not only mitigates the effects of any successful attack by an outsider, but also makes insider attacks more difficult. Guards and firewalls provide this functionality.

Authentication to the user's workstation is recommended. A password, hardware/software token, or biometric device should be used, depending upon the level of assurance required. See Section 6.1, Firewalls, for more information on this issue.

Technology Gaps

The only government off-the-shelf (GOTS) solution supporting the remote access user is the AT&T Secure Telephone Unit (STU)-III 1910 Secure Data Device (SDD). The SDD runs at data transfer rates much lower than those of modems available in today's commercial market. A cumbersome device, the 1910 is actually heavier and larger than the laptop it supports. There is a consensus in the user population that there is no technology available today. No technology currently provides a high enough level of assurance to pass classified data over the PSTN to and from a classified enclave at the same level of performance that is available in non-encrypting commercial off-the-shelf (COTS) modems. This gap is certainly noticeable when comparing capabilities with the 56 Kbps modems on the market today.

In general, there is a technology gap in high-assurance security solutions applicable to remote access in the COTS environment. In particular, little commercial work is being done on media encryptors, although several file encryption products are available. File encryptors are not widely available for non-Windows operating systems. A few commercial encrypting modems are available, but high-assurance encrypting modems are not commercially available. In addition, secure remote access servers and communication servers are not widely available. Support for top secret remote access will require additional features that are not available in today's commercial marketplace, at least at an acceptable risk level. Workstation integrity and configuration guidance are also issues. Future versions of this Framework will address these gaps in more detail.

6.2.7.2 Case 2: Remote Access to Secret Enclave via ISP Connection

This section will be provided in a future release of the Framework.

6.2.7.3 Case 3: Remote Access to Unclassified Enclave via Direct Connection

The recommended solution for this case involves implementing a RADIUS server within the enclave and configuring each remote workstation with a RADIUS client. When a remote workstation requests access to the network, RADIUS-based authentication is used.

- **Media Encryption.** In this scenario, all information is unclassified. Therefore media encryption is not necessary for information stored on the remote workstation. File encryption may be desired for protection of unclassified information that is sensitive or not sensitive.

- **Workstation Integrity.** An unclassified remote access workstation will also likely have access to the Internet. There may be a requirement for the remote workstation to download files from the Internet or to exchange files with the unclassified enclave. Downloading files from the Internet poses a risk to the workstation's integrity. The workstation should have a robust and updated virus scanning capability. Additionally, the workstation connecting to the enclave poses a risk to the integrity of the enclave if precautions are not taken to check for viruses on the workstation. Again, to protect the integrity of the workstation and the enclave, virus scanning should be resident on the remote workstation.
- **Enclave and Connection Protection.** The enclave is vulnerable to unintentional virus insertion through the remote workstation. Although RADIUS-based authentication of remote workstations prevents unauthorized remote workstations from gaining access to the enclave's network, there is still a risk of valid workstations being lost or compromised.

All workstations should be equipped with a robust user-to-workstation authentication mechanism. Although in the case of workstation theft or compromise, this mechanism alone may not provide adequate assurance that the workstation cannot be used to access the enclave. A way of mitigating the risk of such access is by implementing an incident report procedure for reporting lost or compromised remote workstations and by installing and maintaining an intrusion detection system. If a lost or compromised workstation is reported in a timely manner, the RADIUS server can be configured to deny access from that compromised workstation. If the compromised workstation establishes a connection to the network before the compromise is reported and mitigated, an intrusion detection system will identify anomalous behavior and alert administrators to the possibility of a compromised workstation.

Although the user information in this scenario is unclassified, there still may be a requirement to provide confidentiality for the connection. A VPN solution can be established across the remote connection. A layer 2 mechanism, such as L2TP, or a layer 3 mechanism such as IPSec may be implemented to provide confidentiality. These technologies are discussed in further detail in Section 5.3.

- **Authentication Mechanism.** Authentication between the remote workstation and the home enclave is achieved by using the RADIUS protocol. The RADIUS protocol relies on a shared secret between the RADIUS client and the RADIUS server. MD5 is used to hash the shared secret, the user password, and other fields in the RADIUS message. The strength of the authentication is based on protecting the shared secret.

Authentication to the user's workstation also is recommended. A password, hardware/software token, or biometric device should be used, depending on the level of assurance required.

6.2.7.4 Case 4: Remote Access to Unclassified Enclave via ISP Connection

The recommended solution for this scenario involves implementing an IPSec-compliant firewall or other boundary protection device. Remote workstations must be configured with an IPSec-compliant network card, software, or other component. This case also involves implementing a RADIUS server within the enclave and configuring each remote workstation with a RADIUS client. In this scenario, the remote workstation usually uses the PSTN to establish a connection to the ISP. The ISP then interfaces with the Internet, which interfaces with the enclave. The remote workstation establishes an IPSec-secured connection over the PSTN that terminates at the enclave ISP-compliant firewall or boundary protection device.

- **Media Encryption.** In this scenario, all user information is unclassified. Therefore, media encryption for information stored on the remote client is not necessary. File encryption may be desired for protection of unclassified information that is sensitive or not sensitive.
- **Workstation Integrity.** An unclassified remote workstation also will likely have access to the Internet. There may be a requirement for the remote workstation to download files from the Internet or to exchange files with the unclassified home enclave. Downloading files from the Internet poses a risk to the workstation's integrity. The Internet-connected workstation connecting to the enclave poses a risk to the integrity of the enclave if precautions are not taken to check for viruses. Therefore, to protect the integrity of the workstation and the enclave, a robust and updated virus scanning capability should be resident on the remote workstation.
- **Enclave and Connection Protection.** The enclave is vulnerable to unintentional virus insertion through the remote workstation. Although RADIUS-based authentication of remote workstations prevents unauthorized remote workstations from gaining access to the enclave's network, there is still a risk of valid workstations being lost or compromised.

All workstations should be equipped with a robust user-to-workstation authentication mechanism. Although in the case of workstation theft or compromise, this mechanism alone may not provide adequate assurance that the workstation will not be used to access the enclave. A way of mitigating the risk of such access is by implementing an incident report procedure for reporting lost or compromised remote workstations and by installing and maintaining an intrusion detection system. If a lost or compromised workstation is reported in a timely manner, the RADIUS server can be configured to deny access from that compromised workstation. If the compromised workstation succeeds in establishing a connection to the network before the compromise is reported and mitigated, an intrusion detection system will identify anomalous behavior and alert administrators to the possibility of a compromised workstation.

Although the user information in this scenario is unclassified, there still may be a

requirement for confidentiality. If confidentiality is required, the IPSec client on the remote workstation can use the ESP feature of IPSec to encrypt the IP payload.

- **Authentication Mechanism.** Authentication between the remote workstation and the home enclave is achieved by using the authentication header of IPSec. The IPSec authentication header relies on a shared secret using either a symmetric encryption algorithm (i.e., Data Encryption Standard [DES]), or a one-way hashing algorithm (e.g., MD5, HA).

Authentication to the user's workstation also is recommended. A password, hardware/software token, or biometric device should be used, depending on the level of assurance required.

UNCLASSIFIED

Remote Access
IATF Release 3.1—September 2002

This page intentionally left blank.

6.3 Guards

Guards enable users to exchange data between private and public networks, which is normally prohibited because of information confidentiality. A combination of hardware and/or software guards is used to allow secure local area network (LAN) connectivity between enclave boundaries operating at different security classification levels (i.e., one private and the other public). Guard technology can bridge across security boundaries by providing some of the interconnectivity required between systems operating at different security levels. Several types of guards exist. These protection approaches employ various processing, filtering, and data-blocking techniques in an attempt to provide data sanitization (e.g., downgrade) or separation between networks. Some approaches involve human review of the data flow and support data flow in one or both directions. Information flowing from public to private networks is considered an upgrade. This type of transfer may not require a review cycle, but should always require a verification of the integrity of the information originating from the public source system and network. This section discusses guards, the environment and mannerism in which they are most suited for implementation, how they can be used to counteract attacks made on the enclave, and the variety of guards and their functions.

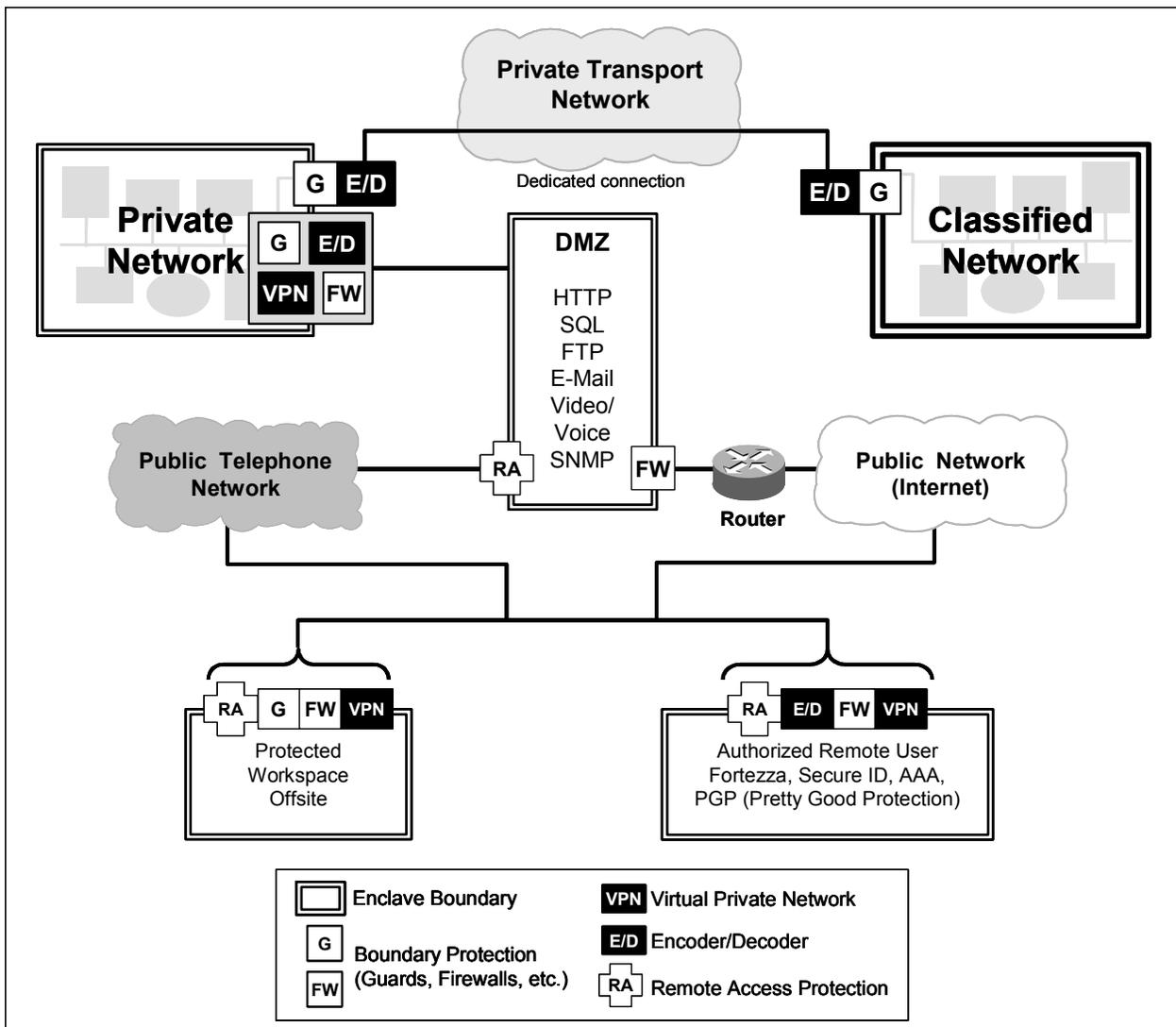
A guard is a device used to defend the network boundary by employing the following functions and properties:

- Typically subjected to high degree of assurance in its development.
- Supports fewer services.
- Services are at the application level only.
- May support application data filtering (review).
- May support sanitization of data.
- Typically used to connect networks with differing levels of trust (provides regrading of data).

6.3.1 Target Environment

The guard is designed to provide a secure information path for sharing data between multiple system networks operating at different security levels. The overall system that employs a guard is illustrated in Figure 6.3-1. The system is composed of a server, workstations, malicious code detection, a firewall, and/or filtering routers all configured to allow transfer of information among communities of users operating at different security levels. The server and workstation components may implement a hardware- or software-based authentication scheme to authenticate to the guard. The firewall component is usually commercial off-the-shelf (COTS) hardware and/or software that filters the network traffic and is configured to forward only authorized packets. A commercial filtering router may also be used to perform this function. The firewall's primary function is to provide barriers against successful penetration of the low side LAN by

unauthorized external users. The firewall hides the networks behind it and supplements the guard. The firewall restricts access to all traffic other than the traffic being scrutinized by the guard. Virtual private networks (VPN) can also be employed using either a firewall or other encryption device. To ensure the security of the overall system, all users, managers, and system administrators must exercise the security policies and practices of the organization. Some considerations include valid personnel approval for access to all information stored and/or processed on the system; formal access approval process for, and signed nondisclosure agreements for all information stored or processed on the system; valid need-to-know process for some of the information stored or processed by the system. Communication links, data communications, and data networks of the system must protect the network determined by the sensitivity level of data on that particular network.



latf_6_3_1_0027

Figure 6.3-1. Guard Environment

The guard can be configured to function in different directions.

- The private to public bidirectional mode facilitates data to move from private to public after the review process for releasability to the lower network classification. Data moving from low to high need not undergo the review process for releasability, but processing, filtering, and blocking should occur to identify viruses and other malicious code transfers. Private network users would be allowed to push public data to public network users, and in turn, users on the public network could push public data to users on the private network. Private network users would also be allowed to view and pull data that exists on the public network.
- The private to public unidirectional mode allows data to move from private to public after the review process for releasability to the lower network classification. No transfer is permitted from the lower network to the private network. Private network users would send data to be downgraded to the public level, which would then be pushed to a server on the public network for subsequent pull by users on the public network.
- The peer-to-peer mode allows communications between networks bridged by the guard at the same security level (e.g., private and private releasable)—that is, all the screening the guard normally performs on private to public transfers in the private to public configuration is performed in both directions. Standard operating procedures must be implemented so that appropriately cleared personnel from each side can administer the guard screening criteria databases. This configuration allows private network users to downgrade data to the private-releasable level and to push that data to a server on the private-releasable network for subsequent pull by users on the private-releasable network.

6.3.2 Requirements

This section addresses the functional requirements of the communication, releasability, and network access capabilities.

6.3.2.1 Communication Requirements

Requirements for communication include the following:

- The guard shall allow users on the private networks to communicate with only specified hosts on the public networks.
- The guard shall prohibit workstations to be used as a pass-through or gateway device from either the private or public sides for any communications, including mail.
- The guard shall send public data to one of the public networks or private networks using the appropriate router.
- Routers shall be configured to restrict the types of network services that may pass through them as well as the sources and destinations of service requests.

- The guard shall transfer the appropriate data from the private network to the public network.
- The guard shall allow protocols to pass through it.
- The guard shall allow only authorized users to send and/or receive a message by performing access control on both the source and destination addresses of the message.

6.3.2.2 Releasability Requirements

Current requirements for releasability include the following:

- The guard shall allow only a properly labeled message to pass from the private level to the public level.
- The guard shall support a policy that allows only attachments that have been reviewed for security level at the user's workstation to pass from the private-to-public side.
- The guard shall allow only selected application attachments to pass through it—this capability will be configurable to support a variety of application packages.
- The guard shall perform word and/or phrase search.
- The guard shall support rule-based sanitization (i.e., message content modification) of messages from high levels through low levels.
- The guard shall ensure that only allowed data is distributed.
- The guard shall validate proper message construction, including configurable verification of message content.
- The guard shall remove classification labels, which were inserted into the e-mail body and attachments prior to delivery to the other side.

6.3.2.3 Access Requirements

Current access requirements for file transfers include the following:

- The guard shall run on a trusted platform.
- The guard shall prevent message flow directly between the private side wide area network (WAN) and the guard in either direction.
- The guard shall support a programmable set of security identification (ID) labels per flow.
- The guard shall ensure that the security level of a message subsumes (is equal to or greater than) the security level of its attachment(s).

- The guard shall protect against unauthorized disclosure of private side information.
- The guard shall provide safeguards to protect the private side from attacks (including penetration, malicious code, and denial of service) from the public side.
- The guard shall support user authentication and encryption capabilities.
- The guard shall perform audit all security-related functions.
- The guard shall provide an access control mechanism to limit access to the controls and provide separate roles for the security administration, system operator, and mail administration functions. Thus, a supporter authorized to function in one area will be prevented from performing functions in another, unless specifically given permission to do so.
- The guard shall prevent disclosure or release data to unauthorized consumers.
- The guard shall provide a secure bridge for passing messages between networks of differing levels of security.
- The guard shall strip off the digital signature as the message passes through the guard.
- The guard shall restrict source routing. Source routing, which is a form of addressing, can alter the routing of a message from its normal route.
- The guard shall journal/log all passed and/or failed messages.

6.3.3 Potential Attacks

The focus within this category is on attacks into an enclave by malicious e-mail, file, or message transfers. Guards can be implemented to provide a high level of assurance for networks by preventing certain types of malicious messages from entering the enclave. The types of attacks are categorized into three sections: Section 6.3.3.1, Active Attacks; Section 6.3.3.2, Distribution Attacks; and Section 6.3.3.3, Insider Attacks. For more information related to attacks, please refer to Chapter 4.2, Adversaries, Threats (Motivations/Capabilities), and Attacks.

6.3.3.1 Active Attacks

Active attacks attempt to breach security features or exploit data in transit, whether it be e-mail, file, or message transfers. Some firewall technologies and e-mail systems that perform content filtering will help establish a level of trust for messages that are signed but not encrypted. Messages may be signed and/or encrypted at the user level and/or the organizational level. However, a digital signature on a message does not increase the safety level for the message contents. Active attacks may include the insertion of malicious code or the theft of data. Examples of active attacks in regard to the transmission of messages and files are listed below. For further description of network-based attacks, please refer to Section 4.2.1.4.2, Network-Based Vulnerabilities and Active Attacks.

UNCLASSIFIED

Guards

IATF Release 3.1—September 2002

- **Modification of Data in Transit.** Modifications are not necessarily always malicious or intentional. A modification could be the conversion of spaces to tabs or vice versa within an e-mail or real-time message. A network-based modification could also be the occurrence of a complete violation of standards. Internet e-mail standards necessary for the secure transmission of messages from one domain to another are Pretty Good Privacy (PGP); Multipurpose Internet Mail Extensions (MIME); and Secure Multipurpose Internet Mail Extensions (S/MIME). Although instant/real-time messaging do not yet have interoperable standards established, protocols must be established to ensure that the messages have not been intercepted and corrupted.
- **Insertion of Data.** Reinsertion of previous messages.
- **Inserting and Exploiting Malicious Code** (e.g., Trojan horse, trap door, virus, and worm).
- Defeating login mechanisms into e-mail accounts, messaging accounts, or file storage servers.
- **Session Hijacking.** In the case of e-mail, file or real-time message transfers unauthorized access could be gained into a communications channel with malicious intent.
- Denial of service.
- Establishment of unauthorized network connections.
- **Masquerading as an Authorized User.** An attacker would use the identification of a trusted entity to gain unauthorized access to information either by e-mail, real-time messaging, or requesting file transfers.
- Manipulation of data on the private side.
- Decrypting weakly encrypted traffic.
- **Misrepresentation or information “faking” through Internet relay attacks.** Third-party mail relay occurs when a mail server processes and delivers e-mail from an external client. In this manner, mail appears to originate from that mail server’s site and not the original site. Spam e-mail is generally distributed this way, at the mail owner’s expense. Intruders can spam e-mails with embarrassing content or by flooding a site with e-mails. Damage caused by spamming includes not only the loss of reputation of the system that has been identified with the attack e-mail but also the loss of connectivity to large parts of the Internet that have blocked sites from spamming. E-mail servers will become clogged, mail can be lost or delivered late, and cleanup costs will be incurred to remove spammed mail without destroying legitimate mail.
- **Monitoring Plain Text Messages.** Plain text messages are not encrypted, and therefore not secure in any manner. Once intercepted, plain text messages can be easily read.

6.3.3.2 Distribution Attacks

Distribution attacks can occur anytime during the transfer of a guard's software and/or hardware. The software or hardware could be modified during development or before production. The software is also susceptible to malicious modification during production or distribution. Section 6.3.4.2 discusses methods in which these attacks could be prevented. For additional information, please refer to Section 4.2.1.4.4, Hardware/Software Distribution Vulnerabilities and Attacks. Also, refer to Table 4-3, Examples of Specific Modification Attacks.

6.3.3.3 Insider Attacks

Although an enclave must be protected from outside intruders, it must also be protected from attacks from inside the enclave. Interception or attacks to messages can occur during transit from the insider level. The originators' and recipients' mail system administrators are able to look at e-mail messages and files that are being sent. E-mail messages that bounce back usually have a copy sent to the e-mail system administrator to help determine the reason behind the bouncing; therefore, the administration has bounced messages brought to his/her attention with full viewing privileges to the message that is attempting to be sent. An insider attack occurs when someone located within the boundaries of the enclave intercepts or modifies data or security mechanisms without authorization.

Unauthorized access could also be gained into the overhead portion of a covert channel. The use of a covert channel is a vulnerable point of attack as a result of the transport overhead not being completely defined and therefore being susceptible to exploitation. The physical theft of data is another threat within the enclave. For further detail, please refer to Section 4.2.1.4.3, Insider Vulnerabilities and Attacks.

6.3.4 Potential Countermeasures

For all efforts aimed at attacking an enclave through the unauthorized access or modification to e-mail messages, real-time message transfers, or file transfers, measures must be in place to prevent these attacks from penetrating the boundaries of an enclave. In the case of attacks that originate from inside the enclave, precautionary measures also need to be taken in areas vulnerable to attacks, including the physical theft and unauthorized access to data. The following subsections address measures that can be taken to counteract attacks against an enclave and information transfers among enclaves. These countermeasures are placed into three categories: Section 6.3.4.1, Boundary Protection Via Guards; Section 6.3.4.2, Distribution Attack Countermeasures; and Section 6.3.4.3, Insider Attack Countermeasures.

6.3.4.1 Boundary Protection Via Guards

Guards can be implemented to protect the enclave and the messages passing within and through the enclave boundaries. Guards enable users to exchange information between either networks of the same or differing classification levels. Traffic analysis is a means by which traffic can be

UNCLASSIFIED

Guards

IATF Release 3.1—September 2002

monitored. Traffic analysis can be conducted to help identify traffic patterns (i.e., origination and destination endpoints for traffic), and thus aid in the discovery of the endpoints of unauthorized network connections. Enclave boundaries need protection from the establishment of unauthorized network connections. The responsibility lies with the management and administration of the local network to prohibit unauthorized connections between networks of different classification levels and to enforce this policy through nontechnical means.

The following bulleted items list the type of attack and the countermeasure that can be used to prevent that attack from occurring.

- **Modification of Data in Transit.** The countermeasure to this attack is to use digital signatures or keyed hash integrity checks to detect unauthorized modification to the data in transit. E-mail, real-time messaging, and file transfers are all susceptible to interception and modification while in transit.
- **Insertion of Data.** Many countermeasures exist for the malicious insertion of data. They include the use of time stamps and sequence numbers, along with cryptographic binding of data to a user identity, to prevent the replay of previously transmitted legitimate data. Data separation or partitioning techniques, such as those used by guards and firewalls, deny or restrict direct access and the ability to insert data during transit.
- **Inserting and Exploiting Malicious Code (Trojan horse, trap door, virus, and worm).** Implement a guard and employ strong authentication in order to filter and block incoming messages that are not from authenticated parties. To help ensure that mail is neither modified during transit nor forged, technologies and products such as PGP and S/MIME can be used to encrypt and sign messages on a regular basis. Real-time messaging protocols are necessary to also ensure authentication among parties.
- **Defeating Login Mechanisms.** The most appropriate countermeasure for this attack is the cryptographic authentication of session establishment requests. This effort pertains to logging into an e-mail account or to obtaining access to a file server or messaging channel.
- **Session Hijacking.** The countermeasure for this attack is continuous authentication through digital signatures affixed to packets, or at the application layer, or both.
- **Denial of Service.** Countermeasures that can be taken against these attacks include having a guard to filter out bad source Internet Protocol (IP) addresses, filter Internet Control Message Protocol (ICMP) echo responses or limit echo traffic, and guard against all incoming User Datagram Protocol (UDP) service requests. A nontechnical countermeasure would be to subscribe to the certification and accreditation (C&A) Computer Emergency Response Team (CERT) mailing list (www.cert.org) in order to receive notifications every time a new Internet weakness emerges. [2]
- **Establishment of Unauthorized Network Connections.** A nontechnical countermeasure lies with the management and administration of the local network to prohibit and enforce the policy against unauthorized connections between networks of different security levels. Commercial tools also are available for system administration

personnel to use for detecting unauthorized connections. Unauthorized connections would allow for otherwise prohibited access to e-mail and data files and for real-time message interception.

- **Masquerading as an Authorized User.** The appropriate countermeasure is to use cryptographic authentication in conjunction with time stamps or sequence numbers to prevent any recording and/or replay of authentication data, whether it be e-mail, real-time messaging, or file transfers. Another countermeasure to prevent stealing an authentic session is to cryptographically bind authentication data to the entire session or transaction.
- **Manipulation of Data on the Private Side.** The appropriate countermeasure is to permit only authorized users to access the data, through file transfers, on the private side using cryptographic authentication and data separation techniques.
- **Decrypting Weekly Encrypted Traffic.** To ensure that unauthorized persons cannot access e-mail messages, real-time messages, or files in transit, adequate encryption algorithms and sound key management processes must be observed.
- **Misrepresentation or Information “Faking” Through Internet Relay Attacks.** The countermeasure for these spamming attacks would involve the use of a guard to filter the messages and therefore block malicious messages, whether they are e-mail messages or real-time messages, from entering the enclave.
- **Monitoring Plain Text Messages.** The monitoring of messages can be counteracted by denying access to the data by unauthorized users. Access denial is possible by encrypting the data or by using other data separation techniques that will restrict those who are unauthorized from obtaining access to the data contained within a file.

6.3.4.2 Distribution Attack Countermeasures

During the development, manufacturing, and distribution stages, technical and nontechnical measures must be taken to avoid the malicious modification of guard software and hardware. The following lists the stage at which an attack could occur and the countermeasure to prevent such an attack.

- **Modification of Software or Hardware During Development, Prior to Production.** Strong development processes and criteria are essential during this phase as a countermeasure for threats. Continuous risk management through processes, methods, and tools is also necessary. The following Web site link contains a collection of software engineering processes, methods, tools, and improvement references, <http://www.sei.cmu.edu/managing/managing.html>. [3] Subsequent third-party testing and evaluation of software should also be conducted to ensure that the software and hardware have not been modified. High-assurance methods and criteria should be followed, such as the Trusted Product Evaluation Program (TPEP) and Common Criteria. Please refer to <http://www.radium.ncsc.mil/tpep/tpep.html> for program details. [4]

- **Malicious Software Modification During Production and/or Distribution.** The countermeasures for threats during this phase are high-assurance configuration control, cryptographic signatures over tested software products, use of tamper detection technologies during packaging, use of authorized couriers and approved carriers, and use of blind-buy techniques.

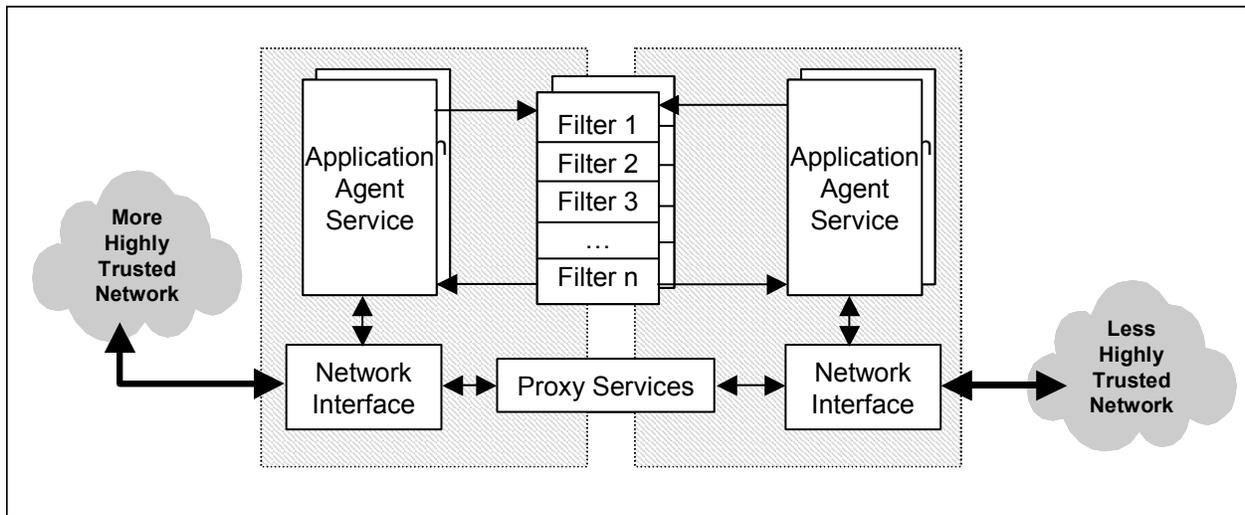
6.3.4.3 Insider Attack Countermeasures

Technical and nontechnical countermeasures must both be taken to prevent against attacks originating within the boundaries of an enclave. The following are the types of insider attacks that can occur and the countermeasure that must be taken to prevent the attack.

- **Modification of Data or Modification of Security Mechanisms by Insiders.** The primary technical countermeasure is to implement auditing procedures of all actions taken by users that could pose a threat to security. Audit logs will need to be generated and timely, diligent reviews and analysis must be conducted. Nontechnical countermeasures include personnel security and physical procedures.
- **Physical Theft of Data.** Appropriate nontechnical countermeasures include personnel security and physical security procedures, which inhibit actual removal of data, either in printed form or on storage media.
- **Covert Channels.** The countermeasure against a covert channel between networks of different classification levels is a trusted guard function that examines network header fields and network messages for possible unauthorized information.

6.3.5 Guard Technology Assessment

Guards are usually used to enable connectivity that is normally prohibited because the information requires confidentiality. Where a firewall is usually used to restrict or scrutinize information flow on an already existing link to LAN or WAN circuits, guards allow the transfer of information between segments operating at different security classification levels (one private and the other public). A combination of hardware and software components is designed to allow this connectivity between segments. Most guard implementations use a dual network approach, which physically separates the private and public sides from each other. As shown in Figure 6.3-2, guards are application specific; therefore, all information will enter and exit by first passing through the Application Layer, Layer 7, of the open systems interconnection (OSI) model. In addition, most guard processors are high-assurance platforms that host some form of trusted operating system and trusted networking software.



iatf_6_3_2_0028

Figure 6.3-2. Dual Network Approach

A guard can be a fully automated (without any human intervention) multilevel security (MLS) guard system that permits one-way or bidirectional transfers of data among multiple LAN systems operating at different security or releasability levels. Guards can concurrently review and sanitize multiple binary and American Standard Code for Information Interchange (ASCII) files and virtually any complicated data format. Almost any data type that can be “packaged” into a file can be transferred through certain guards, including structured query language (SQL), HyperText Transfer Protocol (HTTP), UDP, Simple Mail Transfer Protocol (SMTP)/e-mail attachments, and others. The guard controls the automated information flow among multiple LAN systems according to security rule filters. When implemented in conjunction with a firewall, a higher degree of security for protecting the enclave is achieved.

This section is further broken down to discuss guard technological areas that can be used to protect the enclave:

- Authenticated Parties Technologies.
- Confidentiality and Integrity.
- Data Processing, Filtering, and Blocking Technologies.

This categorization allows for a high-level assessment of system assurance so that a determination can be made as to the level of security robustness a network will require. These three categories of potential protection approaches are explained in more detail in the following subsections.

