National Institute of Standards and Technology
Technology Administration
U.S. Department of Commerce

An Introduction to Computer Security: The NIST Handbook

Special Publication 800-12

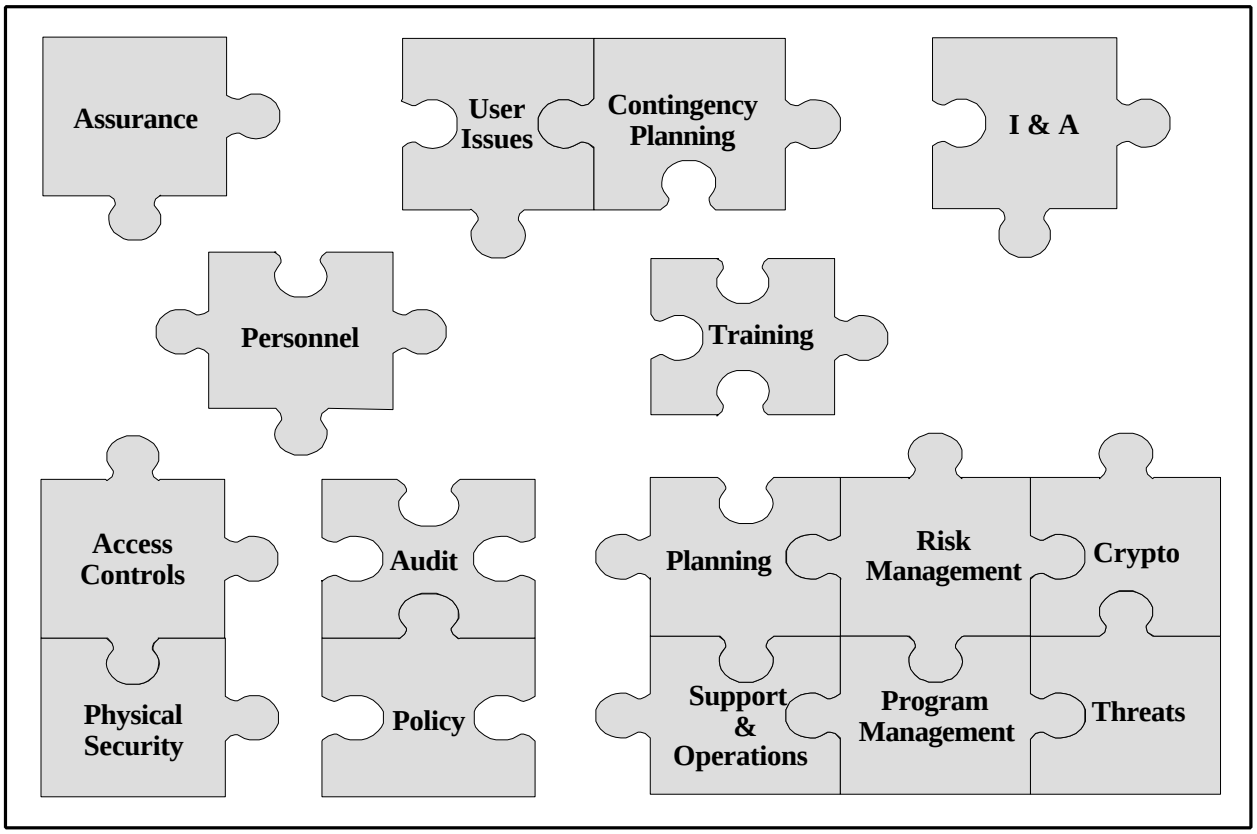


Table of Contents

I. INTRODUCTION AND OVERVIEW

Chapter 1

INTRODUCTION

Chapter 2

ELEMENTS OF COMPUTER SECURITY

Chapter 3

ROLES AND RESPONSIBILITIES

Chapter 4

COMMON THREATS: A BRIEF OVERVIEW

II. MANAGEMENT CONTROLS

Chapter 5

COMPUTER SECURITY POLICY

Chapter 6

COMPUTER SECURITY PROGRAM MANAGEMENT

Chapter 7

COMPUTER SECURITY RISK MANAGEMENT

Chapter 8

SECURITY AND PLANNING IN THE COMPUTER SYSTEM LIFE CYCLE

Chapter 9

ASSURANCE

III. OPERATIONAL CONTROLS

Chapter 10

PERSONNEL/USER ISSUES

Chapter 11

PREPARING FOR CONTINGENCIES AND DISASTERS

Chapter 12

COMPUTER SECURITY INCIDENT HANDLING

Chapter 13

AWARENESS, TRAINING, AND EDUCATION

Chapter 14

SECURITY CONSIDERATIONS IN COMPUTER SUPPORT AND OPERATIONS

Chapter 15

PHYSICAL AND ENVIRONMENTAL SECURITY

IV. TECHNICAL CONTROLS

Chapter 16

IDENTIFICATION AND AUTHENTICATION

Chapter 17

LOGICAL ACCESS CONTROL

Chapter 18

AUDIT TRAILS

Chapter 19

CRYPTOGRAPHY

V. EXAMPLE

Chapter 20

ASSESSING AND MITIGATING THE RISKS TO A HYPOTHETICAL COMPUTER SYSTEM

Acknowledgments

NIST would like to thank the many people who assisted with the development of this handbook. For their initial recommendation that NIST produce a handbook, we thank the members of the Computer System Security and Privacy Advisory Board, in particular, Robert Courtney, Jr. NIST management officials who supported this effort include: James Burrows, F. Lynn McNulty, Stuart Katzke, Irene Gilbert, and Dennis Steinauer.

In addition, special thanks is due those contractors who helped craft the handbook, prepare drafts, teach classes, and review material:

> Daniel F. Sterne of Trusted Information Systems (TIS, Glenwood, Maryland) served as Project Manager for Trusted Information Systems on this project. In addition, many TIS employees contributed to the handbook, including: David M. Balenson, Martha A. Branstad, Lisa M. Jaworski, Theodore M.P. Lee, Charles P. Pfleeger, Sharon P. Osuna, Diann K. Vechery, Kenneth M. Walker, and Thomas J. Winkler-Parenty.

Additional drafters of handbook chapters include:

> Lawrence Bassham III (NIST), Robert V. Jacobson, International Security Technology, Inc. (New York, NY) and John Wack (NIST).

Significant assistance was also received from:

> Lisa Carnahan (NIST), James Dray (NIST), Donna Dodson (NIST), the Department of Energy, Irene Gilbert (NIST), Elizabeth Greer (NIST), Lawrence Keys (NIST), Elizabeth Lennon (NIST), Joan O'Callaghan (Bethesda, Maryland), Dennis Steinauer (NIST), Kibbie Streetman (Oak Ridge National Laboratory), and the Tennessee Valley Authority.

Moreover, thanks is extended to the reviewers of draft chapters. While many people assisted, the following two individuals were especially tireless:

> Robert Courtney, Jr. (RCI) and Steve Lipner (MITRE and TIS).

Other important contributions and comments were received from:

> Members of the Computer System Security and Privacy Advisory Board, and the Steering Committee of the Federal Computer Security Program Managers' Forum.

Finally, although space does not allow specific acknowledgement of all the individuals who contributed to this effort, their assistance was critical to the preparation of this document.

Disclaimer: Note that references to specific products or brands is for explanatory purposes only; no endorsement, explicit or implicit, is intended or implied.

I. INTRODUCTION AND OVERVIEW

Chapter 1

INTRODUCTION

1.1 Purpose

This handbook provides assistance in securing computer-based resources (including hardware, software, and information) by explaining important concepts, cost considerations, and interrelationships of security controls. It illustrates the benefits of security controls, the major techniques or approaches for each control, and important related considerations.[1]

The handbook provides a broad overview of computer security to help readers understand their computer security needs and develop a sound approach to the selection of appropriate security controls. It does not describe detailed steps necessary to implement a computer security program, provide detailed implementation procedures for security controls, or give guidance for auditing the security of specific systems. General references are provided at the end of this chapter, and references of "how-to" books and articles are provided at the end of each chapter in Parts II, III and IV.

The purpose of this handbook is not to specify requirements but, rather, to discuss the benefits of various computer security controls and situations in which their application may be appropriate. Some requirements for federal systems[2] are noted in the text. This document provides advice and guidance; no penalties are stipulated.

1.2 Intended Audience

The handbook was written primarily for those who have computer security responsibilities and need assistance understanding basic concepts and techniques. Within the federal government,[3] this includes those who have computer security responsibilities for *sensitive* systems.

---

[1] It is recognized that the computer security field continues to evolve. To address changes and new issues, NIST's Computer Systems Laboratory publishes the *CSL Bulletin* series. Those bulletins which deal with security issues can be thought of as supplements to this publication.

[2] Note that these requirements do not arise from this handbook, but from other sources, such as the Computer Security Act of 1987.

[3] In the Computer Security Act of 1987, Congress assigned responsibility to NIST for the preparation of standards and guidelines for the security of sensitive *federal* systems, excluding classified and "Warner Amendment" systems (unclassified intelligence-related), as specified in 10 USC 2315 and 44 USC 3502(2).

For the most part, the concepts presented in the handbook are also applicable to the private sector.[4] While there are differences between federal and private-sector computing, especially in terms of priorities and legal constraints, the underlying principles of computer security and the available safeguards – managerial, operational, and technical – are the same. The handbook is therefore useful to anyone who needs to learn the basics of computer security or wants a broad overview of the subject. However, it is probably too detailed to be employed as a user awareness guide, and is not intended to be used as an audit guide.

Definition of Sensitive Information

Many people think that sensitive information only requires protection from unauthorized disclosure. However, the Computer Security Act provides a much broader definition of the term "sensitive" information:

any information, the loss, misuse, or unauthorized access to or modification of which could adversely affect the national interest or the conduct of federal programs, or the privacy to which individuals are entitled under section 552a of title 5, United States Code (the Privacy Act), but which has not been specifically authorized under criteria established by an Executive Order or an Act of Congress to be kept secret in the interest of national defense or foreign policy.

The above definition can be contrasted with the long-standing confidentiality-based information classification system for national security information (i.e., CONFIDENTIAL, SECRET, and TOP SECRET). This system is based only upon the need to protect classified information from unauthorized disclosure; the U.S. Government does not have a similar system for unclassified information. No governmentwide schemes (for either classified or unclassified information) exist which are based on the need to protect the integrity or availability of information.

1.3 Organization

The first section of the handbook contains background and overview material, briefly discusses of threats, and explains the roles and responsibilities of individuals and organizations involved in computer security. It explains the executive principles of computer security that are used throughout the handbook. For example, one important principle that is repeatedly stressed is that only security measures that are cost-effective should be implemented. A familiarity with the principles is fundamental to understanding the handbook's philosophical approach to the issue of security.

The next three major sections deal with security controls: Management Controls[5] (II), Operational Controls (III), and Technical Controls (IV). Most controls cross the boundaries between management, operational, and technical. Each chapter in the three sections provides a basic explanation of the control; approaches to implementing the control, some cost

[4] As necessary, issues that are specific to the federal environment are noted as such.

[5] The term *management controls* is used in a broad sense and encompasses areas that do not fit neatly into operational or technical controls.

considerations in selecting, implementing, and using the control; and selected interdependencies that may exist with other controls. Each chapter in this portion of the handbook also provides references that may be useful in actual implementation.

- The *Management Controls* section addresses security topics that can be characterized as managerial. They are techniques and concerns that are normally addressed by management in the organization's computer security program. In general, they focus on the management of the computer security program and the management of risk within the organization.

- The *Operational Controls* section addresses security controls that focus on controls that are, broadly speaking, implemented and executed by people (as opposed to systems). These controls are put in place to improve the security of a particular system (or group of systems). They often require technical or specialized expertise – and often rely upon management activities as well as technical controls.

- The *Technical Controls* section focuses on security controls that the computer system executes. These controls are dependent upon the proper functioning of the system for their effectiveness. The implementation of technical controls, however, always requires significant operational considerations – and should be consistent with the management of security within the organization.

Finally, an example is presented to aid the reader in correlating some of the major topics discussed in the handbook. It describes a hypothetical system and discusses some of the controls that have been implemented to protect it. This section helps the reader better understand the decisions that must be made in securing a system, and illustrates the interrelationships among controls.

1.4 Important Terminology

To understand the rest of the handbook, the reader must be familiar with the following key terms and definitions as used in this handbook. In the handbook, the terms *computers* and *computer systems* are used to refer to the entire spectrum of information technology, including application and support systems. Other key terms include:

Computer Security: The protection afforded to an automated information system in order to attain the applicable objectives of preserving the integrity, availability and confidentiality of information system resources (includes hardware, software, firmware, information/data, and telecommunications).

Integrity: In lay usage, information has integrity when it is timely, accurate, complete, and consistent. However, computers are unable to provide or protect all of these qualities.

Location of Selected Security Topics

Because this handbook is structured to focus on computer security controls, there may be several security topics that the reader may have trouble locating. For example, no separate section is devoted to mainframe or personal computer security, since the controls discussed in the handbook can be applied (albeit in different ways) to various processing platforms and systems. The following may help the reader locate areas of interest not readily found in the table of contents:

Topic	**Chapter**
Accreditation	8. Life Cycle 9. Assurance
Firewalls	17. Logical Access Controls
Security Plans	8. Life Cycle
Trusted Systems	9. Assurance Security features, including those incorporated into trusted systems, are discussed throughout.
Viruses & Other Malicious Code	9. Assurance (Operational Assurance section) 12. Incident Handling
Network Security	Network security uses the same basic set of controls as mainframe security or PC security. In many of the handbook chapters, considerations for using the control is a networked environment are addressed, as appropriate. For example, secure gateways are discussed as a part of Access Control; transmitting authentication data over insecure networks is discussed in the Identification and Authentication chapter; and the Contingency Planning chapter talks about data communications contracts. For the same reason, there is not a separate chapter for PC, LAN, minicomputer, or mainframe security.