6.3.5.1 Authenticated Parties Technologies

Approaches for protecting the enclave that are included within this category are those that mandate the use of cryptographic authentication mechanisms before allowing access. Authentication allows two parties that intend to exchange data to identify themselves to one

UNCLASSIFIED

Guards

IATF Release 3.1—September 2002

another and positively authenticate their identities. Hence, they become mutual trusting parties. The data flowing between these trusting parties is at the lower security level. Authenticated access is widely available and is supported by a large number of standards and protocols. Authentication protects the enclaves of private users that are separated from public network users through an enclave boundary protection device, such as a guard and/or firewall. In such a topology, public network users might use digital signature technology to authenticate themselves to private network users. In addition, the guard might incorporate access control list (ACL) mechanisms to make access decisions governing the set of users that is authorized to release information from the private network. The ACLs can also be used to restrict the set of public network users that are authorized to push data up to the private network. The enclave boundary protection system might also perform content review of the data submitted for release. Protection approaches that use authenticated parties are discussed below.

User and document authentication can be achieved with the digital signature and FORTEZZA technologies. Guards can check data packets for digital signatures or user identification and authentication (I&A). Based on this information, guards can accept or deny traffic from entering the enclave. The enclave boundary protection system cannot perform the functions of inspecting the contents of the message or verify the digital signature if the message is encrypted. Messages must be able to be decrypted before processing through the guard so that the guard will be able to perform filtering on the message contents.

Digital Signature

The digital signature, which is the result of encrypting a document using the private key of the signer, can be applied to spreadsheets, Word documents, e-mail messages, portable document format (PDF) files, and others. A digital signature is a string of numbers that is the representation of the document. Using a digital signature ensures that the contents of a document cannot be altered; doing so would invalidate the signature. A digital signature is unique to both the signer and the document; therefore, user and document authentication can be achieved. However, the signature cannot provide confidentiality to the data contents.

An important note is the difference between the digital signature and a digitized signature. A digitized signature is simply the visual form of a handwritten signature to an electronic image. A digitized signature can be forged, duplicated, and cannot be used to determine if information has been altered after signature.

Hardware Tokens

Hardware tokens, which can be used to identify and authenticate users, include One-Time Only Passwords, FORTEZZA, and smart cards (the latter two are addressed in more detail below). One-Time Only Passwords protect against unauthorized access by providing dynamic user authentication. A personal identification number (PIN) along with a code that changes very frequently (e.g., every 30 to 60 seconds) is requested from the user for I&A. A guard will process this information to permit or deny access. By requiring two factors of authentication,

greater protection is provided against unauthorized access than with the traditional fixed password.

FORTEZZA

FORTEZZA is a registered trademark held by the National Security Agency (NSA) that is used to describe a family of security products that provides data integrity, originator authentication, nonrepudiation, and confidentiality. FORTEZZA is an “open system,” allowing for seamless integration with most data communication hardware platforms, operating systems, software application packages and computer network configurations and protocols. This technology uses a cryptographic device: a personal computer (PC) card called the FORTEZZA crypto card. This card contains the user’s unique cryptographic key material and related information and executes the public key cryptologic algorithms. The FORTEZZA card enables users to encrypt, decrypt, archive data, and generate digital signatures. The card uses the Secure Hash Algorithm, Digital Signature Standard, Digital Signature Algorithm, and the Key Exchange Algorithm. A guard can identify and authenticate the originator of a message based on a digital signature. However, a guard must be able to decrypt traffic before determining permissibility into an enclave. If a guard is unable to decrypt data, then the information will be denied from passing through the guard and entering the enclave.

Smart Cards

The use of smart cards is another technological method in which users can be identified and authenticated. A smart card is a plastic card embedded with a computer chip that stores and exchanges data between users. Smart cards provide the tamperproof storage of user and account identity and add to system security for exchanging data across any type of network. They can serve as a means for network system, application, or file access because smart cards can be used to obtain access to a computer or even e-mail accounts. Insertion of the card or proximity to an antenna is required to be able to “read” the information on the card using a smart card reader that can be attached to a computer. Users can be authenticated and granted access based on preset privileges. A guard can authenticate and identify users and thus determine access privileges into an enclave based on the information provided from the smart card.

Secure Sockets Layer

Secure Sockets Layer (SSL) is a popular security protocol for implementing privacy and authentication between communicating applications. This transport layer security protocol enables the encryption and authentication of arbitrary applications. The protocol prevents eavesdropping, tampering with information, and forging of information sent over the Internet.

The SSL protocol includes a lower level protocol (called the SSL Record Protocol) that encapsulates higher level security protocols. The SSL Handshake Protocol is one such encapsulated protocol. It allows communicating parties to authenticate one another and to establish cryptographic algorithms and keys at the start of a communication session. For more information about SSL, please visit <http://welcome.to/ssl>. [5]

Connections using SSL have three properties:

- The communication is private. The initial handshake uses public key cryptography to define a secret key. The secret key is then used with symmetric cryptography to encrypt all communications.
- Clients and servers can authenticate one another during the handshake using public key cryptography.
- The entire communication is protected against tampering or insertion of data. Each datagram has a message authentication code that is a keyed hash value.

The SSL protocol can be used for network access between clients on the private side and servers on the public side. By checking a server's identity, confidence is obtained that the server is trusted to some degree. A policy requiring that SSL be used for all network access between private and public networks would effectively permit access to only those servers on the public side that are able to authenticate using SSL. However, the goal should not only be authentication; rather, the goal should be access control, with authentication being a means to implement access control. This is accomplished by maintaining a list of public servers and directories that, once authenticated, can be accessed by private clients. That ACL is best maintained by an enclave boundary protection system such as a guard.

6.3.5.2 Confidentiality and Integrity

Confidentiality and Integrity can be assured through the following technologies: FORTEZZA, COTS Encryption, Audit Logs, and Operating System.

FORTEZZA

In addition to the I&A features of FORTEZZA, the cryptographic features of the "FORTEZZA Crypto Card" are employed to offer confidentiality and integrity. The integrity protection is provided primarily when data served from a server or client is key hashed (via the Secure Hash Algorithm Federal Information Processing Standards Publication [FIPS PUB] 180). [6] Confidentiality is accomplished with preencryption of the data to be served from the server, and the encryption/decryption of all data passed from a server to a client and from a client to a server (via the Key Exchange Algorithm and SKIPJACK Algorithm FIPS PUB 185). [7] These cryptographic features also include not only digital signature capabilities, but also associated key and certificate management infrastructure support. FORTEZZA encryption and decryption functions include the following:

- Interface to and function with any government-certified FORTEZZA Cryptographic Card for encryption and decryption.
- Do not corrupt the integrity of a file's data content.

- Ensure that the resultant decrypted file retains the original file's attributes (e.g., if the original file was read-only, then when that file is decrypted after being encrypted, it shall retain the read-only attribute).
- Be able to encrypt and decrypt files of all types.
- Inform the user if the encryption and decryption process succeeded or failed.
- Verify that any signature on the certificate is valid (based on the public key from the issuer's certificate).
- Allow the originator to select the type of protection to be applied to the message: signed-only, encrypted-only, or signed and encrypted.

Commercial Off-the-Shelf Encryption

Some guard products incorporate COTS encryption algorithms, such as triple Data Encryption Standard (DES). Although these algorithms are not suitable to protect classified information, they may be used to segregate communities of interest in a protected environment. For example, two users with different privileges at the same classification level may use a commercial encryption algorithm to logically and reliably segregate their traffic. Other organizations that do not possess classified traffic, but rather sensitive traffic, may allow commercial algorithms to provide data confidentiality. In either scenario, commercial encryption may be used on the enclave side of the guard to provide logical data separation.

Audit Logs

Audit logs maintain a record of system activity by system and application processes and by user activity of systems and applications. In conjunction with appropriate tools and procedures, audit logs can assist in detecting security violations, performance problems, and flaws in applications and ensure data integrity. A computer system may have several audit trails, each devoted to a particular type of activity. Auditing is a review and analysis of management, operational, and technical controls. The auditor can obtain valuable information about activity on a computer system from the audit trail. Audit trails improve the accountability and integrity of the computer system. For example, audits can be used in concert with access controls to identify and provide information about users suspected of improper modification of data (e.g., introducing errors into a database). An audit trail may record "before" and "after" versions of records. (Depending on the size of the file and the capabilities of the audit logging tools, this may be very resource intensive.) Comparisons can then be made between the actual changes made to records and what was expected. This can help management determine if errors were made by the user, by the system or application software, or by some other source.

Operating System

A guard cannot provide any degree of assurance if it is installed on an operating system with well-known vulnerabilities. To be effective, guard software must be developed on a trusted

operating platform. Additionally, the guard software must make effective use of the security mechanisms and services offered by the operating system. Part of the guard development process should be documenting how the guard uses the operating system in an effective manner. Guards built on insecure operating systems should not be considered.

The operation and security level of a guard is dependent on the operating system. The platform must be a trusted operating system with high-level security mechanisms. Hackers who become frustrated while trying to penetrate the guard will try to attack the underlying operating system in hopes of gaining access into the enclave. The operating system must have segmentation of processes to minimize the risk from hacker attempts. Segmentation of processes is the separation of system calls at the operating system level. This segmentation allows applications to use restricted portions of the operating system and denies the user's ability to penetrate different security levels—that is, a separate login and password is required for different command levels of the operating system. Usually, each security level of the operating system will have a limited command set in compliance with the security policy of the operating system. The system administrator should therefore hold a clearance that is at least equal to that of the highest network connected to the guard.

In an MLS environment, the strength of some guards remains within the user workstations and the gateways. Each user workstation and gateway must be installed with a trusted operating system. Guards trust users to make decisions regarding the classification and sensitivity of information. The trusted operating systems control access to information displayed on a user workstation and control the movement of information out of the multilevel network (MLN). The MLN must use a trusted operating system, defined as an operating system accredited to maintain the trust between sensitive information and the authorized users. In the MLN architecture, an authentication server controlling user logins and monitoring network system activity enhances this service.

6.3.5.3 Processing, Filtering, and Blocking Technologies

Protection approaches that fit logically within this category use various processing, filtering, and data-blocking techniques in an attempt to provide data sanitization or separation between private network data/users and public network data/users. Data originating from the private network is implicitly labeled as private data, though it may be asserted to be data of a lower sensitivity level by a private network user. Enclave boundary protection devices such as a guard may perform automated processing and filtering techniques. If such tests are successfully passed, the data is actually regraded by automated means. In the reverse direction, such approaches often incorporate data blocking techniques (typically in firewalls but also in guards) to regulate the transfer of data from public network users to private network users. Use of certain protocols may be blocked and/or data may be processed or filtered in an attempt to eliminate or identify viruses and other malicious code transfers.

Information passed between public and private networks may be encoded as binary information in some applications (e.g., imagery, the size of the piece of information to be processed may be

very large). The guard will have to reconstruct the entire message from multiple packets, which requires large working memory space. Then, the guard must pass the information through filtering and processing rules. With large files, this action may take a nontrivial amount of time. If any of the imagery files are time sensitive (i.e., used as part of a training exercise that requires commands to be issued based on the imagery files), the guard may add delay that degrades the usability of the information.

Note that data transfer between private and public networks involves risks, and one must take steps to mitigate risk. Processing, filtering, and blocking techniques involve inexact attempts to filter private data from outgoing transmission through content checking against a predefined list of prohibited strings. Scanning and detecting virus-infected executables, and blocking executables are also conducted. Because an almost infinite number of possible executables exist and malicious ones can be detected only through prior knowledge of their existence, the problem of detecting “maliciousness” in an arbitrary executable is not computable. Furthermore, the problem is exacerbated by the exist of many executables that users wish to allow to cross the network boundary (e.g., Java applets, Active X controls, JavaScript, and Word macros) and that they would therefore not wish to filter out or block. Only by performing a detailed risk management tradeoff analysis, wherein operational needs are weighed against security concerns, can these issues be resolved.

Protection approaches that use processing, filtering, and blocking technologies rely on processing to allow information flow between two networks while attempting to detect and block the leakage of classified data and attacks. Such approaches include ACLs, malicious code detection, content checking, application/attachment checking, and public to private replication. These approaches are discussed in the following subsections.

Access Control Lists

The ACLs enable users to selectively access information. The ACLs identify which users are permitted access to secure files, databases, programs, and administrative power. Discretionary Access Control (DAC) is used to restrict access to a file. Only those users specified by the owner of the file are granted access permission to that file. Mandatory Access Control (MAC) occurs when the security policy is dictated by the system and not by the object owner. Before access can be permitted or denied, I&A of the user must be available. Guards use the I&A presented by the user to determine if an ACL applies to that user. For example, if an ACL requires authentication via digital signature, then permission will be denied immediately to all users who do not authenticate with a digital signature. When a user authenticates with a digital signature, access permission will be granted if that user is on that ACL.

Malicious Code Detection

Although not a part of the guard itself, malicious code detection is integral to providing the high-assurance level associated with guards. Attachments opened by the guard must be sent to the malicious code detector to scan for known macro viruses or other malicious code. Files that are reassembled must also be scanned for known malicious code. The high assurance that can be

provided by a guard can be undermined easily if the guard is allowed to pass information containing malicious code.

Content Checking

Content checking service scans internal and external e-mail to detect and remove content security threats. Dirty word search filters, which are configurable, may be applied to search for specific words and send rejection messages back to the originators' system. A dirty word search scans messages for certain security-sensitive words, as defined by a word list. The content checking feature can be adequately defined, developed, and verified to evaluate the contents of the data to be transferred through the guard to ensure that no information at a sensitive level is transferred to a lower level system.

Application/Attachment Checking

Part of the application layer assurance offered by guards is application checking. This mechanism protects against attachments possessing improper file extensions. For example, the security policy for the organization may allow Microsoft Word attachments to pass through its mail guard. However, simply inspecting the file extension to verify that it is ".doc" is not enough to assure that the file is actually a Word file. The guard must launch its version of Microsoft Word and attempt to actually open the file. If the file cannot be opened, it either has errors or is mislabeled, and it should not be allowed to pass through the guard. If the file can be opened, it should be passed to a gateway malicious code checker to check for macro viruses. If no macro viruses are found and its message passes all other content checking filters, the attachment may be allowed to pass through the guard.

Public to Private Replication

Public to private replication allows users on a private network to receive data that originates on a public network, without having to explicitly request that the data be sent from the public servers. Replication can be used for network access by pushing data from a public network to a private network. It can give the private network any application that passes messages from one host to another. The primary security property of replication is the prevention of data flows from a private to a public network.

A common example of this technology is a database replication. If a node on a private network requires access to a database on a public server, the database can be duplicated on another server that is reachable by the private network. The guard controls the information flow between the replicated database and the private node. The private node may only have read privileges to the database, and not be able to write, depending on the security policy for the private network. The ability to write to the database would be dependent on the guards' private network and the guards' ability to reliably downgrade information. Other examples of replication are File Transfer Protocol (FTP), e-mail, and Web Push protocols.

Replication does not reduce the potential risk that data replicated into the private network may be hostile executable code. To mitigate this risk, a guard would have to be implemented so that data could be first replicated in this network guard. The guard inspects the data for potentially hostile code and ensures that the data passes this inspection before being forwarded into a private network.

To prevent data leakage from private networks to a public network, replication does not allow a direct back channel to send message acknowledgments from a private network to the public network; doing so would allow a large covert channel. The replication acts as an intermediary, sending acknowledgments to the public sender, and receiving acknowledgments from the private recipient. The public sender cannot determine with precision the timing of the acknowledgments sent from the private side. Hence, the intermediate buffer within the replication process reduces the bandwidth of the back channel. This action disconnects any direct communication from private networks to a public network.

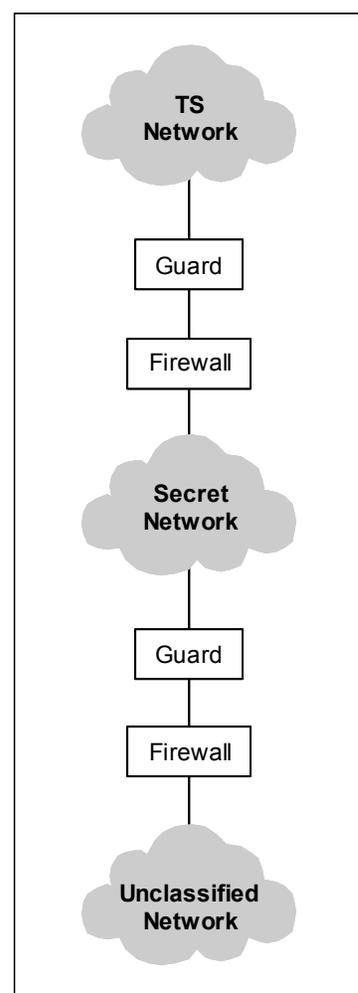
6.3.5.4 Cascading

Cascading occurs when two or more guards are used to connect three different networks containing information of three or more different levels. For example, if a top secret and secret network establish an agreement and a connection and the secret network has a preexisting connection to an unclassified network, the possibility exists for a path between the top secret and unclassified network. Please refer to Figure 6.3-3. The security policy for each guard needs to be examined to determine if a possible connection exists between the top secret and the unclassified network. Possible methods to reduce the risk associated with cascading are to allow different services through the two guards or restrict each user to interact with a single guard. When establishing a connection between two different networks using a guard, the connections each network have to other networks needs to be considered.

6.3.6 Selection Criteria

When selecting a guard, the following should be taken into consideration:

- The guard should send and receive e-mail between the private network and the public network.
- The guard should conform to standards used in the wider community.
- The guard should allow users to send and receive attachments in both directions.



iatf_6_3_3_0029

Figure 6.3-3.
Cascading Protection

UNCLASSIFIED

Guards

IATF Release 3.1—September 2002

- The guard should provide a user-friendly and seamless e-mail capability that passes messages with transit times comparable to those of a commercial electronic Message Transfer Agent (MTA).
- The guard should run on a trusted platform.
- The guard should only permit e-mail protocols (SMTPs) to pass through the guard.
- The guard should allow only authorized users to send and/or receive a message by performing access control on both the source and destination addresses of the message.
- The guard should prevent message flow directly between the high side WAN and the guard in either direction.
- The guard should allow only a properly labeled message to pass from the private level to the public level; each message must include a classification label.
- The guard should ensure that the security level of a message subsumes (is equal to or greater than) the security level of its attachment(s).
- The guard should protect against unauthorized disclosure of information from a private network.
- The guard should provide safeguards to protect the private side from attacks (including penetration, malicious code, and denial of service) from the public side.
- The guard should allow word or phrase search.
- The guard should support user digital signatures and encryption applications.
- The guard should support a digital signature or encryption capability.
- The guard should audit all security-related functions.
- The guard should provide an access control mechanism to limit access to the guard's controls and provide separate roles for the security administration, system operator, and mail administration functions.
- The guard should provide rules-based sanitization (i.e., message content modification) of fixed format messages from high levels through low levels.
- The guard should ensure that only allowed data is distributed.
- The guard should validate the proper message construction, including configurable verification of message content.
- The guard should provide secure bridge for passing messages between networks of differing levels of security.
- The guard should downgrade high-level data from designated communications channels according to validated rules.

- The guard should verify that the data meets a set of rigorously controlled criteria.
- The guard should prevent disclosure or release data to unauthorized consumers.
- The guard should communicate with only specified hosts on the public networks.
- The guard should prevent workstations from being used as a pass-through or gateway device from the public sides for any communications, including mail.

6.3.7 Framework Guidance

6.3.7.1 Case 1: File Transfers From a Top Secret to a Secret Network

This case study represents a situation in which a user on a secret network must obtain files from a user on a top secret network. Major risks are involved when connecting differing LANs. Therefore, when data files are to be transferred between networks of differing classification levels, the requirement arises for a guard that can recognize the FTP. Please refer to the Internet Engineering Task Force Request for Comment (RFC) 959 for additional information about the FTP, <http://www.ietf.org/rfc/rfc0959.txt?number=959>. [8] The guard's function is to permit communication between different classification boundaries while preventing the leakage of sensitive information. Included with the risks of connecting networks of differing classifications is the accidental or malicious release of data from one network to another. Therefore, when files must be transferred from a top secret network to a secret network, a guard can ensure that only permissible files are released. To be capable of this function, a guard should be able to process files regardless of type (e.g., graphic interchange format [GIF], Moving Pictures Expert Group [MPEG] file format, hypertext markup language [HTML]). The file will be subject to review by the established application checking policy to scan the contents and verify the sensitivity level. The guard will then downgrade files to allow releasability of the file to a lower sensitivity level user. Downgrading only occurs if the file's content meets the requirements of the sensitivity level of the network for which the data is being delivered. Downgrading is the change of a classification label to a lower level without changing the contents of the data.

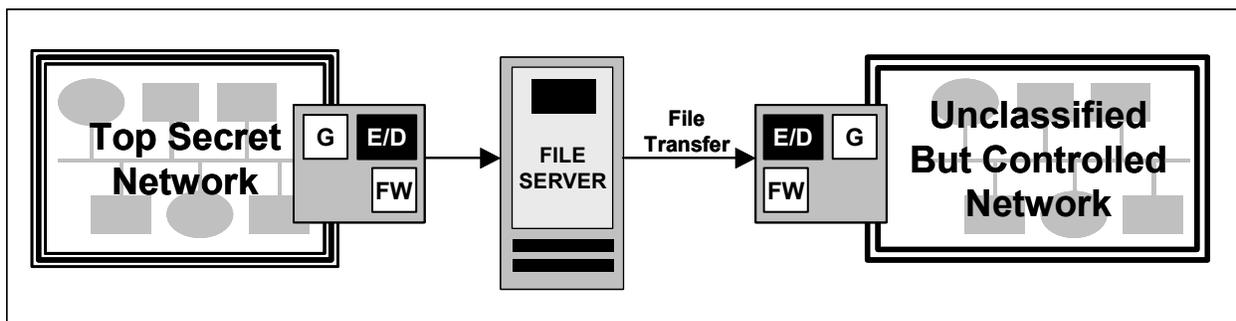
In addition, limits must be placed as to which users have permission to release files from the top secret network and which users on the secret network have permission to obtain these files. The originator of a file will have permission granted through an ACL kept by the guard to release files to the lower level network, secret. In return, the recipient must also have permission granted to access files that were released from the top secret network. Data owners must be able to restrict access to their data, and the system must also be able to deny access. DAC is the access control mechanism that allows the file owners to grant or deny access to users. The file owner can also specify an ACL to assign access permission to additional users or groups. MAC is a system-enforced access control mechanism that uses clearances and sensitivity labels to enforce security policy. MAC associates information requested from a user with the user's accessible security level. If data is classified as top secret, the information owner cannot make the information available to users who do not have access to top secret data. When access is

UNCLASSIFIED

Guards

IATF Release 3.1—September 2002

restricted, authentication and authorization policies must be in place. Authentication verifies the claimed identity of users from a preexisting label. Authorization is the determination of privileges a user has to grant permission for access of requested information. Authentication and authorization must be performed for all users requesting sensitive files from a user, as shown in Figure 6.3-4. Files may be stored on a server, making the files available to users on the secret networks who have permission to access the files. The server that allows the release of files shall be a COTS product that receives files and places them in a directory so that they will be accessible to authorized users. A guard must also be configurable to allow changes to be made to a database. Changes made to the master database of downgraded data shall be applied to replicated databases in near real time.



iatf_6_3_4_0030

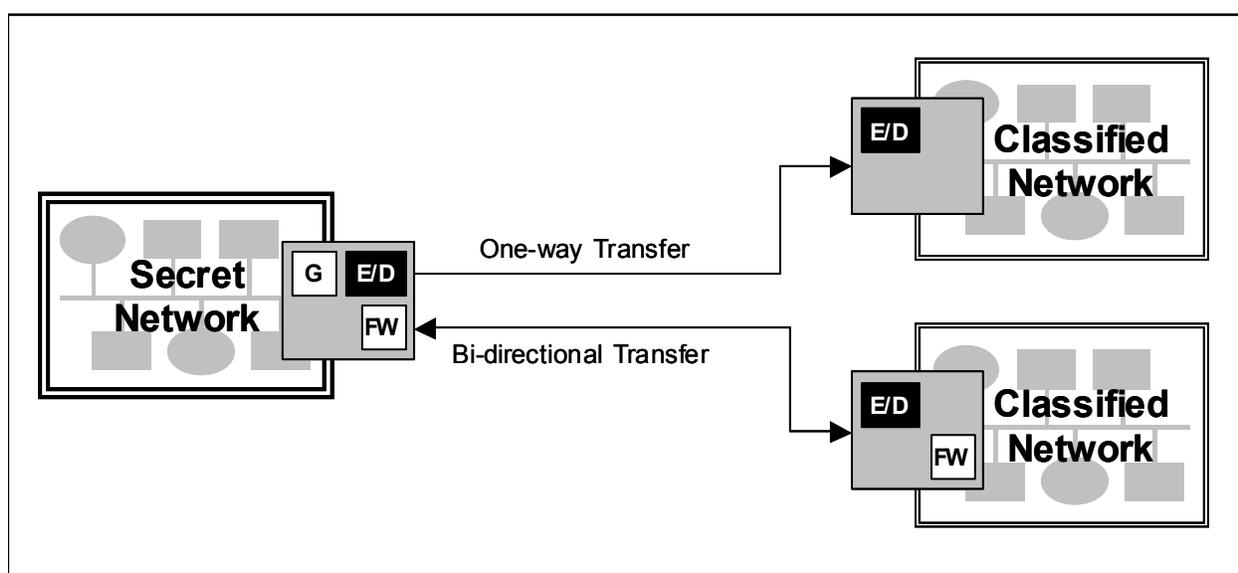
Figure 6.3-4. File Transfers

In keeping with the established releasability policy for file transfers, the guard will release the data to the lower level (secret) network based on the match of the content label and the security attributes of the recipient. The releasability policy followed by the guard shall adhere to the following:

- The guard shall allow only a very small set of users on the top secret network to release files.
- The guard shall maintain an ACL of these users and check the list every time a file is submitted for release.
- Only files of a specific format (plain text or HTML) shall be releasable.
- Strict audit logs shall be kept on the guard of all released files.
- Released files shall be scanned for content.
- Images contained within a file shall be reviewed.
- All files shall be authenticated (for example, digital signatures).

6.3.7.2 Case 2: Releasability From Secret to Unclassified Networks

When opening communication channels between secret and unclassified networks, a determination shall be made as to whether a bidirectional flow of information through a guard will be allowed. Guards differ in that some support only one-way transfers of information, whereas others support a bidirectional flow of information. Releasing information from a secret to an unclassified network can be performed through e-mail transmissions. Therefore, a mail guard is required, as shown in Figure 6.3-5, and can be coupled with a firewall to further enhance the security measures taken to protect the secret enclave.



iatf_6_3_5_0031

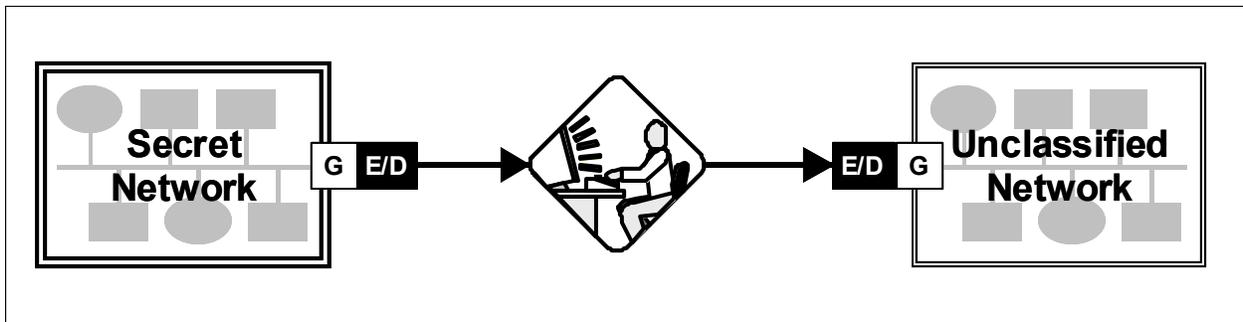
Figure 6.3-5. Secret to Unclassified Releasability

The mail guard enforces the policy for release of messages from the secret network. This policy may include the following:

- Content filtering/dirty word search.
- Malicious code checking.
- Message format check.
- Envelope filtering to determine if a sender and receiver are permitted to send and receive messages.
- Authentication (for example, cryptographic digital signatures).
- Message journaling/logging.
- Allowance or disallowance of attachments.

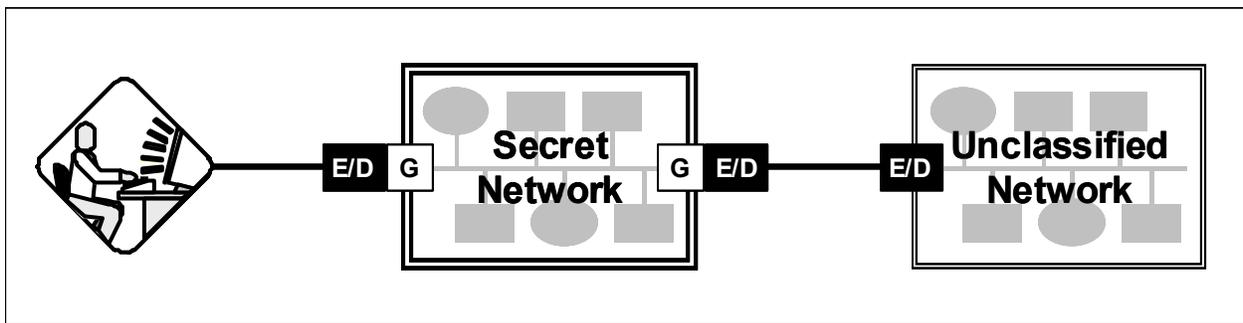
- Review of attachment.
- Allowance or disallowance of mail receipts.
- Allowance and disallowance of sending blind carbon copies of messages.
- Maintenance and review audit logs of all mail message transfers for questionable actions.

Although the goal is to have a guard that has full functionality and can automatically review all information, a human reviewer may also be placed to review messages before the guard receives and reviews messages. A user can manually review messages by being placed between the guards of two separate networks, as shown in Figure 6-3-6. Or, as shown in Figure 6.3-7, a human reviewer can review information before the guard for verification that the sensitivity level of the information can be released to the unclassified network.



latf_6_3_6_0032

Figure 6.3-6. Human Reviewer-Man in the Middle



latf_6_3_7_0033

Figure 6.3-7. Releasability Human Verification

The human reviewer has the release authority over a message with respect to allowing or rejecting the sending of the message. The established security policy may require that all messages are reviewed or only rejected messages are reviewed, or perhaps messages might not need to be manually approved. The functionality goal of a guard is to allow a fully automated review process. A process without a human reviewer must have fully automated guards that are able to check content, check attachments to e-mail messages, have a configurable security filter,

perform dirty word searches, and have imagery processing capabilities. Dirty word searches are looking for words or codes that could be used to disclose sensitive information.

Encrypted messages must be able to be decrypted before processing through the guard, allowing the message to be released. Guards with decryption capability (which may be through embedded FORTEZZA cards) will decrypt a copy of a message and, upon release approval, pass the original message to the recipient and discard the decrypted copy. If a message cannot be decrypted, then the guard must reject that message. A rejection notice policy shall be established to address the handling of message rejection notices. The rejection notice policy may have notices sent to only the mail administrator of the secret network or may also allow rejection notices to be sent to the user. A policy shall also be established as to the allowance of mail receipts.

Confirmation that recipients have received an e-mail can be equally important as the security measures taken to protect the information contained within the e-mail. Mail receipts, however, cannot always be relied on because some e-mail servers will not allow receipts out of their own e-mail system. Therefore, when sending e-mail through a guard, rules must be established regarding the allowance of return receipts. Automatic return receipts may not be part of the guard's security policy. However, once a recipient verifies that the appropriate message was received, a signed receipt can be generated and sent to the guard for filtering and then forwarded to the originator. In place of return receipts, servers capable of providing automatic tracking capabilities can be used to confirm document receipt.

Remote access capabilities pose a risk as a backdoor mechanism to gain access into a network. Therefore, for this scenario, the guard security mechanism would be most effective if coupled with a firewall. A firewall will protect the LAN from Internet or modem attacks by blocking direct access. Besides maintaining network access controls, the firewall will also maintain extensive audit records detailing successful and unsuccessful attempts to access the system. Once connected and authenticated, a dial-in user then has the same Internet services as local users. Internet connectivity is an inherent risk because it opens up channels of additional risk when connecting secret networks to unclassified networks. Therefore, a guard must be able to recognize Web-based protocols (i.e., HTTP) to mitigate risk for access into the networks.

Another important means of communicating for business is real-time messaging. Therefore, guards should be able to support real-time and instant messaging. When communicating by real-time messaging, messages should be ensured against corruption, tampering, recording, and nonplayback.

6.3.8 Technology Gaps

6.3.8.1 High Volume of Binary Data

Some applications require that information be passed in a binary representation. Examples of these applications are voice, imagery, and video. Binary data is more difficult to perform content checking on and to pass through filter rules. Guard technology needs to become faster to allow

large amounts of binary files and streaming binary information to pass through the high-assurance mechanisms to which other information is subject.

6.3.8.2 Quality of Service

Quality of service (QoS) is being deployed in networks to support real-time applications, such as voice, video, and for other applications that might have strict latency requirements. Several different approaches exist for supporting QoS in IP networks. Although multiple approaches exist for providing QoS in an IP network, the guard that is implemented must support the QoS strategy for the organization.

Guards must support QoS mechanisms provided by the network. All incoming traffic is passed through the guard. If the QoS mechanism is not supported by the guard, end-to-end QoS that is required by the application cannot be supported.

6.3.8.3 High Speed Across Optical and Other Networks

Most guards are designed to work in IP networks. However, many different types of networks could make use of guard technology, including all optical networks and asynchronous transfer mode (ATM) networks. These networks typically operate at speeds in excess of those of IP networks. In addition to adding the proper interface to the guard, the filtering mechanisms within the guard must be capable of the speeds on the optical network. Furthermore, optical and ATM networks are very sensitive to delays. If the guard is incapable of supporting the bandwidth requirements of a connection, communications through the guard may be degraded to a point where further connections cannot be accepted.

6.3.8.4 HyperText Markup Language Browsing

Today's network environment uses HTML traffic for a variety of applications. Having a guard that supported HTML browsing for Internet or internal HTML would greatly increase the functionality of organizations.

To support HTML, a guard would have to allow requests (i.e., domain name server [DNS] queries, requests for Web pages) to pass through the guard. When the response returns, the guard must intercept the message and perform its checking before it is allowed to pass back to the user. All this must happen in real time to allow for human interaction and viewing behind the guard.

References

1. Reserved.
2. CERT® Coordination Center. 17 July 2000 www.cert.org.
3. Software Engineering Management Practices. Carnegie Mellon Software Engineering Institute. 18 July 2000. 12 June 2000 <http://www.sei.cmu.edu/managing/managing.html>.
4. Trusted Product Evaluation Program. 12 June 2000. <http://www.radium.ncsc.mil/tpep/tpep.html>.
5. Lashley Brian and Andrzej Tarski. SSL <http://welcome.to/ssl>.
6. Federal Information Processing Standards Publications (FIPS) Pub 180. Secure Hash Standard 17 Apr 96 <http://www.itl.nist.gov/fipspubs/by-num.htm>.
7. Federal Information Processing Standards Publications (FIPS) 185. Escrowed Encryption Standard. 09 Feb 94 <http://www.itl.nist.gov/fipspubs/by-num.htm>.
8. Postal, J. and J. Reynolds. “File Transfer Protocol (FTP)”. RFC 959, ISI, 1985 October. <http://www.ietf.org/rfc/rfc0959.txt?number=959>.

Additional References

- a. Computer Advisory Incident Capability. Department of Energy. 6 June 2000, <http://ciac.llnl.gov/ciac/bulletins/I-005c.shtml>. Enter at <http://ciac.llnl.gov>, then navigate to: <<http://ciac.llnl.gov/ciac/bulletins/I-005c.shtml>>.
- b. Digital Signature Trust Co. 3 July 2000. <http://www.digsigtrust.com/>.
- c. Reserved.
- d. National Institute of Standards and Technology (NIST) FIPS 186. FACT SHEET ON DIGITAL SIGNATURE STANDARD. Online posting May 1994. 3 July 2000 http://www.nist.gov/public_affairs/releases/digsigst.htm.
- e. NetworkWorldFusion News. 20 June 2000. <http://www.nwfusion.com/news/tech/0906tech.html>.
- f. Stronghold Webserver Administration Guide Chapter 6 SSL Authentication and Encryption. 22 June 2000 http://mclean2.his.com/docs/Administration_Guide/SSL.html.
- g. Stronghold Webserver Administration Guide Chapter 6 SSL Authentication and Encryption. 22 June 2000 <http://developer.netscape.com/docs/manuals/security/ssl/contents.htm>.
- h. The Source of JAVA™ Technology. Smart Card Overview. 5 July 2000. <http://www.java.sun.com/products/javacard/smartcards.html>.
- i. Smart Card Basics.com. 5 July 2000 <<http://www.smartcardbasics.com/security.html>>.

UNCLASSIFIED

Guards

IATF Release 3.1—September 2002

- j. Hulme, George V. “Secure Document Delivery Gains Favor.” *InformationWeek*. 17 July, 2000.

6.4 Network Monitoring Within Enclave Boundaries and External Connections

A fundamental tenet of the defense-in-depth strategy is to prevent cyber attacks from penetrating networks and to detect and to respond effectively to mitigate the effects of attacks that do. As discussed above, an integral aspect of the defense-in-depth strategy embraced by this Framework is enclave boundary protection, which often takes the form of firewalls and virtual private networks (VPN). While these technologies offer perimeter and access controls, “authorized” internal and remote users can attempt probing, misuse, and malicious activities within an enclave. Firewalls do not monitor authorized users’ actions, nor do they address internal (insider) threats. Firewalls also must allow some degree of access, which may open the door for external vulnerability probing and the potential for attacks.

Detect and respond capabilities are complex structures that run the gamut of intrusion and attack detection, characterization, and response. The various detection aspects of detect and respond are actually measurement services. Intrusion detection, network scanning, and the like are measurement functions that determine the effectiveness of the deployed protection systems and procedures on a continuous or periodic basis. In themselves, detection capabilities are not protection measures. The respond aspect can initiate changes to existing protection systems (e.g., configuration changes in a firewall to block an attacker’s Internet Protocol [IP] address) or deploy additional protection measures (e.g., placement of another firewall appliance). The local environments (within enclaves) are the logical location for network-based sensors. This section addresses sensors that operate in near real time. Specific network monitoring technologies addressed in the Framework are shown in Figure 6.4-1. Section 6.5, Network Scanners Within Enclave Boundaries, addresses sensors that typically operate off-line. Section 7.2, Host-Based Detect and Respond Capabilities Within Computing Environments, provides similar guidance for host-based sensors.

Local environments have the option to implement as much or as little above the sensors as they believe is prudent, obtaining services and support from the infrastructure as necessary. Section 8.2 of the Framework provides an in-depth discussion of the various detect and respond processes and functions in the context of a supporting information assurance (IA) infrastructure capability. It also offers guidance on technologies for processes beyond the sensors,

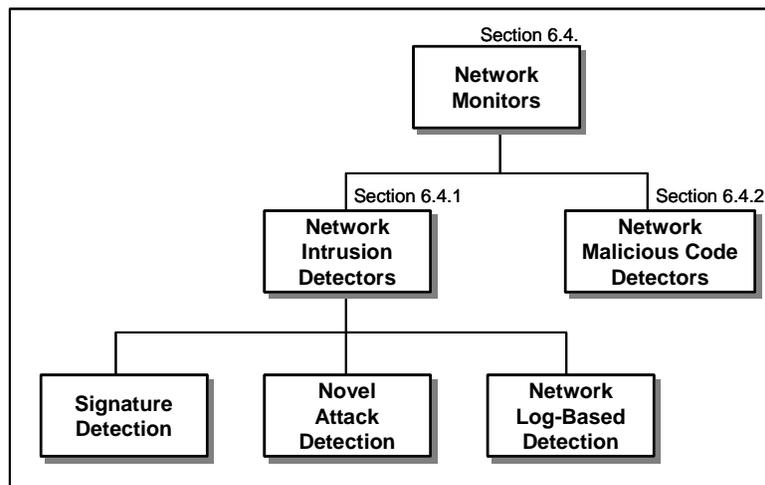


Figure 6.4-1. Breakdown of Network Monitor Technologies

iatf_6_4_1_0001

but recognizes that these processes may be implemented at any level in a network hierarchy, including a local enclave environment.

Network monitors, including network intrusion detection and network malicious code detection technology areas, are covered in this section. The section provides an overview of each relevant technology, general considerations for their use, the rationale for selecting available features, deployment considerations, and a perspective on how these technologies are typically bundled into products. The section concludes with sources for additional information and a list of the references used in developing this guidance.

6.4.1 Network Intrusion Detection

The goal of an intrusion detection system (IDS) is to identify and potentially stop unauthorized use, misuse, and abuse of computer systems by both internal network users and external attackers in near real time. Because this section of the Framework addresses network-based monitoring, these discussions center on operations using network information. As discussed in Section 7.2, Host-Based Detect and Respond Capabilities Within Computing Environments, similar structures and technologies are also available for performing comparable functions using host-based information.

6.4.1.1 Technology Overview

Normally, a dedicated computer is deployed for each network IDS on each network or network segment being monitored. A network interface card (NIC) is placed into promiscuous mode, enabling the IDS software to watch all traffic passing from computer to computer on that particular network. The IDS software looks for signs of abuse (e.g., malformed packets, incorrect source or destination addresses, and particular key words).

A network-based IDS bases its attack detection on a comparison of the parameters of the user's session and the user's commands with a rules-base of techniques used by attackers to penetrate a system. These techniques, referred to as "attack signatures," are what network-based IDSs look for in the behavior of network traffic. An attack signature can be any pattern or sequence of patterns that constitutes a known security violation. The patterns are monitored on the network data. The level of sophistication of an intrusion can range from a single event, events that occur over time, and sequential events that together constitute an intrusion.

Detection Approaches

There are three basic technology approaches for performing network intrusion detection:

- **Signature detection approach** typically incorporates search engines that seek to identify known intrusion or attack signatures.
- **Novel attack detection** is based on identifying abnormal network behavior that could be indicative of an intrusion.

- **Network log-based detection** monitors for attacks using audit logs of network components.

Signature Detection Approach. This approach utilizes traffic analysis to compare session data with a known database of popular attack signatures. These IDSs act like a “sniffer” of network traffic on the network, caching network traffic for analysis. Typically, they do not introduce path delays while they are processing traffic and therefore do not impact network or application performance. Vendors refer to this operation as “real time.” Northcutt offers the perspective that “one of the great marketing lies in intrusion detection is ‘real time.’ What marketers mean by real time is that intrusion detection analysts are supposed to respond to beeps and alarms.” [“Network Intrusion Detection An Analyst’s Handbook,” by Stephen Northcutt, New Riders Publishing, 1999]

This technology examines the traffic against a predefined set of rules or attack signatures, typically using one of these techniques:

- **Pattern expression or bytecode matching.** The ability to determine regular behavior patterns to distinguish abnormal patterns, as well as determine if the traffic being monitored matches a predefined attack signature.
- **Frequency or threshold crossing.** The ability to establish a predefined threshold; if the threshold is exceeded, an intrusion is assumed.

There are two basic signature-based options: one, referred to as a “static signature IDS,” which uses a built-in attack signature base and a second, “dynamic signature IDS,” which relies on signature information that can be loaded dynamically into the IDS. Some product vendors provide routine updates of attack signatures. Some IDS tools give the customer the capability to customize attack signatures.

Novel Attack Detection. This relatively new detection strategy monitors Transmission Control Protocol (TCP) Dump data and attempts to filter out activities that are considered normal behavior. The genesis for this approach was to implement a sensor that would allow an analyst to evaluate large quantities of network information and select anomalous behavior. Unlike signature detection techniques, in which the sensor has to have a priori knowledge of specific attack scripts, this technique relies on screening by an analyst and can detect a variety of probes and attacks that other detection approaches miss. Initial versions dealt with packet header information only. Later versions capture the full packet content.

Network Log-Based Detection. This detection technique focuses on the monitoring of audit logs from network devices. It has two major components. One is a catalog of audited events that are considered “bad” behavior. The catalog could include attack profiles, suspicious activity profiles, and activities defined as unacceptable. The second component is an audit trail analysis module. Audit trails come from a chronological record of activities on a system. The analysis module examines the monitored system’s audit trail for activity that matches activity in the catalog; when a match occurs, intrusive activity is assumed. Audit-based systems may also

provide the ability to identify and track additional activity by an individual who is suspected of intrusive behavior.

IDS Tuning Options

Typically, an IDS provides capabilities for selecting which attacks are being monitored. Depending on the specific implementation of an IDS, it is often possible to select which attacks will be monitored, what the response will be for each detected intrusion, specific source and/or destination addresses (to be monitored or excluded), and characterizations of the class (indication of the importance or severity) of each alarm. This capability, to configure the monitoring screen, is critical to optimize the monitoring capability of an IDS. In this way, it is possible to focus the sensor on specific events of interest and the response that the IDS will have on detection of events.

Response Options

While the sensors detect and collect information about intrusions, it is the analyst who interprets the results. Some network IDS technologies offer automated response features to various alarms. In addition to logging the session and reporting, as indicated below, some have the option to terminate the connection, shun an address that was the source of the detected intrusion, throttle the amount of traffic allowed through a port, or even close down a site's operation. In some cases, the IDS can accomplish these operations itself; in others, it works in conjunction with a network interface device (e.g., firewall, router, or gateway) to achieve the desired result.

Reporting Mechanisms

When it detects a threat, a network IDS generally sends an alert to a centralized management console where alert information can be recorded and brought to the attention of an administrator. Some of the network IDS technologies offer additional reporting capabilities. Some can automatically send an e-mail message over the network to alert an operator to the alarm condition. Others can initiate a message to a pager.

6.4.1.2 General Considerations for Use

Network IDS technologies are an important aspect of an enclave's defensive posture. Table 6.4-1 provides a synopsis of advantages and disadvantages of using network-based IDS technology.

Table 6.4-1. Network-Based IDS Considerations

Advantages	Disadvantages
<p>Provides real-time measure of the adequacy of an infrastructure's network protection measures.</p> <p>Network-level sensors can monitor and detect network attacks (e.g., SYN flood and packet storm attacks).</p> <p>The insertion of a network-level sensor does not affect existing data sources from a performance and reliability standpoint.</p> <p>Well-placed network sensors are designed to provide an integrated, enterprise wide view, at the management console, of any large-scale attack.</p> <p>Operator expertise and training only required for the single network IDS platform.</p>	<p>Some network-based systems can infer from network traffic what is happening on hosts, yet they cannot tell the outcome of the commands executed on the host.</p> <p>Network-based monitoring and intrusion detection becomes more difficult on modern switched networks. Switched networks establish a network segment for each host; therefore, network-based sensors are reduced to monitoring a single host. Network switches that support a monitoring or scanning port can at least partially mitigate this issue.</p> <p>Network-based sensors cannot scan protocols or content if network traffic is encrypted.</p> <p>Must be used on each network segment because they are unable to see across routers and switches.</p> <p>Current network-based monitoring technologies cannot handle high-speed networks.</p>

The network-based IDS is typically deployed in the middle of a communications path between client and server and has access to data at all layers of communication. This process allows this type of sensor to do extensive analysis for attack detection and provide detection in near real time. Since a network IDS runs on an independent computer, there is no impact on the performance of other network resources.

Today, network traffic is often encrypted through mechanisms such as VPNs. A network IDS simply watches information traversing a network and is typically not capable of decrypting the packets. In these cases, the encryption blinds the IDS to any attacks that may occur. This type of sensor relies on passive protocol analysis causing it to “fail open.” This leaves the network available and vulnerable and leaves the IDS itself open to potential compromise.

Throughput is another concern. If only one network IDS computer was to monitor an entire network, that one computer would have to be capable of scanning every single network packet. At modest throughput levels (e.g., 50 Mb/s), most network IDSs can keep pace with the incoming stream of data. However, as network bandwidth increases and network loads reach higher rates (100 Mbps and beyond), one or even several network IDS computers may not be able to keep up with the flow of traffic.

6.4.1.3 Important Features

When selecting a network IDS, there are a number of features that should be considered. This section identifies these important features. The section that follows discusses rationales for the selection of these features.

UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections
IATF Release 3.1—September 2002

Detection

- Detection approach used by the network IDS.
- Does it perform packet fragmentation/reassembly?
- Which threshold adjustments can be made to the IDS?

Signatures

- Number of events/signatures that can be stored.
- How often the signatures can be updated.
- Is the update static (manual) or dynamic (automated)?
- Are user-defined attack signatures allowed; if so, are the scripting tools easy to use?

Operations

- Can it protect itself from unauthorized modifications?
- Does it recover from system crashes?

Response Options

- Does it offer provisions for reconfiguring firewalls?
- Does it have session closing and reset capabilities?
- Does it have address blocking (shunning) capabilities?
- Can it execute program scripts on alarm?

Reporting Options

- Does it report in real time to a workstation?
- Can network and host-based IDSs report to the same analyst console?
- Is the reporting interval configurable?
- Can IDS notify personnel using e-mail or pagers?
- Is the amount/type of information reported to a management station configurable?

Performance

- Network compatibility.
- Number of packets that can be processed over an interval (packet size/bandwidth).
- Rate of false positives (identification of a nonintrusive activity as intrusive).
- Rate of false negatives (failure to identify an intrusive activity).

Platform

- Operating system.
- Type of platform required to host network IDS.
- Processing burden for anticipated network traffic load.

Console Considerations

- **Operator Interface.** Type of command and monitoring provisions available to an operator.
- **Mark as Analyzed.** Ability to clear or mark selected alarms that have been reviewed
- **Drill Down.** Ability to provide additional information for selected events.
- **Correlation.** Tools to correlate events based on source, destination, type.
- **Report Generation.** Ability to generate reports upon event detection and as periodic summary reports.

6.4.1.4 Rationale for Selecting Features

Detect and respond capabilities exemplify the necessity of integrating operations and personnel considerations with the selection of technology solutions, consistent with the overall defense-in-depth philosophy. As indicated earlier, network monitoring does not itself offer protection from intrusions or attacks. It should really be considered instrumentation that monitors (and “measures”) the effectiveness of a network’s existing protection structures. It is up to operators (personnel and operations) to interpret the outputs of the IDS and initiate an appropriate response. If full-time operators¹ are not available to interpret and formulate responses based on the IDS outputs, then IDS implementations will not typically add real value. In this case, it is likely that IDS deployments should not be considered. Otherwise, when selecting features for an IDS, there are a number of factors to be considered, based on how the IDS is intended to be used, whether full- or part-time operators will be available, and the skills of the operators to interpret the results.

Detection

The type of detection mechanism is one primary consideration when selecting a network IDS technology. Another important consideration is the anticipated skills of the attacker. Signature-based detection, which is the traditional method used in network IDS technologies, typically lacks the ability to detect new (or modified) versions of attack strings. While many intrusions (typical of novices) use standard attack sequences (often downloaded from hacker bulletin boards), an accomplished adversary will have the capability to create new attacks or modify old attacks and thus thwart traditional signature detection mechanisms. Anomaly and misuse detection approaches have greater flexibility for identifying new or modified attacks (since they monitor network usage or behavior). But they are more complex to operate and not necessarily as responsive to traditional attack strings. These are also the only mechanisms currently available to monitor actions of otherwise authorized users for inadvertent or intentional misuse.

¹ Ideally operators should be available on a 24x7 basis. The number of operators will depend on the traffic loads and anticipated numbers of incidents. It is not uncommon to experience hundreds of thousands of intrusion alerts per day, and each must be investigated to determine which, if any, are serious threats.

UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections
IATF Release 3.1—September 2002

The ability of the various detection schemes to correctly identify intrusions is a fundamental consideration. The rate of false positives (alerts resulting from normal traffic) and false negatives (failure to identify a real intrusion attempt) should be considered. While the technologies are continually being refined for improved performance, there are inherent features that may limit performance (e.g., anomaly detectors have been known to generate significantly higher false positive indications).

As always, any decision is based on level of risk, anticipated performance, cost (for purchase, deployment, and operation), and operational impact. The Framework recommends consideration for deployment of multiple attack detection schemes, ideally from different vendor sources. In this way, there is a greater likelihood of detection by at least one of the mechanisms deployed.

Signatures

If a signature-based IDS is selected, it is desirable to have as many signatures as possible used for detection. However, there is usually an inverse relationship among the number of signatures, the response time for possible detection. The amount of traffic that can be monitored is also typically reduced when a large signature set is employed. Since the lists of possible attacks change frequently, it is strongly recommended that the IDS be capable of dynamically loading signatures. It is usually operationally more feasible and efficient if the downloading is handled on an enterprise (or at least site) basis. Most vendors that offer dynamic loading of signatures provide periodic updates to their signature base. While the update periods differ among vendors, a good rule of thumb is the more often the better. If operators have the skills to create custom signatures, then having the ability to support user-defined attacks is also desirable, particularly if custom attacks are found in one of your sites.

Operations

It is desirable for the IDS to be easily configurable according to the security policies of the information system that is being monitored. Consideration should also be given to the IDS's ability to adapt to changes in system and user behavior over time (e.g., new applications being installed, users changing from one activity to another, or new resources becoming available that cause changes in system resource usage patterns).

By their nature, IDS sensors are located where intrusions are anticipated. Thus, it is important that an adversary not be capable of modifying the IDS to render it ineffective. It is desirable that the IDS be able to perform self-monitoring, detect unauthorized modifications, and notify an attendant console. To simplify recovery of operations after an intrusion, it is also desirable that the IDS be able to recover from system crashes, either accidental or due to malicious activity, and upon startup, be able to recover its previous state and resume its operation unaffected.

Response Options

Many available solutions offer automated response options that seem on the surface to be very desirable. They imply that little or no human interaction is involved, as the devices can provide

an immediate response. There are serious pitfalls to consider, however, before these options are deployed. First, it is not uncommon for a network IDS to find thousands (and possibly hundreds of thousands) of events daily, depending on where it is employed, characteristics of the normal network traffic load, and many other factors. Often, the number of false positives may be high, giving rise to frequent unwarranted indications of intrusions. Automated responses that terminate connections, shun addresses, throttle traffic, or actually shut down a facility can often cause severe denial-of-service (DOS) threats to the network. It is strongly recommended that automated options not be used if there is a concern that they may cause DOS on the networks they are trying to defend.

Reporting Options

Most network-based IDSs report alarms to an operator console. (See discussion of console features, below.) The desirable level and frequency of reporting is based primarily on the availability and skills of the operators. Some network IDS technologies offer the option of paging or sending e-mail messages to notify personnel of alarms. While these sound desirable, they have the potential to give rise to operational issues. With an IDS detecting thousands of alarms a day, these features have the potential for overloading e-mail servers (creating a DOS threat themselves) or paging operators extremely frequently at all times of the day and night. Most often, these features are not recommended.

Performance

Network IDS performance varies due to the speed of the network, the amount of traffic, the number of nodes being protected, the number of attack signatures employed, and the power of the platform on which the IDS resides. IDSs may be overtaxed on busy networks. However, multiple IDSs can be placed on a given segment to subdivide host protection, thereby increasing performance and overall protection. For instance, high-speed networks employing asynchronous transfer mode (ATM), which uses packet fragmentation to improve efficiency over high-bandwidth communications, do pose problems in terms of performance and response.

Platform

A major issue for the selection of a network-based IDS is the type of computer skills (e.g., UNIX, NT) required for operators. Operators will likely need these skills to perform installation, configuration, adjustment, and maintenance. Since a network-based IDS usually is located on its own platform, the platform will have to be acquired and maintained, so it may be useful to select a technology that functions on the types of platforms used within the enterprise.

Console Considerations

As discussed in Section 8.2 of the Framework, the primary function of the console is to serve as an aid in the characterization and analysis of the many alarms that will be identified. Operators will have to not only identify alarms that were unwarranted, those that do not offer serious risks

to the network, and those that do, but also gain a first-order understanding of the source and impact of possible attacks.

Operator Interface. The type of interface that is operationally desired tends to be driven directly by operator preference. Novices typically prefer a graphical user interface (GUI) with intuitive operations, pull-down screens, and substantial aids available. Skilled operators may prefer command string operations, tailored screen options, and options for operator customization. It is best if operators can get a hands-on trial evaluation of the console capabilities prior to final selection.

Mark as Analyzed. Operators will typically be faced with large numbers of alarms that have to be analyzed and cleared. A capability that is usually critical is the ability to selectively keep track of alarms that have been reviewed.

Drill Down. Many network IDS consoles display a high level characterization of events in order to display the large number of alarms that are detected. Operators will usually have to access additional details about each alarm to be able to characterize it properly. It is very desirable for the console to be able to provide the additional levels of information when requested by the operator. As with the operator interface, the types of information desired will typically depend on the skills of the operators.

Correlation. In the same vein as drill-down features, operators will require tools for correlating events (e.g., based on source, destination, type of alarms, and events) in order to identify and properly characterize intrusions and attacks. This is particularly necessary in situations where the incidents are distributed in time or location. The ability of the console to integrate the reporting of various network-based and host-based IDSs and other relevant events is a strong plus, if the operators will use the additional information. Again, as with the operator interface, the types of tools desired will typically depend on the skills of the operators.

Report Generation. The type of reporting options will depend predominantly on the type of information operators will want to perform their characterization, and the organization's need for reporting to higher levels (e.g., periodic summary reports). It is always desirable to select a console that is capable of generating and disseminating reports with little extra effort beyond the hour-to-hour and minute-to-minute responsibilities that the operators will have otherwise.

6.4.1.5 Considerations for Deployment

Network architectures present another major challenge for a network IDS. Network switches, which segregate network traffic into specific individual "subnets," reduce network loads across an organization by implementing a form of "need to know" policy among connected computers. Network switches only allow traffic to enter a subnet if it is meant for a computer within that subnet; similarly, they only allow packets out of a subnet that are destined for a computer outside its particular realm.

A network IDS can see only traffic available on the segments on which it is installed. As long as the network IDS is placed on critical segments, it will be able to measure the effectiveness of the

security protection mechanisms for the most critical systems and applications. Within an enclave environment, there are a number of possible locations to consider in deploying a network IDS, as depicted in Figure 6.4-2. The challenge is to identify where the traffic of most interest (i.e., that most likely to be used as an attack channel) can be monitored.

The external gateways are an obvious candidate in that they allow the IDS to see all of the traffic destined for the enclave. If IDSs are placed outside the firewall, they have access to the raw wide area network (WAN) traffic (e.g., Internet) without the benefit of filtering by the firewall. If network encryption is used on that traffic, this will offer little if any value. Placing the IDS inside the firewall resolves network encryption issues but will not give any indication of the effectiveness of the firewall operation. Placing sensors at both points and correlating the output of the alarm causing packets that are detected outside but blocked by the firewall could provide this additional perspective. Note that these locations provide monitoring either for external traffic that is destined for the enclave or for internal traffic that is destined for the WAN. IDSs in these locations do not monitor traffic that is only internal to the enclave.

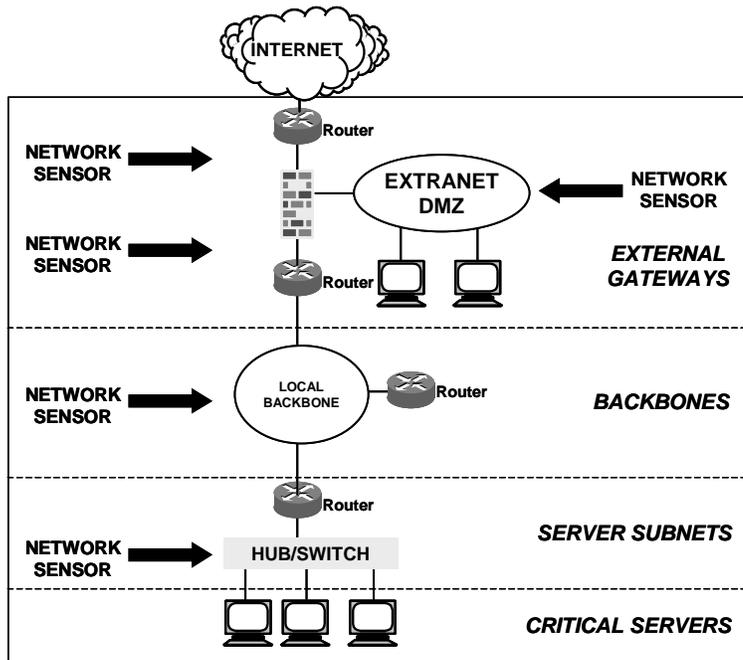


Figure 6.4-2. Network IDS Deployment Options

iattf_6_4_2_0015

If an extranet (or what may be referred to as a demilitarized zone, or DMZ) is deployed, an IDS on that segment of the network could offer monitoring of traffic from outsiders to assets structured for an isolated segment of the enclave.

The network backbone represents another deployment option. This option does provide access to internal traffic on the backbone. However, at this point in the network, consideration should be given to the traffic speeds and switching technologies employed on those backbones. In some cases (e.g., ATM, Fiber Distributed-Data Interface [FDDI]) the switching technologies and transmission speeds make currently available IDS technologies impractical.

A final placement option is on server subnets. This is typically a good option if hubs are used, so that all traffic on the subnet is available at each hub port. If switches are used rather than hubs, this is still a good option if there is a spanning port available (that allows access to all traffic). If not, the IDS will not have access to all the traffic through the switch and will be ineffective unless deployed between a host and a switch (or “onto” a host).

There is always a trade-off between the possible deployment locations and the number of IDSs to be deployed. Factors to consider include the workload of the operators needed to analyze and characterize the alarms that each IDS generates; the complexity of correlating the alarms that multiple monitors will generate for the same event; and the costs associated with purchase, installation, operation, and maintenance of the various deployment options.

6.4.1.6 Considerations for Operation

As discussed above, most IDS technologies provide the capability to tune the sensor to improve its performance for specific deployments. When an IDS is first deployed, it is prudent to operate the technology for some period depending on the complexity of the deployment to complete this tuning. This provides a means for determining that the IDS is capable of detecting alarms, and that the IDS is installed on the network as intended (by verifying network addresses that are monitored and the direction of traffic).

Tuning enables the IDS to preclude the detection of authorized traffic patterns that might otherwise cause false positive alarm indications. There are two fundamental approaches for tuning. The first approach is to have knowledge a priori of the traffic sources that could trigger false alarms. This could include the addresses of servers (that expect significant traffic), network management station locations (that normally sweep the network), and computers that are remotely located. The IDS can then be configured (tuned) to preclude these from causing an alarm.

While it is desirable to have such information ahead of time, it is often not available. The other approach is to run the IDS and have it find alarms. As alarms are detected, an analyst determines if indeed they reflect an intrusion or a false positive based on normal operation. This form of “discovery” also gives operators an opportunity to become familiar with the technology before it goes on-line operationally.

Tuning should not be thought of as strictly an installation process. This process should be done on a regular basis to refine detection mechanisms and focus them on real intrusions and to reduce false positives throughout IDS operation.

6.4.2 Malicious Code (or Virus) Detectors

Malicious code can attack authorized local area network (LAN) users, administrators, and individual workstation/personal computer users in numerous ways, such as modifying data in transit, replaying (inserting data), exploiting data execution, inserting and exploiting malicious code, exploiting protocols or infrastructure bugs, and modifying malicious software during production and/or distribution.

Over the past decade, malicious code (also commonly referred to as computer viruses²) has gone from an academic curiosity to a persistent, worldwide problem. Viruses can be written for and spread on virtually any computing platform. Typically, viruses are written to affect client personal computers. However, if the personal computer is connected to other machines on a LAN, it is possible for the virus to invade these machines as well. See Section 6.6, Malicious Code Protection, for detailed descriptions of the various types of malicious code, potential malicious code attacks and countermeasures, and requirements for malicious code detection products and technologies.

6.4.2.1 Technology Overview

Malicious code scanning technologies prevent and/or remove most types of malicious code. The use of malicious code scanning products with current virus definitions is crucial in preventing and/or detecting attacks by all types of malicious code.