Therefore, in the computer security field, integrity is often discussed more narrowly as having two facets: *data integrity* and *system integrity*. "Data integrity is a requirement that information and programs are changed only in a specified and authorized manner."[6] System integrity is a requirement that a system "performs its intended function in an unimpaired manner, free from

[6] National Research Council, *Computers at Risk*, (Washington, DC: National Academy Press, 1991), p. 54.

deliberate or inadvertent unauthorized manipulation of the system."[7] The definition of *integrity* has been, and continues to be, the subject of much debate among computer security experts.

Availability: A "requirement intended to assure that systems work promptly and service is not denied to authorized users."[8]

Confidentiality: A requirement that private or confidential information not be disclosed to unauthorized individuals.

1.5 Legal Foundation for Federal Computer Security Programs

The executive principles discussed in the next chapter explain the need for computer security. In addition, within the federal government, a number of laws and regulations mandate that agencies protect their computers, the information they process, and related technology resources (e.g., telecommunications).[9] The most important are listed below.

- The *Computer Security Act of 1987* requires agencies to identify sensitive systems, conduct computer security training, and develop computer security plans.

- The *Federal Information Resources Management Regulation* (FIRMR) is the primary regulation for the use, management, and acquisition of computer resources in the federal government.

- *OMB Circular A-130* (specifically Appendix III) requires that federal agencies establish security programs containing specified elements.

Note that many more specific requirements, many of which are agency specific, also exist.

Federal managers are responsible for familiarity and compliance with applicable legal requirements. However, laws and regulations do not normally provide detailed instructions for protecting computer-related assets. Instead, they specify requirements – such as restricting the availability of personal data to authorized users. This handbook aids the reader in developing an effective, overall security approach and in selecting cost-effective controls to meet such requirements.

---

[7] National Computer Security Center, Pub. NCSC-TG-004-88.

[8] *Computers at Risk*, p. 54.

[9] Although not listed, readers should be aware that laws also exist that may affect nongovernment organizations.

References

Auerbach Publishers (a division of Warren Gorham & Lamont). *Data Security Management*. Boston, MA. 1995.

British Standards Institute. *A Code of Practice for Information Security Management*, 1993.

Caelli, William, Dennis Longley, and Michael Shain. *Information Security Handbook*. New York, NY: Stockton Press, 1991.

Fites, P., and M. Kratz. *Information Systems Security: A Practitioner's Reference*. New York, NY: Van Nostrand Reinhold, 1993.

Garfinkel, S., and G. Spafford. *Practical UNIX Security*. Sebastopol, CA: O'Riley & Associates, Inc., 1991.

Institute of Internal Auditors Research Foundation. *System Auditability and Control Report*. Altamonte Springs, FL: The Institute of Internal Auditors, 1991.

National Research Council. *Computers at Risk: Safe Computing in the Information Age*. Washington, DC: National Academy Press, 1991.

Pfleeger, Charles P. *Security in Computing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

Russell, Deborah, and G.T. Gangemi, Sr. *Computer Security Basics*. Sebastopol, CA: O'Reilly & Associates, Inc., 1991.

Ruthberg, Z., and Tipton, H., eds. *Handbook of Information Security Management*. Boston, MA: Auerbach Press, 1993.

Chapter 2

ELEMENTS OF COMPUTER SECURITY

This handbook's general approach to computer security is based on eight major elements:

1. Computer security should support the mission of the organization.

2. Computer security is an integral element of sound management.

3. Computer security should be cost-effective.

4. Computer security responsibilities and accountability should be made explicit.

5. System owners have computer security responsibilities outside their own organizations.

6. Computer security requires a comprehensive and integrated approach.

7. Computer security should be periodically reassessed.

8. Computer security is constrained by societal factors.

Familiarity with these elements will aid the reader in better understanding how the security controls (discussed in later sections) support the overall computer security program goals.

2.1 Computer Security Supports the Mission of the Organization.

The purpose of computer security is to protect an organization's valuable resources, such as information, hardware, and software. Through the selection and application of appropriate safeguards, security helps the organization's mission by protecting its physical and financial resources, reputation, legal position, employees, and other tangible and intangible assets. Unfortunately, security is sometimes viewed as thwarting the mission of the organization by imposing poorly selected, bothersome rules and procedures on users, managers, and systems. On the contrary, well-chosen security rules and procedures do not exist for their own sake – they are put in place to protect important assets and thereby support the overall organizational mission.

Security, therefore, is a means to an end and not an end in itself. For example, in a private- sector business, having good security is usually secondary to the need to make a profit. Security, then, *ought to* increase the firm's ability to make a profit. In a public-sector agency, security is usually secondary to the agency's service provided to citizens. Security, then, *ought to* help improve the service provided to the citizen.

To act on this, managers need to understand both their organizational mission and how each information system supports that mission. After a system's role has been defined, the security requirements implicit in that role can be defined. Security can then be explicitly stated in terms of the organization's mission.

The roles and functions of a system may not be constrained to a single organization. In an interorganizational system, each organization benefits from securing the system. For example, for electronic commerce to be successful, each of the participants requires security controls to protect their resources. However, good security on the buyer's system also benefits the seller; the buyer's system is less likely to be used for fraud or to be unavailable or otherwise negatively affect the seller. (The reverse is also true.)

2.2 Computer Security is an Integral Element of Sound Management.

Information and computer systems are often critical assets that support the mission of an organization. Protecting them can be as critical as protecting other organizational resources, such as money, physical assets, or employees.

However, including security considerations in the management of information and computers does not completely eliminate the possibility that these assets will be harmed. Ultimately,

This chapter draws upon the OECD's *Guidelines for the Security of Information Systems*, which was endorsed by the United States. It provides for:

Accountability - The responsibilities and accountability of owners, providers and users of information systems and other parties...should be explicit.

Awareness - Owners, providers, users and other parties should readily be able, consistent with maintaining security, to gain appropriate knowledge of and be informed about the existence and general extent of measures...for the security of information systems.

Ethics - The Information systems and the security of information systems should be provided and used in such a manner that the rights and legitimate interest of others are respected.

Multidisciplinary - Measures, practices and procedures for the security of information systems should take account of and address all relevant considerations and viewpoints....

Proportionality - Security levels, costs, measures, practices and procedures should be appropriate and proportionate to the value of and degree of reliance on the information systems and to the severity, probability and extent of potential harm....

Integration - Measures, practices and procedures for the security of information systems should be coordinated and integrated with each other and other measures, practices and procedures of the organization so as to create a coherent system of security.

Timeliness - Public and private parties, at both national and international levels, should act in a timely coordinated manner to prevent and to respond to breaches of security of information systems.

Reassessment - The security of information systems should be reassessed periodically, as information systems and the requirements for their security vary over time.

Democracy - The security of information systems should be compatible with the legitimate use and flow of data and information in a democratic society.

organization managers have to decide what the level of risk they are willing to accept, taking into account the cost of security controls.

As with many other resources, the management of information and computers may transcend organizational boundaries. When an organization's information and computer systems are linked with external systems, management's responsibilities also extend beyond the organization. This may require that management (1) know what general level or type of security is employed on the external system(s) or (2) seek assurance that the external system provides adequate security for the using organization's needs.

2.3 Computer Security Should Be Cost-Effective.

The costs and benefits of security should be carefully examined *in both monetary and non-monetary terms* to ensure that the cost of controls does not exceed expected benefits. Security should be appropriate and proportionate to the value of and degree of reliance on the computer systems and to the severity, probability and extent of potential harm. Requirements for security vary, depending upon the particular computer system.

In general, security is a smart business practice. By investing in security measures, an organization can reduce the frequency and severity of computer security-related losses. For example, an organization may estimate that it is experiencing significant losses per year in inventory through fraudulent manipulation of its computer system. Security measures, such as an improved access control system, may significantly reduce the loss.

Moreover, a sound security program can thwart hackers and can reduce the frequency of viruses. Elimination of these kinds of threats can reduce unfavorable publicity as well as increase morale and productivity.

Security benefits, however, do have both direct and indirect costs. Direct costs include purchasing, installing, and administering security measures, such as access control software or fire-suppression systems. Additionally, security measures can sometimes affect system performance, employee morale, or retraining requirements. All of these have to be considered in addition to the basic cost of the control itself. In many cases, these additional costs may well exceed the initial cost of the control (as is often seen, for example, in the costs of administering an access control package). Solutions to security problems should not be chosen if they cost more, directly or indirectly, than simply tolerating the problem.

2.4 Computer Security Responsibilities and Accountability Should Be Made Explicit.

The responsibilities and accountability[10] of owners, providers, and users of computer systems and other parties[11] concerned with the security of computer systems should be explicit.[12] The assignment of responsibilities may be internal to an organization or may extend across organizational boundaries.

Depending on the size of the organization, the program may be large or small, even a collateral duty of another management official. However, even small organizations can prepare a document that states organization policy and makes explicit computer security responsibilities. This element does *not* specify that individual accountability must be provided for on all systems. For example, many information dissemination systems do not require user identification and, therefore, cannot hold users accountable.

2.5 Systems Owners Have Security Responsibilities Outside Their Own Organizations.

If a system has external users, its owners have a responsibility to share appropriate knowledge about the existence and general extent of security measures so that other users can be *confident* that the system is adequately secure. (This does not imply that all systems must meet any minimum level of security, but does imply that system owners should inform their clients or users about the nature of the security.)

In addition to sharing information about security, organization managers "should act in a timely,

---

[10] The difference between responsibility and accountability is not always clear. In general, *responsibility* is a broader term, defining obligations and expected behavior. The term implies a proactive stance on the part of the responsible party and a causal relationship between the responsible party and a given outcome. The term *accountability* generally refers to the ability to hold people responsible for their actions. Therefore, people could be responsible for their actions but not held accountable. For example, an anonymous user on a system is responsible for not compromising security but cannot be held accountable if a compromise occurs since the action cannot be traced to an individual.

[11] The term *other parties* may include but is not limited to: executive management; programmers; maintenance providers; information system managers (software managers, operations managers, and network managers); software development managers; managers charged with security of information systems; and internal and external information system auditors.

[12] Implicit is the recognition that people or other entities (such as corporations or governments) *have* responsibilities and accountability related to computer systems. These are responsibilities and accountabilities are often shared among many entities. (Assignment of responsibilities is usually accomplished through the issuance of policy. See Chapter 5.)

coordinated manner to prevent and to respond to breaches of security" to help prevent damage to others.[13] However, taking such action should *not* jeopardize the security of systems.

2.6 Computer Security Requires a Comprehensive and Integrated Approach.

Providing effective computer security requires a comprehensive approach that considers a variety of areas both within and outside of the computer security field. This comprehensive approach extends throughout the entire information life cycle.

2.6.1 Interdependencies of Security Controls

To work effectively, security controls often depend upon the proper functioning of other controls. In fact, many such interdependencies exist. If appropriately chosen, managerial, operational, and technical controls can work together synergistically. On the other hand, without a firm understanding of the interdependencies of security controls, they can actually undermine one another. For example, without proper training on how and when to use a virus-detection package, the user may apply the package incorrectly and, therefore, ineffectively. As a result, the user may mistakenly believe that their system will always be virus-free and may inadvertently spread a virus. In reality, these interdependencies are usually more complicated and difficult to ascertain.

2.6.2 Other Interdependencies

The effectiveness of security controls also depends on such factors as system management, legal issues, quality assurance, and internal and management controls. Computer security needs to work with traditional security disciplines including physical and personnel security. Many other important interdependencies exist that are often unique to the organization or system environment. Managers should recognize how computer security relates to other areas of systems and organizational management.

2.7 Computer Security Should Be Periodically Reassessed.

Computers and the environments they operate in are dynamic. System technology and users, data and information in the systems, risks associated with the system and, therefore, security requirements are ever-changing. Many types of changes affect system security: technological developments (whether adopted by the system owner or available for use by others); connecting to external networks; a change in the value or use of information; or the emergence of a new

[13] Organisation for Economic Co-operation and Development, *Guidelines for the Security of Information Systems*, Paris, 1992.

threat.

In addition, security is *never* perfect when a system is implemented. System users and operators discover new ways to intentionally or unintentionally bypass or subvert security. Changes in the system or the environment can create new vulnerabilities. Strict adherence to procedures is rare, and procedures become outdated over time. All of these issues make it necessary to reassess the security of computer systems.

2.8 Computer Security is Constrained by Societal Factors.

The ability of security to support the mission of the organization(s) may be limited by various factors, such as social issues. For example, security and workplace privacy can conflict. Commonly, security is implemented on a computer system by identifying users and tracking their actions. However, expectations of privacy vary and can be violated by some security measures. (In some cases, privacy may be mandated by law.)

Although privacy is an extremely important societal issue, it is not the only one. The flow of information, especially between a government and its citizens, is another situation where security may need to be modified to support a societal goal. In addition, some authentication measures, such as retinal scanning, may be considered invasive in some environments and cultures.

The underlying idea is that security measures should be selected and implemented with a recognition of the rights and legitimate interests of others. This many involve balancing the security needs of information owners and users with societal goals. However, rules and expectations change with regard to the appropriate use of security controls. These changes may either increase or decrease security.

The relationship between security and societal norms is not necessarily antagonistic. Security can enhance the access and flow of data and information by providing more accurate and reliable information and greater availability of systems. Security can also increase the privacy afforded to an individual or help achieve other goals set by society.

References

Organisation for Economic Co-operation and Development. *Guidelines for the Security of Information Systems.* Paris, 1992.

Chapter 3

ROLES AND RESPONSIBILITIES

One fundamental issue that arises in discussions of computer security is: "Whose responsibility is it?" Of course, on a basic level the answer is simple: computer security is the responsibility of everyone who can affect the security of a computer system. However, the specific duties and responsibilities of various individuals and organizational entities vary considerably.