There are several basic categories of antivirus (AV) technologies:

- **Preinfection Prevention Products.** A first level of defense against malicious code, used before a system has been attacked
- **Infection Prevention Products.** Used to stop replication processes and prevent malicious code from initially infecting the system.
- **Short-Term Infection Detection Products.** Used to detect an infection very soon after the infection has occurred
- **Long-Term Infection Detection Products.** Used to identify specific malicious code on a system that has already been infected for some time, usually removing the malicious code and returning the system to its prior functionality.

See Section 6.6.5.2, Viruses and E-Mail, for a more detailed description of the types of malicious code detection technologies.

6.4.2.2 Important Features

When selecting AV technologies, there are a number of features that should be considered. This section identifies important features for selection. The section that follows discusses the rationale for the selection of these features. Additional factors to consider when selecting a malicious code detection product can be found in Section 6.6.6, Selection Criteria.

² Throughout the remainder of this section, the term *virus* will be used to encompass the broader class of malicious code and delivery mechanisms.

Detection Capabilities

- Data integrity checks.
- Perimeter-level scanning for e-mail and Web traffic.
- Does tool exploit malicious mobile code?
- Real-time virus scanning.
- On-demand virus scanning.
- Network packet monitoring.
- Different strains of polymorphic viruses.
- Viruses residing in encrypted messages, compressed files.
- Viruses in different languages (e.g., JAVA, ActiveX, and Visual Basic).
- Trojan horses and worms.

Updates

- Can tool upgrade an existing version?
- Are regular updates available?
- Frequency of update releases.
- Response mechanisms.
- Quarantine at the server level.
- Quarantine at the console level.
- Supply network-based responders.
- Send alerts to network or system administrators.
- Send alerts (in the case of e-mail borne viruses) to sender and receiver(s).

Platform Considerations

- What platforms does the tool run on?
- Does tool allow cross-platform support?

6.4.2.3 Rationale for Selecting Features

When selecting AV products, two important guidelines must be followed. The “best” product may not be good enough by itself. Also, since data security products operate in different ways, one product may be more useful than another in different situations. The following categories provide a rationale for evaluating the features of specific technology offerings. Rating each product according to these categories will allow an organization to choose the best malicious code detection product for its needs.

Detection Capabilities

As discussed in Section 6.6.5.2, Viruses and E-mail, most computer-virus scanners use pattern-matching algorithms that can scan for many different signatures at the same time. Malicious code detection technologies have to include scanning capabilities that detect known and unknown worms and Trojan horses. Most AV products search hard disks for viruses, detect and

UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections
IATF Release 3.1—September 2002

remove any that are found, and include an auto-update feature that enables the program to download profiles of new viruses so that it will have the profiles necessary for scanning. The virus signatures these programs recognize are quite short: typically, 16 to 30 bytes out of the several thousand that make up a complete virus. It is more efficient to recognize a small fragment than to verify the presence of an entire virus, and a single signature may be common to many different viruses.

Updates

Maintaining an effective defense against virus and hostile code threats involves far more than the ability to produce perfect detection rates at a given point in time. With an average of nearly 300 new viruses discovered each month, the actual detection rate of AV software can decline rapidly if not kept current. This AV protection should be updated regularly. As new viruses are discovered, corresponding cures are developed to update protections. These updates should not be ignored. AV systems should do these updates automatically, reliably, and through a centrally controlled management Framework. To stay current, these scanning programs must be updated when new virus strains are found and AV codes are written. Most computer-virus scanners use pattern-matching algorithms that can scan for many different signatures at the same time. This is why enterprise-class AV solutions must be able to offer timely and efficient upgrades and updates across all client and server platforms.

Often, in large enterprise environments, a typical acquisition and deployment strategy is to deploy one brand of AV software at end-user workstations and a different vendor's product in the e-mail, file, and application server environments. This broadens the spectrum of coverage because in any given instance, one vendor is typically ahead of another in releasing the latest round of virus signature discoveries.

Response Mechanisms

Once malicious code has been detected, it must be removed. One technique is simply to erase the infected program, but this is a harsh method of elimination. Most AV programs attempt to repair infected files rather than destroy them. If a virus-specific scanning program detects an infected file, it can usually follow a detailed prescription, supplied by its programmers, for deleting virus code and reassembling a working copy of the original file. There are also generic techniques that work well for known and unknown viruses. One method is to gather a mathematical fingerprint for each program on the system. If a program subsequently becomes infected, this method can reconstitute a copy of the original. Most tools perform scanning for viruses, but all do not detect and remove Trojan horses, worms, and malicious mobile code upon all levels of entry. Most currently available AV tools do not have the same capabilities when responding across a network. Additional countermeasures related to malicious code can be found in Section 6.6.4, Potential Countermeasures.

Platform Considerations

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing or interfering with other applications running on the same system.

6.4.2.4 Considerations for Deployment

Defense in depth dictates that any virus protection must be implemented across the enterprise. This means installing and managing AV software on every system. Some advocate installing AV software only on edge devices, such as servers, firewalls, and gateways. But defense against viruses is only as good as its weakest link, and if one system can be compromised, then the entire enterprise is at risk.

Centralized management for the AV capabilities with a common set of policies is strongly recommended. Though some vendor offerings cater to end-users who are being held responsible for following security mandates, this can lead to more and varied security holes. What most often happens is that end users (or many of them), when their sessions are interrupted with a pop-up screen telling them their files are about to be scanned or that they are about to receive an AV update, tend to override the update manually, because it is distracting.

6.4.2.5 Considerations for Operation

Most AV technologies provide a means for sending responses or alerts at the server level, and some at the console level. It is always desirable to notify anyone that may have been infected that malicious code has been detected. This should include system and network administrators. If malicious code is encountered in e-mail transactions, it is desirable to notify the sender and recipient. If it is found on a file system that knows the file owner, he or she should be notified. In general, anyone that could be notified should be.

6.4.3 Discussion of Typical Bundling of Capabilities

At one point, network monitors were offered as stand-alone devices. Vendors may prefer to offer these technologies as appliances, sold with what is otherwise a commercial off-the-shelf (COTS) computer system, at an inflated price. There are also a number of offerings that combine these monitors with firewalls, routers, vulnerability scanners, and the like as a means for vendors to leverage existing market positions to gain market share in related areas. Another trend that is becoming popular is for larger vendors to offer integrated architecture approaches, in which they combine a number of related technologies as a bundled offering. Vendors tend to prefer custom rather than standard interfaces to preclude the merging of other vendor offerings. This offers a so-called “complete solution”; however, it tends to lock the buyer into one

particular product suite. While this often sounds attractive, it is often valuable to be able to integrate various technologies together in order to take advantage of the detection capabilities available from the different implementations.

There is a natural linkage of these monitoring technologies with Enterprise Security Management (ESM) systems, and vendors have been talking about the integration for some time. However, there is little evidence to suggest that this integration will be realized in the foreseeable future.

6.4.4 Beyond Technology Solutions

While the focus of the Information Assurance Technical Framework (IATF) is on technology solutions, there are important operational aspects of effective network monitoring that are also critical to an effective IA solution. The Framework recommends the following guidance:

Operational Planning

- Develop intrusion detection and AV-related requirements as an integral part of the enterprise security policy.
- Assess the ability of system administration personnel to perform intrusion detection and related vulnerability scanning.
- Consult with experienced intrusion detection and vulnerability scanning personnel regarding the best approach.
- Consult with the appropriate legal council regarding affected personnel rights and procedures, as discussed below.
- Provide for adequate technical and legal training of all involved personnel.
- Acquire software and expertise from a high-integrity vendor.
- Perform network monitoring consistent with the enterprise security policy.
- Tightly couple vulnerability scanning and intrusion detection activities.

Intrusion Detection Activities

- Look for intrusion evidence based on found vulnerabilities; use intrusion evidence to find and correct vulnerabilities.
- Provide and monitor bogus sites/services/information. Possibly monitor intrusions through known vulnerabilities to satisfy prosecution requirements in conjunction with the appropriate legal authorities.
- Perform intrusion responses that are approved by the appropriate authority.

Network Malicious Code Detection Activities

- Select and deploy virus scanning capabilities that are consistent with the location, functions, and capabilities.
- Acquire or download the appropriate AV software from a high-integrity source, and acquire any necessary hardware (such as an ancillary firewall dedicated to virus scanning of incoming or outgoing traffic).
- Institute enterprise wide AV training and procedures.
- Scan consistently based on time and/or events.
- Follow up on all indications of potential contamination (as defined in the security policy and AV procedures for the enterprise).
- Update AV software and hardware as appropriate (e.g., consistent with new releases of AV software and specific experiences throughout the enterprise).

General Activities

- Archive (within any legal constraints) audit and intrusion information, and correlate with vulnerability scan information.
- Keep authorities apprised of all activities, ensuring that any legal rights are not violated.
- Regularly repeat steps, as appropriate.

Privacy Concerns

Organizations may own the intellectual property of employees and may also legally restrict computer activities to those approved by management. A common practice is to present this warning to all computer users as part of the normal login message. This does not mean that ALL managers in an enterprise own ALL of the transactions of ALL of the employees. Especially unclear is how to handle the conflict that arises between privacy and monitoring. Use of IDSs and system monitoring tools requires caution. Sniffers that search for key words in messages (e.g., “attack,” “weakness,” or “confidentiality”) as a standard set of “watchwords” may find them used in an appropriate manner depending on the type of correspondence. Audit trail reports may contain full command strings (including parameters). Knowing that an employee is sending several messages to a particular department (e.g., Human Resources) may be an infringement on his or her privacy. It is important to refer privacy concerns to the appropriate legal and policy organizations for the enterprise prior to deployment and use of these technologies.

6.4.5 For More Information

The source materials used in the preparation of this section provide an excellent base of knowledge of relevant technologies from which to draw. A number of additional sources of

information exist. This section of the Framework focuses on on-line sources since they tend to offer up-to-date information. These include the following.

6.4.5.1 IATF Executive Summaries

An important segment of the IATF is a series of “Executive Summaries” that are intended to provide summary implementation guidance for specific situations. These offer important perspectives on the application of specific technologies to realistic operational environments. While these are still being formulated, they will be posted on the IATF Web site <http://www.iatf.net> as they become available. [1]

6.4.5.2 Protection Profiles

The National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11 provides the national policy that governs the acquisition of IA and IA-enabled information technology products for national security telecommunications and information systems. This policy mandates that, effective January 2001, preference be given to products that are in compliance with one of the following:

- International Common Criteria for Information Security Technology Evaluation Mutual Recognition Arrangement.
- National Security Agency (NSA)/National Institute of Standards and Technology (NIST) National Information Assurance Partnership (NIAP).
- NIST Federal Information Processing Standard (FIPS) validation program.

After January 2002, this requirement is mandated. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance references this same NSTISSP No. 11 as an acquisition policy for the Department.

- The International Common Criteria and NIAP initiatives base product evaluations on Common Criteria Protection Profiles.
- NSA and NIST are working to develop a comprehensive set of protection profiles for use by these initiatives. An overview of these initiatives, copies of the Protection Profiles, and status of various products that have been evaluated are available at the NIST Web site <http://niap.nist.gov/> [2]

6.4.5.3 Independent Third Party Reviewers of Relevant Vendor Technologies

- ICOSA Net Security Page www.icsa.net

UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections
IATF Release 3.1—September 2002

- Talisker's Intrusion Detection Systems www.networkintrusion.co.uk/
- Network Computing—The Technology Solution Center
www.nwc.com/1023/1023f12.html
- Paper on CMDS Enterprise 4.02 <http://www.Intrusion.com/Products/enterprise.shtml>
(ODS Networks has changed its name to Intrusion.com)
- PC Week On-Line www.zdnet.com/pcweek/reviews/0810/10sec.html

6.4.5.4 Overview of Relevant Research Activities

- Coast Home page – Purdue University www.cs.purdue.edu/coast
- UC Davis <http://seclab.cs.ucdavis.edu/>

6.4.5.5 Overview of Selected Network Monitor Vendor Technologies

- Axent Technologies <http://www.axent.com/>
- cai.net <http://www.cai.net/>
- Cisco Connection Online www.cisco.com
- CyberSafe Corporation <http://www.cybersafe.com>
- Internet Security Systems www.iss.net
- Network ICE www.networkice.com

UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections
IATF Release 3.1—September 2002

References

1. Information Assurance Technical Framework (IATF) <http://www.iatf.net>
2. National Institute of Standards and Technology <http://niap.nist.gov/>.

Additional References

- a. Amoroso, Edward, Intrusion Detection. Intrusion.Net Books. 1999.
- b. Escamilla, Terry. Intrusion Detection, Network Security Beyond the Firewall. Wiley Computer publishing. 1998.
- c. Northcutt, Stephen. Network Intrusion Detection, An Analyst's Handbook. New Riders Publishing. 1999.
- d. Snapp, Steven R., et al. A System for Distributed intrusion Detection. IEEE CH2961-1/91/0000/0170. 1999
- e. Balasubramanian, J. S., et al. An Architecture for Intrusion Detection Using Autonomous Agents. COAST Technical Report. 11 June 1998.
- f. AXENT Technologies, Inc. Intruder Alert 3.5 IDS Review Guide, May 2000.
- g. AXENT Technologies, Inc. Everything You Need to Know About Intrusion Detection, 1999.
- h. Schneider, Sondra, et al. Life After IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- i. Graham, Robert. New Security Trends for Open Networks. SC Magazine. October 1999.
- j. SC Magazine. Intrusion Detection. June 2000.
- k. Information Assurance Technology Analysis Center (IATAC). Tools Report on Intrusion Detection. Defense Technical Information Center. December 1999.
- l. Maes, V. How I Chose an IDS. Information Security Magazine. Volume 2, Number 9. September 1999.
- m. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Trade Study Report Intrusion Detection System. Report No. 0017-UU-TE-000621. April 14, 2000.
- n. Information Assurance Technology Analysis Center (IATAC). Tools Report on Vulnerability Analysis Information. Defense Technical Information Center. March 15, 2000.
- o. Ulsch, Macdonnell and Joseph Judge. Bitter-Suite Security. Information Security Magazine. Volume 2, Number 1. January 1999.

UNCLASSIFIED

Network Monitoring Within Enclave Boundaries and External Connections
IATF Release 3.1—September 2002

- p. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Baseline Tool Assessment Task Anti-Virus Trade Study Report. Report No. 0017-UU-TE-000623. April 13, 2000.
- q. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance.
- r. National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11. National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products. January 2000.

6.5 Network Scanners Within Enclave Boundaries

As discussed in Section 6.4, Network Monitoring Within Enclave Boundaries and External Connections, on-line network monitoring technologies provide a critical layer of defense within enclave boundary protection. In addition to the network monitoring technologies, another class of technologies, referred to as network scanners, can also be deployed to improve overall security posture. The framework makes a distinction between these scanners and network monitoring devices. Monitors typically operate in near real time and have network traffic (or related characteristics) as their focus. Monitors tend to measure the effectiveness of the network's protection services that are subject to attempted exploitation. This is somewhat of an "after the fact" measure, not a preventive measure. Scanners, on the other hand, are preventive measures. Typically, they operate periodically (or on demand) and examine systems for vulnerabilities that an adversary could exploit, measuring the effectiveness of the system's infrastructure protection.

The local environment is the logical place for addressing these network assessment technologies. Scanning can be performed at the network boundary or at the host level. This segment of the Information Assurance Technical Framework (IATF) specifically considers network vulnerability scanner and War Dialer technologies that are germane to the enclave environment. Please refer to Section 7.2, Host-Based Detect and Respond Capabilities Within Computing Environments, for guidance on the use of similar technologies that are more suitable for deployment at the host level.

Unlike the near-real-time network monitoring technologies addressed in Section 6.4, Network Monitoring Within Enclave Boundaries and External Connections, network assessment technologies are typically executed in a periodic or on-demand manner, providing perspectives on the posture of a local environment. Section 8.2, Detect and Respond as a Supporting Element, of the framework provides a perspective on an overall detect and response infrastructure; however, because these assessments typically focus on the local level, they tend not to interact with or be particularly relevant to a broader network infrastructure.

6.5.1 Network Vulnerability Scanners

Periodic or on-demand network assessment tools are adept at finding security holes at boundary-point devices or on network hosts within an enclave environment, hopefully before an attacker does. They accomplish this effort by discovering known vulnerabilities in host or network system components and improper configurations visible from the network that create the potential for unauthorized access or exploitation or are counter to enterprise policies.

6.5.1.1 Technology Overview

Vulnerability analysis tools help automate the identification of vulnerabilities in a network or system. Network-based vulnerability scanners take an inventory of all devices and components within the network infrastructure. These scanners operate over a network “against” target nodes by probing and examining the network components and hosts to identify vulnerabilities that are typically visible to their network connection. They seek to identify network services that allow uncontrolled access, contain buffer control vulnerabilities, violate possible trust privileges, and contain weaknesses in network component (e.g., router, firewall, and Web server) configurations.

A scanner probes for weaknesses by comparing data about a network configuration with its database of known vulnerabilities. Network components, the network configuration, and the various versions of the software controlling the network are examined and compared with this database. Network vulnerability scanners fall within one or more of the following classes.

Simple Vulnerability Identification and Analysis

A number of tools have been developed that perform relatively limited security checks. These tools may automate the process of scanning Transmission Control Protocol/Internet Protocol (TCP/IP) ports on target hosts, attempting to connect to ports running services with well-known vulnerabilities and recording the response. They also may perform secure configuration checks for specific system features. The user interface of these tools is likely to be command-line based, and the reporting may include limited analysis and recommendations. The tools are likely to be freeware.

Comprehensive Vulnerability Identification and Analysis

More sophisticated vulnerability and analysis tools have been developed that are fairly comprehensive in terms of the scope of vulnerabilities addressed, the degree of analysis performed, and the extent of recommendations made to mitigate potential security risks. Many of these tools also provide a user-friendly graphical user interface (GUI).

Password Crackers

Password cracker tools attempt to match encrypted forms of a dictionary list of possible passwords with encrypted passwords in a password file. This is possible because the algorithm used to encrypt an operating system’s passwords is public knowledge. An attacker or insider would run these tools after successfully gaining access to the system in order to acquire a higher privilege level, such as root. These tools allow operators to verify compliance with password selection policies. Many tools from the previous category have integrated password-cracking modules.

Risk Analysis Tools

Risk analysis tools typically provide a framework for conducting a risk analysis but do not actually automate the vulnerability identification process. These tools may include large databases of potential threats and vulnerabilities along with a mechanism to determine, based on user input (typically query/response scripts), cost-effective solutions to mitigate risks. The vulnerabilities identified using a vulnerability analysis tool may be input into a risk analysis tool to assist in determining the overall risk to the system, or conversely, vulnerabilities predicted by a risk analysis tool can be specifically targeted for confirmation using vulnerability scanning tools.

6.5.1.2 General Considerations for Use

Network vulnerability scanners operate across the network to identify weaknesses in a connected system's security scheme, exploitation of which would negatively affect the confidentiality, integrity, or availability of the system or its information. These scanners are easy to install and can run a wide variety of attacks on a network to determine the network's resilience to each attack. However, a scanner only takes a snapshot of the network and does not operate in real time, often requiring post-capture analysis to understand and implement any mitigation approaches that may be required. Typically, local area network (LAN) administrators do not use scanners on a day-to-day basis.

Scanners work either by examining attributes of objects or by emulating an attacker. To act as a hacker, a scanner can execute a variety of attack scripts. Because these can look (and act) like real attacks, it is important to consider what and when scans are performed. Otherwise, it is possible that the scanner could have as much impact on the network as an actual incident. Coordination with network operations staff is critical, particularly in environments that implement real-time intrusion detection techniques. However, another use of such scanners is a "live" test and readiness evaluation of intrusion detection and incident response processes and procedures for an enterprise environment.

The vulnerability scanner will detect only objects it is configured to scan. If the scanner is not configured and set up properly, there may be vulnerabilities that are not identified. Therefore, using these tools may be of less value than performing no scans at all, because it may offer a false sense of the adequacy of the network's resiliency to attacks.

6.5.1.3 Important Features

When considering the selection of a network-based vulnerability scanner, a number of features should be considered. This section identifies important features for selection. The section that follows discusses the rationale for the selection of these features.

Scanning Capabilities

- Does the tool offer an ability to add custom scanning routines to look for site- or technology-specific weaknesses of concern?

Response Mechanisms

- Automatic shutoff of vulnerable ports of entry.

User Interfaces

- Does the tool have a GUI for number entry, dialing status, and call results?
- Can reports be viewed in real time?

Reporting Capabilities

- Does the tool offer automatic alerting when new non-network ports are detected?
- Are all system answers logged in a database or file?
- Is there an updated database of network numbers with which to compare newly identified numbers?
- Does the database automatically combine logged information and place it in a report format?
- Does the tool provide suggested mitigation approaches for discovered vulnerabilities?

Platform Compatibility

- What are the platforms (operating systems) on which the tool will run?
- Does it use executables?
- Does it support scripts or macros?

6.5.1.4 Rationale for Selecting Features

The type and level of detail of information provided varies greatly among tools. Although some can identify only a minimal set of vulnerabilities, others can perform a greater degree of analysis and provide detailed recommended mitigation approaches. The selected scanner technologies should cover the full range of vulnerabilities for the given environment and system platforms. In addition, the technologies should offer a comprehensive library of vulnerabilities, periodically updated by the vendor. Capabilities including grouping of nodes into scan groups and customized scan options may be valuable for larger environments.

Some scanner technologies offer features that are useful depending on the training and skill levels of the operators that will be using them. Depending on the planned usage of the scanner

and the skills of the operators available, it is often desirable to select technologies that can be tuned to ignore some false positives. It is also desirable to select features that enable the scanner to be tuned for important application environments, such as database environments, Web server environments, file server environments, firewalls, etc., since such profiles may differ based on the functions provided.

Scanning Capabilities

The type and level of detail of information provided varies greatly among tools. Although some can identify only a minimal set of vulnerabilities, others can perform a greater degree of analysis and provide detailed recommended mitigation approaches.

Response Mechanisms

Assessment tools will continue to evolve in usability, with some vendors offering click-and-fix solutions. The assessment software flags vulnerabilities in terms of the risk posed to the network and the ease of the fix. Some technologies can generate trouble tickets to trigger a manual response. They may offer an ability to change policies in firewalls and other enclave boundary defense mechanisms. Some identify patches that should be installed. Some offer to obtain and install patches. Although installing patches is feasible, allowing the security administrator the ability to undertake these tasks and the difficulty of undoing configuration changes should leave customers wary of this function. Such features should be considered in light of an environment's existing configuration management policies and procedures.

User Interfaces

Most scanners enable the operator to configure what network elements are to be scanned and when the scans are to occur. Typically, scanners are preconfigured with lists of vulnerabilities and can operate without customization. Some technologies allow operators to customize the vulnerabilities the scanner will investigate. Usually the results are sorted into a file that can be accessed upon demand to review the results. More recently developed tools provide user-friendly front ends and sophisticated reporting capabilities.

Reporting Capabilities

Old products inundated customers with phonebook-size reports on all the various vulnerabilities that the network faced. New products have database interfaces that prioritize vulnerabilities and allow network managers to deal with the network's problems in a logical manner. Many generate reports that are Web-enabled with hot-links and other "labor savers."

Platform Compatibility

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The vulnerability scanner software should function properly and perform its duties without failing.

Source

- Has the tool been developed by the Government (or under government sponsorship); if so, is it reserved; can your organization obtain authorization for its use?
- Is the tool available from a reputable vendor?
- Is the tool in the public domain (e.g., freeware from the Internet); if so, is source code available?

6.5.2 War Dialers

Firewalls and other enclave boundary protection devices can create a level of defense against network attacks that adversaries have to defeat. However, as the trend continues toward borderless networks, machines with attached modems are often scattered throughout organizations. When modems are installed on telephone lines connected to the data network, firewalls are no longer the only access port to the network, and thus cannot detect or control ALL of the data traffic that is traveling in or out of the network. The result is that “back doors” are created that offer alternative, unprotected portals for adversaries to exploit, as depicted in Figure 6.5-1. Analysts estimate that the bulk of damaging hacks on corporate networks come over modem connections that are not secure. One technology, called War Dialers, is a specific form of network vulnerability scanner.

6.5.2.1 Technology Overview

Most commonly, War Dialers are associated with hackers. Most hackers target organizations because they rarely control the dial-in ports as strictly as a firewall. One way of combating intrusions by hackers is to use the same type of scanning tool as a defensive mechanism.

A War Dialer consists of software that dials a specific range of telephone numbers looking for modems that provide a login prompt. The tools, at a minimum, record the modem numbers and login screen, but can also be configured to attempt brute force, dictionary-based login attempts. Visibility into telephone networks is provided by identifying modem, fax, or voice tones and characterizing security behaviors. This process allows identification of network vulnerabilities.

War Dialers call a given list or range of telephone numbers and record those that answer with handshake tones. Those handshake tones may be characterized as entry points to computer or telecommunications systems. Some of these programs have become quite sophisticated, and can now detect modem, fax, or private branch exchange (PBX) tones and log each one separately. A block of specified numbers is attempted and any modems found in that block are noted.

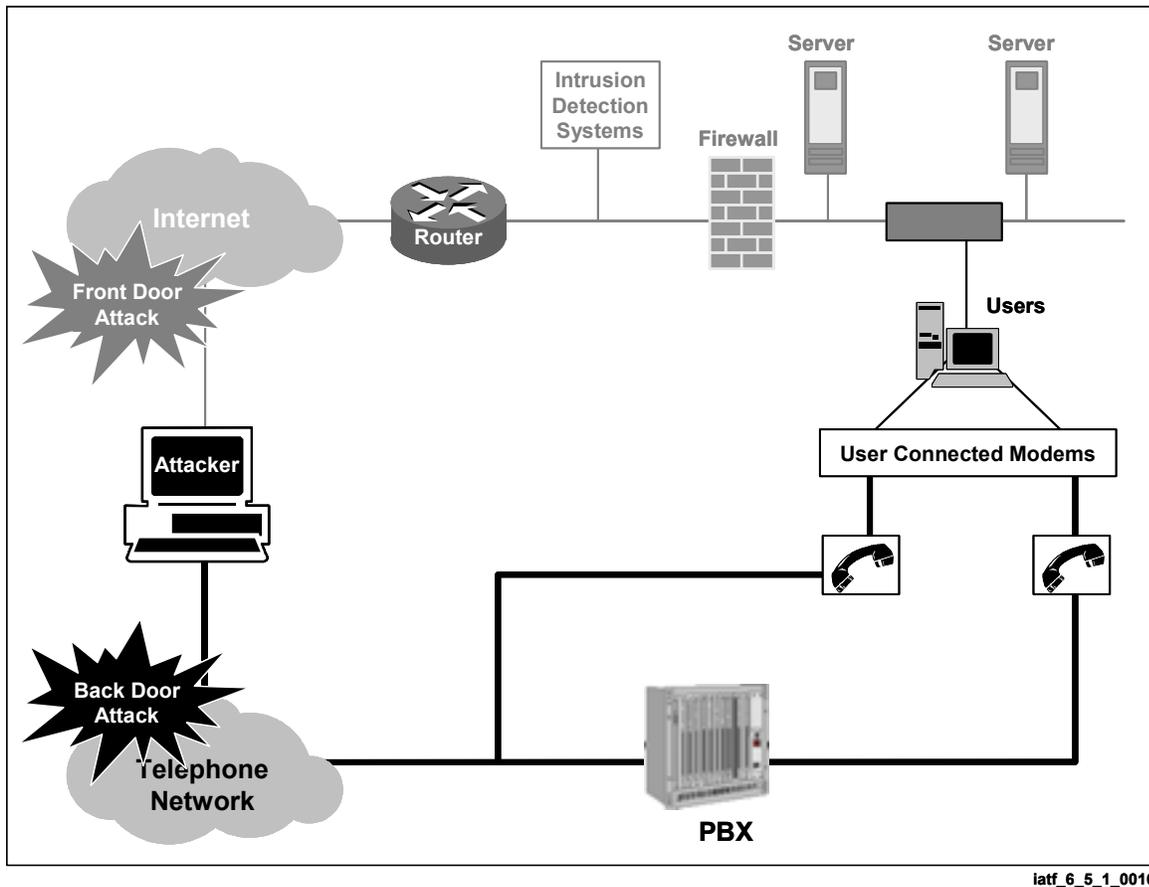


Figure 6.5-1. Back-Door Attacks Through Telephone Networks

6.5.2.2 General Considerations for Use

Remote access to most organizations' information systems is usually performed through ordinary telephone lines. The lack of visibility into telephone networks makes it possible for any user to connect to an entire private data network via a modem. These telephone lines must be thought of as ports of entry for possible network attacks and intrusions. When an enclave does not deploy protection mechanisms that effectively secure or monitor telephone networks, intruders can gain access to proprietary information; existing security systems remain blind to unauthorized activity. War Dialers are an effective way to identify unsecured modems. Along with a strong modem policy describing the need for modem registration and PBX controls, War Dialer scanning can help an organization defend itself against such dangers. Use of this type of technology can help an enterprise to identify those vulnerable back doors before an attack occurs. Once identified, those back doors can be closed or some type of security plan created to preclude use of that particular point of entry.

6.5.2.3 Important Features

When selecting a War Dialer technology, a number of features should be considered. This section identifies important features for selection. The section that follows discusses the rationale for the selection of these features.

Scanning Capabilities

- Identification of every dial-up system.
- Facsimile machine detection.
- Multi-modem scanning.
- Brute force username and/or password guessing (code cracking).
- Support terminal emulation to allow tool to enable access to mainframe computers.
- Built-in knowledge of various dial-in authentication technologies.

Response Mechanisms

- Automatic shutoff of vulnerable ports of entry (interface to telephone network).

User Interfaces

- Does the tool have a GUI for number entry, dialing status, and call results?
- Can reports be viewed in real time?

Reporting Capabilities

- Automatic alerting when new non-network ports are detected.
- Are all system answers logged in a database or file?
- Is there an updated database of network numbers with which to compare newly identified numbers?
- Does the database automatically combine logged information and place it in a report format?

Platform Compatibility

- What platforms (operating systems) will the tool run on?
- Does it use executables?
- Does it support scripts or macros?

Source

- Has the tool been developed by the Government (or under government sponsorship); if so, is it reserved; can your organization obtain authorization for its use?

- Is the tool available from a reputable vendor?
- Is the tool in the public domain (e.g., freeware from the Internet); if so, is source code available?

6.5.2.4 Rationale for Selecting Features

War Dialers identify known modems, modem banks, and communication servers; compare discovered modem configuration data against predefined modem configurations; and alert administration when a vulnerable port of entry has been detected. The major discriminator is how well each product performs these functions.

It is often difficult to determine the true nature of the features that are provided in a particular technology offering (beyond strict vendor claims). It is always advisable to seek test results of reputable, independent third-party laboratories. When these are available, they should be an important consideration in any technology selection. A number of organizations provide these types of results.

Scanning Capabilities

It is important that the War Dialer be capable of uncovering and characterizing all back doors on the network, because each represents a potential unprotected portal for an adversary. Thus, beyond simply identifying when a modem responds to an incoming call on each telephone line specified, it is possible to uncover when computers serving as facsimile machines and modem banks are encountered. Further, the ability to emulate a terminal (to enable access to mainframe computers) and apply password cracking mechanisms provides valuable information regarding how susceptible the identified parts actually are, supporting efforts to prioritize those that require earlier resolution. The more extensive scanning capabilities a tool offers the more thorough and reliable report it can provide on the actual posture of the network.

Response Mechanisms

For the most part, War Dialers report on back doors they have uncovered. However, technologies are available that can automatically shut off vulnerable ports of entry. Care should always be taken when selecting any automated response. In this case, shutting down a remote access port may have negative effects on operational capabilities.

User Interfaces

Most scanners enable the operator to enter telephone numbers and provide dialing status and call results. Usually the results are stored in a file that can be accessed upon demand to review the results. Depending on the skills of the intended operator, it may be desirable to select a tool that offers a user-friendly interface. Recently developed tools provide a user-friendly user interface for number entry, dialing status, and call results.

Reporting Capabilities

Again, based on the intended manner in which the War Dialer is operated, it may be desirable to select features that provide automatic alerting when new non-network ports are detected. If reports of the results of War Dialer scans are required by the organization, consideration should be given to technologies that offer the capability for the database to automatically combine logged information and place it in a report format. If the enterprise allows selected remote access ports to remain operational, operators may be concerned primarily with new ports that were not reported previously. In this situation, consideration should be given to technologies that are able to update the database of network numbers with which to compare newly identified numbers.

It is important to ensure that the selected technology logs all system answers in a database or file. If the operator will be monitoring the results of the War Dialer assessment during its operation, it will be important to consider technologies where reports can be viewed in real time.

Platform Compatibility

The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing.

Source

A number of War Dialers have been developed by the Government (or under government sponsorship). If one of these is selected, it may be reserved for use only by selected communities. In these situations, it is necessary to determine if your organization can obtain authorization for its use.

A wide array of War Dialers are available as freeware or shareware. These are regarded as hacker tools and are an open source via the Internet. Many commercial scanners dial only predetermined numbers in a telemarketing atmosphere. Commercial products are preferred because they tend to offer technical support mechanisms; typically, no reliable means exist for support for freeware and/or shareware. Overall, the functions are the same, but technical support, better reporting styles, and more attractive GUIs can be found with the commercial products offered today.

Care should be taken when using any software obtained from the public domain (e.g., freeware from the Internet). The software should be scanned carefully for potential malicious code. If source code is not available, the software's use is *NOT* recommended.

6.5.3 Considerations for Deployment

The same considerations that apply to placement of network monitors, discussed in Section 6.4, Network Monitoring Within Enclave Boundaries and External Connections, are in general applicable in deploying network scanners. Network switches, which segregate network traffic

into specific individual “subnets,” reduce network loads across an organization by implementing a form of “need-to-know” policy among connected computers. Network switches allow traffic to enter a subnet only if it is meant for a computer within that subnet; similarly, packets are only allowed out of a subnet that are destined for a computer outside its particular realm.

Network scanners only can find vulnerabilities that they can see based on the segments on which they are installed. As long as the network scanner is placed on critical segments, it will be able to measure the effectiveness of the security protection mechanisms for the most critical systems and applications. Within an enclave environment, a number of possible locations should be considered in deploying a network scanner. The challenge is to identify the locations where the potential vulnerabilities are of most interest. This is often considered from the view of potential attacker sources that are of concern. For example, if the concern is for hackers from the Internet, the scanner should be structured to look at the network from that vantage point. If the concern is for insider threats, that vantage point should be considered. Because the scanners can operate on demand, they can be used in one location and then moved to another to determine the overall security posture of a network environment.

6.5.4 Considerations for Operation

Assessment frequency is a factor of how often network changes are made and the security policy for the enterprise. Depending on the organization, assessments may take place quarterly, monthly, weekly, or even daily. Some service providers offer scanning services on a subscription basis, ensuring that assessments occur regularly.

6.5.5 Discussion of Typical Bundling of Capabilities

At one point, network monitors were offered as stand-alone devices. Vendors may prefer to offer these technologies as appliances, sold with what is otherwise a commercial off-the-shelf (COTS) computer system, at an inflated price. A number of offerings combine these monitors with firewalls, routers, vulnerability scanners, and the like as a means for vendors to leverage existing market positions to gain market share in related areas. Another trend that is becoming popular is for larger vendors to offer integrated architecture approaches, in which they combine a number of related technologies as a bundled offering. Vendors tend to prefer custom rather than standard interfaces to preclude the merging of other vendor offerings. This offers a so-called “complete solution”; however, it tends to lock the buyer into one particular product suite. Although this often sounds attractive, it is valuable to be able to incorporate various technologies to take advantage of the detection capabilities available from the different implementations.

There is a natural linkage of these monitoring technologies with Enterprise Security Management (ESM) systems, and vendors have been discussing the integration for some time. However, there is little evidence to suggest that this integration will be realized in the foreseeable future.

6.5.6 Beyond Technology Solutions

Although the focus of the IATF is on technology solutions, operational aspects of effective network scanning are critical to an effective information assurance (IA) solution. Network scanning is the primary means of assessing the security of the network. The functions performed by the scanner should be tailored to the network configuration and environment, together with the applications performed by the protected network. The framework recommends the following guidance for network scanners:

- Develop network scanning requirements as an integral part of the enterprise security policy.
- Scan your network consistent with the guidance listed for intrusion detection and response, using the best available scanners.
- Assess the results in light of your security policy.
- Adjust and counter identified deficiencies relative to your policy. This may include patches, changes in configuration, changes in procedures, or better enforcement of procedures such as the use of good passwords that change frequently.
- Repeat the process regularly.

6.5.7 For More Information

The list of reference materials used in preparing this section provides an excellent base of knowledge from which to draw on relevant technologies. A number of additional sources of information exist. This section of the framework focuses on on-line sources because they tend to offer up-to-date information. These include the following.

6.5.7.1 IATF Executive Summaries

An important segment of the IATF is a series of “Executive Summaries” that provides summary implementation guidance for specific situations. These summaries offer important perspectives on the application of specific technologies to realistic operational environments. Although these are still being formulated, they will be posted on the IATF Web site www.iatf.net as they become available. [1]

6.5.7.2 Protection Profiles

The National Security Telecommunications and Information Systems Security Policy (NSTISSP) Number 11 provides the national policy that governs the acquisition of IA and IA-enabled information technology products for national security telecommunications and information systems. This policy mandates that, effective January 2001, preference be given to products that are in compliance with one of the following.

- International Common Criteria for Information Security Technology Evaluation Mutual Recognition Arrangement.
- National Security Agency/National Institute of Standards and Technology (NSA/NIST) National Information Assurance Partnership (NIAP).
- NIST Federal Information Processing Standard (FIPS) validation program.

After January 2002, this requirement is mandated. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance references this same NSTISSP Number 11 as an acquisition policy for the Department.

The International Common Criteria and NIAP initiatives base product evaluations on Common Criteria Protection Profiles. NSA and NIST are developing a comprehensive set of protection profiles for use by these initiatives. An overview of these initiatives, copies of the Protection Profiles, and the status of various products that have been evaluated are available at the NIST Web site [http://niap.nist.gov/\[2\]](http://niap.nist.gov/[2])

6.5.7.3 Independent Third Party Reviewers of Relevant Vendor Technologies

- ICSA Net Security Page www.icsa.net
- Talisker's Intrusion Detection Systems www.networkintrusion.co.uk/
- Network Computing—The Technology Solution Center www.nwc.com/1023/1023f12.html
- Paper on CMDS Enterprise 4.02 www.ods.com/downloads/docs/Cmds-us.pdf (ODS Networks has changed its name to Intrusion.com)
- PC Week On-Line www.zdnet.com/pcweek/reviews/0810/10sec.html

6.5.7.4 Overview of Relevant Research Activities

- Coast Home page—Purdue University www.cs.purdue.edu/coast
- UC Davis www.seclab.cs.ucdavis.edu/cidf
- UC Davis www.seclab.cs.ucdavis.edu

6.5.7.5 Overview of Selected Network Scanner Vendor Technologies

- Axent Technologies www.axent.com
- cai.net <http://www.cai.net/>

Network Scanners Within Enclave Boundaries
IATF Release 3.1—September 2002

- Cisco Connection Online www.cisco.com
- CyberSafe Corporation www.cybersafe.com
- Internet Security Systems www.iss.net
- Network ICE www.networkice.com

6.5.7.6 Overview of Selected War Dialer Technologies

- VerTTeX Software www.verttex.com
- The Hackers Choice www.infowar.co.uk/thc/toneloc
- AT&T Information Security Center www.att.com/isc/docs/war_dialer_detection.pdf

References

1. Information Assurance Technical Framework (IATF) <http://www.iatf.net>.
2. National Institute of Standards and Technology <http://niap.nist.gov/>.

Additional References

- a. Amoroso, Edward, Intrusion Detection. Intrusion. Net Books. 1999.
- b. Escamilla, Terry. Intrusion Detection, Network Security Beyond the Firewall. Wiley Computer publishing. 1998.
- c. Northcutt, Stephen. Network Intrusion Detection, An Analyst's Handbook. New Riders Publishing. 1999.
- d. Concurrent Technologies Corporation. Attack Sensing, Warning, and Response (ASW&R) Baseline Tool Assessment Task War Dialer Trade Study Report. Report No. 0017-UU-TS-000480. March 23, 2000.
- e. King, Nathan A. Sweeping Changes for Modem Security. Information Security Magazine. Volume 3, Number 6. June 2000.
- f. Ulsch, Macdonnell and Joseph Judge. Bitter-Suite Security. Information Security Magazine. Volume 2, Number 1. January 1999.
- g. Department of Defense (DoD) Chief Information Officer (CIO) Guidance and Policy Memorandum No. 6-8510, Guidance and Policy for Department of Defense Global Information Grid Information Assurance.
- h. National Security Telecommunications and Information Systems Security Policy (NSTISSP) No. 11. National Policy Governing the Acquisition of Information Assurance (IA) and IA-Enabled Information Technology (IT) Products. January 2000.

UNCLASSIFIED

Network Scanners Within Enclave Boundaries
IATF Release 3.1—September 2002

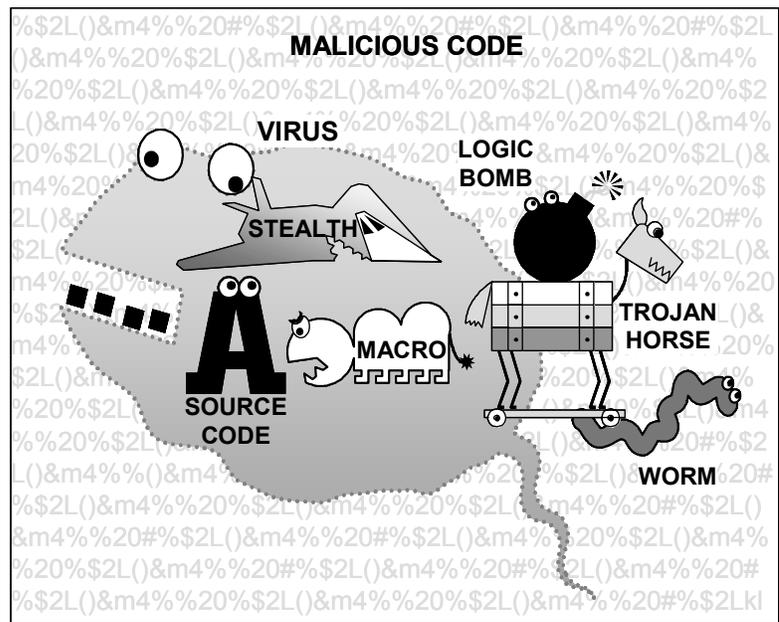
This page intentionally left blank.

6.6 Malicious Code Protection

The objective in this section of the framework is to elucidate the importance of defense from destructive malicious code. Information is provided regarding malicious code protection techniques and how malicious code infiltrates a system. Detection and recovery tactics are described as well as different types of malicious code scanners used to protect systems.

Malicious code protection allows authorized local area network (LAN) users, administrators, and individual workstation/personal computer users to safely conduct daily functions in a secure manner. Commonly, many people misuse the word virus assuming it means anything that infects their computer and causes damage. The correct term for this is really malicious code. A virus is simply a computer program created to infect other systems/programs with copies of itself. Worms are similar to viruses; however, they do not replicate and the intent is usually destruction. Logic bombs contain all types of malicious code and activate when certain conditions are met. Viruses, worms, and logic bombs can also be concealed within source code disguised as innocent programs like graphic displays and games. These apparently innocent programs are called Trojan horses. The relationship among these different types of malicious code is illustrated in Figure 6.6-1.

The quantity of new malicious code introduced into the computing environment has increased exponentially. This situation has occurred for several reasons. Computer users have become increasingly proficient and sophisticated, and software applications have become increasingly complex. Some brands of software are now widely used, thus their bugs and security loopholes are often known to intelligent users capable of writing destructive code. With the widespread use of personal computers that lack effective malicious code protection mechanisms, it is relatively easy for knowledgeable users to author malicious software and dupe unsuspecting users into copying or downloading it. In addition, since virus information and source code is readily available through the Internet and other sources, creating viruses has become a relatively simple task.

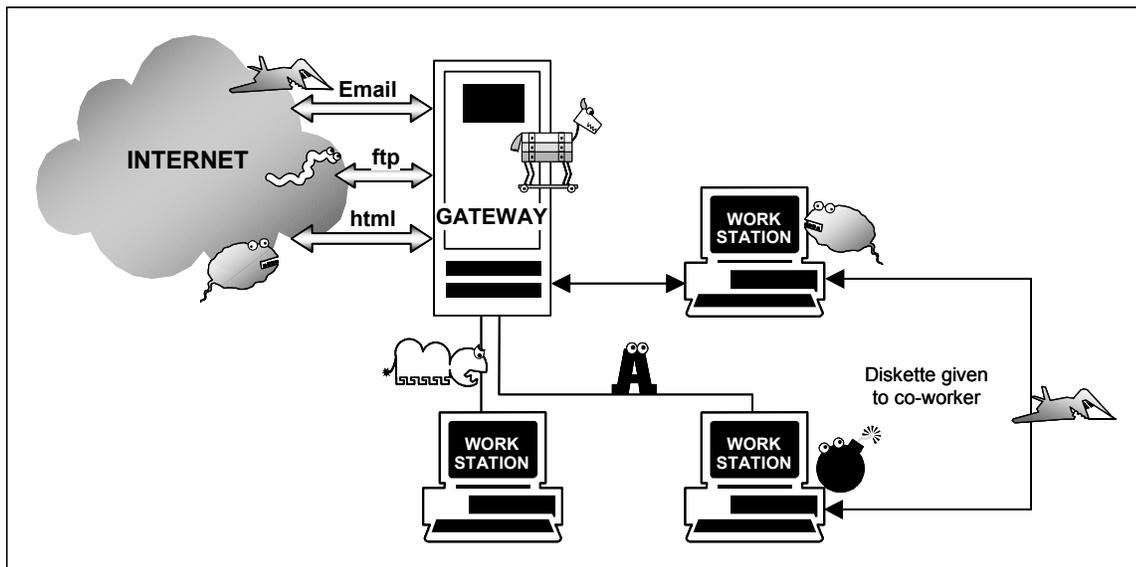


iatf_6_6_1_0018

Figure 6.6-1. Malicious Code Relationship

6.6.1 Target Environment

Malicious codes protection typically is provided at two places in the architecture: at the gateway and at workstations that access information services. Malicious code can infiltrate and destroy data through network connections if allowed beyond the gateway or through individual user workstations. Today, the majority of individual users keep all data files on networks or shared file systems instead of on diskettes. Therefore, the continual application of protection of network connections at the gateway is essential. Malicious code usually enters existing networks through the gateway by means of security loopholes or e-mail attachments. Its intent is to cripple the network and individual workstations. Malicious code can also attack the network through protocols, typically, File Transfer Protocol (FTP), Hypertext Transfer Protocol (HTTP), and Simple Mail Transfer Protocol (SMTP) (e-mail). The individual user workstation is then subsequently infected. In Figure 6.6-2 below, a simplified network is illustrated with several workstations connected to a single gateway, and through that, to the Internet. Although a single user can bring an infected disk to work, infecting his or her workstation and eventually the entire network, the majority of infections by malicious code result from file sharing across different protocols. Malicious codes attacking individual user workstations are primarily macro viruses and other less potentially destructive viruses. These viruses typically enter systems through e-mail attachments; however, their primary intent is not destruction.



iatf_6_6_2_0017

Figure 6.6-2. Sources of Malicious Code Infections

6.6.2 Malicious Code Protection Requirements

Malicious Code Detection System Requirements

The following have been identified as representative malicious code detection system requirements from a customer's perspective of needs.

The malicious code detection system shall—

- Allow access to all services available on the wide area networks (WAN) using any of the existing and emerging networking technologies and applications.
- Be able to locate the source and type of an infection, be able to react to such intrusions, and be able to fully reconstitute the system following damage caused by intrusions.
- Have minimal operational effect on the user.
- Have minimal operational effect on performance of the associated components.
- Have appropriate documentation for its use and upgradability and contain all currently available references and resources.
- Allow automatic malicious code prevention programs to run in the background.
- Allow a disaster recovery plan to recover data if necessary.
- Provide adequate scanning tools to be able to contain an identified virus by isolating affected systems and media.
- Have appropriate means to trace all incoming and outgoing data, including e-mail, FTP transactions, and Web information.
- Be able to, in the event the Internet is unavailable for any reason, still have access to virus updates from the manufacturer or vendor of the antivirus product.
- Monitor usage as required by the administrator.
- Scan for malicious software at the enclave boundary and at individual workstations.
- Log and analyze source-routed and other packets; react to or restrict malicious code attacks.
- Allow a rapid disconnect from the network in the event of a detected malicious code attack.

Configuration/Management Requirements

The following have been identified as representative configuration and/or management requirements for malicious code detection systems.

The malicious code detection system shall—

- Be updated with regard to relevant security issues (malicious code detection, system vulnerability) so maximum protection is provided.
- Be capable of preventing worm programs from infecting networks by allowing the administrator to disable the network mail facility from transferring executable files.
- Be configured by the administrator to filter all incoming data, including e-mail, FTP transactions, and Web information, for all types of malicious code.
- Allow the administrator to automatically create policy for network usage that details what sort of computing activity will and will not be tolerated.
- Allow regular backups of all system data by the administrator.
- Provide adequate controls such as strong user authentication and access control mechanisms on network connections for the administrator.
- Be capable of setting additional passwords or authentication for select files and accounts accessed from network ports.
- Be capable of placing restrictions on types of commands used on networks and in select files.
- Deny access to system manager accounts from network ports, if possible.
- Monitor usage of the network during odd hours, if possible, and create a log of all activity for the system administrator.
- Provide no more than one administrator account (i.e., not give other users administrator privileges).

6.6.3 Potential Attack Mechanisms

Malicious code can attack authorized LAN users, administrators, and individual workstation/personal computer users in numerous ways, such as modifying data in transit, replaying (inserting previously collected data), exploiting data execution, inserting and exploiting malicious code, exploiting protocols or infrastructure bugs, and modifying malicious software during production and/or distribution. (See Sections 4.2.1.4.2, Network-Based Vulnerabilities and Active Attacks, and 4.2.1.4.4, Hardware/Software Distribution.)

6.6.3.1 Viruses and Worms

The operating system (OS) is software that controls all inputs and outputs to the system and manages the execution of programs. A virus or worm can infect the OS in two ways: by completely replacing one or more OS programs or by attaching itself to existing OS programs and altering functionality. Once a virus or worm has altered or changed OS functionality, it can

control many OS processes that are running. To avoid detection, the virus or worm usually creates several hidden files within the OS source code or in “unusable” sectors. Since infections in the OS are difficult to detect, they have deadly consequences on systems relying on the OS for basic functions.

Macro Viruses

Application programs on a system provide users with significant functionality. A macro virus can easily infect many types of applications such as Microsoft Word and Excel. To infect the system, these macro viruses attach themselves to the application initialization sequence. When an application is executed, the virus’ instructions execute before control is given to the application. These macro viruses move from system to system through e-mail file sharing, demonstrations, data sharing, and disk sharing. Viruses that infect application programs are the most common and can lie dormant for a long time before activating. Meanwhile, the virus replicates itself, infecting more and more of the system.

6.6.3.2 Logic Bombs

After a logic bomb has been activated, it can maliciously attack a system in the following ways: halt machine, make garbled noise, alter video display, destroy data on disk, exploit hardware defects, cause disk failure, slow down or disable OS. It can also monitor failures by writing illegal values to control ports of video cards, cause keyboard failure, corrupt disks and release more logic bombs and/or viruses (indirect attacks). These attacks make logic bombs an extremely destructive type of malicious code.

6.6.3.3 Trojan Horses

Trojan horses are another threat to computer systems. Trojan horses can be in the guise of anything a user might find desirable, such as a free game, mp3 song, or other application. They are typically downloaded via HTTP or FTP. Once these programs are executed, a virus, worm, or other type of malicious code hidden in the Trojan horse program is released to attack the individual user workstation and subsequently a network.

6.6.3.4 Network Attacks

With the number of networks increasing exponentially, potential threats to these networks are numerous and devastating. The most common attack is to deny service by generating large volumes of Transmission Control Protocol/Internet Protocol (TCP/IP) traffic. The target site is rendered “unavailable” to the rest of the Internet community. The next level of denial-of-service (DOS) attacks is the distributed DOS-attack where several machines on the target site are exploited. Distributed DOS attacks are the most effective and insidious because they generate more traffic from other sources, making it much harder to identify the attacker’s source, and subsequently more difficult to resolve. An example of a distributed DOS attack was the attack by “coolio” in February 2000, which caused the crash of numerous Web sites in the United

States, including Ebay, CNN, Yahoo!, and E*Trade. This attack involved sending Internet Control Message Protocol (ICMP) echo request datagrams (ping packets) to the broadcast address of networks using a faked or “spoofed” IP address of the host to be attacked. The IP host responds to these ICMP echo requests on either the nominal address or the broadcast address of its interfaces. When the broadcast address of a network was pinged, all active hosts on that network responded, and for any one request, there were many replies. This amplification makes distributed DOS attacks very powerful and causes large networks to crash.

6.6.3.5 Trapdoors

Trapdoors provide easy access for system administrators and authorized personnel to a system or a system’s resources. Individuals can usually gain this access without a password. When these trapdoors are exploited, however, threats to a computer system are created. Authorized or unauthorized users with knowledge of trapdoors, can plant various types of malicious code into sensitive areas of a system. Therefore, the first layer of defense, prevention of malicious code, is bypassed, and the system must rely on detection and removal mechanisms to rid the system of the newly introduced malicious code.

6.6.3.6 Insider Attacks

Traditionally, insiders are a primary threat to computer systems. Insiders have legitimate access to the system and usually have specific goals and objectives. They can affect availability of system resources by overloading processing or storage capacity, or by causing the system to crash. Insiders can plant Trojan horses in sensitive data files, which attack the integrity of the entire system. Insiders can also exploit bugs in the OS by planting logic bombs or by causing systems to crash. All of these attacks by insiders are difficult to prevent, as legitimate access is essential to all users for crucial daily functions.

6.6.3.7 Connection/Password Sniffing

Other threats to the integrity of a system include connection and password “sniffing.” A “sniffer” is malicious software or hardware that monitors all network traffic, unlike a standard network station that only monitors network traffic sent explicitly to it. Software sniffers can be a real threat to a network because they are “invisible” and easily fit on all workstations and servers. The specific threat presented by sniffers is their ability to catch all network traffic, including passwords or other sensitive information sent in plain text. An added threat to network security is that detecting sniffers on other machines is extremely difficult.

6.6.4 Potential Countermeasures

This section is subdivided into six types of countermeasures that can be applied to prevent and/or remove malicious code: malicious code scanning products, electronic security (access constraint countermeasures), trapdoor access constraints, network security, connection and password sniffing countermeasures, and physical security.

6.6.4.1 Malicious Code Scanning Products

Malicious code scanning products are used to prevent and/or remove most types of malicious code, including viruses, worms, logic bombs, and Trojan horses, from a system. The use of malicious code scanning products with current virus definitions is crucial in preventing and/or detecting attacks by all types of malicious code.

6.6.4.2 Electronic Security

Electronic security typically refers to access constraint mechanisms used to prevent malicious code from being introduced into a system, intentionally or unintentionally, by authorized users. Unintentional system infiltration is the primary reason to implement access constraint mechanisms. If a set number of attempts to input a password correctly is exceeded, the system administrator must be contacted immediately. The system or system administrator should ensure that users change their passwords frequently and should not allow the use of dictionary words. This prevents easy decryption of passwords. Checksums can also be used; however, they only pertain to some strains of viruses. All of these electronic security measures protect against employees' intentionally or inadvertently deploying malicious code into a system or network.

The following are additional access constraint countermeasure requirements:

- **Provide data separation.** For data that is allowed access to the protected network workstation, steps should be taken to constrain the portion of the system that can be affected in case of a malicious code attack.
- **Employ application-level access control.** Access restrictions may also be implemented within a workstation or at various points within a LAN to provide additional layers and granularity of protection against authorized and unauthorized malicious code attacks.

6.6.4.3 Trapdoor Access/Distribution

To protect against unauthorized use of trapdoors to introduce malicious code, reliable companies should be used when considering software and hardware purchases. When inputting data, only use reliable inputting individuals and use monitoring devices to monitor them. Reliable system administrators should remove passwords immediately after an employee leaves a company. All of these prevention techniques are crucial to prevent malicious code from infiltrating systems through trapdoors.

6.6.4.4 Network Security

A boundary protection mechanism at the gateway must be used within a network. The requirements for a boundary protection mechanism are mentioned in the following sections of the Information Assurance Technical Framework (IATF): Section 6.1, Firewalls, Section 6.3,

Guards, and Section 8.2, Intrusion Detection. The requirements in these sections describe a boundary protection mechanism for network security.

There are also several ways to protect a network against distributed DOS attacks by malicious code. Secure hosts on the network by replacing “rlogin” and “rexec” commands with “ssh” or other encrypted commands. Also, disallow IP spoofing to keep hosts from pretending to be others. Do not allow ICMP to broadcast and multicast addresses from outside the network. These few preventive methods will help prevent distributed DOS attacks.

6.6.4.5 Connection and Password Sniffing Countermeasures

Although sniffing of Internet traffic is difficult to stop, there are several ways to defend a system and make sniffing difficult. First, use an encryption mechanism (e.g., Secure Sockets Layer [SSL]) to allow encryption of message transmissions across Internet protocols whenever possible. Also, encrypt e-mail through the use of Pretty Good Privacy (PGP) and Secure Multi-Purpose Internet Mail Extensions (S/MIME). Although e-mail is sent encrypted, when e-mail is read it must be unencrypted. If mail programs allow attachments to automatically run, malicious code can still infect a system. The malicious code will be encrypted with the rest of the message and activate when you read the decrypted message. Also, implement “ssh” or other encrypted commands instead of insecure remote login. To stop password sniffers, use secure remote access and smart cards to keep passwords private. To protect a LAN from sniffing, replace a hub with a switch, which is extremely effective in practice. Although sniffers can still access the LAN, it becomes more difficult for them to do so.

6.6.4.6 Physical Security

To be physically secure against potential infections by malicious code, the system must be protected from physical attack. It is necessary to use a monitoring system to authenticate users to restrict physical access. Once access is granted, users’ actions must be monitored.

6.6.4.7 Detection Mechanism

The detection mechanism enables users to detect the presence of malicious code, respond to its presence, and recover data or system files, if possible.

Detect

The objectives for detection are to discover attacks at or inside the protected boundary as well as to facilitate tracking and prosecuting of adversaries. Malicious code detection involves the continual probing of internal networks for the existence of services or applications infected by malicious code. This may be done routinely to assist in the selection of additional appropriate countermeasures, to determine the effectiveness of implemented countermeasures, or to detect all

types of malicious code. The following are typical security capability requirements associated with malicious code detection and system probing.

- Provide centralized operation.
- Provide automated reports.
- Recommend corrective action.
- Archive significant security events.
- Display and record status in real time.

Respond

To respond to the presence of detected malicious code within a system or network, malicious code scanning must be performed. The following are typical security capability (counter-measure) requirements.

- Detect occurrence of infection and locate malicious software, e.g., a virus found in local memory.
- Perform scanning automatically, e.g., run continual malicious code scans throughout the day on systems.
- Implement scanning at the network gateway and at network components such as the desktop.
- Identify specific malicious code, e.g., macro virus.
- Remove malicious code from all infected systems so it cannot infect further, e.g., boot from uninfected write-protected boot diskette, then remove the malicious code from the system.
- Correct all effects of malicious code and restore system to original state, e.g., check all diskettes with files that may have been in disk drives during virus residency; reload files as appropriate.
- Reload program backups in cases where malicious code cannot be completely identified or where removal is not possible.
- Perform manually initiated scanning regularly, e.g., scan for malicious code after any Internet downloads.

Recover

To recover data from the infection of malicious code, first concentrate on the specific area infected. The recovery process will take longer if malicious code has been in the system for a longer time. The number of computers that have been infected is also important as it affects time and resources for recovery. There are four stages in the infection process, and each stage requires a different amount of time and resources for recovery.

UNCLASSIFIED

- 1) Local Memory Infection is the first stage of the infection process of a malicious code. If malicious code is caught in the first few hours before an appropriate host is found and replication begins, the following straightforward approach can be applied:
 - a) Power down,
 - b) Cold reboot with a clean, write-protected diskette,
 - c) Run a utility program to check hard disk and remove the few infected files, and
 - d) Locate and destroy the source containing the malicious code.
- 2) Local Disk Storage Infection is the second stage of the infection process. If an infection goes undetected, malicious code will infect an increasing number of programs and data files over time. In this case, the removal process becomes more complicated and several things could happen. If data and program files have been destroyed, it is possible that a complete reformat of the infected media will be required for recovery. File backups can also be dangerous due to the risk of reinfection during the restoration process. Total data loss may occur.
- 3) Shared File System Infection is the third stage of the infection process of malicious code. The risk of malicious code infecting the network attached to a computer is very high. If the infection is widespread, it is possible that a reformat of the entire medium will be required for recovery. Many things could happen during the recovery process. Again, file backups can be dangerous due to the risk of reinfection during the restoration process. One complication is numerous computers attached to the infected network will also be infected. The malicious code must be removed simultaneously from all workstations as well as the network. Another complication is that other users may have saved the malicious code unknowingly onto a floppy disk that may infect the entire network later.
- 4) System-wide Removable Media Infection is the final stage of the infection process. An infected computer will infect many of the physical disks it contacts. This is an extremely difficult situation to deal with for numerous reasons. Malicious code infects all types of removable media, such as floppy diskettes, removable hard disks, reel and cartridge tapes, etc. Once an infected disk has successfully infected a network computer, the number of infected disks drastically increases. A complication with all the infected disks is the possibility of reinfection after malicious code has been discovered and removed. Although scanning devices would have been updated since the original infection and would catch many possible reinfections, new malicious code, like the polymorphic virus that changes itself after each infection, could still compromise the network. Malicious code could also reach client sites and computers.

6.6.4.8 Administrative Countermeasures

Administrative concerns regarding infection by malicious code include training, policy, and coping with fears about malicious code and computers. “Viruses affect the emotional relationships that many people develop with their computer. Viruses could change the very nature of computing, from an essentially logical, predictable function to one fraught with

uncertainty and danger.” It is crucial for administrators to minimize stress due to computer viruses while not blaming employees.

Administrators can combat fears about malicious code and computers in many ways. The staff should be educated and motivated with regard to malicious code protection, detection, and recovery. A review of computer security with a risk analysis of exposure to infection and likely consequences should be conducted. A corporate policy with information about malicious code should be distributed to all staff. In addition, special briefing sessions should be held for all staff involved with computing functions. Administrators need to institute prevention programs that incorporate safe computing practices that should be posted at all terminals. Regular training sessions on safe computing should be scheduled. Administrators should also have a disaster recovery plan that is practiced on worst-case scenarios. Twenty-four-hour emergency phone numbers should be displayed. Most employees should also be cautioned to avoid overreaction and deploy backup facilities to minimize consequential damage.

6.6.5 Technology Assessment

Before describing malicious code detection products, it is important to understand the different types of malicious code.

6.6.5.1 Types of Malicious Code

Viruses

There are several classes of viruses, which range from innocuous to catastrophic. An understanding of each class is crucial to understanding the evolutionary process of an infiltrating virus. Innocuous viruses reside in unobtrusive areas of the system and cause no noticeable disruption. These viruses infect diskettes and other media that come into contact with the system but intend no damage. Humorous viruses cause aggravating events to occur, humorous messages to appear, or graphic images to be displayed. Although irritating, these viruses intend no damage and are commonly used for jokes. Potentially the most disruptive and difficult to detect are the data-altering viruses that alter system data. The viruses modify data file numeric information in spreadsheets, database systems, and other applications, such as changing all occurrences of the number three to the number eight. Catastrophic viruses erase critical system files and immediately cause widespread destruction. The viruses scramble key information tables and/or remove all information on all disks, including shared and network drives.

There are two main phases in the lifecycle of a virus.