This chapter presents a brief overview of roles and responsibilities of the various officials and organizational offices *typically* involved with computer security.[14] They include the following groups:[15]

- senior management
- program/functional managers/application owners,
- computer security management,
- technology providers,
- supporting organizations, and
- users.

This chapter is intended to give the reader a basic familiarity with the major organizational elements that play a role in computer security. *It does not describe all responsibilities of each in detail, nor will this chapter apply uniformly to all organizations.* Organizations, like individuals, have unique characteristics, and no single template can apply to all. Smaller organizations, in particular, are not likely to have separate individuals performing many of the functions described in this chapter. Even at some larger organizations, some of the duties described in this chapter may not be staffed with full-time personnel. What is important is that these *functions* be handled in a manner appropriate for the organization.

As with the rest of the handbook, *this chapter is not intended to be used as an audit guide*.

---

[14] Note that this includes groups *within* the organization; outside organizations (e.g., NIST and OMB) are not included in this chapter.

[15] These categories are generalizations used to help aid the reader; if they are not applicable to the reader's particular environment, they can be safely ignored. While all these categories may not exist in a particular organization, the functionality implied by them will often still be present. Also, some organizations may fall into more than one category. For example, the personnel office both supports the computer security program (e.g., by keeping track of employee departures) and is also a user of computer services.

3.1 Senior Management

> Senior management has ultimate responsibility for the security of an organization's computer systems.

Ultimately, responsibility for the success of an organization lies with its senior managers. They establish the organization's computer security program and its overall program goals, objectives, and priorities in order to support the mission of the organization. Ultimately, the head of the organization is responsible for ensuring that adequate resources are applied to the program and that it is successful. Senior managers are also responsible for setting a good example for their employees by following all applicable security practices.

3.2 Computer Security Management

The *Computer Security Program Manager* (and support staff) directs the organization's day-to-day management of its computer security program. This individual is also responsible for coordinating all security-related interactions among organizational elements involved in the computer security program – as well as those external to the organization.

3.3 Program and Functional Managers/Application Owners

Program or Functional Managers/Application Owners are responsible for a program or function (e.g., procurement or payroll) including the supporting computer system.[16] Their responsibilities include providing for appropriate security, including management, operational, and technical controls. These officials are usually assisted by a technical staff that oversees the actual workings of the system. This kind of support is no different for other staff members who work on other program implementation issues.

Also, the program or functional manager/application owner is often aided by a *Security Officer* (frequently dedicated to that system, particularly if it is large or critical to the organization) in developing and implementing security requirements.

3.4 Technology Providers

System Management/System Administrators. These personnel are the managers and technicians who design and operate computer systems. They are responsible for implementing technical security on computer systems and for being familiar with security technology that relates to their system. They also need to ensure the continuity of their services to meet the needs of functional

[16] The functional manager/application owner may or may not be the *data owner*. Particularly within the government, the concept of the data owner may not be the most appropriate, since citizens ultimately own the data.

What is a Program/Functional Manager?

The term *program/functional manager* or *application owner* may not be familiar or immediately apparent to all readers. The examples provided below should help the reader better understand this important concept. In reviewing these examples, note that computer systems often serve more than one group or function.

Example 1. A personnel system serves an entire organization. However, the Personnel Manager

3.5 Supporting Functions[17]

The security responsibilities of managers, technology providers and security officers are supported by functions normally assigned to others. Some of the more important of these are described below.

Audit. Auditors are responsible for examining systems to see whether the system is meeting stated security requirements, including system and organization policies, and whether security controls are appropriate. Informal audits can be performed by those operating the system under review or, if impartiality is important, by outside auditors.[18]

Physical Security. The physical security office is usually responsible for developing and enforcing appropriate physical security controls, in consultation with computer security management, program and functional managers, and others, as appropriate. Physical security should address not only central computer installations, but also backup facilities and office environments. In the government, this office is often responsible for the processing of personnel background checks and security clearances.

Disaster Recovery/Contingency Planning Staff. Some organizations have a separate disaster recovery/contingency planning staff. In this case, they are normally responsible for contingency planning for the organization as a whole, and

Who Should Be the Accrediting Official?

The Accrediting Officials are agency officials who have authority to accept an application's security safeguards and approve a system for operation. The Accrediting Officials must also be authorized to allocate resources to achieve acceptable security and to remedy security deficiencies. Without this authority, they cannot realistically take responsibility for the accreditation decision. In general, Accreditors are senior officials, who may be the Program or Function Manager/Application Owner. For some very sensitive applications, the Senior Executive Officer is appropriate as an Accrediting Official. In general, the more sensitive the application, the higher the Accrediting Officials are in the organization.

Where privacy is a concern, federal managers can be held personally liable for security inadequacies. The issuing of the accreditation statement fixes security responsibility, thus making explicit a responsibility that might otherwise be implicit. Accreditors should consult the agency general counsel to determine their personal security liabilities.

Note that accreditation is a formality unique to the government.

Source: NIST FIPS 102

[17] Categorization of functions and organizations in this section as supporting is in no way meant to imply any degree of lessened importance. Also, note that this list is not all-inclusive. Additional supporting functions that can be provided may include configuration management, independent verification and validation, and independent penetration testing teams.

[18] The term *outside auditors* includes both auditors external to the organization as a whole and the organization's internal audit staff. For purposes of this discussion, both are outside the management chain responsible for the operation of the system.

normally work with program and functional mangers/application owners, the computer security staff, and others to obtain additional contingency planning support, as needed.

Quality Assurance. Many organizations have established a quality assurance program to improve the products and services they provide to their customers. The quality officer should have a working knowledge of computer security and how it can be used to improve the quality of the program, for example, by improving the integrity of computer-based information, the availability of services, and the confidentiality of customer information, as appropriate.

Procurement. The procurement office is responsible for ensuring that organizational procurements have been reviewed by appropriate officials. The procurement office cannot be responsible for ensuring that goods and services meet computer security expectations, because it lacks the technical expertise. Nevertheless, this office should be knowledgeable about computer security standards and should bring them to the attention of those requesting such technology.

Training Office. An organization has to decide whether the primary responsibility for training users, operators, and managers in computer security rests with the training office or the computer security program office. In either case, the two organizations should work together to develop an effective training program.

Personnel. The personnel office is normally the first point of contact in helping managers determine if a security background investigation is necessary for a particular position. The personnel and security offices normally work closely on issues involving background investigations. The personnel office may also be responsible for providing security-related exit procedures when employees leave an organization.

Risk Management/Planning Staff. Some organizations have a full-time staff devoted to studying all types of risks to which the organization may be exposed. This function should include computer security-related risks, although this office normally focuses on "macro" issues. Specific risk analyses for specific computer systems is normally not performed by this office.

Physical Plant. This office is responsible for ensuring the provision of such services as electrical power and environmental controls, necessary for the safe and secure operation of an organization's systems. Often they are augmented by separate medical, fire, hazardous waste, or life safety personnel.

3.6 Users

Users also have responsibilities for computer security. Two kinds of users, and their associated responsibilities, are described below.

Users of Information. Individuals who use information provided by the computer can be considered the "consumers" of the applications. Sometimes they directly interact with the system (e.g., to generate a report on screen) – in which case they are also users of the system (as discussed below). Other times, they may only read computer-prepared reports or only be briefed on such material. Some users of information may be very far removed from the computer system. Users of information are responsible for letting the functional mangers/application owners (or their representatives) know what their needs are for the protection of information, especially for its integrity and availability.

Users of Systems. Individuals who directly use computer systems (typically via a keyboard) are responsible for following security procedures, for reporting security problems, and for attending required computer security and functional training.

References

Wood, Charles Cresson. "How to Achieve a Clear Definition of Responsibilities for Information Security." DATAPRO Information Security Service, IS115-200-101, 7 pp. April 1993.

Chapter 4

COMMON THREATS: A BRIEF OVERVIEW

Computer systems are vulnerable to many threats that can inflict various types of damage resulting in significant losses. This damage can range from errors harming database integrity to fires destroying entire computer centers. Losses can stem, for example, from the actions of supposedly trusted employees defrauding a system, from outside hackers, or from careless data entry clerks. Precision in estimating computer security-related losses is not possible because many losses are never discovered, and others are "swept under the carpet" to avoid unfavorable publicity. The effects of various threats varies considerably: some affect the confidentiality or integrity of data while others affect the availability of a system.

This chapter presents a broad view of the risky environment in which systems operate today. The threats and associated losses presented in this chapter were selected based on their prevalence and significance in the current computing environment and their expected growth. This list is not exhaustive, and some threats may combine elements from more than one area.[19] This overview of many of today's common threats may prove useful to organizations studying their own threat environments; however, the perspective of this chapter is very broad. Thus, threats against particular systems could be quite different from those discussed here.[20]

To control the risks of operating an information system, managers and users need to know the vulnerabilities of the system and the threats that may exploit them. Knowledge of the threat[21] environment allows the system manager to implement the most cost-effective security measures. In some cases, managers may find it more cost-effective to simply tolerate the expected losses. Such decisions should be based on the results of a risk analysis. (See Chapter 7.)

---

[19] As is true for this publication as a whole, this chapter does not address threats to national security systems, which fall outside of NIST's purview. The term "national security systems" is defined in National Security Directive 42 (7/5/90) as being "those telecommunications and information systems operated by the U.S. Government, its contractors, or agents, that contain classified information or, as set forth in 10 U.S.C. 2315, that involves intelligence activities, involves cryptologic activities related to national security, involves command and control of military forces, involves equipment that is an integral part of a weapon or weapon system, or involves equipment that is critical to the direct fulfillment of military or intelligence missions."

[20] A discussion of how threats, vulnerabilities, safeguard selection and risk mitigation are related is contained in Chapter 7, Risk Management.

[21] Note that one protects against threats that can exploit a vulnerability. If a vulnerability exists but no threat exists to take advantage of it, little or nothing is gained by protecting against the vulnerability. See Chapter 7, Risk Management.

4.1 Errors and Omissions

Errors and omissions are an important threat to data and system integrity. These errors are caused not only by data entry clerks processing hundreds of transactions per day, but also by all types of users who create and edit data. Many programs, especially those designed by users for personal computers, lack quality control measures. However, even the most sophisticated programs cannot detect all types of input errors or omissions. A sound awareness and training program can help an organization reduce the number and severity of errors and omissions.

Users, data entry clerks, system operators, and programmers frequently make errors that contribute directly or indirectly to security problems. In some cases, the error is the threat, such as a data entry error or a programming error that crashes a system. In other cases, the errors create vulnerabilities. Errors can occur during all phases of the systems life cycle. A long-term survey of computer-related economic losses conducted by Robert Courtney, a computer security consultant and former member of the Computer System Security and Privacy Advisory Board, found that 65 percent of losses to organizations were the result of errors and omissions.[22] This figure was relatively consistent between both private and public sector organizations.

Programming and development errors, often called "bugs," can range in severity from benign to catastrophic. In a 1989 study for the House Committee on Science, Space and Technology, entitled *Bugs in the Program,* the staff of the Subcommittee on Investigations and Oversight summarized the scope and severity of this problem in terms of government systems as follows:

> As expenditures grow, so do concerns about the reliability, cost and accuracy of ever-larger and more complex software systems. These concerns are heightened as computers perform more critical tasks, where mistakes can cause financial turmoil, accidents, or in extreme cases, death.[23]

Since the study's publication, the software industry has changed considerably, with measurable improvements in software quality. Yet software "horror stories" still abound, and the basic principles and problems analyzed in the report remain the same. While there have been great

---

[22] Computer System Security and Privacy Advisory Board, *1991 Annual Report* (Gaithersburg, MD), March 1992, p. 18. The categories into which the problems were placed and the percentages of economic loss attributed to each were: 65%, errors and omissions; 13%, dishonest employees; 6%, disgruntled employees; 8%, loss of supporting infrastructure, including power, communications, water, sewer, transportation, fire, flood, civil unrest, and strikes; 5%, water, not related to fires and floods; less than 3%, outsiders, including viruses, espionage, dissidents, and malcontents of various kinds, and former employees who have been away for more than six weeks.

[23] House Committee on Science, Space and Technology, Subcommittee on Investigations and Oversight, *Bugs in the Program: Problems in Federal Government Computer Software Development and Regulation*, 101st Cong., 1st sess., 3 August 1989, p. 2.

improvements in program quality, as reflected in decreasing errors per 1000 lines of code, the concurrent growth in program size often seriously diminishes the beneficial effects of these program quality enhancements.

Installation and maintenance errors are another source of security problems. For example, an audit by the President's Council for Integrity and Efficiency (PCIE) in 1988 found that every one of the ten mainframe computer sites studied had installation and maintenance errors that introduced significant security vulnerabilities.[24]

4.2 Fraud and Theft

Computer systems can be exploited for both fraud and theft both by "automating" traditional methods of fraud and by using new methods. For example, individuals may use a computer to skim small amounts of money from a large number of financial accounts, assuming that small discrepancies may not be investigated. Financial systems are not the only ones at risk. Systems that control access to any resource are targets (e.g., time and attendance systems, inventory systems, school grading systems, and long-distance telephone systems).

Computer fraud and theft can be committed by insiders or outsiders. Insiders (i.e., authorized users of a system) are responsible for the majority of fraud. A 1993 *InformationWeek*/Ernst and Young study found that 90 percent of Chief Information Officers viewed employees "who do not need to know" information as threats.[25] The U.S. Department of Justice's Computer Crime Unit contends that "insiders constitute the greatest threat to computer systems."[26] Since insiders have

---

[24] President's Council on Integrity and Efficiency, *Review of General Controls in Federal Computer Systems*, October, 1988.

[25] Bob Violino and Joseph C. Panettieri, "Tempting Fate," *InformationWeek*, October 4, 1993: p. 42.

[26] Letter from Scott Charney, Chief, Computer Crime Unit, U.S. Department of Justice, to Barbara Guttman, NIST. July 29, 1993.