- 1) The first phase, replication, could last a few weeks to several years. In this phase, viruses typically remain hidden and do not interfere with normal system functions. Viruses also actively seek out new hosts to infect such as attaching themselves to other software programs or infiltrating the OS. A virus that is attached to an executable program executes its instructions before passing control to the program (see Figure 6.6-3). These viruses are hard to detect because they only infect a small number of programs on a disk and the user does not suspect.
- 2) During the second phase, activation, the beginning of gradual or sudden destruction of the system, occurs. Typically, the decision to activate is based on a mathematical formula with criteria such as date, time, number of infected files, and others. The possible damage at this stage could include destroyed data, software or hardware conflicts, space consumption, and abnormal behavior.

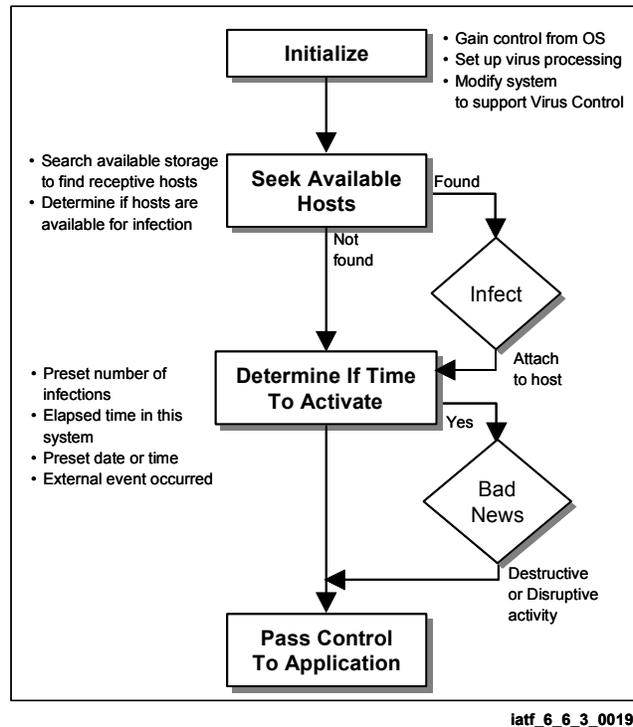


Figure 6.6-3. Virus Execution

LAN users, administrators, and individual workstation/personal computer users should scan for viruses because of the unrealized potential for harm. Numerous viruses make major computing disasters inevitable. Extraordinary damage caused by these viruses can result in loss of man-hours, disruption of normal activities, and wasted monetary resources. Therefore, the unrealized potential for harm is the main reason why malicious code scanning and prevention are extremely important.

Macro Viruses

The 1995 advent of macro programming for applications like MS Word and Excel automated repetitive keystroke functions, but also created an effective new way for viruses to spread. Word and Excel data files had previously been data-only files, like text-only e-mail messages—unable to harbor viruses because they did not include executable code.

Virus writers soon discovered these applications' macros could also be used to create viruses. At the same time, sharing of documents and spreadsheet files via e-mail became increasingly commonplace between users both within and between companies—creating the most effective virus carrier ever. Among the factors contributing to the dominance of macro viruses is the Visual BASIC for Applications (VBA) programming language, which makes it as easy for virus

writers to create time-robbing macro viruses as it does for users to create legitimate timesaving macro commands.

Once the macro-infected file is accessed, it replaces one of the Word or Excel standard macros with an infected version that can then infect all documents it comes into contact with. Macro viruses usually disable the macro menu selection, making users unable to see what macros are executing.

Today, macro viruses like ILOVEYOU are the most prevalent computer viruses in the wild—accounting for the vast majority of virus encounters in corporations. Today's widespread sharing of macro-enabled files, primarily through e-mail attachments, is rapidly increasing along with the associated macro virus threat.

Table 6.6-1, Comparison of Macro Viruses, describes the current impact of several macro viruses compared to an older virus, and the associated costs to corporations.

Table 6.6-1. Comparison of Macro Viruses

Virus	Year	Type	Time to Become Prevalent	Estimated Damages
Jerusalem, Cascade, Form	1990	Executable file, boot sector	3 Years	\$50 million for all viruses over 5 years
Concept	1995	Word macro	4 months	\$60 million
Melissa	1999	E-mail enabled Word macro	4 days	\$93 million to \$385 million
I Love You	2000	E-mail enabled Visual Basic script/word macro	5 hours	\$700 million

Polymorphic Viruses

Polymorphic viruses alter their appearance after each infection. Such viruses are usually difficult to detect because they hide themselves from antivirus software. Polymorphic viruses alter their encryption algorithm with each new infection. Some polymorphic viruses can assume over two billion different guises. This means antivirus software products must perform heuristic analysis, as opposed to spectral analysis that can find simpler viruses.

There are three main components of a polymorphic virus: a scrambled virus body, a decryption routine, and a mutation engine. In a polymorphic virus, the mutation engine and virus body are both encrypted. When a user runs a program infected with a polymorphic virus, the decryption routine first gains control of the computer, then decrypts both the virus body and the mutation engine. Next, the decryption routine transfers control of the computer to the virus, which locates a new program to infect. At this point, the virus makes a copy of itself and the mutation engine in random access memory (RAM). The virus then invokes the mutation engine, which randomly generates a new decryption routine capable of decrypting the virus yet bearing little or no resemblance to any prior decryption routine. Next, the virus encrypts the new copy of the virus body and mutation engine. Finally, the virus appends the new decryption routine, along with the

UNCLASSIFIED

Malicious Code Protection
IATF Release 3.1—September 2002

newly encrypted virus and mutation engine, onto a new program. As a result, not only is the virus body encrypted, but also the virus decryption routine varies from infection to infection. This confuses a virus scanner searching for the telltale sequence of bytes that identifies a specific decryption routine. Therefore, with no fixed signature to scan for, and no fixed decryption routine, no two infections look alike.

A good way to contain a polymorphic virus is to set up false data directories or repositories to fool the attacker into thinking he or she has reached exploitable data. This can significantly reduce the risk of being attacked. The polymorphic virus executes in these false data directories, and is fooled into believing it has infected the entire system. In reality, the directories are either deleted or nonexistent, and the virus is thus unable to infect the system.

Stealth Viruses

Stealth viruses attempt to hide their presence from both the OS and the antivirus software. Some simple techniques include hiding the change in date and time as well as hiding the increase in file size. Stealth viruses sometimes encrypt themselves to make detection even harder. Stealth viruses also enter systems through simple download procedures. Unsuspecting users can do little against this type of infection except download files only from trusted sources.

Worms

Worms are constructed to infiltrate legitimate data processing programs and alter or destroy the data. Although worms do not replicate themselves as viruses do, the resulting damage caused by a worm attack can be just as serious as a virus, especially if not discovered in time. However, once the worm invasion is discovered, recovery is much easier because there is only a single copy of the worm program to destroy since the replicating ability of the virus is absent.

A prevalent worm, "Ska," is a Windows e-mail and newsgroup worm. An e-mail attachment disguised as "Happy99.exe" will display fireworks when executed the first time. After execution, every e-mail and newsgroup posting sent from the machine will cause a second message to be sent. Since people receive "Happy99.exe" from someone they know, people tend to trust this attachment, and run it. Then the worm causes damage by altering functionality of the WSOCK32 dynamic library link (DLL) file. Now the worm can actively attack other users on the network by placing itself on the same newsgroups or same e-mail addresses to which the user was posting or mailing.

Trojan Horses

A Trojan horse is an apparently harmless program or executable file, often in the form of an e-mail message, that contains malicious code. Once a Trojan horse gets into a computer or network, it can unleash a virus or other malicious code, take control of the computer infrastructure, and compromise data or inflict other damage. The Melissa virus that struck in 1999 is a good example of a harmful Trojan horse. Attached to a harmless-looking e-mail message, the virus accessed Microsoft Outlook, replicated itself, and sent itself to many other

users listed in the recipient's e-mail address book. The resulting e-mail-sending flurry caused many Microsoft Exchange servers to shut down while users' mailboxes flooded with bogus messages.

Trojan horses can also be carried via Internet traffic such as FTP downloads or downloadable applets from Web sites. These can not only compromise enterprise computers and networks by rapidly infecting entire networks, but also can invite unauthorized access to applications that results in downtime and costs to business potentially reaching into the millions of dollars.

Logic Bombs

Logic bombs are programs added to an already existing application. Most are added to the beginning of the application they are infecting so they are run every time that application is run. When the infected program is run, the logic bomb is run first and usually checks the condition to see if it is time to run the bomb. If not, control is passed back to the main application and the logic bomb silently waits (see Figure 6.6-4). When the right time does come, the rest of the logic bomb's code is executed. At that time, the hard disk may be formatted, a disk erased, memory corrupted, or anything else. There are numerous ways to trigger logic bombs:

counter triggers, time triggers, replication triggers (activate after a set number of virus reproductions), disk space triggers, and video mode triggers (activate when video is in a set mode or changes from set modes). There are also Basic Input Output System (BIOS) read only memory (ROM) triggers (activate when a set version of BIOS is active), keyboard triggers, antivirus triggers (activate when a virus detects variables declared by virus-protection software such as "SCAN_STRING"), and processor triggers (activate if a program is run on a particular processor).

Logic bombs cannot replicate themselves and therefore cannot infect other programs. However, if the program that is infected is given to someone else and the right conditions are met on that computer it will go off.

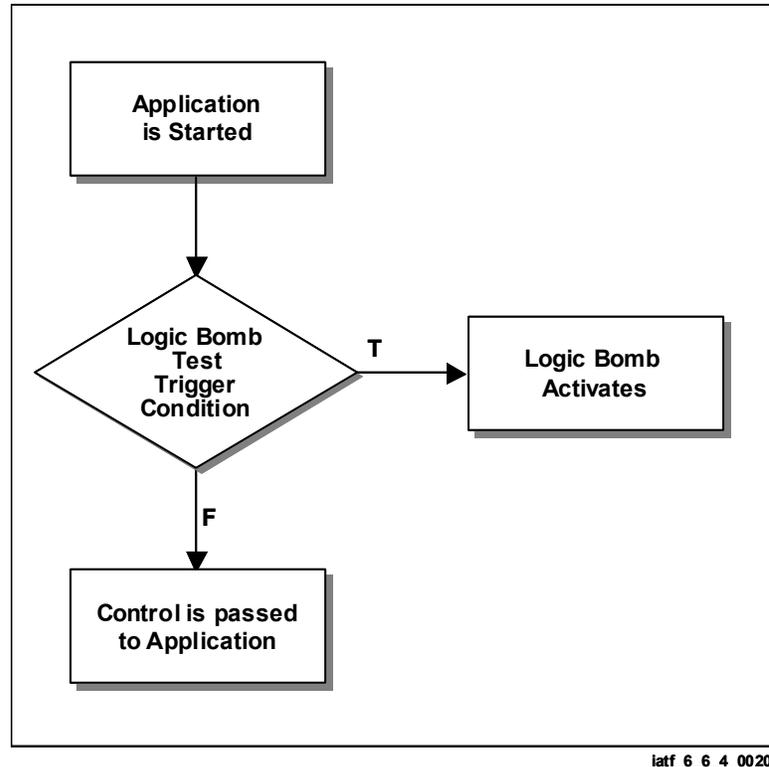


Figure 6.6-4. Logic Bomb Execution

6.6.5.2 Viruses and E-Mail

Today's office worker receives an average of more than 40 e-mail messages each day. Many of these messages have Microsoft Word or Excel data files attached, that may carry macro viruses. Since plain text data cannot carry the executable program code viruses need to copy and spread themselves, the text messages of electronic mail are, by themselves, unable to spread viruses. The virus danger from e-mail stems from attachments containing active executable program files with extensions such as: CLASS, OCX, EXE, COM, and DLL—and from macro-enabled data files. These attachments do not even need to be opened, as many mail clients automatically display all attachments. To prevent attachments from automatically being displayed, simply configure the mail client to prompt the user. Another safeguard is to identify file extensions prior to opening attachments so the infection of many computer systems may be prevented. These attachments could contain malicious code that could be masquerading as another file type.

6.6.5.3 Virus Creation

There are two types of viruses that can be created: simple viruses and complex viruses.

Simple Viruses

Simple viruses do not attempt to hide themselves and are easy to write. Users with little computer knowledge can use Internet programs to create these viruses. Since thousands of sites contain virus source code, users can easily download and use existing viruses to infect systems. Users with slightly more computer knowledge may even alter existing virus source code or combine several viruses to create a new undetectable virus capable of compromising systems.

Complex Viruses

Complex viruses require more source code than simple viruses, which is used to conceal them from systems. Knowledge of assembly language is required to manipulate interrupts so these viruses can remain hidden. While hiding, complex viruses replicate, and will destroy data later. A complex virus is divided into three parts: the replicator, the concealer, and the bomb. The replicator part controls spreading the virus to other files, the concealer keeps the virus from being detected, and the bomb executes when the activation conditions of the virus are satisfied. After these parts are created and put together, the virus creator can infect systems with a virus that current antivirus software cannot detect.

6.6.5.4 Virus Hoaxes

The Internet is constantly being flooded with information about malicious code. However, interspersed among real virus notices are computer virus hoaxes. Virus hoaxes are false reports about nonexistent viruses, often claiming to do impossible things. While these hoaxes do not infect systems, they are still time consuming and costly to handle. Corporations usually spend much more time handling virus hoaxes than handling real virus incidents. The most prevalent

virus hoax today is the “Good Times Hoax” that claims to put your computer’s central processing unit (CPU) in an “nth-complexity infinite binary loop that can severely damage the processor.” In this case, there is no such thing as an nth-complexity infinite binary loop. It is estimated virus hoaxes cost more than genuine virus incidents. No antivirus product will detect hoaxes because they are not viruses, and many panic when they receive a hoax virus warning and assume the worst—making the situation much worse.

6.6.5.5 System Backup

There are two main strategies to follow when performing a system backup.

Workstation Strategy

The best backup strategy for workstations is to back up often. If the workstation is running the Windows OS, there are some simple backup tools already provided. There are also several utilities and programs available from other companies to assist users in performing backups. The following features can make backup chores more bearable: incremental backup, unattended scheduling, and easy, simple restoration. Incremental backup saves changes made since the most recent full or incremental backup. This is important because users who do not want to wait to back up a system can use incremental backup as a substitute for a lengthy full backup. Scheduling uses software automation to execute backup chores without the need for personal interaction. Although a backup medium must be selected and in place, the user does not need to be present for the actual backup. Zip drives and small tape drives are also cost-effective solutions used to back up workstation data.

Network Strategy

The best backup strategy for networks is an approach that combines several features to save time and effort, and still assure complete backups. Execute full backups often. Since backups take up network, server, and/or workstation resources, it is best to run full backups when nobody is working. In addition, open files are skipped during backup and do not get backed up at all until some future time when the file is closed and not being used. Having few to no users holding files open will ensure the greatest backup saturation possible. Full backups are most efficiently executed in the evenings. Store the full backup tape off site. On each of the remaining workdays of the week, using a separate tape for each day, run an incremental backup and store it off site, too. The last full backup of the month should be permanently moved off site and held for archival purposes. Therefore, if a network is attacked by malicious code, these backup techniques will ensure data integrity and allow all systems to be recovered.

6.6.5.6 Types of Malicious Code Detection Products

Most computer malicious code scanners use pattern-matching algorithms that can scan for many different signatures at the same time. Malicious code detection technologies have to include scanning capabilities that detect known and unknown worms and Trojan horses. Most antivirus

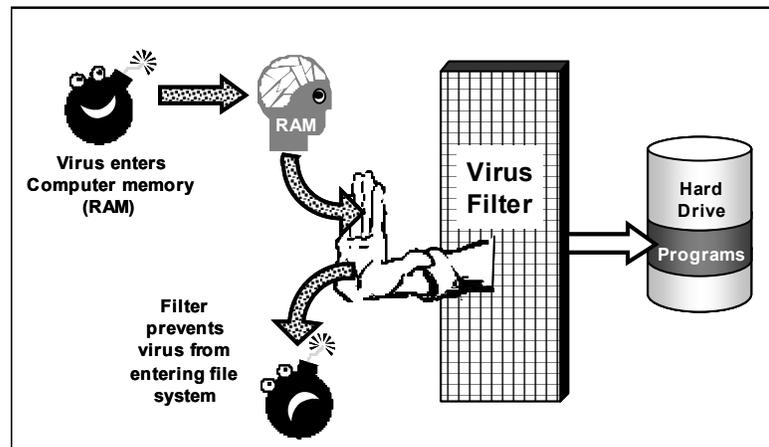
products search hard disks for viruses, detect and remove any that are found, and include an auto-update feature that enables the program to download profiles of new viruses so that it will have the profiles necessary for scanning. The virus like signatures these programs recognize are quite short: typically 16 to 30 bytes out of the several thousand that make up a complete virus. It is more efficient to recognize a small fragment than to verify the presence of an entire virus, and a single signature may be common to many different viruses.

6.6.5.6.1 Pre-Infection Prevention Products

Pre-infection prevention products are used as the first level of defense against malicious code. Before the code actually attacks a system, prevention products should be applied. E-mail filtering products are available that do not allow executable programs or certain file types to be transferred. Also, options in browsers that limit the use of and/or disable Java and ActiveX plug-ins should be implemented. Simply changing browser options allows the user to see hidden files and file extension names. This could prevent opening an infected file masquerading as a normal text file. These essential pre-infection prevention products are the first level of defense against malicious code attacks.

6.6.5.6.2 Infection Prevention Products

Infection prevention products are used to stop the replication processes and prevent malicious code from initially infecting the system. These types of products, protecting against all types of malicious code, reside in memory all the time while monitoring system activity. When an illegal access of a program or the boot sector occurs, the system is halted and the user is prompted to remove the particular type of malicious code. These products act like filters that prevent malicious code from infecting file systems (see Figure 6.6-5).



iatf_6_6_5_0021

Figure 6.6-5. Virus Filter

6.6.5.6.3 Short-Term Infection Detection Products

Short-term infection detection products detect an infection very soon after the infection has occurred. Generally, the specific infected area of the system is small and immediately identified. These products also detect all types of malicious code and work on the principle that all types of malicious code leave traces. Short-term infection detection products can be implemented through vaccination programs and the snapshot technique.

Vaccination Programs

Vaccination programs modify application programs to allow for a self-test mechanism within each program. If the sequence of that program is altered, a virus is assumed and a message is displayed. The drawbacks to this implementation include the fact that the boot segment is very hard to vaccinate, and the malicious code may gain control before the vaccination program can warn the user. The majority of short-term infection detection products use vaccination because it is easier to implement.

Snapshot Technique

The snapshot technique has been shown to be the most effective. Upon installation, a log of all critical information is made. During routine system inspections (snapshots) the user is prompted for appropriate action if any traces of malicious code are found. Typically, these system inspections occur when the system changes: disk insertion, connection to different Web site, etc. This technique is difficult to implement in short-term infection detection products and is not widely used. However, when the snapshot technique is used with vaccination programs, an effective protection against malicious code is established.

6.6.5.6.4 Long-Term Infection Detection Products

Long-term infection detection products identify specific malicious code on a system that has already been infected for some time. They usually remove the malicious code and return the system to its prior functionality. These products seek a particular virus, and remove all instances of it. There are two different techniques used by long-term infection detection products: spectral analysis and heuristic analysis.

Spectral Analysis

Using spectral analysis, long-term infection detection products search for patterns from code trails that malicious code leaves. To discover this automatically generated code, all data is examined and recorded. When a pattern or subset of it appears, a counter is incremented. This counter is used to determine how often a pattern occurs. Using these patterns and the quantity of their occurrence, these products then judge the possible existence of malicious code and remove all instances of it. These products search for irregularities in code and recognize them as particular instances of malicious code.

Heuristic Analysis

Using heuristic analysis, long-term infection detection products analyze code to figure out the capability of malicious code. The underlying principle that governs heuristic analysis is that new malicious code must be identified before it can be detected and subsequently removed. This technique is much less scientific, as educated guesses are created. Because they are guesses, heuristic analysis does not guarantee optimal or even feasible results. However, it is impossible to scientifically analyze each part of all source code. Not only is this unproductive, it is terribly

inefficient. Typically, good educated guesses are all that is needed to correctly identify malicious code in source code.

These long-term infection detection products then remove all instances of the detected malicious code.

DOS file viruses typically append themselves on the end of DOS .EXE files. DOS file viruses can also append themselves to the beginning or end of DOS .COM files (see Figure 6.6-6). Other infection techniques are also possible but less common.

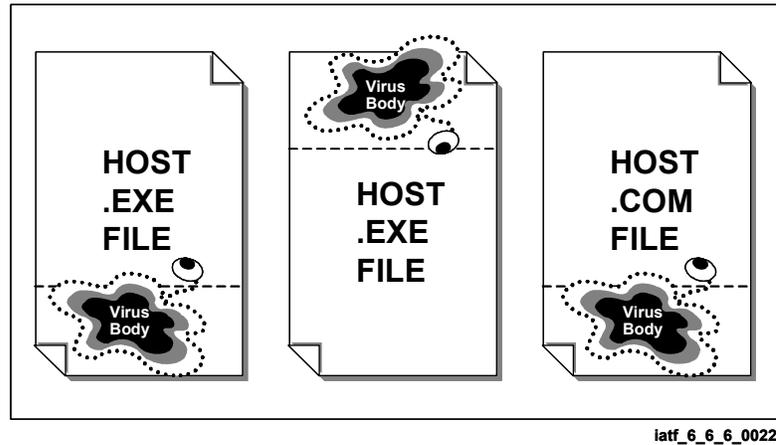


Figure 6.6-6. DOS File Infection

6.6.5.6.5 Interoperability

The different types of products mentioned above must be used together to create effective protection against all types of malicious code. Many layers of defense must be in place for a system to deal effectively with malicious code. If each type of product is implemented in a system, four different levels of defense are created. Before malicious code can attack a system, it must first get to the system through the pre-infection prevention products. If it gets that far, the second layer of defense, prevention products will attempt to stop the malicious code from replicating. If that is not successful, then the detection products will try to locate and remove the infection before it reaches the majority of the system. If the malicious code reaches the entire system, identification products can apply two different techniques to remove the infection. Each of these levels of defense is essential to the prevention of infection and the protection of a system.

Today, commercial software packages combine all the above levels of defense and provide malicious code protection services. With new computer systems connecting to the Internet daily, security problems will also grow at an exponential rate. Unless a well-defined security policy is in place, information technology managers will continue to lose the battle against computer viruses. Computer Emergency Response Team (CERT) statistics show the number of virus attacks rose from 3,734 in 1998 to 9,859 in 1999. In the first quarter of 2000, the CERT has reported 4,266 incidents. Despite the fact that antivirus applications are essential for the detection of known viruses, no mail filter or malicious code scanner can defend against a new mail worm attack. The recent “Love Bug” virus was caught quickly and still did a wealth of damage. It seems to only be a matter of time before crackers figure out how to send e-mail worms that infect systems without opening attachments. While not sophisticated enough to stop new viruses from entering systems, antivirus application makers are producing software that can prevent the damaging, data-altering effects of the malicious code.

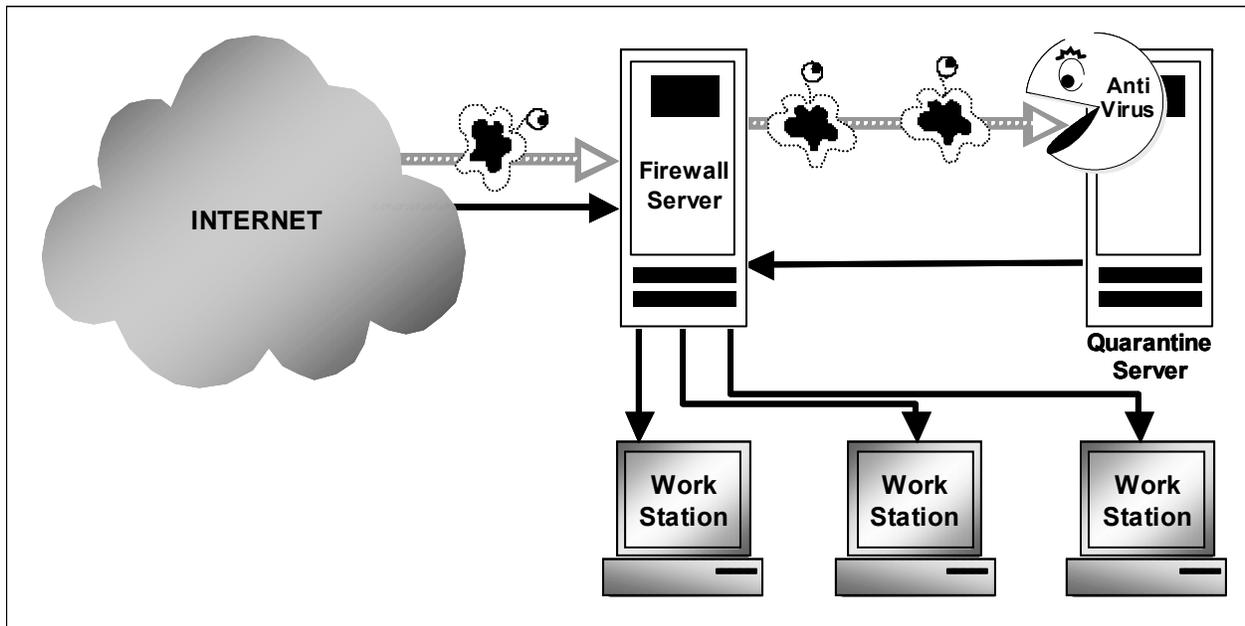
6.6.5.7 Protection at the Workstation

There are numerous ways to protect a workstation from malicious code attacks. The implementation of pre-infection prevention, infection prevention, infection detection, and infection identification products provide four separate levels of defense and are essential in protecting a workstation. Although this is the best way to protect a workstation, other techniques can be applied. New malicious code protection products introduce a “sandbox” technology allowing users the option to run programs such as Java and ActiveX in quarantined sub-directories of systems. If malicious code is detected in a quarantined program, the system simply removes the associated files, protecting the rest of the system. Another protection mechanism is to allow continual virus definition updates that are transparent to the user. Implementing these updates at boot time, or periodically (1 hour, 2 hours, etc.) drastically reduces the chance a system will be infected with newly discovered malicious code. In the past 6 months alone, over 4,000 new viruses have been discovered. Without current virus definition updates, a system is left vulnerable to the devastating effects from malicious code.

6.6.5.8 Protection at the Network Gateway

When protecting a network, a number of issues must be considered. A common technique used in protecting networks is to use a firewall with Intelligent Scanning Architecture (ISA). (Figure 6.6-7) In this technique, if a user attempts to retrieve an infected program via FTP, HTTP, or SMTP, it is stopped at the quarantine server before it reaches the individual workstations. The firewall will only direct suspicious traffic to the antivirus scanner on the quarantine server. This technique scales well since LAN administrators can add multiple firewall or gateway scanners to manage network traffic for improved performance. In addition, users cannot bypass this architecture, and LAN administrators do not need to configure clients at their workstations.

Other useful scanning techniques for a network include continuous, automated malicious code scanning using numerous scripts. Simple commands can be executed and numerous computers in a network can be scanned for possible infections. Other scripts can be used to search for possible security holes through which future malicious code could attack the network. Only after fixing these security holes can a network withstand many attacks from malicious code.



latf_6_6_7_0023

Figure 6.6-7. Intelligent Scanning Architecture (ISA)

6.6.6 Selection Criteria

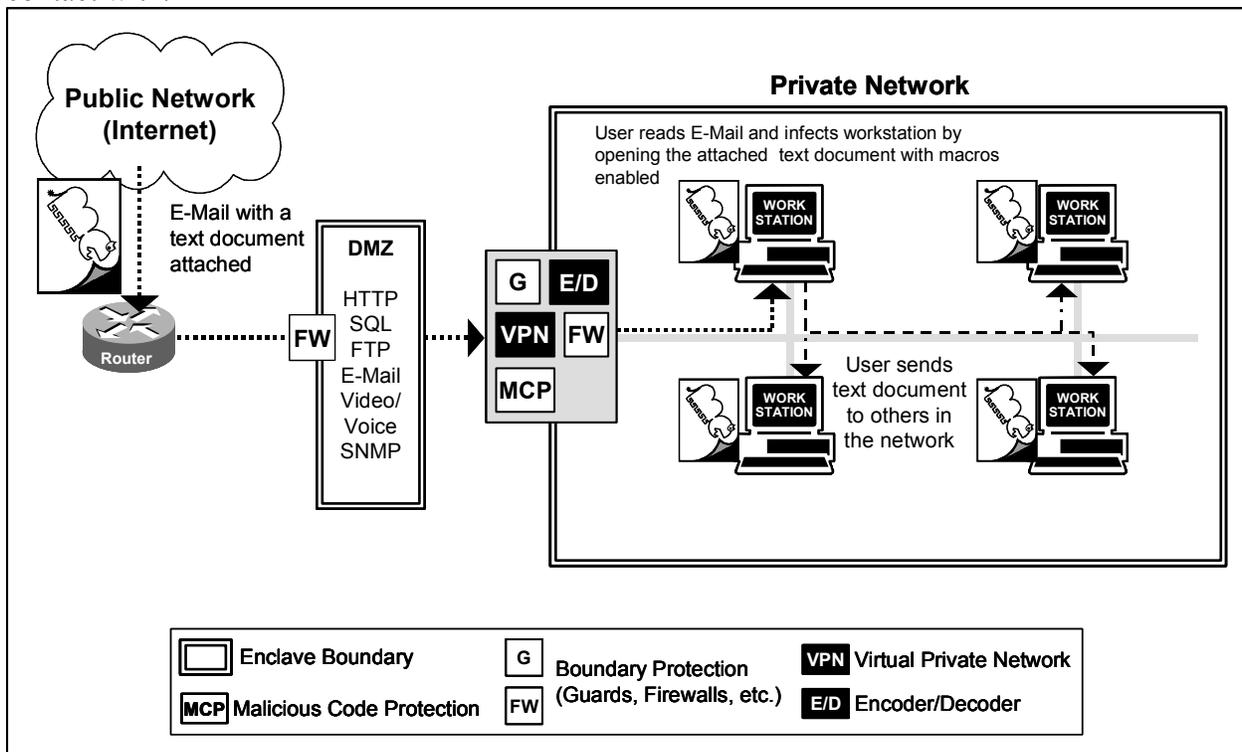
When selecting antivirus products, two important guidelines must be followed. The “best” product may not be good enough by itself. In addition, since data security products operate in different ways, one product may be more useful than another in different situations. When selecting a particular malicious code protection product, its installation must be considered. Is the program shipped on compact disk (CD) or on 1.44MB disks? Does the installation itself operate smoothly? There should be no questions without answers when properly installing a product. This product should be easy to use, providing clear and uncluttered menu systems as well as meaningful screen messages.

Help systems are essential, as users need current information regarding all types of malicious code. The trend is to provide on-line help; however, manuals should also be provided with the product. The malicious code protection product should be compatible with all hardware and software and should not create conflicts. The company that produces the product should be stable and able to provide necessary local technical support for all questions and problems. The product should be fully documented, that is, all messages and error codes should be deciphered and full installation guides and how-to manuals should be provided. The computers to run this software must meet the hardware and software requirements specified by the manufacturer. The malicious code protection software should function properly and perform its duties without failing. Rating each of these categories will allow a company to choose the best malicious code protection product for its needs.