[27] "Theft, Power Surges Cause Most PC Losses," *Infosecurity News*, September/October, 1993, 13.

4.3 Employee Sabotage

Common examples of computer-related employee sabotage include:

- destroying hardware or facilities,
- planting logic bombs that destroy programs or data,
- entering data incorrectly,
- "crashing" systems,
- deleting data,
- holding data hostage, and
- changing data.

Employees are most familiar with their employer's computers and applications, including knowing what actions might cause the most damage, mischief, or sabotage. The downsizing of organizations in both the public and private sectors has created a group of individuals with organizational knowledge, who may retain potential system access (e.g., if system accounts are not deleted in a timely manner).[28] The number of incidents of employee sabotage is believed to be much smaller than the instances of theft, but the cost of such incidents can be quite high.

Martin Sprouse, author of *Sabotage in the American Workplace*, reported that the motivation for sabotage can range from altruism to revenge:

> As long as people feel cheated, bored, harassed, endangered, or betrayed at work, sabotage will be used as a direct method of achieving job satisfaction – the kind that never has to get the bosses' approval.[29]

4.4 Loss of Physical and Infrastructure Support

The loss of supporting infrastructure includes power failures (outages, spikes, and brownouts), loss of communications, water outages and leaks, sewer problems, lack of transportation services, fire, flood, civil unrest, and strikes. These losses include such dramatic events as the explosion at the World Trade Center and the Chicago tunnel flood, as well as more common events, such as broken water pipes. Many of these issues are covered in Chapter 15. A loss of infrastructure often results in system downtime, sometimes in unexpected ways. For example, employees may not be able to get to work during a winter storm, although the computer system may be functional.

4.5 Malicious Hackers

The term *malicious hackers*, sometimes called *crackers*, refers to those who break into computers

---

[28] Charney.

[29] Martin Sprouse, ed., *Sabotage in the American Workplace: Anecdotes of Dissatisfaction, Mischief and Revenge* (San Francisco, CA: Pressure Drop Press, 1992), p. 7.

without authorization. They can include both outsiders and insiders. Much of the rise of hacker activity is often attributed to increases in connectivity in both government and industry. One 1992 study of a particular Internet site (i.e., one computer system) found that hackers attempted to break in at least once every other day.[30]

The hacker threat should be considered in terms of past and potential future damage. Although current losses due to hacker attacks are significantly smaller than losses due to insider theft and sabotage, the hacker problem is widespread and serious. One example of malicious hacker activity is that directed against the public telephone system.

Studies by the National Research Council and the National Security Telecommunications Advisory Committee show that hacker activity is not limited to toll fraud. It also includes the ability to break into telecommunications systems (such as switches), resulting in the degradation or disruption of system availability. While unable to reach a conclusion about the degree of threat or risk, these studies underscore the ability of hackers to cause serious damage.[31, 32]

The hacker threat often receives more attention than more common and dangerous threats. The U.S. Department of Justice's Computer Crime Unit suggests three reasons for this.

- First, the hacker threat is a more recently encountered threat. Organizations have always had to worry about the actions of their own employees and could use disciplinary measures to reduce that threat. However, these measures are ineffective against outsiders who are not subject to the rules and regulations of the employer.

- Second, organizations do not know the purposes of a hacker – some hackers browse, some steal, some damage. This inability to identify purposes can suggest that hacker attacks have no limitations.

- Third, hacker attacks make people feel vulnerable, particularly because their identity is unknown. For example, suppose a painter is hired to paint a house and, once inside, steals a piece of jewelry. Other homeowners in the neighborhood may not feel threatened by this crime and will protect themselves by not doing business with that painter. But if a burglar breaks into the same house and steals the same

---

[30] Steven M. Bellovin, "There Be Dragons," *Proceedings of the Third Usenix UNIX Security Symposium*.

[31] National Research Council, *Growing Vulnerability of the Public Switched Networks: Implication for National Security Emergency Preparedness* (Washington, DC: National Academy Press), 1989.

[32] Report of the National Security Task Force, November 1990.

piece of jewelry, the entire neighborhood may feel victimized and vulnerable.[33]

4.6 Industrial Espionage

Industrial espionage is the act of gathering proprietary data from private companies or the government[34] for the purpose of aiding another company(ies). Industrial espionage can be perpetrated either by companies seeking to improve their competitive advantage or by governments seeking to aid their domestic industries. Foreign industrial espionage carried out by a government is often referred to as economic espionage. Since information is processed and stored on computer systems, computer security can help protect against such threats; it can do little, however, to reduce the threat of authorized employees selling that information.

Industrial espionage is on the rise. A 1992 study sponsored by the American Society for Industrial Security (ASIS) found that proprietary business information theft had increased 260 percent since 1985. The data indicated 30 percent of the reported losses in 1991 and 1992 had foreign involvement. The study also found that 58 percent of thefts were perpetrated by current or former employees.[35] The three most damaging types of stolen information were pricing information, manufacturing process information, and product development and specification information. Other types of information stolen included customer lists, basic research, sales data, personnel data, compensation data, cost data, proposals, and strategic plans.[36]

Within the area of economic espionage, the Central Intelligence Agency has stated that the main objective is obtaining information related to technology, but that information on U.S. Government policy deliberations concerning foreign affairs and information on commodities, interest rates, and other economic factors is also a target.[37] The Federal Bureau of Investigation concurs that technology-related information is the main target, but also lists corporate proprietary information, such as negotiating positions and other contracting data, as a target.[38]

---

[33] Charney.

[34] The government is included here because it often is the custodian for proprietary data (e.g., patent applications).

[35] The figures of 30 and 58 percent are not mutually exclusive.

[36] Richard J. Heffernan and Dan T. Swartwood, "Trends in Competitive Intelligence," *Security Management* 37, no. 1 (January 1993), pp. 70-73.

[37] Robert M. Gates, testimony before the House Subcommittee on Economic and Commercial Law, Committee on the Judiciary, 29 April 1992.

[38] William S. Sessions, testimony before the House Subcommittee on Economic and Commercial Law, Committee on the Judiciary, 29 April 1992.

4.7 Malicious Code

Malicious code refers to viruses, worms, Trojan horses, logic bombs, and other "uninvited" software. Sometimes mistakenly associated only with personal computers, malicious code can attack other platforms.

A 1993 study of viruses found that while the number of known viruses is increasing exponentially, the number of virus incidents is not.[39] The study concluded that viruses are becoming more prevalent, but only "gradually."

> The rate of PC-DOS virus incidents in medium to large North American businesses appears to be approximately 1 per 1000 PCs per quarter; the number of infected machines is perhaps 3 or 4 times this figure if we assume that most such businesses are at least weakly protected against viruses.[40, 41]

Actual costs attributed to the presence of malicious code have resulted primarily from system outages and staff time involved in repairing the systems. Nonetheless, these costs can be significant.

Malicious Software: A Few Key Terms

Virus: A code segment that replicates by attaching copies of itself to existing executables. The new copy of the virus is executed when a user executes the new host program. The virus may include an additional "payload" that triggers when specific conditions are met. For example, some viruses display a text string on a particular date. There are many types of viruses, including variants, overwriting, resident, stealth, and polymorphic.

Trojan Horse: A program that performs a desired task, but that also includes unexpected (and undesirable) functions. Consider as an example an editing program for a multiuser system. This program could be modified to randomly delete one of the users' files each time they perform a useful function (editing), but the deletions are unexpected and definitely undesired!

Worm: A self-replicating program that is self-contained and does not require a host program. The program creates a copy of itself and causes it to execute; no user intervention is required. Worms commonly use network services to propagate to other host systems.
Source: NIST Special Publication 800-5.

4.8 Foreign Government Espionage

In some instances, threats posed by foreign government intelligence services may be present. In addition to possible economic espionage, foreign intelligence services may target unclassified

[39] Jeffrey O. Kephart and Steve R. White, "Measuring and Modeling Computer Virus Prevalence," *Proceedings, 1993 IEEE Computer Society Symposium on Research in Security and Privacy* (May 1993): 14.

[40] Ibid.

[41] Estimates of virus occurrences may not consider the strength of an organization's antivirus program.

systems to further their intelligence missions. Some unclassified information that may be of interest includes travel plans of senior officials, civil defense and emergency preparedness, manufacturing technologies, satellite data, personnel and payroll data, and law enforcement, investigative, and security files. Guidance should be sought from the cognizant security office regarding such threats.

4.9 Threats to Personal Privacy

The accumulation of vast amounts of electronic information about individuals by governments, credit bureaus, and private companies, combined with the ability of computers to monitor, process, and aggregate large amounts of information about individuals have created a threat to individual privacy. The possibility that all of this information and technology may be able to be linked together has arisen as a specter of the modern information age. This is often referred to as "Big Brother." To guard against such intrusion, Congress has enacted legislation, over the years, such as the Privacy Act of 1974 and the Computer Matching and Privacy Protection Act of 1988, which defines the boundaries of the legitimate uses of personal information collected by the government.

The threat to personal privacy arises from many sources. In several cases federal and state employees have sold personal information to private investigators or other "information brokers." One such case was uncovered in 1992 when the Justice Department announced the arrest of over two dozen individuals engaged in buying and selling information from Social Security Administration (SSA) computer files.[42] During the investigation, auditors learned that SSA employees had unrestricted access to over 130 million employment records. Another investigation found that 5 percent of the employees in one region of the IRS had browsed through tax records of friends, relatives, and celebrities.[43] Some of the employees used the information to create fraudulent tax refunds, but many were acting simply out of curiosity.

As more of these cases come to light, many individuals are becoming increasingly concerned about threats to their personal privacy. A July 1993 special report in *MacWorld* cited polling data taken by Louis Harris and Associates showing that in 1970 only 33 percent of respondents were

---

[42] House Committee on Ways and Means, Subcommittee on Social Security, *Illegal Disclosure of Social Security Earnings Information by Employees of the Social Security Administration and the Department of Health and Human Services' Office of Inspector General: Hearing,* 102nd Cong., 2nd sess., 24 September 1992, Serial 102-131.

[43] Stephen Barr, "Probe Finds IRS Workers Were `Browsing' in Files," *The Washington Post,* 3 August 1993, p. A1.

concerned about personal privacy. By 1990, that number had jumped to 79 percent.[44]

While the magnitude and cost to society of the personal privacy threat are difficult to gauge, it is apparent that information technology is becoming powerful enough to warrant fears of both government and corporate "Big Brothers." Increased awareness of the problem is needed.

References

House Committee on Science, Space and Technology, Subcommittee on Investigations and Oversight. *Bugs in the Program: Problems in Federal Government Computer Software Development and Regulation*. 101st Congress, 1st session, August 3, 1989.

National Research Council. *Computers at Risk: Safe Computing in the Information Age*. Washington, DC: National Academy Press, 1991.

National Research Council. *Growing Vulnerability of the Public Switched Networks: Implication for National Security Emergency Preparedness.* Washington, DC: National Academy Press, 1989.

Neumann, Peter G. *Computer-Related Risks*. Reading, MA: Addison-Wesley, 1994.

Schwartau, W. *Information Warfare*. New York, NY: Thunders Mouth Press, 1994 (Rev. 1995).

Sprouse, Martin, ed. *Sabotage in the American Workplace: Anecdotes of Dissatisfaction, Mischief, and Revenge*. San Francisco, CA: Pressure Drop Press, 1992.

[44] Charles Piller, "Special Report: Workplace and Consumer Privacy Under Siege," *MacWorld*, July 1993, pp. 1-14.

II. MANAGEMENT CONTROLS

Chapter 5

COMPUTER SECURITY POLICY

In discussions of computer security, the term *policy* has more than one meaning.[45] *Policy* is senior management's directives to create a computer security program, establish its goals, and assign responsibilities. The term *policy* is also used to refer to the specific security rules for particular systems.[46] Additionally, *policy* may refer to entirely different matters, such as the specific managerial decisions setting an organization's e-mail privacy policy or fax security policy.

> *Policy* means different things to different people. The term "policy" is used in this chapter in a broad manner to refer to important computer security-related decisions.

In this chapter the term *computer security policy* is defined as the "documentation of computer security decisions" – which covers all the types of policy described above.[47] In making these decisions, managers face hard choices involving resource allocation, competing objectives, and organizational strategy related to protecting both technical and information resources as well as guiding employee behavior. Managers at all levels make choices that can result in policy, with the scope of the policy's applicability varying according to the scope of the manager's authority. In this chapter we use the term *policy* in a broad manner to encompass all of the types of policy described above – regardless of the level of manager who sets the particular policy.

Managerial decisions on computer security issues vary greatly. To differentiate among various kinds of policy, this chapter categorizes them into three basic types:

- *Program policy* is used to create an organization's computer security program.
- *Issue-specific policies* address specific issues of concern to the organization.

---

[45] There are variations in the use of the term *policy*, as noted in a 1994 Office of Technology Assessment report, *Information Security and Privacy in Network Environments*: "*Security Policy* refers here to the statements made by organizations, corporations, and agencies to establish overall policy on information access and safeguards. Another meaning comes from the Defense community and refers to the rules relating clearances of users to classification of information. In another usage, *security policies* are used to refine and implement the broader, organizational security policy...."

[46] These are the kind of policies that computer security experts refer to as being *enforced* by the system's technical controls as well as its management and operational controls.

[47] In general, policy is set by a manager. However, in some cases, it may be set by a group (e.g., an intraorganizational policy board).

- *System-specific policies* focus on decisions taken by management to protect a particular system.[48]

Procedures, standards, and guidelines are used to describe how these policies will be implemented within an organization. (See following box.)

Tools to Implement Policy:
Standards, Guidelines, and Procedures

Because policy is written at a broad level, organizations also develop standards, guidelines, and procedures that offer users, managers, and others a clearer approach to implementing policy and meeting organizational goals. Standards and guidelines specify technologies and methodologies to be used to secure systems. Procedures are yet more detailed steps to be followed to accomplish particular security-related tasks. Standards, guidelines, and procedures may be promulgated throughout an organization via handbooks, regulations, or manuals.

Organizational standards (not to be confused with American National Standards, FIPS, Federal Standards, or other national or international standards) specify uniform use of specific technologies, parameters, or procedures when such uniform use will benefit an organization. Standardization of organizationwide identification badges is a typical example, providing ease of employee mobility and automation of entry/exit systems. Standards are normally compulsory within an organization.

Guidelines assist users, systems personnel, and others in effectively securing their systems. The nature of guidelines, however, immediately recognizes that systems vary considerably, and imposition of standards is not always achievable, appropriate, or cost-effective. For example, an organizational guideline may be used to help develop system-specific standard procedures. Guidelines are often used to help ensure that specific security measures are not overlooked, although they can be implemented, and correctly so, in more than one way.

Procedures normally assist in complying with applicable security policies, standards, and guidelines. They are detailed steps to be followed by users, system operations personnel, or others to accomplish a particular task (e.g., preparing new user accounts and assigning the appropriate privileges).

Some organizations issue overall computer security manuals, regulations, handbooks, or similar documents. These may mix policy, guidelines, standards, and procedures, since they are closely linked. While manuals and regulations can serve as important tools, it is often useful if they clearly distinguish between policy and its implementation. This can help in promoting flexibility and cost-effectiveness by offering alternative implementation approaches to achieving policy goals.

[48] A *system* refers to the entire collection of processes, both those performed manually and those using a computer (e.g., manual data collection and subsequent computer manipulation), which performs a function. This includes both application systems and support systems, such as a network.

Familiarity with various types and components of policy will aid managers in addressing computer security issues important to the organization. Effective policies ultimately result in the development and implementation of a better computer security program and better protectio n of

[49] No standard terms exist for various types of policies. These terms are used to aid the reader's understanding of this topic; no implication of their widespread usage is intended.

Responsibilities. Once the computer security program is established, its management is normally assigned to either a newly created or existing office.[50]

> Program policy establishes the security program and assigns program management and supporting responsibilities.

The responsibilities of officials and offices throughout the organization also need to be addressed, including line managers, applications owners, users, and the data processing or IRM organizations. This section of the policy statement, for example, would distinguish between the responsibilities of computer services providers and those of the managers of applications using the provided services. The policy could also establish operational security offices for major systems, particularly those at high risk or most critical to organizational operations. It also can serve as the basis for establishing employee accountability.

At the program level, responsibilities should be specifically assigned to those organizational elements and officials responsible for the implementation and continuity of the computer security policy.[51]

Compliance. Program policy typically will address two compliance issues:

1. General compliance to ensure meeting the requirements to establish a program and the responsibilities assigned therein to various organizational components. Often an oversight office (e.g., the Inspector General) is assigned responsibility for monitoring compliance, including how well the organization is implementing management's priorities for the program.

2. The use of specified penalties and disciplinary actions. Since the security policy is a high-level document, specific penalties for various infractions are normally not detailed here; instead, the policy may authorize the creation of compliance structures that include violations and specific disciplinary action(s).[52]

---

[50] The program management structure should be organized to best address the goals of the program and respond to the particular operating and risk environment of the organization. Important issues for the structure of the computer security program include management and coordination of security-related resources, interaction with diverse communities, and the ability to relay issues of concern, trade-offs, and recommended actions to upper management. (See Chapter 6, Computer Security Program Management.)

[51] In assigning responsibilities, it is necessary to be specific; such assignments as "computer security is everyone's responsibility," in reality, mean no one has specific responsibility.

[52] The need to obtain guidance from appropriate legal counsel is critical when addressing issues involving penalties and disciplinary action for individuals. The policy does not need to restate penalties already provided

Those developing compliance policy should remember that violations of policy can be unintentional on the part of employees. For example, nonconformance can often be due to a lack of knowledge or training.

5.2 Issue-Specific Policy

Whereas program policy is intended to address the broad organizationwide computer security program, issue-specific policies are developed to focus on areas of current relevance and concern (and sometimes controversy) to an organization. Management may find it appropriate, for example, to issue a policy on how the organization will approach contingency planning (centralized vs. decentralized) or the use of a particular methodology for managing risk to systems. A policy could also be issued, for example, on the appropriate use of a cutting-edge technology (whose security vulnerabilities are still largely unknown) within the organization. Issue-specific policies may also be appropriate when new issues arise, such as when implementing a recently passed law requiring additional protection of particular information. Program policy is usually broad enough that it does not require much modification over time, whereas issue-specific policies are likely to require more frequent revision as changes in technology and related factors take place.

In general, for issue-specific and system-specific policy, the issuer is a senior official; the more global, controversial, or resource-intensive, the more senior the issuer.

5.2.1 Example Topics for Issue-Specific Policy[53]

Both new technologies and the appearance of new threats often require the creation of issue-specific policies.

There are many areas for which issue-specific policy may be appropriate. Two examples are explained below.

Internet Access. Many organizations are looking at the Internet as a means for expanding their research opportunities and communications. Unquestionably, connecting to the Internet yields many benefits – and some disadvantages. Some issues an Internet access policy may address include who will have access, which types of systems may be connected to the network, what types of information may be transmitted via the network, requirements for user authentication for Internet-connected systems, and the use of firewalls and secure gateways.

---

for by law, although they can be listed if the policy will also be used as an awareness or training document.

[53] Examples presented in this section are not all-inclusive nor meant to imply that policies in each of these areas are required by all organizations.

E-Mail Privacy. Users of computer e-mail systems have come to rely upon that service for informal communication with colleagues and others. However, since the system is typically owned by the employing organization, from time-to-time, management may wish to monitor the employee's e-mail for various reasons (e.g., to be sure that it is used for business purposes only or if they are suspected of distributing viruses, sending offensive e-mail, or disclosing organizational secrets.) On the other hand, users may have an expectation of privacy, similar to that accorded U.S. mail. Policy in this area addresses what level of privacy will be accorded e-mail and the circumstances under which it may or may not be read.

Other potential candidates for issue-specific policies include: approach to risk management and contingency planning, protection of confidential/proprietary information, unauthorized software, acquisition of software, doing computer work at home, bringing in disks from outside the workplace, access to other employees' files, encryption of files and e-mail, rights of privacy, responsibility for correctness of data, suspected malicious code, and physical emergencies.

5.2.2 Basic Components of Issue-Specific Policy

As suggested for program policy, a useful structure for issue-specific policy is to break the policy into its basic components.

Issue Statement. To formulate a policy on an issue, managers first must define the issue with any relevant terms, distinctions, and conditions included. It is also often useful to specify the goal or justification for the policy – which can be helpful in gaining compliance with the policy. For example, an organization might want to develop an issue-specific policy on the use of "unofficial software," which might be defined to mean any software not approved, purchased, screened, managed, and owned by the organization. Additionally, the applicable distinctions and conditions might then need to be included, for instance, for software privately owned by employees but approved for use at work, and for software owned and used by other businesses under contract to the organization.

Statement of the Organization's Position. Once the issue is stated and related terms and conditions are discussed, this section is used to clearly state the organization's position (i.e., management's decision) on the issue. To continue the previous example, this would mean stating whether use of unofficial software as defined is prohibited in all or some cases, whether there are further guidelines for approval and use, or whether case-by-case exceptions will be granted, by whom, and on what basis.

Applicability. Issue-specific policies also need to include statements of applicability. This means clarifying where, how, when, to whom, and to what a particular policy applies. For example, it could be that the hypothetical policy on unofficial software is intended to apply only to the organization's own on-site resources and employees and not to contractors with offices at other

locations. Additionally, the policy's applicability to employees travelling among different sites and/or working at home who need to transport and use disks at multiple sites might need to be clarified.

Roles and Responsibilities. The assignment of roles and responsibilities is also usually included in issue-specific policies. For example, if the policy permits unofficial software privately owned by employees to be used at work with the appropriate approvals, then the approval authority granting such permission would need to be stated. (Policy would stipulate, who, by position, has such authority.) Likewise, it would need to be clarified who would be responsible for ensuring that only approved software is used on organizational computer resources and, perhaps, for monitoring users in regard to unofficial software.

Compliance. For some types of policy, it may be appropriate to describe, in some detail, the infractions that are unacceptable, and the consequences of such behavior. Penalties may be explicitly stated and should be consistent with organizational personnel policies and practices. When used, they should be coordinated with appropriate officials and offices and, perhaps, employee bargaining units. It may also be desirable to task a specific office within the organization to monitor compliance.

Points of Contact and Supplementary Information. For any issue-specific policy, the appropriate individuals in the organization to contact for further information, guidance, and compliance should be indicated. Since positions tend to change less often than the people occupying them, specific positions may be preferable as the point of contact. For example, for some issues the point of contact might be a line manager; for other issues it might be a facility manager, technical support person, system administrator, or security program representative. Using the above example once more, employees would need to know whether the point of contact for questions and procedural information would be their immediate superior, a system

Some Helpful Hints on Policy

To be effective, policy requires visibility. Visibility aids implementation of policy by helping to ensure policy is fully communicated throughout the organization. Management presentations, videos, panel discussions, guest speakers, question/answer forums, and newsletters increase visibility. The organization's computer security training and awareness program can effectively notify users of new policies. It also can be used to familiarize new employees with the organization's policies.

Computer security policies should be introduced in a manner that ensures that management's unqualified support is clear, especially in environments where employees feel inundated with policies, directives, guidelines, and procedures. The organization's policy is the vehicle for emphasizing management's commitment to computer security and making clear their expectations for employee performance, behavior, and accountability.

To be effective, policy should be consistent with other existing directives, laws, organizational culture, guidelines, procedures, and the organization's overall mission. It should also be integrated into and consistent with other organizational policies (e.g., personnel policies). One way to help ensure this is to coordinate policies during development with other organizational offices.

administrator, or a computer security official.

Guidelines and procedures often accompany policy. The issue-specific policy on unofficial software, for example, might include procedural guidelines for checking disks brought to work that had been used by employees at other locations.

5.3 System-Specific Policy

Program policy and issue-specific policy both address policy from a broad level, usually encompassing the entire organization. However, they do not provide sufficient information or direction, for example, to be used in establishing an access control list or in training users on what actions are permitted. System-specific policy fills this need. It is much more focused, since it addresses only one system.

Many security policy decisions may apply only at the system level and may vary from system to system within the same organization. While these decisions may appear to be too detailed to be policy, they can be extremely important, with significant impacts on system usage and security. These types of decisions can be made by a *management official*, not by a technical system administrator.[54] (The impacts of these decisions, however, are often analyzed by technical system administrators.)

To develop a cohesive and comprehensive set of security policies, officials may use a management process that derives security rules from security goals. It is helpful to consider a two-level model for system security policy: security objectives and operational security rules, which together comprise the system-specific policy. Closely linked and often difficult to distinguish, however, is the implementation of the policy in technology.

> System-specific security policy includes two components: security objectives and operational security rules. It is often accompanied by implementing procedures and guidelines.

5.3.1 Security Objectives

The first step in the management process is to define security objectives for the specific system. Although, this process may start with an analysis of the need for integrity,

> **Sample Security Objective**
>
> Only individuals in the accounting and personnel departments are authorized to provide or modify information used in payroll processing.

[54] It is important to remember that policy is not created in a vacuum. For example, it is critical to understand the system mission and how the system is intended to be used. Also, users may play an important role in setting policy.

availability, and confidentiality, it should not stop there. A security *objective* needs to more specific; it should be concrete and well defined. It also should be stated so that it is clear that the objective is achievable. This process will also draw upon other applicable organization policies.

Security objectives consist of a series of statements that describe meaningful actions about explicit resources. These objectives should be based on system functional or mission requirements, but should state the security actions that support the requirements.

Development of system-specific policy will require management to make trade-offs, since it is unlikely that all desired security objectives will be able to be fully met. Management will face cost, operational, technical, and other constraints.

5.3.2 Operational Security Rules

After management determines the security objectives, the rules for operating a system can be laid out, for example, to define authorized and unauthorized modification. Who (by job category, organization placement, or name) can do what (e.g., modify, delete) to which specific classes and records of data, and under what conditions.

Sample Operational Security Rule

Personnel clerks may update fields for weekly attendance, charges to annual leave, employee addresses, and telephone numbers. Personnel specialists may update salary information. No employees may update their own records.

The degree of specificity needed for operational security rules varies greatly. The more detailed the rules are, *up to a point*, the easier it is to know when one has been violated. It is also, *up to a point*, easier to automate policy enforcement. However, overly detailed rules may make the job of instructing a computer to implement them difficult or computationally complex.

In addition to deciding the level of detail, management should decide the degree of formality in documenting the system-specific policy. Once again, the more formal the documentation, the easier it is to enforce and to follow policy. On the other hand, policy at the system level that is too detailed and formal can also be an administrative burden. In general, good practice suggests a reasonably detailed formal statement of the access privileges for a system. Documenting access controls policy will make it substantially easier to follow and to enforce. (See Chapters 10 and 17, Personnel/User Issues and Logical Access Control.) Another area that normally requires a detailed and formal statement is the assignment of security responsibilities. Other areas that should be addressed are the rules for system usage and the consequences of noncompliance.

Policy decisions in other areas of computer security, such as those described in this handbook, are often documented in the risk analysis, accreditation statements, or procedural manuals. However,

any controversial, atypical, or uncommon policies will also need formal statements. Atypical policies would include any areas where the system policy is different from organizational policy or from normal practice within the organization, either more or less stringent. The documentation for a typical policy contains a statement explaining the reason for deviation from the organization's standard policy.