6.6.7 Cases

6.6.7.1 Case 1: Macro Virus Attack

Within a network environment, macro virus attacks are increasing exponentially. In Figure 6.6-8 below, a macro virus has infected an enclave via an e-mail attachment sent by an outsider. This e-mail attachment is a text document that enables macros. The e-mail recipient has e-mailed this document to his coworkers and saved it to diskette to view at home. A macro virus initiates when the document is opened and macros are enabled. As soon as the document is opened, the macro virus infects standard macros in the word processing program. After altering functionality of these standard macros, this virus replicates and infects many of the documents it comes into contact with.



latf_6_6_8_0024

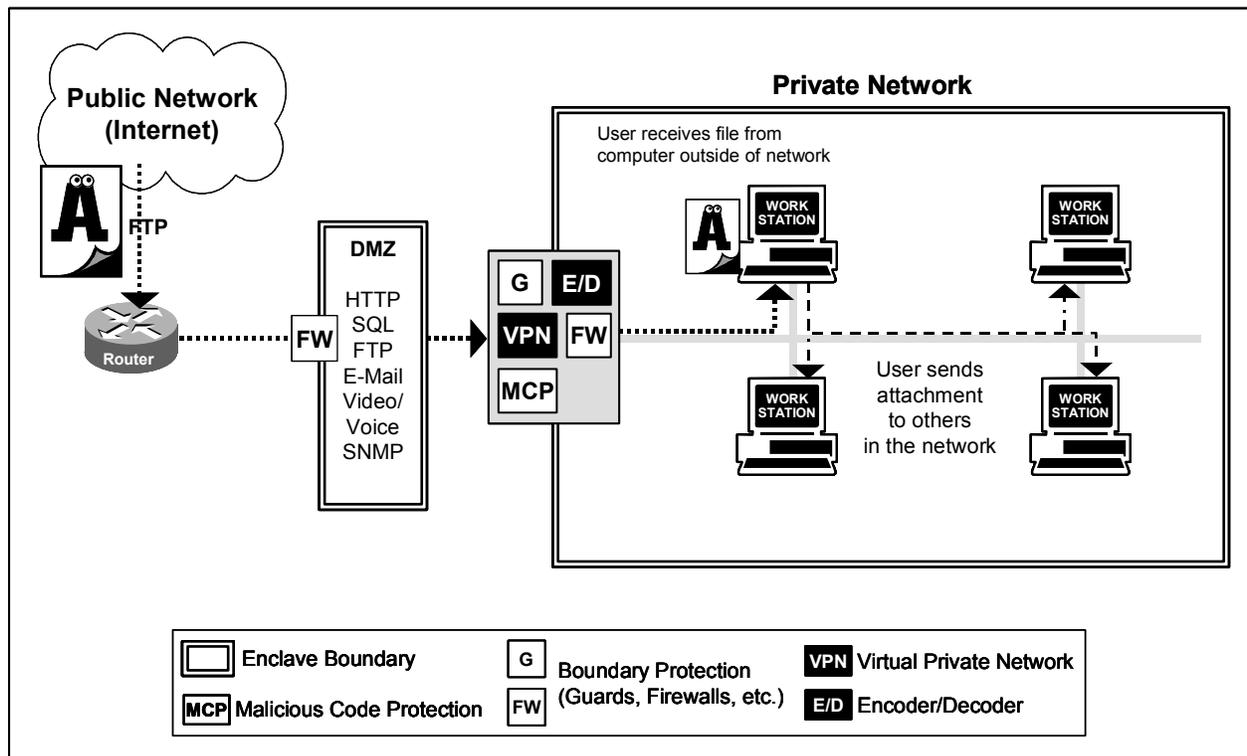
Figure 6.6-8. Macro Virus Infection

6.6.7.2 Case 2: Polymorphic Virus Attack

Polymorphic viruses represent the upper echelon of computer viruses. Today's polymorphic viruses are very difficult to detect using conventional antivirus search engines because they possess the ability to mutate themselves and conceal their digital identity as they spread. The unique ability of this form of virus to change its signature to avoid detection makes it virtually undetectable, and therefore potentially disastrous in nature.

Polymorphic viruses infect enclaves in much the same way as macro viruses. In Figure 6.6-9 below, a polymorphic virus enters a system through FTP, as an unsuspecting user retrieves a single file from a computer outside the network. The user then sends this file via an e-mail attachment to other coworkers throughout the network.

Once that file is accessed by any user, the polymorphic virus begins its programming and begins to replicate by e-mailing itself to the entire address book on its newfound host. It continuously changes its digital signature to escape the detection capabilities if any antivirus application is resident.



iatf_6_6_9_0025

Figure 6.6-9. Polymorphic Virus Infection

6.6.7.3 Case 3: Trojan Horse Attack

There exists a growing threat from another type of malicious software, the Trojan horse. In Figure 6.6-10 below, a Trojan horse has been embedded into an existing network. A user downloaded a program that he thought was useful. However, after executing it, he realized it was not exactly what he needed. So, he deleted the file off of his computer. This unsuspecting user did not realize that the program downloaded was a Trojan horse that embedded itself into the network as a sniffer program after it was executed. Although this event occurred several weeks ago, there have been no problems in the network until now, when employees are noticing forged e-mails being sent to various clients.

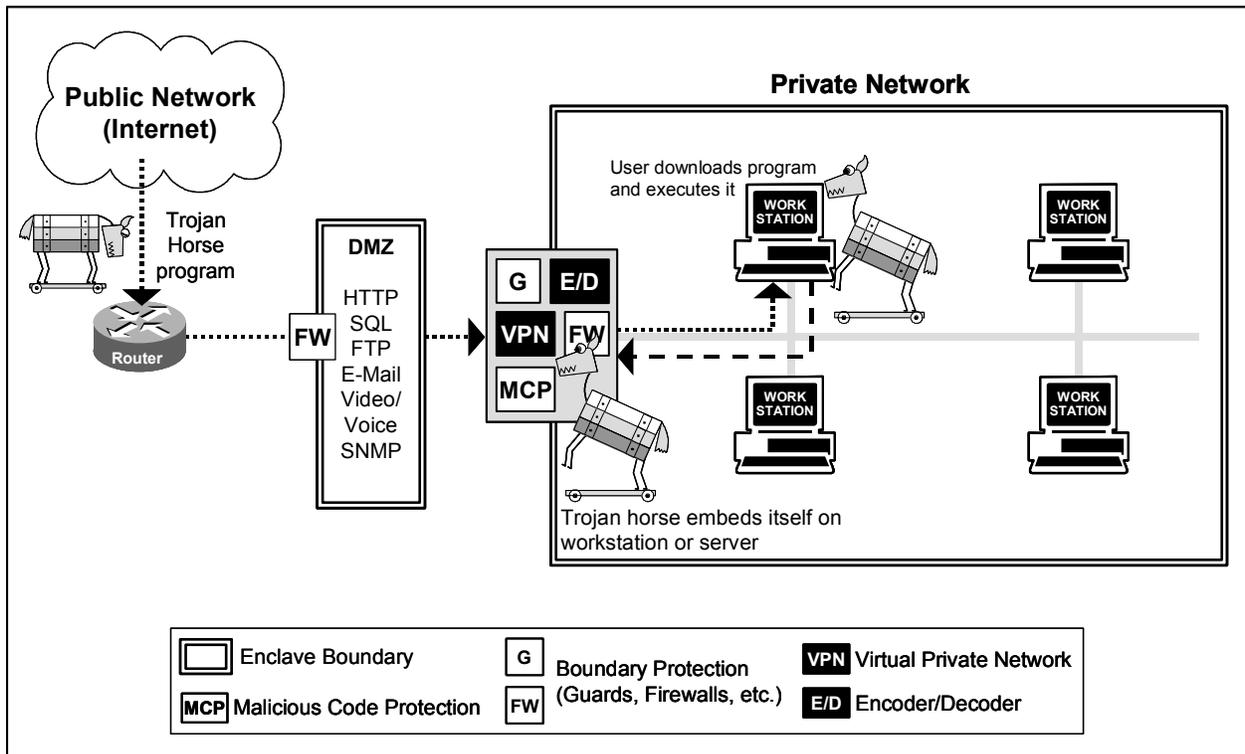


Figure 6.6-10. Trojan Horse Infection

6.6.8 Framework Guidance

In this section, guidance is provided on solutions that can be implemented so system infiltration by malicious code does not occur. Guidance will also be provided to detect and remove malicious code if it infects a system. Also, restoration guidance for the compromised system will be described.

6.6.8.1 Case 1: Macro Virus Attack

There are many ways to prevent, detect, respond to, and restore from macro virus attacks. The first level of defense is prevention so the macro virus does not reach the system. In a network environment, the first contact with the macro virus will be at the gateway. If the network is configured properly and using ISA (see Section 6.6.5.8, Protection at the Network Gateway), the macro virus should be stopped at the quarantine server. It is crucial to have current virus definition updates in the malicious code detection software on the quarantine server. These updates should occur continually, and should be transparent to the user. Implementing these updates at boot time, or periodically (hourly) drastically reduces the chance a system will be infected by a newly discovered macro virus. So, these updates prevent new macro viruses from infecting the entire network. If the macro virus is not stopped at the gateway, individual workstations should detect the presence of the macro virus and remove it. At the next layer of

defense, the individual user workstation will scan all incoming e-mail attachments for the presence of malicious code. If the malicious code detection software discovers the macro virus, the file is simply deleted and the system and network are preserved. If virus updates are automatic, virus definitions for the quarantine server and the individual workstation should be the same at the time of original system infiltration. In this case, the detection software at the workstation will probably detect the macro virus. If virus updates are not automatic, the individual user workstation will probably not detect the presence of the macro virus. This is because most users do not update their virus definitions as quickly as the system administrator of the quarantine server does. However, if this new macro virus has infected many workstations during a time frame of several days, the possibility of vendors discovering this macro virus and updating their virus definitions increases. Once this macro virus is detected by an individual workstation, the system administrator should automatically be notified.

If the macro virus does infect the network by infecting workstations, the virus must be detected and removed. Typically, new macro viruses are detected when a user notices abnormal computer behavior and that abnormality is investigated. Another way to detect viruses is through automatic virus scanning with virus software definition updates. Once the presence of the macro virus is detected, it is essential to update all virus definition updates in all copies of malicious code protection software throughout the network. Then, several methods can be applied to remove all instances of the macro virus. If the infection has occurred recently (within a few hours), short-term infection detection products should be used. Using the snapshot technique, or vaccination programs, all instances of the macro virus are detected and then removed. If the infection is not recent, long-term infection detection products should be used. Using spectral and/or heuristic analysis, all instances of the macro virus are detected and then removed.

However, if the macro virus has fully infected network workstations, the macro virus removal will then allow for the data recovery process to begin. By practicing simple system backup procedures (see Section 6.6.5.5, System Backup), applications and data can be restored from tape backups with minimal data loss. After updating malicious code definitions for all malicious code protection software, the reconstituted network is then ready to proceed with daily functions. Any damage caused by the macro virus is removed and the system is restored to its prior functionality.

If the unsuspecting user places the macro virus on his or her home computer via diskette, many problems can occur. Not only can the home computer become infected, but the network could also be reinfected. After modifying the infected file at home, the user can bring the file back to the office and infect his individual workstation. However, since the virus definitions should have been updated, the malicious code protection at the workstation should identify the virus and remove it. The user should then scan the home computer and remove all infections on that computer as well.

6.6.8.2 Case 2: Polymorphic Virus Attack

Polymorphic viruses increasingly represent serious threats to computer networks. Prevention, detection, containment, and recovery from potentially lethal polymorphic computer viruses

should be an important task of every user, network administrator, and senior management officer. Establishment of an adhered to antivirus computer policy is a must for all those requiring any degree of protection for their systems against polymorphic virus attacks.

To successfully prevent polymorphic viruses from entering into a computer system, potential vulnerabilities must be identified and eliminated. Attackers often look to exploit the most obvious vulnerability of a computer network. Inadequate security mechanisms allow unauthorized users entry into computer systems, potentially allowing data to be compromised, replaced, or destroyed. Determent of attackers can be accomplished by having a predetermined computer protection plan in place. Also, contingency plans will enable the containment of and eventual recovery from a polymorphic virus attack. Another technique for preventing polymorphic virus attacks is to set up false data directories or repositories to fool the attacker. (See Section 6.6.5.1, Types of Malicious Code, Polymorphic Viruses.) Preparation for any incident of an attack and knowledge of how a given attack might occur is all part of the strategic virus protection plan that should be implemented prior to operation of a computer network.

Detection of polymorphic viruses becomes exponentially easier when the polymorphic virus signature is cataloged in an antivirus definition table and updated regularly to all systems gateways. This can happen in one of two ways. A user can notice altered functionality on a workstation, and after technicians investigate the problem, the polymorphic virus is finally discovered. Then, technicians inform vendors who update the virus definitions for others. A user can also remove the polymorphic virus after vendors have updated their virus definitions by downloading the newest virus definitions and scanning the entire system. Establishment of an updating policy not only for system gateways, but also for individual computer workstations, greatly increases the likelihood of preventing a polymorphic virus from entering and replicating itself on a given network.

Recovery methodologies are integral to the overall readiness of an antivirus prevention plan. Even the best prepared plans sometimes fail. Having written procedures in place to recover from a catastrophic event could mean the difference between a company surviving or going out of business. Recovery consists of virus-free tape backups of recent data, providing an environment free from all viruses, and restoring the network to pre-virus infection operation. There are inexpensive software applications that unobtrusively track disk activity in such a way that they can return a system to precisely the way it was prior to a computer virus incident. Backing up data or implementation of a mirroring solution is key to having a ready alternative source of providing information to users on a moment's notice. Unless uniformly adopted throughout the entire organization, a plan will have little chance of ever becoming successful. Dedicated personnel responsible for predetermined actions in anticipated situations are crucial for the protection of computer systems.

6.6.8.3 Case 3: Trojan Horse Attack

Eradication of a Trojan horse encompasses many of the same procedures taken to eradicate macro and polymorphic viruses (see Sections 6.6.8.1, Case 1: Macro Virus Attack, and 6.6.8.2, Case 2: Polymorphic Virus Attack). This is because the Trojan horse can contain a virus inside

UNCLASSIFIED

Malicious Code Protection
IATF Release 3.1—September 2002

of the apparently harmless program. However, in this case, something else must be done to rid the network of the sniffer program hidden inside the Trojan horse. There is no one solution to prevent, detect, or remove sniffers. Since sniffer programs are extremely difficult to detect, the first level of defense against them is to make sniffing difficult. The network should use a switch instead of a hub to prevent sniffing of internal user passwords. By using an encryption mechanism for message transmissions and e-mail transactions, sniffing of important data such as passwords can be prevented. The use of “ssh” or other encrypted commands can help keep passwords private. Another precaution against password sniffing is the use of 1 time passwords. It does an attacker no good to sniff a password that is only valid during a very short time period.

In this case, the presence of sniffers is suspected since numerous forged e-mails have occurred. By applying the above measures of encryption and secure commands, sniffers can be rendered ineffective as passwords become much harder to decipher. It is also a good practice to change passwords often, or have the system administrator force users to change their passwords periodically to decrease the chance sniffer program users have time to decrypt encrypted passwords.

Also, it cannot be stressed enough how important it is to establish a complete and comprehensive malicious code protection backup system. If sniffer program users gain unauthorized access to the network, user applications and data files could be deleted. The only countermeasure in this case is to change all passwords and restore the system to prior functionality from full system backups. However, when systems are restored the sniffer must not be restored also.

References

1. "A Clear and Present Danger," Information Week. May 22, 2000, p.166.
2. "AINT Misbehaving: a Taxonomy of Anti-Intrusion Techniques" SANS Institute Resources Intrusion Detection FAQ. Ver. 1.33.
3. Bassham, Lawrence E. and Polk W. Timothy, "Threat Assessment of Malicious Code and Human Computer Threats," NIST – Computer Security Division, October 1992.
4. "Batten Down The Digital Hatches!" Forbes. June 12, 2000 p.246.
5. CIAC, "H-05 Internet Hoaxes: PKZ300, Irina, goot Times, Deeyenda, Ghost," U.S. Department of Energy, Nov 20, 1996.
6. Chess, David., "The Future of Viruses on the Internet," Virus Bulletin International Conference In San Francisco, October 1997.
7. "DANGEROUS 'LOVE': Recent virus attacks prompt enhanced security measures," Computer Reseller News. May 29, 2000, p.45.
8. "Don't fall for a Virus Hoax," Sophos Virus Info, 23 Nov. 1999.
9. F-Secure, "Security Risks for the Road Warrior," Wed. July 12, 2000.
10. "Frost & Sullivan Awards Internet Security Systems the 2000 Market Engineering Marketing Strategy Award," Press Release. June 28, 2000.
11. Gabrielson, Bruce C., "Computer Viruses," INFOSEC Engineering, AFCEA Seminar, Burke, VA. Sept. 1994.
12. "An Introduction to Computer Viruses (and other Destructive Programs)," McAfee Network Security and Management.
13. Ludwig, Mark., The Giant Black Book of Computer Viruses, Show Low, AZ, 1995.
14. McAfee, John., Computer Viruses, Worms, Data Diddlers, Killer Programs, and Other Threats to Your System, Fifth Avenue, NY, 1989.
15. Micro, Trend., "Eliminating Viruses in the Lotus Notes Environment," 1999.
16. "Securing dot-com – New viruses, distributed security threats pose perpetual challenges to IT," eWeek, June 26, 2000 p.1.
17. Slade, Robert M., "Antiviral Protection Comparison Reviews," 1995.
18. Wack, John P. & Carnahan, Lisa J., "Computer Viruses and Related Threats: A Management Guide," NIST Special Publication.
19. "Understanding Symantec's Anti-virus Strategy for Internet Gateways," The Symantec Enterprise Papers, Volume XXX.

UNCLASSIFIED

Malicious Code Protection
IATF Release 3.1—September 2002

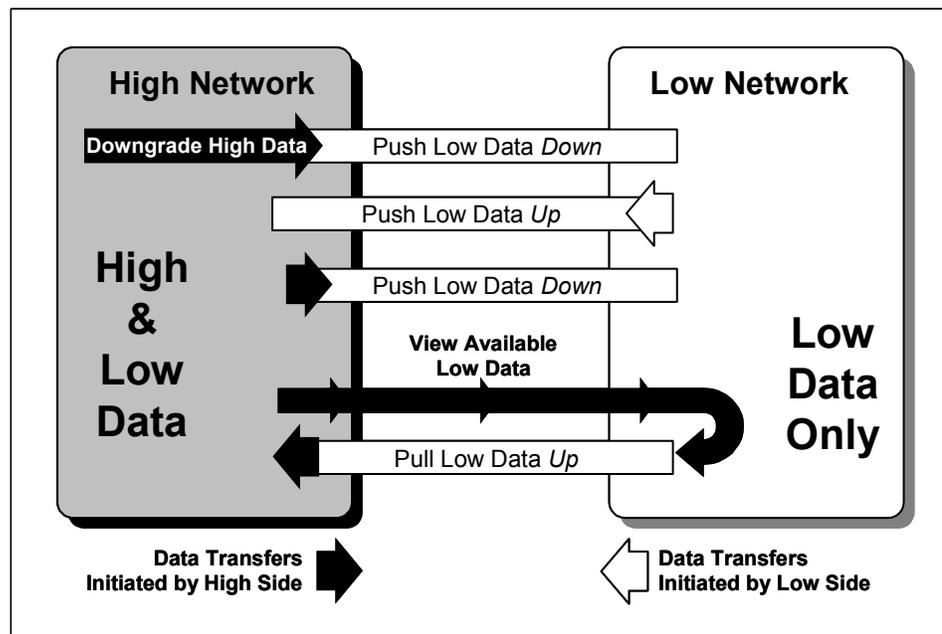
20. “Understanding and Managing Polymorphic Viruses,” The Symantic Enterprise Papers, Volume XXX.
21. “What Virus Is Lurking?—Better not touch that E-mail.” Computer Reseller News. June 5, 2000 p.1.

6.7 Multilevel Security

6.7.1 High-to-Low

The High-to-Low category is a subcategory of multilevel security (MLS). The goal of this category is to provide solutions giving installations the ability to connect networks of unlike classification (in generic terms, the classifications can be described as “High” and “Low”), as depicted in Figure 6.7-1. Given that the classifications of the data on the two networks are ordered, i.e., one is higher than the other is, users would have the ability to exchange Low data between the High and low networks. This ability is in spite of the fact that neither the High network nor the Low network has the ability to label the data. All data on the High side is considered to be High data. Users on the High network must explicitly designate data as Low and then request that it be transferred to the Low network.

This is a flow of Low data from High to Low. Likewise, Low data may flow from Low to High as a result of a user on the Low network sending data to the High network (e.g., in an e-mail message), or a user on the High network requesting data from the Low network (e.g., through a HyperText Transfer Protocol [HTTP] request to a Web server on the Low side.



iattf_6_7_1_0003

Figure 6.7-1. High-to-Low Concepts

In no case is it desirable for High data to cross between the two networks in either direction. There are three primary statements within the policy for High-to-Low. First, the High data on the High network must never cross to the Low network. Second, the High network must be protected from attacks that could cause High data to be leaked to, modified by, or destroyed by users on the Low network. Third, High network resources may not be utilized, modified, destroyed, or made unavailable by unauthorized Low network users.

These requirements apply to all High-to-Low connections, regardless of the actual classifications. Possible scenarios include Secret-to-Unclassified, Secret U.S.-to-Secret

Releasable, Top Secret-to-Secret, and High-to-Low connections that are not formally classified such as (Unclassified but Controlled)-to-Unclassified Internet. It is the intention of this framework to specify requirements in a form that is generic enough to address all popular network services, e.g., e-mail, HTTP, File Transfer Protocol (FTP), database. The requirements will be phrased in terms of “pushing” and “pulling” data between the two networks.

6.7.1.1 Target Environment

There are three target environments that this framework will address:

- 1) Allow users on the High network to push Low data to users on the Low network, and allow users on the Low network to push Low data to users on the High network.
- 2) Allow users on the High network to downgrade data to Low, and push that data to a server on the Low network for subsequent pull by users on the Low network.
- 3) Allow users on the High network to view and import (pull) data that exists on the Low network.

In the remainder of this framework, the above three capabilities will be referred to, respectively, as—

- Communication.
- Releasability.
- Network access.

6.7.1.2 Consolidated Functional Requirements

6.7.1.2.1 Requirements for Communication

Current requirements are—

- Send and receive electronic mail between the High network and the Low network.
- E-mail must conform to standards used in the wider community.
- E-mail must allow users to send and receive attachments in both directions.

Anticipated requirements are—

- Enable users to use Chat as a means of communication between High and Low network users.
- Enable Internet telephony between High network users and Low network users as the technology becomes available.
- Enable video teleconferencing between High network users and Low network users.

6.7.1.2.2 Requirements for Releasability

Current requirements are—

- Enable authorized users on the High network to designate and push—e.g. FTP, e-mail, HTTP Post, etc.—data to the Low network that is releasable to users on the Low network.
- Enable authorized users on the Low network to access the released data using Web technology, FTP, database access techniques.
- Released data may be restricted to certain users, or it may be made publicly available.
- Released data may be text, video, images, audio, or executable software.

6.7.1.2.3 Requirements for Access

Current requirements are—

- Users on the High network must be able to access the vast information resources on the Low network.
- Access methods may be HTTP, FTP, Gopher, Wide Area Information Service (WAIS), SQL, or Web Push. With Web Push, as a result of a previous High-to Low-access request, information is pushed onto the High network from the Low network.

6.7.1.3 Attacks and Potential Countermeasures

The following section itemizes previously identified attacks that were explained in Chapter 3, System Security Methodology, of this document, and matches these attacks with potential countermeasures that may be included in solutions addressing the High-to-Low requirement category.

6.7.1.3.1 Passive Attacks

- **Traffic Analysis.** As of now, no technical countermeasure has been identified that is appropriate for inclusion in High-to-Low requirement category solutions.
- **Monitoring Plaintext.** The appropriate countermeasure to this attack is to deny access to the data by unauthorized users by encrypting the data or by using other data separation techniques that will restrict unauthorized release of data. (Note that utilizing encryption is possible only when both parties have access to the same algorithms and keys and the same capability to encrypt and decrypt the data properly.)
- **Decrypting Weakly Encrypted Traffic.** Countermeasures are to use adequate encryption algorithms and maintain sound key management.

6.7.1.3.2 Network-Based Attacks

- **Modification of Data in Transit.** The countermeasure to this attack is to use digital signatures or keyed hash integrity checks to detect unauthorized modification to the data in transit.
- **Insertion of Data.** There are many countermeasures to the malicious insertion of data. They include the use of timestamps and sequence numbers, along with cryptographic binding of data to a user identity, to prevent replay of previously transmitted legitimate data. Data separation or partitioning techniques, such as those used by firewalls and guards deny or restrict direct access and the ability to insert data by Low-side agents into the High-side network.
- **Insertion of Code.** Virus scanning by High-side users and enclave protection devices attempts to detect incoming viruses. Cryptographically authenticated access controls may be utilized to allow data only from authorized sources to enter the High network. Audit and intrusion detection techniques may detect breaches in established security policy and anomalies.
- **Defeating Login Mechanisms.** The most appropriate countermeasure for this is cryptographic authentication of session establishment requests.
- **Session Hijacking.** The countermeasure for this is continuous authentication through digital signatures affixed to packets, or at the application layer, or both.
- **Establishment of Unauthorized Network Connections.** There is no technical countermeasure for this. It is incumbent on the management and administration of the local network to prohibit unauthorized connections between High and Low networks, and to enforce that policy through nontechnical means. Various commercial tools may be utilized by system administrator personnel to detect such connections.
- **Masquerading as an Authorized User.** The appropriate countermeasure is to use cryptographic authentication in conjunction with timestamps or sequence numbers to prevent replay of authentication data. Another countermeasure to prevent stealing an authentic session is to cryptographically bind authentication data to the entire session/transaction.
- **Manipulation of Data on the High Side.** The appropriate countermeasure is to permit only authorized users to access the data on the High side using cryptographic authentication and data separation techniques.

6.7.1.3.3 Insider Attacks

- **Modification of Data or Modification of Security Mechanisms by Insiders.** The primary technical countermeasure is to implement auditing of all security relevant actions taken by users. Auditing must be supported by timely, diligent review and analysis of the audit logs generated. Other countermeasures to these attacks are nontechnical and

therefore not addressed by the High-to-Low requirement category solutions. Nontechnical countermeasures include personnel security and physical procedures.

- **Physical Theft of Data.** Again, the countermeasures to these attacks are nontechnical and therefore not addressed by the High-to-Low requirement category solutions. Appropriate nontechnical countermeasures include personnel security and physical security procedures, which inhibit actual removal of data, either in printed form or on storage media.
- **Covert Channels.** The countermeasure against a covert channel between the High and Low networks is a trusted guard function that examines network header fields and network messages for possible unauthorized information.

6.7.1.3.4 Development and Production/Distribution Attacks

- **Modification of Software During Development, Prior to Production.** The countermeasures for threats during this phase include use of strong development processes/criteria such as Trusted Software Development Methodology and subsequent evaluation of software by third-party testing using high assurance methods and criteria such as the Trusted Product Evaluation Program (TPEP) and Common Criteria testing.
- **Malicious Software Modification During Production and/or Distribution.** The countermeasures for threats during this phase include high assurance configuration control, cryptographic signatures over tested software products, use of tamper detection technologies during packaging, use of authorized couriers and approved carriers, and use of blind-buy techniques.

6.7.1.4 Technology Assessment

This section discusses general technology areas that can be used in system solutions to address the functional and related security requirements associated with the High-to-Low requirement category. Section 6.3.1.5, Requirement Cases, proposes various system-level solutions that build upon these general technology areas. The proposed security countermeasures included in each system solution result from our analysis of user target environments; functional requirements applicable to the *communications*, *releasability*, and *network access* requirements, and attacks and potential countermeasures as discussed in previous sections.

The framework divides the technology of protection between High and Low networks into three categories:

- 1) Data Separation Technologies
- 2) Authenticated Parties Technologies
- 3) Data Processing, Filtering, and Blocking Technologies.

This categorization allows us to make some high-level assessment of system assurance provided for groups of similar solutions, thereby ordering solutions in terms of security robustness. These three generic categories of potential solutions are explained in more detail in subsequent paragraphs of this section.

6.7.1.4.1 Data Separation Technologies

System solutions that would logically fit into this technology category would allow users who are located in High-side protected enclave environments to have access to both High network and Low network data, but prohibit pushing and pulling of data between these two networks. Typically, solutions in this category rely upon physical separation of data (from user interface to redundant distribution networks) in order to provide data segregation between High and Low applications.

In most cases High-side users are restricted from using sophisticated automated means that allow for the storage or manipulation of Low-side generated data on the High network. In addition, High-side users are also restricted from directly extracting Low data from the High network applications, or using a broad range of applications to move the extracted data to the Low network.

All of the proposed solutions that are included in this category do provide for the data transfer techniques previously described as *communications*, *releasability*, and *network access*, but do so only within networks of the same level.

For *communications* exchanges, typical solutions in this category allow access for High-side users to redundant network access points, which are individually connected to both networks, i.e., High network users have access to two network access points, one for the High network and one for the Low network. Users may have two processors with shared monitors and keyboards, or several users may be provided access to a shared Low network interface located in a centralized location. Likewise, for both *releasability* and *network access* exchanges, users on the High network side will interface to logically separated network interfaces.

The economics of solutions that fit into this category must be examined and a tradeoff analysis completed that compares the savings resulting from greatly simplified security mechanisms and reduced complexity of security management infrastructure and personnel support with the cost of redundant local networks and network management. The primary advantage of data separation solutions is that all of the solutions in this category provide the highest degree of system-level security, and may in fact be the only solutions that are acceptable for very high assurance networking requirements. These are very secure system topologies, providing the best protection from both passive and network attacks.

These solutions do not allow data to flow between the High network and the Low network. Hence, they are robust in preventing attack of the High network and leakage of High data to the Low network. The only true data separation technology is physical isolation of the network. Any connection between the two networks will create the potential for at least minimal leakage

via covert channels, as well as the operational risk of attacks from Low to High. Solutions here include—

- Isolated Networks.
- Secure Network Computers.
- Starlight Interactive Link.
- Compartmented Mode Workstations (CMW).

Each of these is discussed below.

Isolated Networks

This solution is simply to maintain two networks, one for High data and one for Low data. The two networks are never to be connected together. This would require redundant infrastructures, at additional cost. However, the cost can be justified in environments where users cannot tolerate the risk that the High data might be compromised or the High network attacked.

The number of workstations on each network is a function of the need within the organization to have individuals with access to both networks. Perhaps the Low network can be accessed via shared workstations if it is not necessary for all users to have access from their desktops.

The specific capabilities addressed by this solution are communication and network access. Automated releasability to the Low network of data created on the High network is not addressed by this technique. Regrading, and subsequent release to a co-located Low network computer, of information contained on the High network computer may be performed by overt human intervention, e.g., human review and retyping of data on the Low network computer or optical scanning. Communication and network access are addressed by allowing the user who has access to a terminal for each network to exchange electronic mail, participate in Chat sessions, and perform World Wide Web (WWW) browsing with other parties on either network by using the appropriate terminal.

While many customers wish to avoid using separate networks, this option bears consideration with the increased availability of low-cost personal computers (PC) and network computers. The cost of implementing and operating two separate networks might actually be less than implementing and managing sophisticated network security systems. Furthermore, the richness of the network access will be unimpaired by the security at the boundary of the High network.

Secure Network Computers

Research is being done on a secure network computer that will employ a cryptographic token to separate data on the network. The concept is that the network will be classified for Low data, while having servers connected that process High data. All High data on the network is encrypted to provide separation. The workstations on the network are all *single level at a time* with only volatile memory. They are network computers that accept a cryptographic token to encrypt and decrypt all communications over the network. Depending on the token placed in the network computer at any one time, it will be able to access either High servers or Low servers,

UNCLASSIFIED

Multi-Level Security
IATF Release 3.1—September 2002

but not both. When the token is changed, the volatile memory of the network computer is cleared. Since this is a research project, no commercial products are yet available. Hence, this is identified as a technology gap that is being addressed.

When secure network computers become available, they will allow communication and network access on High networks and Low networks using the same device. They will not allow automated regrading of data, so it would not be possible to forward an e-mail message from the Low network to recipients on the High network. Likewise, the secure network computer does not support automated releasing of Low data from the High network. To release Low data residing on the High network, users would be required to perform a human regrade procedure, using nonautomated methods such as retyping of the data or optical scanning.

Starlight Interactive Link

This is a technology that is being developed in Australia that allows a single monitor, mouse, and keyboard to have access to two different computers. One computer is connected to the High network, and one is connected to the Low network. The technology allows *single level at a time* access to the two networks from a single location. Data does not transfer between the two without human review. It is possible to cut-and-paste data from Low to High only (never High to Low) using the standard X Windows cut and paste capability. This can be done only with human intervention. There is no way to automate the regrading of data. It should be noted that the cut-and-paste Low-to-High capability introduces risk that the data pasted to the High network could contain malicious code.

The implementation employs a one-way fiber optic link with the Low computer. This prohibits data leakage from High to Low. Because of the fiber optic link, data can only flow away from the Low computer to the display; it can never flow from the display to the Low computer.