5.3.3 System-Specific Policy Implementation

Technology plays an important – but not sole – role in enforcing system-specific policies. When technology is used to enforce policy, it is important not to neglect nontechnology- based methods. For example, technical system-based controls could be used to limit the printing of confidential reports to a particular printer. However, corresponding physical security measures would also have to be in place to limit access to the printer output or the desired security objective would not be achieved.

Technical methods frequently used to implement system-security policy are likely to include the use of *logical access controls*. However, there are other automated means of enforcing or supporting security policy that typically supplement logical access controls. For example, technology can be used to block telephone users from calling certain numbers. Intrusion-detection software can alert system administrators to suspicious activity or can take action to stop the activity. Personal computers can be configured to prevent booting from a floppy disk.

Technology-based enforcement of system-security policy has both advantages and disadvantages. A computer system, properly designed, programmed, installed, configured, and maintained,[55] consistently enforces policy within the computer system, although no computer can force users to follow all procedures. Management controls also play an important role – and should not be neglected. In addition, deviations from the policy may sometimes be necessary and appropriate; such deviations may be difficult to implement easily with some technical controls. This situation occurs frequently if implementation of the security policy is too rigid (which can occur when the system analysts fail to anticipate contingencies and prepare for them).

5.4 Interdependencies

Policy is related to many of the topics covered in this handbook:

Program Management. Policy is used to establish an organization's computer security program, and is therefore closely tied to program management and administration. Both program and system-specific policy may be established in any of the areas covered in this handbook. For

[55] Doing all of these things properly is, unfortunately, the exception rather than the rule. Confidence in the system's ability to enforce system-specific policy is closely tied to assurance. (See Chapter 9, Assurance.)

example, an organization may wish to have a consistent approach to incident handling for all its systems – and would issue appropriate program policy to do so. On the other hand, it may decide that its applications are sufficiently independent of each other that application managers should deal with incidents on an individual basis.

Access Controls. System-specific policy is often implemented through the use of access controls. For example, it may be a policy decision that only two individuals in an organization are authorized to run a check-printing program. Access controls are used by the system to implement (or enforce) this policy.

Links to Broader Organizational Policies. This chapter has focused on the types and components of computer security policy. However, it is important to realize that *computer* security policies are often *extensions* of an organization's *information* security policies for handling information in other forms (e.g., paper documents). For example, an organization's e-mail policy would probably be tied to its broader policy on privacy. Computer security policies may also be extensions of other policies, such as those about appropriate use of equipment and facilities.

5.5 Cost Considerations

A number of potential costs are associated with developing and implementing computer security policies. Overall, the major cost of policy is the cost of implementing the policy and its impacts upon the organization. For example, establishing a computer security program, accomplished through policy, does not come at negligible cost.

Other costs may be those incurred through the policy development process. Numerous administrative and management activities may be required for drafting, reviewing, coordinating, clearing, disseminating, and publicizing policies. In many organizations, successful policy implementation may require additional staffing and training – and can take time. In general, the costs to an organization for computer security policy development and implementation will depend upon how extensive the change needed to achieve a level of risk acceptable to management.

References

Howe, D. "Information System Security Engineering: Cornerstone to the Future." *Proceedings of the 15th National Computer Security Conference*. Baltimore, MD, Vol. 1, October 15, 1992. pp. 244-251.

Fites, P., and M. Kratz. "Policy Development." *Information Systems Security: A Practitioner's Reference*. New York, NY: Van Nostrand Reinhold, 1993. pp. 411-427.

II. Management Controls

Lobel, J. "Establishing a System Security Policy." *Foiling the System Breakers*. New York, NY: McGraw-Hill, 1986. pp. 57-95.

Menkus, B. "Concerns in Computer Security." *Computers and Security*. 11(3), 1992. pp. 211-215.

Office of Technology Assessment. "Federal Policy Issues and Options." *Defending Secrets, Sharing Data: New Locks for Electronic Information*. Washington, DC: U.S Congress, Office of Technology Assessment, 1987. pp. 151-160.

Office of Technology Assessment. "Major Trends in Policy Development." *Defending Secrets, Sharing Data: New Locks and Keys for Electronic Information*. Washington, DC: U.S. Congress, Office of Technology Assessment, 1987. p. 131-148.

O'Neill, M., and F. Henninge, Jr. "Understanding ADP System and Network Security Considerations and Risk Analysis." *ISSA Access*. 5(4), 1992. pp. 14-17.

Peltier, Thomas. "Designing Information Security Policies That Get Results." *Infosecurity News*. 4(2), 1993. pp. 30-31.

President's Council on Management Improvement and the President's Council on Integrity and Efficiency. *Model Framework for Management Control Over Automated Information System*. Washington, DC: President's Council on Management Improvement, January 1988.

Smith, J. "Privacy Policies and Practices: Inside the Organizational Maze." *Communications of the ACM*. 36(12), 1993. pp. 104-120.

Sterne, D. F. "On the Buzzword `Computer Security Policy.'" In *Proceedings of the 1991 IEEE Symposium on Security and Privacy*, Oakland, CA: May 1991. pp. 219-230.

Wood, Charles Cresson. "Designing Corporate Information Security Policies." *DATAPRO Reports on Information Security*, April 1992.

Chapter 6

COMPUTER SECURITY PROGRAM MANAGEMENT

Computers and the information they process are critical to many organizations' ability to perform their mission and business functions.[56] It therefore makes sense that executives view computer security as a management issue and seek to protect their organization's computer resources as they would any other valuable asset. To do this effectively requires developing of a comprehensive management approach.

> OMB Circular A-130, "Management of Federal Information Resources," requires that federal agencies establish computer security programs.

This chapter presents an organizationwide approach to computer security and discusses its important management function.[57] Because organizations differ vastly in size, complexity, management styles, and culture, it is not possible to describe one ideal computer security program. However, this chapter does describe some of the features and issues common to many federal organizations.

6.1 Structure of a Computer Security Program

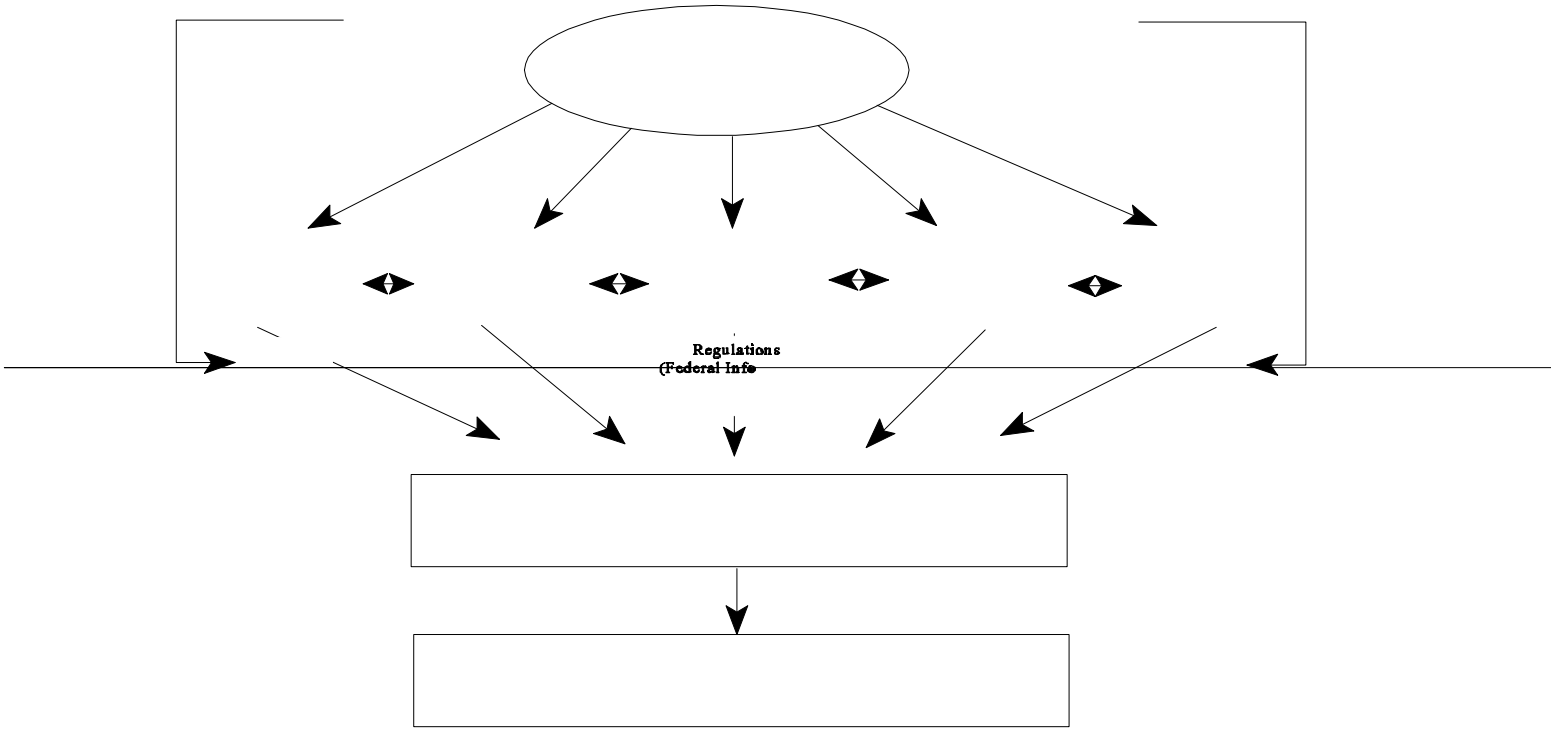
Many computer security programs that are distributed throughout the organization have different elements performing various functions. While this approach has benefits, the distribution of the computer security function in many organizations is haphazard, usually based upon history (i.e., who was available in the organization to do what when the need arose). Ideally, the distribution of computer security functions should result from a planned and integrated management philosophy.

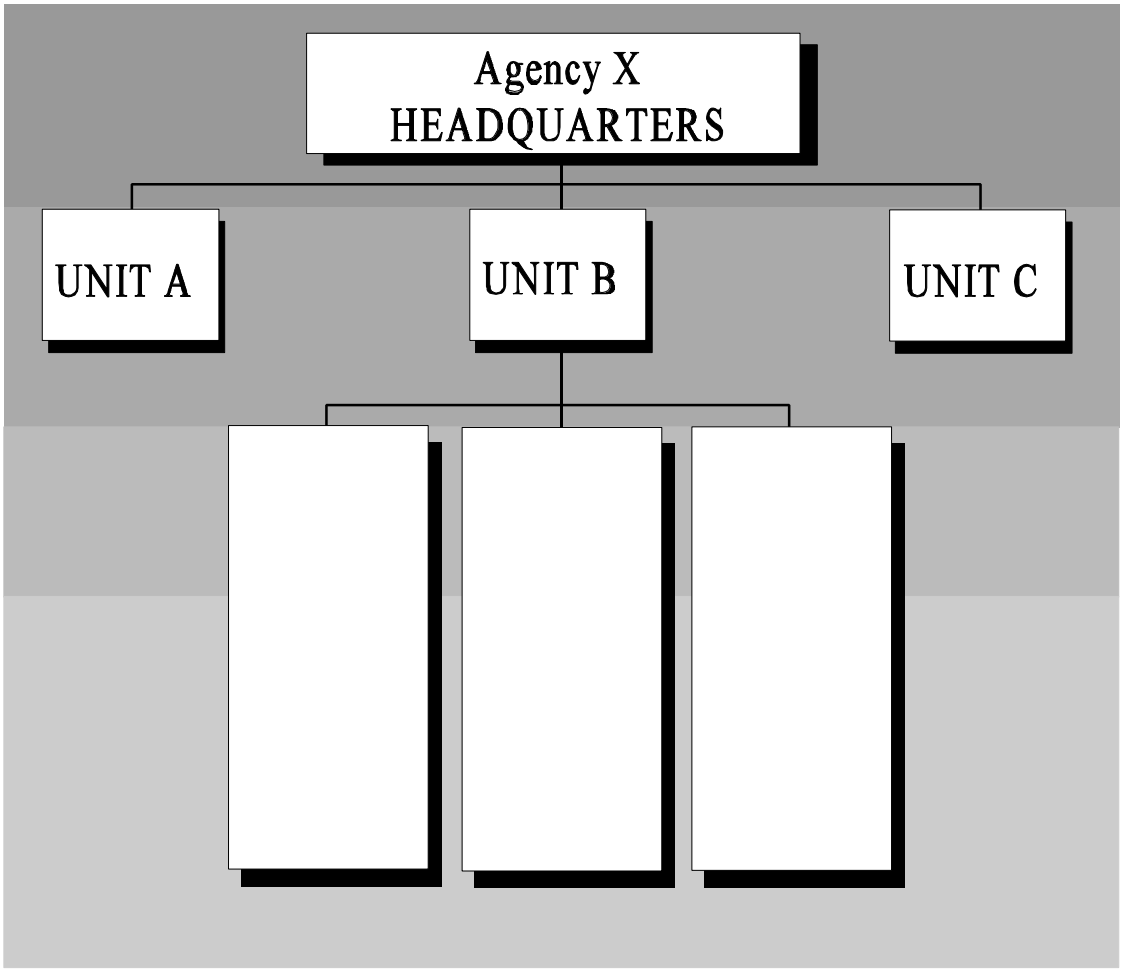
Managing computer security at multiple levels brings many benefits. Each level contributes to the overall computer security program with different types of expertise, authority, and resources. In general, higher-level officials (such as those at the headquarters or unit levels in the agency described above) better understand the organization as a whole and have more authority. On the other hand, lower-level officials (at the computer facility and applications levels) are more familiar with the specific requirements, both technical and procedural, and problems of the systems and

---

[56] This chapter is primarily directed at federal agencies, which are generally very large and complex organizations. This chapter discusses programs which are suited to managing security in such environments. They may be wholly inappropriate for smaller organizations or private sector firms.

[57] This chapter addresses the management of security programs, not the various activities such as risk analysis or contingency planning that make up an effective security program.