The Starlight Interactive Link supports communication and network access from a single location. It does not support automated releasability from the High network to the Low network.

Since the Starlight Interactive Link is not yet a commercial product, it is identified as a technology gap.

Compartmented Mode Workstations

Another solution in the data separation class is to use CMWs or higher assurance workstations, if available. These could be judiciously allocated to the users who need to access both the High network and the Low network. With this approach, each user is then able to access both the High network and the Low network.

The specific capabilities addressed by this solution are communication, network access, and releasability. Communication and network access are addressed by allowing the user who has access to a CMW, which is connected to each network, to exchange electronic mail, participate in Chat sessions, and perform WWW browsing with other parties on either network by using a window dedicated to the proper network. Releasability and communication between the High

network and the Low network are addressed by the CMW *cut-and-paste* and *downgrade capability*. This operation allows users to highlight information in a High window and use the cut or copy command to place it in a buffer for review. The resulting information is then downgraded, appropriately classification marked, and displayed to the user in a Low window for visual review and release.

Cut and paste between sensitivity levels is an action that requires the CMW to be configured with this privilege; it is not allowed by default. If the CMW is not configured with this privilege, complete logical data separation is achieved.

6.7.1.4.2 Authenticated Parties Technologies

System solutions that would logically fit within this category are solutions that mandate the use of cryptographic authentication mechanisms prior to allowing access. Examples of actions that could be governed by this technology are—

- Allowing High users to access servers on the Low network when the servers can be authenticated.
- Allowing High users to release data from the High network based on their authenticated identity.
- Allowing Low data to enter the High network when the Low data is cryptographically bound to an authorized individual through a digital signature.

Authenticated access is widely available and is supported by a large number of standards and protocols. It allows two parties that intend to exchange data to identify themselves to one another and positively authenticate their identities. Hence, they become mutual trusting parties. The data that flows between these trusting parties is at the level of the lower party. This paradigm is applicable to the previously discussed modes of data exchange: *communication*, *releasability*, and *network access*.

Authenticated access solutions typically address *communication* data exchanges by use of digital signatures for electronic mail messaging applications, e.g., Message Security Protocol (MSP) or Secure/Multipurpose Internet Mail Extension (S/MIME). Such solutions typically involve the concept of protected enclaves for the system-high users that are separated from the system-low network users by some sort of enclave boundary protection device such as a guard or firewall. In such a topology, Low network users might utilize digital signature technology to authenticate themselves to High network users. Also, the guard might incorporate access control list (ACL) mechanisms to make access decisions governing the set of users that are authorized to release information from the High network. Access control lists can also be used to restrict the set of Low network users that are authorized to push data up to the High network.

Likewise, authentication solutions are applicable to *releasability* data exchanges in that the releaser can digitally sign data to be released. Again, enclave boundary protection systems such as guards might utilize ACLs that would regulate who in the system-high network is authorized

UNCLASSIFIED

Multi-Level Security
IATF Release 3.1—September 2002

to release data from the High-side network. The enclave boundary protection system might also perform content review of the data submitted for release.

Lastly, authentication solutions are applicable to *network access* data exchanges typically through the use of commercial off-the-shelf (COTS) protocols such as Secure Sockets Layer (SSL), Secure HyperText Transfer Protocol (S-HTTP), SOCKS, Secure Electronic Transaction (SET), and Internet Protocol Security (IPSec) for Web access, database access, FTP access, etc.

It is logical to conclude that security is enhanced if parties that are mutually trusting create a closed virtual community. The downside of these types of solutions is that, in general, they mandate that both parties have compatible security mechanisms to strongly authenticate themselves to one another. Therefore, the implication is that the number of Low network resources that are accessible is greatly reduced to include only those that are “security enabled.” In the case of *network access* requirements, the requirement to be security enabled may greatly reduce the availability of access to public information resources.

It must also be noted that authentication solution topologies normally necessitate a very restrictive policy whereby activity is allowed only with other parties that are authenticated as part of the closed, and therefore trusted, community. Conversely, if the community is opened by a single party who interacts with another party outside of that community, then the entire community is potentially vulnerable to attack.

While authentication technologies are widely available, they have yet to become fully mature. For a discussion of hurdles that must be overcome, see Section 6.3.1.4, Technology Gaps.

Solutions using Authenticated Parties include the following:

- Authentication between clients and servers using SSL.
- Host-to-host authentication using IPSec with the Authentication header.
- Authentication at the application layer via digital signatures.

These are discussed below.

Authentication between Clients and Servers Using SSL

SSL[1] is becoming a popular security protocol for implementing privacy and authentication between communicating applications. It is a transport layer security protocol, enabling the encryption and authentication of arbitrary applications. The protocol prevents eavesdropping, tampering with information, and forging of information sent over the Internet.

The SSL protocol includes a lower level protocol (called the SSL Record Protocol) that encapsulates higher level security protocols. The SSL Handshake Protocol is one such encapsulated protocol. It allows communicating parties to authenticate one another and to establish cryptographic algorithms and keys at the start of a communication session.

Connections using SSL have three properties:

- The communication is private. The initial handshake uses public key cryptography to define a secret key. The secret key is then used with symmetric cryptography to encrypt all communications.
- Clients and servers can authenticate one another during the handshake using public key cryptography.
- The entire communication is protected against tampering or insertion of data. Each datagram has a Message Authentication Code that is a keyed hash value.

The SSL protocol can be used for network access between clients on the High side and servers on the Low side. This can give confidence that the server is trusted to some degree. A policy requiring that SSL be used for all network access between High and Low would effectively permit access only to servers on the Low side that have the ability to authenticate using SSL. However, such a policy might not be useful if there are some Low servers that have the ability to authenticate, but should not be included within the set of servers to which access is allowed. The goal should be, not just authentication. Rather, the goal should be but access control, with authentication used as a means to implement access control. This is accomplished by maintaining a list of Low servers that, once authenticated, can be accessed by High clients. That list is best maintained by an enclave boundary protection system, e.g., guards.

If an enclave boundary protection system is in use, SSL can be used between the enclave boundary and the Low server. If the SSL is between an enclave boundary protection system and the Low server, then guarding, filtering, and blocking technologies can also be applied to allow access to only those Low servers that are on an access control list. The enclave boundary protection system would keep a list of servers to which network access is allowed, and would enforce the policy that no network access is allowed to any other servers. SSL could also be used as a basis for communication via e-mail, Chat, Whiteboarding, or other protocols, since it is a transport layer protocol and is independent of the application. Since SSL also gives the capability to encrypt all application layer data, the communication between the enclave boundary and the Low server is private.

SSL can also be used between the client on the High network and the enclave boundary. This allows the enclave boundary protection system to maintain a list of High clients that are authorized to communicate with users on the Low network, to access information on the Low network, and to release information to the Low network.

Using SSL for end-to-end encryption and authentication from High clients to Low servers limits the effectiveness of an enclave boundary protection system. In this case, the enclave boundary protection system cannot see the application layer information being communicated between the client and the server. Therefore it can make access control decisions only on information in the transport layer and layers lower than the transport layer. Thus, a tradeoff must be made between end-to-end security and the access control capabilities of an enclave boundary protection system. However, the benefits of using an enclave boundary system to enforce access control can be argued to outweigh the loss of uninterrupted end-to-end encryption and authentication.

For High-to-Low, the optimal use of SSL is to have two SSL connections meeting at the enclave boundary protection system. One connection is between the High host and the enclave boundary; another is between the enclave boundary and the Low host. This allows the enclave boundary protection system to perform filtering, authentication, access control, and auditing of all traffic passing from High to Low. To perform this function, the enclave boundary system would use a proxy that effectively glues two separate SSL sessions together.

Host-to-Host Authentication Using IPSec With the Authentication Header

Like SSL, the IPSec security protocols allow encryption and authentication of all information above the network layer in the Transmission Control Protocol (TCP)/IP stack. Unlike SSL, IPSec resides at a lower layer in the communication stack, and has the capability to completely encapsulate IP packets, including the source and destination addresses. Where SSL can be described as a process-to-process security protocol, IPSec is sometimes referred to as a host-to-host security protocol.

In connections between High networks and Low networks, IPSec can be useful in authenticating the hosts at the communication endpoint, and in providing privacy of the data being transmitted. Since IPSec is at a lower layer in the communication stack than SSL, IPSec can help in prevention of spoofed IP addresses.

IPSec is of little use in High-to-Low connections without an enclave boundary protection system at the point where the High network is connected to the Low network. The enclave boundary protection system is needed to perform access control between High and Low. At the same time, the enclave boundary protection system is rendered useless if IPSec with encryption is used between the High host and the Low host, since the communications would be encrypted with a key private to those two endpoints. For High-to-Low, the best use of IPSec is between the Low host and the enclave boundary protection system, and also between the High host and the enclave boundary protection system. This allows the enclave boundary protection system to authenticate both endpoints of the communication, although it creates a complexity in key management for the enclave boundary protection system. Since most enclave boundary protection systems that are suitable for High-to-Low do not perform IPSec, this is considered a technology gap.

Authentication at the Application Layer via Digital Signatures

Current High-to-Low solutions for electronic mail have the capability for digital signatures to identify the originator of e-mail messages. These solutions also depend heavily on a mail guard for enclave boundary protection. Like SSL and IPSec, the enclave boundary protection system cannot perform the functions of inspecting the content of the message or verifying the digital signature if the message is encrypted. The currently available e-mail solutions allow the guard to decrypt a copy of outgoing messages in order to perform filtering on the contents of those messages.

Authentication at the application layer using digital signatures allows the enclave boundary protection system to determine the individual who is responsible for the traffic passing from High to Low, and then to make an access control decision to allow or disallow the traffic. Since the digital signature is based on public key cryptography, a public key infrastructure must be in place to enable this solution.

6.7.1.4.3 Processing, Filtering, and Blocking Technologies

Solutions that logically fit within this solution category utilize various processing, filtering, and data blocking techniques in an attempt to provide data sanitization or separation between High network data/users and Low network data/users. Data originating from the High network is assumed to be High data though it may be asserted to be Low data by a High network user. Automated processing and filtering techniques may be performed by enclave boundary protection devices such as a guard, and if such tests are successfully passed, the data is actually regraded by automated means. In the reverse direction, such solutions often incorporate data blocking techniques (typically in firewalls but also in guards) to regulate the transfer of data from Low network users to High network users. Use of certain protocols may be blocked and/or data may be processed or filtered in an attempt to eliminate or identify viruses and other malicious code transfers.

The technology categories of data separation and authenticated parties do not allow users to use automated means to transfer data between the High and the Low network. The only technology that allows automated data regrading and transfer is processing, filtering, and blocking. Hence, this technology is the linchpin of High-to-Low. Without processing, filtering, and blocking techniques, there are no automated mechanisms supporting the regrading of information from High networks to Low networks. Data separation and authenticated parties technologies are restricted to allowing information transfer between networks only by means of human intervention such as retyping or optical scanning.

It must be emphasized that data transfer between High and Low involves risk, and one must take steps to mitigate risk. If data separation via a technology described in any of the other solution categories is not possible, then processing, filtering, and blocking must be considered. It must, however, be recognized by implementing organizations that these techniques involve inexact attempts to filter High data from outgoing transmission through content checking against a pre-defined list of prohibited strings. It also involves scanning for and detecting virus-infected executables, and blocking executables. Since there are an almost infinite number of possible executables, and malicious ones can be detected only through prior knowledge of their existence, the problem of detecting “maliciousness” in an arbitrary executable is not computable. This is exacerbated by the fact that there are many executables that users wish to allow to cross the network boundary (e.g., Java applets, Active X controls, JavaScript, Word macros) and that they would therefore not wish to filter out or block. Only by performing a detailed risk management tradeoff analysis wherein operational needs are weighed against security concerns can these issues be resolved.

UNCLASSIFIED

Multi-Level Security
IATF Release 3.1—September 2002

Solutions using processing, filtering, and blocking employ some type of processing to allow information flow between the two networks but attempt to detect and block attacks and High data leakage. Solutions here include—

- I-Server for Communication, Network Access, and Releasability.
- Mail Guard.
- Low-to-High Replication.

Each of these is discussed below.

I-Server for Communication, Network Access, and Releasability

This solution uses a special purpose computer, dual-homed at the boundary between the High network and the Low network. The solution is identified as a technology gap due to the nonexistence of commercial products that have this capability. The technology needed to develop such products is well understood, however. The computer, called an *Intermediate Server*, is a remote host that users on the High network can log in to and execute browsers and Internet client software. The *I-server* is ideally a trusted computer with the ability to keep data of differing classifications separated. It also has the ability to protect itself against attack from the outside. Malicious code that might execute as part of Java applets or Active X controls would not be able to damage the I-server or the High network due to rigid design constraints.

The I-server is protected by a robust architecture that prevents tampering or modification of the operating system. This architecture also constrains the processes that are running any hostile executables to their own address space, and gives them no privileges to observe or modify files. The High network is protected by the remote location of the I-server, keeping potentially hostile code off of the High workstations and servers. Only the display of the information retrieved from the Low network is sent to the High network.

The specific capabilities addressed by this solution are communication, network access, and releasability. Communication is addressed by allowing the user on the High network to exchange electronic mail with users on the Low network, and to participate in Chat sessions with parties on the Low network. Network access is addressed by allowing users on the High network to perform WWW browsing via the I-server, and to access FTP servers on the Low network via the I-server. Releasability is addressed by allowing users on the High network to upload files to be released to the I-server, applying filters to determine that the information is indeed releasable, and then sending the released files to external servers.

The I-server architecture enables indirect accesses to the Low network. The I-server is a trusted computer that has MLS capability with high assurance. The I-server is connected both to the Low network and to the High network. Users on the High network log onto the I-server at the Low level. Browsers and other Internet clients, e.g., Simple Mail Transfer Protocol (SMTP), FTP, and Telnet, execute on the I-server, and all information retrieved from the Low network stays on the I-server at the Low level. That information can be viewed by the user on the High network who requested it. The viewing is done through a terminal emulation protocol between the I-server and the user workstation on the High network. Since the I-server is a trusted

computer that can protect itself from attack, the threat posed by malicious executables is greatly diminished.

The following are the steps a user would perform to browse the Low network from the High network through an I-server—

- Log in to the I-server at the Low level.
- Authenticate to the I-server via password or other authentication mechanism.
- Run the Web client available on the I-server.
- Type in the Universal Resource Locators (URL)/IP address desired or select from your personal set of bookmarks/favorites or select entries from an address book.
- See the responses through terminal emulation at the user's workstation and, if desired, save them on the I-server for future reference. Files saved on the I-server will be saved at the Low level.

Note that the steps above do not include a means for a user to pull data retrieved from the Low network to his or her workstation on the High network. Since pulling of data from the Low network could create an avenue for attack, the I-server prohibits this pulling. To allow this pulling of information through the I-server would bring along the inherent risks of pulling data from untrusted sources on the Low network. If pulling of data is a user requirement, then procedures and policies must be in place to mitigate risk of pulling hostile executables. One such policy would be to allow pulling of only ASCII text and to prohibit use of decoding software (such as UUdecode) on that text.

The main security weakness of the I-server is the potential for leakage of data from the workstation on the High network that is untrusted, to the Low process executing on behalf of the user on the I-server. This could occur through a covert channel in the terminal emulation protocol and be driven by a Trojan horse on the user's workstation. It would also require collusion at the receiving end (the Low process on the I-server). This vulnerability would be difficult to exploit, and therefore is considered lower risk than would be present if the HTTP protocol were being sent end-to-end between the workstation on the High network and the server on the Low network.

Mail Guard

This solution is readily available with both commercial and government-developed products. The guard is deployed at the boundary of the High network and the Low network. The guard performs filtering and control of mail messages passing High to Low and Low to High. The filtering is based on the headers of the mail messages, e.g., sender, recipient, presence of signature; as well as the contents of the mail message, e.g., encryption of contents, presence of prohibited words or phrases. At this time the solution only addresses communication via electronic mail. Guards are typically used in conjunction with "authenticated parties"

technology. This adds some strength to the relative weakness of content filtering employed by a guard.

Current mail guards are very flexible, allowing implementation of a wide variety of message acceptance and message release policies. It is possible to configure mail guards to be very liberal in these policies. Policy makers must pay strict attention to policy decisions to assure that policies are not so liberal as to negate the usefulness of the mail guard.

Low-to-High Replication

Low-to-High replication allows users on the High network to receive data that originates on the Low network, without having to explicitly request that the data be sent from the Low servers. Replication can be used for network access, pushing data from the Low network to the High network. It cannot be used for releasability or for communication, because its primary security property is the prevention of data flows from High to Low.

Replication can give the High network any application that passes messages from one host to another. Examples are database replication, FTP, electronic mail, and Web Push protocols.

To prevent data leakage from High to Low, replication does not allow a direct back channel to send message acknowledgements from the High network to the Low network. To do so would allow quite a large covert channel. The replication acts as an intermediary, sending acknowledgements to the Low sender, and receiving acknowledgements from the High recipient. The Low sender cannot determine with precision the timing of the acknowledgements sent from the High side. Hence, the bandwidth of the back channel is reduced by the intermediate buffer within the replication process. This disconnects any direct communication from High to Low.

Replication does not mitigate the potential risk that data replicated into the High network might be hostile executable code. Mitigation of this risk would require that data be replicated first in a network guard that inspects the data for potentially hostile code, making sure the data passes this inspection before being forwarded into the High network.

6.7.1.5 Requirements Cases

This section is intended to address the connection of High-to-Low networks for purposes of communication, network access, and releasability. These are general, functional requirements that have been articulated by various customers. Presently, only the Secret-to-Unclassified network connection scenario has been analyzed in detail. There are other connection scenarios where similar requirements appear to be appropriate. The additional scenarios we are aware of are Top Secret-to-Compartmented-Top Secret, Top Secret-to-Secret, and Secret U.S.-to-Secret (Allied). These other scenarios are under analysis, and their requirements will be presented in future versions of the framework if they are found to be different from the Secret to Unclassified case.

Case 1: Secret-to-Unclassified

Users on the Secret network have a need to connect to the Unclassified network for the purposes of communication, network access, and releasability. For communication, the needed application is electronic mail. Access to the Unclassified network is needed also via Web protocols, using commercially available Web browsers. Finally, Secret users sometimes create large files that are in reality Unclassified. In some cases users have a need to release these Unclassified files to the Unclassified network.

Electronic mail is currently enabled between Secret and Unclassified in many instances through a mail guard, which is sometimes coupled with a COTS firewall. In the Defense Message System, e-mail will be enabled between Secret and Unclassified using a mail guard. The immediate need is to develop the additional capability to use Web-based protocols (i.e., HTTP) to access Web servers on the Unclassified network. Another immediate need is to develop the capability to release large files from Secret to Unclassified (probably using FTP). Current guards do not have the capability to allow network access and releasability. The environmental requirements for the Secret-to-Unclassified connection include—

- Secret users must be able to use COTS software, e.g., browsers and e-mail clients, in accessing information, communicating with users, and releasing information on the Unclassified network.
- Secret users must be able to use the installed base of operating systems, whether they are Windows or Unix.

The new capabilities for access to the Unclassified network and for releasability must coexist with existing capabilities to send and receive e-mail with users on the Unclassified network.

Case 2: Secret U.S.-to-Secret Allied

This section will be provided in a later release of the framework.

Case 3: Top Secret-to-Secret

This section will be provided in a later release of the framework.

6.7.1.6 Framework Guidance

In this section, guidance is provided on the solutions that can be implemented now to perform High-to-Low network connections for the purposes of communication, network access, and releasability.

Case 1: Secret-to-Unclassified

Requirement Considerations

In order to place the framework guidance in a proper perspective, this section delineates the specific security requirements being addressed and discusses issues associated with providing solutions for them.

Communication

- Secret users must be able to send and receive Unclassified electronic mail with communication partners on the Unclassified network.
This requirement opens the possibility of leakage from Secret to Unclassified and also the possibility of attacks being encoded in messages received from the Unclassified network.
- Secret users must get notice of electronic mail that was sent to users on the Unclassified network but could not be delivered, i.e., bounced messages.
- It must be possible to send and receive electronic mail with attachments.
Attachments greatly increase the risk of leakage Secret to Unclassified, and the risk of attack to the Secret network, because it is generally very difficult to determine whether an attachment contains an executable.
- Secret users must be able to participate in live Chat sessions with users on the Unclassified network.
- Secret users must be able to use collaborative technologies such as whiteboarding and video conferencing with users on the Unclassified network.
- Internet Telephony between Secret network users and Unclassified network users must be enabled as the technology becomes available.

Releasability

- Enable Secret users on the Secret network to designate and push, e.g. FTP, e-mail, HTTP Post, etc., data to the Unclassified network that is releasable to users on the Unclassified network.
- Enable Unclassified users on the Unclassified network to access the released Unclassified data using Web technology and FTP database access techniques.
- Access to Unclassified data released from a Secret network may be restricted to specific Unclassified users, or groups of users, or may be made publicly available.
- The format of Unclassified data released from a Secret network may be text, video, images, audio, or executable software.

Network Access

- Secret users on the Secret network must be able to access the vast information resources on the Unclassified network using HTTP, FTP, Gopher, WAIS, SQL, or Web Push.
- When using Web Push as a result of a previous Secret user request to the Unclassified network, Unclassified information is pushed into the Secret network from the Unclassified network.
The implications of these requirements are the dangers in retrieving data from servers. Data could harbor malicious executables. Also, information normally transmitted using the HTTP protocol might give the Unclassified servers a passive intelligence gathering capability.

Secret users must be able to use search engines that reside on the Unclassified network. This effectively means keywords must be sent from the Secret user to the Unclassified search engine.

The main implication of this is that data must be transmitted from Secret to Unclassified via the HTTP Post method. This method allows arbitrary data to be posted to an HTTP server. Measures must be taken to assure that Secret data is not being posted to an Unclassified server.

- The Secret client needs to receive data of arbitrary type and format.
This requirement increases the possibility of attack on the Secret client. The arbitrary format of the data makes it virtually impossible to detect any undesired executable.
- Error conditions sent by Unclassified servers must be received by Secret clients.
- The WWW interface must generate error and warning messages when it is unable to fulfill the request of a Secret client, and the Secret client must receive these messages.

Recommended Security Policies

The security policy for the Secret-to-Unclassified connection must include statements requiring countermeasures for attacks described previously.

For passive attacks the security policy must address:

- **Traffic Analysis.** The guard shall include measures to make all network access requests coming from the Secret network anonymous.
- **Monitoring Plaintext.** Encryption shall be used for all electronic mail passed out of the Secret network. Encryption shall be used between the high workstations and all external hosts receiving data for releasability. Encryption shall be used with all Unclassified hosts that support it (for example, via SSL, IPSec). The minimum size of the encryption key shall be 80 bits.

For network-based attacks the security policy must address the following attacks:

UNCLASSIFIED

Multi-Level Security
IATF Release 3.1—September 2002

- **Modification or Insertion of Data in Transit.** All data in transit shall have either a digital signature or keyed hash algorithms applied. These cryptographic algorithms must be deployed in conjunction with timestamps or sequence numbers to prevent replay of valid data.
- **Insertion of Hostile Executables.** Scanning for viruses and blocking applets and other executables must be performed for all data being transmitted into the Secret network.
- **Defeating Authentication Mechanisms.** Strong cryptographic authentication must be used across the enclave boundary. No Unclassified users shall access the Secret network unless it is done in accordance with the framework guidance for remote access.
- **Session Hijacking.** Continuous authentication along with timestamps or sequence numbers shall be used to prevent session hijacking.
- **Establishment of Unauthorized Network Connections.** Policy shall prohibit connections between the Secret and the Unclassified network other than those providing adequate security countermeasures.
- **Masquerading.** E-mail sender authentication and authorization to release data or to access the Unclassified network shall be handled using digital signature.
- **Manipulation of Data on the Secret Network.** This shall be handled through blocking of executables, and authentication of any users on the Unclassified network that access the Secret network remotely.

The security policy to prevent insider attacks involves procedural, physical, and personnel security. The primary technical countermeasure is to implement audit and intrusion detection systems on the Secret network.

For development, production, and distribution attacks, the vendors of all commercial security products shall use approved configuration control techniques and approved distribution methods.

Recommended Topology

The IATF recommends the topology shown in Figure 6.7-2 for the near-term Secret-to-Unclassified solution.

The figure shows that the only service offered between Secret and Unclassified is e-mail at this time. The guard enforces the policy for release of messages from the Secret user side. This policy can include content filtering, crypto-invocation check, release authority check, message format check, valid receiver check, message nonrepudiation signature, sequence signature, and allow/disallow attachments. The policy for admittance of messages to the Secret network can include all of these elements except crypto-invocation check. The guard will be able to decrypt copies of encrypted messages being released. However, if messages being admitted to the Secret network are encrypted, the guard will not be able to decrypt them. Consequently, the guard will not be able to filter incoming messages that are encrypted.

With minimal work, current mail guards can be modified to allow for releasability for Secret-to-Unclassified networks. It will take considerably more work to enable network access between Secret and Unclassified networks with adequate risk mitigation, because the risks of network access are quite high. The Technology Gaps section outlines a migration path to allow near term Secret-to-Unclassified capability for releasability and midterm capability for network access.

For the near term it is obvious that the guard will remain the linchpin of Secret-to-Unclassified connectivity. Many risks exist that guards will never be able to mitigate. The long-term architectural goals should be to minimize the number of Secret-to-Unclassified connections while working to migrate toward MLS on the desktop workstation and within the servers.

The optimal solution to minimize risk is to move away from Secret-to-Unclassified and move toward MLS. MLS could be implemented on the desktop using CMWs or the Starlight Interactive Link technologies. There are several medium assurance (B2-B3) platforms on the market that are now being used as guard platforms. These could be converted to use as server platforms. Data could be separated on the network cryptographically. The technology exists for MLS; the business case has been the problem. The MLS systems that have been developed by industry have met with a lukewarm reception by government customers. Only if the Government is serious about using MLS will MLS become available.

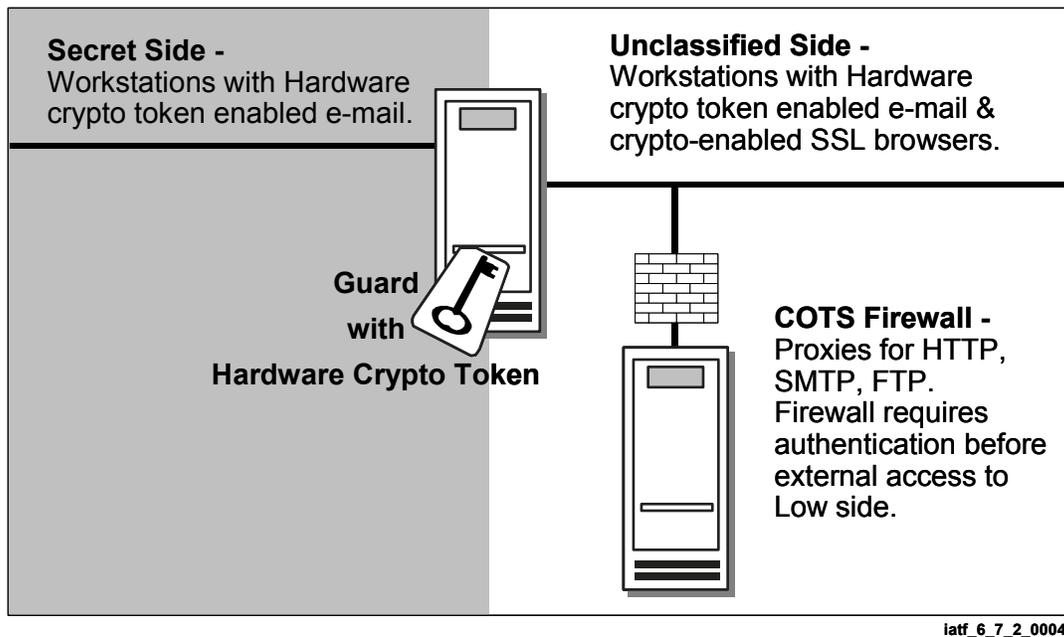


Figure 6.7-2. Recommended Topology

Technology Gaps

This section addresses the near-term technology advances that should be addressed to allow Secret-to-Unclassified releasability, then the midterm advances for Secret-to-Unclassified network access.

UNCLASSIFIED

Multi-Level Security
IATF Release 3.1—September 2002

- a) **Technology Gaps for Communication.** The technology to allow Secret-to-Unclassified communication via electronic mail is readily available. However, the technology to allow Chat, whiteboarding, Internet telephony, and video conferencing across the network boundary is not yet available.
- b) **Technology Gaps for Releasability.** All of the capabilities needed to support releasability are currently technology gaps. However, it is felt that Secret-to-Unclassified releasability can be accomplished within 2 years using the present solution topology shown in Figure 6.7-2. The goal is to allow users on the Secret side to submit files to the guard for downgrading. Then those files should be stored on a releasability server on the Unclassified side, making them available to Unclassified side users. They could also be made available to users outside the firewall, with the firewall and the releasability server performing authentication and controlling dissemination.

This should be accomplished by developing a releasability policy for the guard and then applying the policy to files being mailed to the releasability server. The releasability policy would likely be different from the message release policy applied to regular e-mail. The guard would recognize e-mail destined for the releasability server and would apply the releasability policy. The releasability policy will be more restrictive than the message release policy in the following ways.

- Only a very small set of users on the Secret side shall be allowed to release files to the releasability server.
- The guard shall maintain a list of this set of users and check the list upon each submission of a file to be released.
- All files submitted for release require signatures by two of the authorized individuals; one is a nonrepudiation signature; the other is a sequence signature.
- Only files with specific formats of plain text or HTML shall be releasable.
- Strict audit logs shall be kept on the guard of all files sent to the releasability server.
- Released files shall be scanned for content.

The releasability server should be a COTS product that receives the files and stores them for future publication. Publication occurs when an authorized user on the releasability server unwraps the files from their signed MSP wrappers, and places them in a directory that is accessible to other users. The authorized user of the releasability server must set the appropriate permission on the published files to allow the intended users to access them.

- c) **Technology Gaps for Network Access.** There is considerably more work to be done for network access. A completely new set of filters and proxies must be developed for the guard to recognize HTTP, FTP, Gopher, WAIS, SQL, and Web Push protocols and to apply appropriate policies to these. Work is needed to develop these policies and vet

them to gain confidence that they adequately mitigate risk for network access. Elements of such a policy must include but not be limited to the following.

- HTTP Post is not allowed Secret-to-Unclassified.
- Certain fields within the HTTP protocol that identify the user making the request and the version of the browser being used must be set to arbitrary values, effectively making the Secret user anonymous.
- Executables must be blocked from entering the Secret network as Java applets or Active X controls.
- The guard shall maintain a list of URL to which access is authorized, and enforce the policy that these URLs are the only ones accessible. The guard shall perform stateful filtering of HTTP.
- The guard shall prohibit Secret users from using the FTP PUT command.
- The guard shall maintain a list of users on the Secret network that are allowed to perform network access and network access attempts using SSL.

Case 2: Secret U.S.-to-Secret Allied

This section will be provided in a later release of the framework.

Case 3: Top Secret-to-Secret

This section will be provided in a later release of the framework.

6.7.2 MLS Workstation

This section will be provided in a later release of the framework.

6.7.3 MLS Servers

This section will be provided in a later release of the framework.

6.7.4 MLS Network Components

This section will be provided in a later release of the framework.

UNCLASSIFIED

Multi-Level Security
IATF Release 3.1—September 2002

References

1. Reference: SSL 3.0 Specification, Netscape Communications.
<http://home.netscape.com/eng/ssl3/index.html>.
2. Myong H. Kang, Ira S. Moskowitz, Daniel C. Lee. A Network Pump. Proceedings of the 1995 IEEE Symposium on Security and Privacy, pp 144-154. Oakland, CA.
3. Myong H. Kang, Ira S. Moskowitz, Bruce E. Montrose, James J. Parsonese. A Case Study of Two NRL Pump Prototypes. Proceedings of the 1996 ACSAC Conference. San Diego, CA.
4. Myong H. Kang, Judith N. Froscher, Ira S. Moskowitz. An Architecture for Multilevel Secure Interoperability. Proceedings of the 1997 ACSAC Conference. San Diego, CA.