Regulations
(Federal Info



Figu

computer security within an organization. In the federal government, the organization could consist of a department, agency, or other major operating unit.

As with the management of all resources, central computer security management can be performed in many practical and cost-effective ways. The importance of sound management cannot be overemphasized. There is also a downside to centrally managed computer security programs. Specifically, they present greater risk that errors in judgement will be more widely propagated throughout the organization. As they strive to meet their objectives, managers need to consider the full impact of available options when establishing their computer security programs.

6.2.1 Benefits of Central Computer Security Programs

A central security program should provide two quite distinct types of benefits:

- Increased efficiency and economy of security throughout the organization, and
- the ability to provide centralized enforcement and oversight.

Both of these benefits are in keeping with the purpose of the Paperwork Reduction Act, as implemented in OMB Circular A-130.

> The Paperwork Reduction Act establishes a broad mandate for agencies to perform their information management activities in an efficient, effective, and economical manner... . Agencies shall assure an adequate level of security for all agency automated information systems, whether maintained in-house or commercially.[58]

6.2.2 Efficient, Economic Coordination of Information

A central computer security program helps to coordinate and manage effective use of security-related resources throughout the organization. The most important of these resources are normally *information* and *financial resources*.

Sound and timely information is necessary for managers to accomplish their tasks effectively. However, most organizations have trouble collecting information from myriad sources and effectively processing and distributing it within the organization. This section discusses some of the sources and efficient uses of *computer security* information.

Within the federal government, many organizations such as the Office of Management and

[58] OMB Circular A-130, Section 5; Appendix III, Section 3.

Budget, the General Services Administration, the National Institute of Standards and Technology, and the National Telecommunications and Information Administration, provide information on computer, telecommunications, or information resources. This information includes security-related policy, regulations, standards, and guidance. A portion of the information is channelled through the senior designated official for each agency (see Federal Information Resources Management Regulation [FIRMR] Part 201-2). Agencies are expected to have mechanisms in place to distribute the information the senior designated official receives.

Computer security-related information is also available from private and federal professional societies and groups. These groups will often provide the information as a public service, although some private groups charge a fee for it. However, even for information that is free or inexpensive, the costs associated with personnel gathering the information can be high.

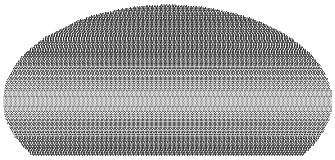
Internal security-related information, such as which procedures were effective, virus infections, security problems, and solutions, need to be shared within an organization. Often this information is specific to the operating environment and culture of the organization.

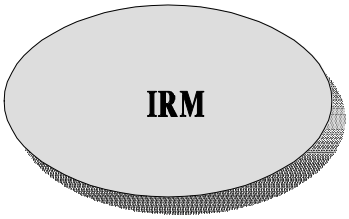
A computer security program administered at the organization level can provide a way to collect the internal security-related information and distribute it as needed throughout the organization. Sometimes an organization can also share this information with external groups. See Figure 6.3.

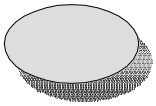
Another use of an effective conduit of information is to increase the central computer security program's ability to influence external and internal policy decisions. If the central computer security program office can represent the entire organization, then its advice is more likely to be heeded by upper management and external organizations. However, to be effective, there should be excellent communication between the system-level computer security programs and the organization level. For example, if an organization were considering consolidating its mainframes into one site (or considering distributing the processing currently done at one site), personnel at the central program could provide initial opinions about the security implications. However, to speak authoritatively, central program personnel would have to actually know the security impacts of the proposed change – information that would have to be obtained from the system-level computer security program.

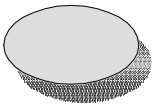
Besides being able to help an organization use information more cost effectively, a computer security program can also help an organization better spend its scarce security dollars. Organizations can develop expertise and then share it, reducing the need to contract out repeatedly for similar services. The central computer security program can help facilitate information sharing.

> An organization's components may develop specialized expertise, which can be shared among components. For example, one operating unit may primarily use UNIX and have developed skills in UNIX security. A second operating unit (with only one UNIX machine), may concentrate on MVS security and rely on the first unit's knowledge and skills for its UNIX machine.




IRM





Besides allowing an organization to share expertise and, therefore, save money, a central computer security program can use its position to consolidate requirements so the organization can negotiate discounts based on volume purchasing of security hardware and software. It also facilitates such activities as strategic planning and organizationwide incident handling and security trend analysis.

6.2.3 Central Enforcement and Oversight

Besides helping an organization improve the economy and efficiency of its computer security program, a centralized program can include an independent evaluation or enforcement function to ensure that organizational subunits are cost-effectively securing resources and following applicable policy. While the Office of the Inspector General (OIG) and external organizations, such as the General Accounting Office (GAO), also perform a valuable evaluation role, they operate outside the regular management channels. Chapters 8 and 9 further discuss the role of independent audit.

There are several reasons for having an oversight function within the regular management channel. First, computer security is an important component in the management of organizational resources. This is a responsibility that cannot be transferred or abandoned. Second, maintaining an internal oversight function allows an organization to find and correct problems without the potential embarrassment of an IG or GAO audit or investigation. Third, the organization may find different problems from those that an outside organization may find. The organization understands its assets, threats, systems, and procedures better than an external organization; additionally, people may have a tendency to be more candid with insiders.

6.3 Elements of an Effective Central Computer Security Program

For a central computer security program to be effective, it should be an established part of organization management. If system managers and applications owners do not need to consistently interact with the security program, then it can become an empty token of upper management's "commitment to security."

Stable Program Management Function. A well-established program will have a program manager recognized within the organization as the central computer security program manager. In addition, the program will be staffed with able personnel, and links will be established between the program management function and computer security personnel in other parts of the organization. A computer security program is a complex function that needs a stable base from which to direct the management of such security resources as information and money. The benefits of an oversight function cannot be achieved if the computer security program is not recognized within an organization as having expertise and authority.

Stable Resource Base. A well-established program will have a stable resource base in terms of personnel, funds, and other support. Without a stable resource base, it is impossible to plan and execute programs and projects effectively.

Existence of Policy. Policy provides the foundation for the central computer security program and is the means for documenting and promulgating important decisions about computer security. A central computer security program should also publish standards, regulations, and guidelines that implement and expand on policy. (See Chapter 5.)

Published Mission and Functions Statement. A published mission statement grounds the central computer security program into the unique operating environment of the organization. The statement clearly establishes the function of the computer security program and defines responsibilities for both the computer security program and other related programs and entities. Without such a statement, it is impossible to develop criteria for evaluating the effectiveness of the program.

Long-Term Computer Security Strategy. A well-established program explores and develops long-term strategies to incorporate computer security into the next generation of information technology. Since the computer and telecommunications field moves rapidly, it is essential to plan for future operating environments.

Compliance Program. A central computer security program needs to address compliance with national policies and requirements, as well as organization-specific requirements. National requirements include those prescribed under the Computer Security Act of 1987, OMB Circular A-130, the FIRMR, and Federal Information Processing Standards.

Intraorganizational Liaison. Many offices within an organization can affect computer security. The Information Resources Management organization and physical security office are two obvious examples. However, computer security often overlaps with other offices, such as safety, reliability and quality assurance, internal control, or the Office of the Inspector General. An effective program should have established relationships with these groups in order to integrate computer security into the organization's management. The relationships should encompass more than just the sharing of information; the offices should influence each other.

Example

Agency IRM offices engage in strategic and tactical planning for both information and information technology, in accordance with the Paperwork Reduction Act and OMB Circular A-130. Security should be an important component of these plans. The security needs of the agency should be reflected in the information technology choices and the information needs of the agency should be reflected in the security program.

Liaison with External Groups. There are many sources of computer security information, such as

NIST's Computer Security Program Managers' Forum, computer security clearinghouse, and the Forum of Incident Response and Security Teams (FIRST). An established program will be knowledgeable of and will take advantage of external sources of information. It will also be a provider of information.

6.4 System-Level Computer Security Programs

While the central program addresses the entire spectrum of computer security for an organization, system-level programs ensure appropriate and cost-effective security for each system.[59] This includes influencing decisions about what controls to implement, purchasing and installing technical controls, day-to-day computer security administration, evaluating system vulnerabilities, and responding to security problems. It encompasses all the areas discussed in the handbook.

System-level computer security program personnel are the local advocates for computer security. The system security manager/officer raises the issue of security with the cognizant system manager and helps develop solutions for security problems. For example, has the application owner made clear the system's security requirements? Will bringing a new function online affect security, and if so, how? Is the system vulnerable to hackers and viruses? Has the contingency plan been tested? Raising these kinds of questions will force system managers and application owners to identify and address their security requirements.

6.5 Elements of Effective System-Level Programs

Like the central computer security program, many factors influence how successful a system-level computer security program is. Many of these are similar to the central program. This section addresses some additional considerations.

Security Plans. The Computer Security Act mandates that agencies develop computer security and privacy plans for sensitive systems. These plans ensure that each federal and federal interest system has appropriate and cost-effective security. System-level security personnel should be in a position to develop and implement security plans. Chapter 8 discusses the plans in more detail.

System-Specific Security Policy. Many computer security policy issues need to be addressed on a system-specific basis. The issues can vary for each system, although access control and the designation of personnel with security responsibility are likely to be needed for all systems. A cohesive and comprehensive set of security policies can be developed by using a process that

[59] As is implied by the name, an organization will typically have several system-level computer security programs. In setting up these programs, the organization should carefully examine the scope of each system-level program. System-level computer security programs may address, for example, the computing resources within an operational element, a major application, or a group of similar systems (either technologically or functionally).

derives security rules from security goals, as discussed in Chapter 5.

Life Cycle Management. As discussed in Chapter 8, security must be managed throughout a system's life cycle. This specifically includes ensuring that changes to the system are made with attention to security and that accreditation is accomplished.

Integration With System Operations. The system-level computer security program should consist of people who understand the system, its mission, its technology, and its operating environment. Effective security management usually needs to be integrated into the management of the system. Effective integration will ensure that system managers and application owners consider security in the planning and operation of the system. The system security manager/officer should be able to participate in the selection and implementation of appropriate technical controls and security procedures and should understand system vulnerabilities. Also, the system-level computer security program should be capable of responding to security problems in a timely manner.

For large systems, such as a mainframe data center, the security program will often include a manager and several staff positions in such areas as access control, user administration, and contingency and disaster planning. For small systems, such as an officewide local-area-network (LAN), the LAN administrator may have adjunct security responsibilities.

Separation From Operations. A natural tension often exists between computer security and operational elements. In many instances, operational components -- which tend to be far larger and therefore more influential -- seek to resolve this tension by embedding the computer security program in computer operations. The typical result of this organizational strategy is a computer security program that lacks independence, has minimal authority, receives little management attention, and has few resources. As early as 1978, GAO identified this organizational mode as one of the principal basic weaknesses in federal agency computer security programs.[60] System-level programs face this problem most often.

This conflict between the need to be a part of system management and the need for independence has several solutions. The basis of many of the solutions is a link between the computer security program and upper management, often through the central computer security program. A key requirement of this setup is the existence of a reporting structure that does not include system management. Another possibility is for the computer security program to be completely independent of system management and to report directly to higher management. There are many hybrids and permutations, such as co-location of computer security and systems management staff but separate reporting (and supervisory) structures. Figure 6.4 presents *one example* of

[60] General Accounting Office, "Automated System Security -- Federal Agencies Should Strengthen Safeguards Over Personal and Other Sensitive Data," GAO Report LCD 78-123, Washington, DC, 1978.

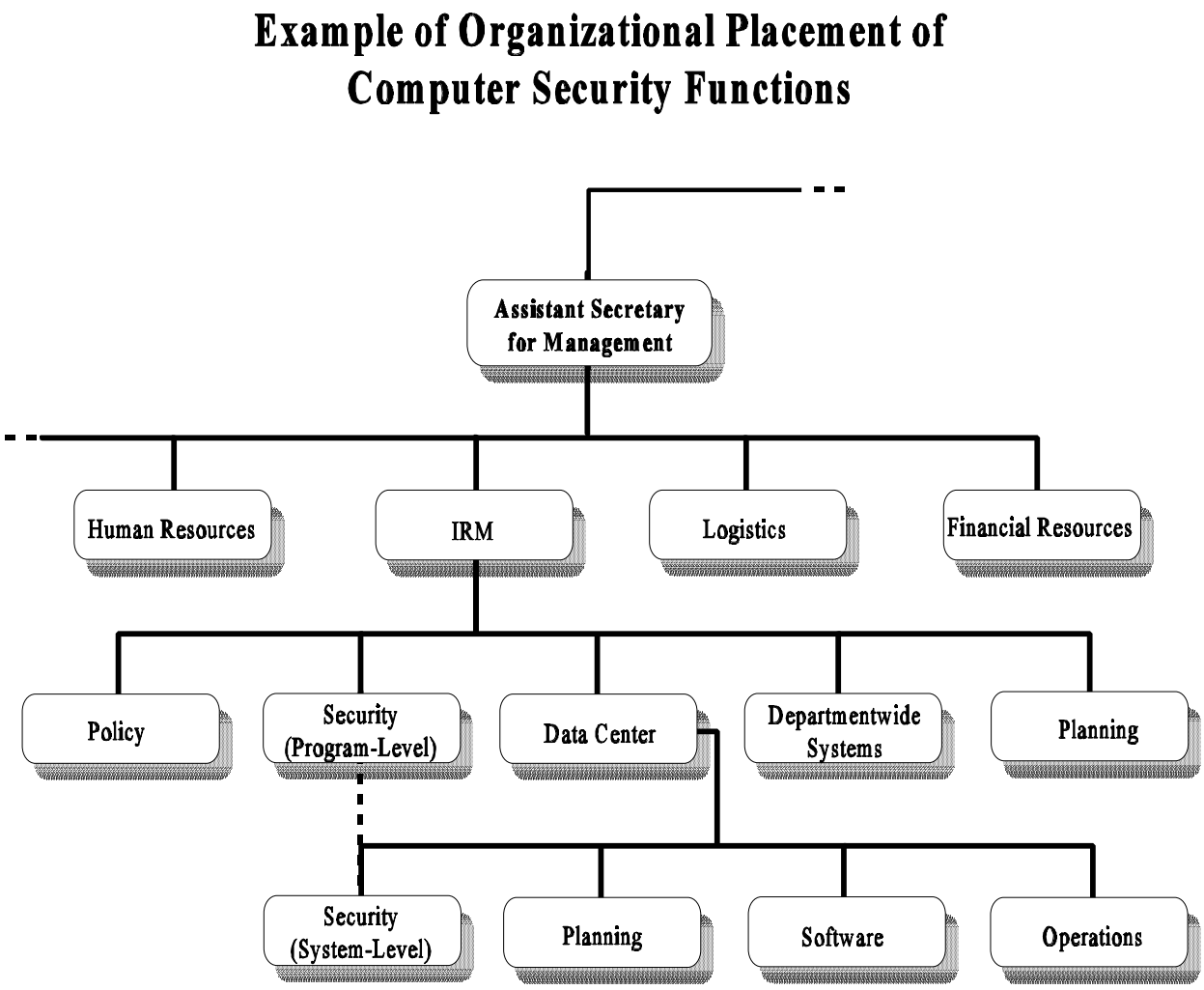


Figure 6.4 illustrates one example of the placement of the computer security program-level and system-level functions. The program-level function is located within the IRM organization and sets policy for the organization as a whole. The system-level function, located within the Data Center, provides for day-to-day security at that site. Note that, although not pictured, other system-level programs may exist for other facilities (e.g., under another Assistant Secretary).

Figure 6.4

placement of the computer security program within a typical Federal agency.[61]

[61] No implication that this structure is ideal is intended.

6.6 Central and System-Level Program Interactions

A system-level program that is not integrated into the organizational program may have difficulty influencing significant areas affecting security. The system-level computer security program implements the policies, guidance, and regulations of the central computer security program. The system-level office also learns from the information disseminated by the central program and uses the experience and expertise of the entire organization. The system-level computer security program further distributes information to systems management as appropriate.

Communications, however, should not be just one way. System-level computer security programs inform the central office about their needs, problems, incidents, and solutions. Analyzing this information allows the central computer security program to represent the various systems to the organization's management and to external agencies and advocate programs and policies beneficial to the security of all the systems.

6.7 Interdependencies

The general purpose of the computer security program, to improve security, causes it to overlap with other organizational operations as well as the other security controls discussed in the handbook. The central or system computer security program will address most controls at the policy, procedural, or operational level.

Policy. Policy is issued to establish the computer security program. The central computer security program(s) normally produces policy (and supporting procedures and guidelines) concerning general and organizational security issues and often issue-specific policy. However, the system-level computer security program normally produces policy for that system. Chapter 5 provides additional guidance.

Life Cycle Management. The process of securing a system over its life cycle is the role of the system-level computer security program. Chapter 8 addresses these issues.

Independent Audit. The independent audit function described in Chapters 8 and 9 should complement a central computer security program's compliance functions.

6.8 Cost Considerations

This chapter discussed how an organizationwide computer security program can manage security resources, including financial resources, more effectively. The cost considerations for a system-level computer security program are more closely aligned with the overall cost savings in having security.

The most significant direct cost of a computer security program is personnel. In addition, many programs make frequent and effective use of consultants and contractors. A program also needs funds for training and for travel, oversight, information collection and dissemination, and meetings with personnel at other levels of computer security management.

References

Federal Information Resources Management Regulations, especially 201-2. General Services Administration. Washington, DC.

General Accounting Office. *Automated Systems Security–Federal Agencies Should Strengthen Safeguards Over Personal and Other Sensitive Data*. GAO Report LCD 78-123. Washington, DC. 1978.

General Services Administration. *Information Resources Security: What Every Federal Manager Should Know*. Washington, DC.

Helsing, C., M. Swanson, and M. Todd. *Executive Guide to the Protection of Information Resources.*, Special Publication 500-169. Gaithersburg, MD: National Institute of Standards and Technology, 1989.

Helsing, C., M. Swanson, and M. Todd. *Management Guide for the Protection of Information Resources.* Special Publication 500-170. Gaithersburg, MD: National Institute of Standards and Technology, 1989.

"Managing an Organization Wide Security Program." Computer Security Institute, San Francisco, CA. (course)

Office of Management and Budget. "Guidance for Preparation of Security Plans for Federal Computer Systems That Contain Sensitive Information." OMB Bulletin 90-08. Washington, DC, 1990.

Office of Management and Budget. *Management of Federal Information Resources*. OMB Circular A-130.

Owen, R., Jr. "Security Management: Using the Quality Approach." *Proceedings of the 15th National Computer Security Conference*. Baltimore, MD: Vol. 2, 1992. pp. 584-592.

Spiegel, L. "Good LAN Security Requires Analysis of Corporate Data." *Infoworld*. 15(52), 1993. p. 49.

II. Management Controls

U.S. Congress. *Computer Security Act of 1987*. Public Law 100-235. 1988.

Chapter 7

COMPUTER SECURITY RISK MANAGEMENT

Risk is the possibility of something adverse happening. Risk management is the process of assessing risk, taking steps to reduce risk to an acceptable level and maintaining that level of risk. Though perhaps not always aware of it, individuals manage risks every day. Actions as routine as buckling a car safety belt, carrying an umbrella when rain is forecast, or writing down a list of things to do rather than trusting to memory fall into the purview of risk management. People recognize various threats to their best interests and take precautions to guard against them or to minimize their effects.

> Management is concerned with many types of risk. Computer security risk management addresses risks which arise from an organization's use of information technology.

Both government and industry routinely manage a myriad of risks. For example, to maximize the return on their investments, businesses must often decide between aggressive (but high-risk) and slow-growth (but more secure) investment plans. These decisions require analysis of risk, relative to potential benefits, consideration of alternatives, and, finally, implementation of what management determines to be the best course of action.

> Risk assessment often produces an important side benefit -- indepth knowledge about a system and an organization as risk analysts try to figure out how systems and functions are interrelated.

While there are many models and methods for risk management, there are several basic activities and processes that should be performed. In discussing risk management, it is important to recognize its basic, most fundamental assumption: computers cannot ever be fully secured. There is always risk, whether it is from a trusted employee who defrauds the system or a fire that destroys critical resources. Risk management is made up of two primary and one underlying activities; risk assessment and risk mitigation are the primary activities and uncertainty analysis is the underlying one.

7.1 Risk Assessment

Risk assessment, the process of analyzing and interpreting risk, is comprised of three basic activities: (1) determining the assessment's scope and methodology; (2) collecting and analyzing

data; and 3) interpreting the risk analysis results.[62]

7.1.1 Determining the Assessment's Scope and Methodology

The first step in assessing risk is to identify the system under consideration, the part of the system that will be analyzed, and the analytical method including its level of detail and formality.

> A risk assessment can focus on many different areas such as: technical and operational controls to be designed into a new application, the use of telecommunications, a data center, or an entire organization.

The assessment may be focused on certain areas where either the degree of risk is unknown or is known to be high. Different parts of a system may be analyzed in greater or lesser detail. Defining the scope and boundary can help ensure a cost-effective assessment. Factors that influence scope include what phase of the life cycle a system is in: more detail might be appropriate for a new system being developed than for an existing system undergoing an upgrade. Another factor is the relative importance of the system under examination: the more essential the system, the more thorough the risk analysis should be. A third factor may be the magnitude and types of changes the system has undergone since the last risk analysis. The addition of new interfaces would warrant a different scope than would installing a new operating system.

Methodologies can be formal or informal, detailed or simplified, high or low level, quantitative (computationally based) or qualitative (based on descriptions or rankings), or a combination of these. No single method is best for all users and all environments.

How the boundary, scope, and methodology are defined will have major consequences in terms of (1) the total amount of effort spent on risk management and (2) the type and usefulness of the assessment's results. The boundary and scope should be selected in a way that will produce an outcome that is clear, specific, and useful to the system and environment under scrutiny.

7.1.2 Collecting and Analyzing Data

> Good documentation of risk assessments will make later risk assessments less time consuming and, if a question arises, will help explain why particular security decisions were made.

Risk has many different components: assets, threats, vulnerabilities, safeguards, consequences, and likelihood. This examination normally includes gathering data about the threatened area *and* synthesizing

[62] Many different terms are used to describe risk management and its elements. The definitions used in this paper are based on the NIST Risk Management Framework.

and analyzing the information to make it useful.

Because it is possible to collect much more information than can be analyzed, steps need to be taken to limit information gathering and analysis. This process is called *screening*. A risk management effort should focus on those areas that result in the greatest consequence to the organization (i.e., can cause the most harm). This can be done by ranking threats and assets.

A risk management methodology does not necessarily need to analyze each of the components of risk separately. For example, assets/consequences or threats/likelihoods may be analyzed together.

Asset Valuation. These include the information, software, personnel, hardware, and physical assets (such as the computer facility). The value of an asset consists of its intrinsic value and the near-term impacts and long-term consequences of its compromise.

Consequence Assessment. The consequence assessment estimates the degree of harm or loss that could occur. *Consequences* refers to the overall, aggregate harm that occurs, not just to the near-term or immediate impacts. While such impacts often result in disclosure, modification, destruction, or denial of service, consequences are the more significant long-term effects, such as lost business, failure to perform the system's mission, loss of reputation, violation of privacy, injury, or loss of life. The more severe the consequences of a threat, the greater the risk to the system (and, therefore, the organization).

Threat Identification. A threat is an entity or event with the potential to harm the system. Typi cal threats are errors, fraud, disgruntled employees, fires, water damage, hackers, and viruses. Threats should be identified and analyzed to determine the likelihood of their occurrence and their potential to harm assets.

In addition to looking at "big-ticket" threats, the risk analysis should investigate areas that are poorly understood, new, or undocumented. If a facility has a well-tested physical access control system, less effort to identify threats may be warranted for it than for unclear, untested software backup procedures.

The risk analysis should concentrate on those threats most likely to occur and affect important assets. In some cases, determining which threats are realistic is not possible until after the threat analysis is begun. Chapter 4 provides additional discussion of today's most prevalent threats.

Safeguard Analysis. A safeguard is any action, device, procedure, technique, or other measure that reduces a system's vulnerability to a threat. Safeguard analysis should include an examination of the effectiveness of the existing security measures. It can also identify new safeguards that could be implemented in the system; however, this is normally performed later in the risk management process.

Vulnerability Analysis. A vulnerability is a condition or weakness in (or absence of) security procedures, technical controls, physical controls, or other controls that could be exploited by a threat. Vulnerabilities are often analyzed in terms of missing safeguards. Vulnerabilities contribute to risk because they may "allow" a threat to harm the system.

The interrelationship of vulnerabilities, threats, and assets is critical to the analysis of risk. Some of these interrelationships are pictured in Figure 7.1. However, there are other interrelationships such as the presence of a vulnerability inducing a threat. (For example, a normally honest employee might be tempted to alter data when the employee sees that a terminal has been left logged on.)

Threats, Vulnerabilities, Safeguards, and Assets

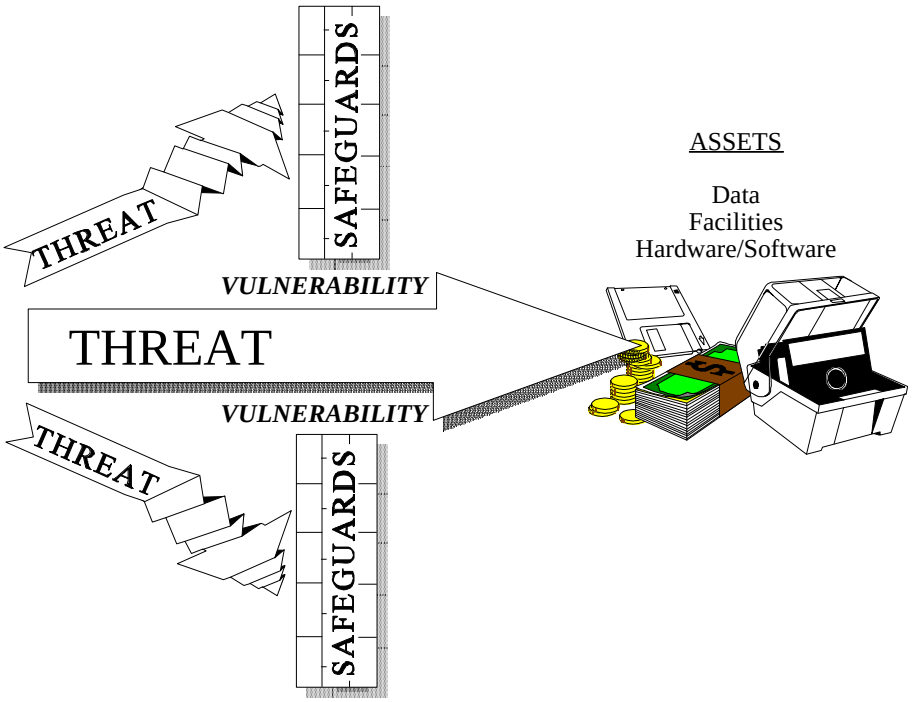


Figure 7.1 Safeguards prevent threats from harming assets. However, if an appropriate safeguard is not present, a vulnerability exists which can be exploited by a threat, thereby puttting assets at risk.

Figure 7.1

Likelihood Assessment. Likelihood is an estimation of the frequency or chance of a threat happening. A likelihood assessment considers the presence, tenacity, and strengths of threats as

well as the effectiveness of safeguards (or presence of vulnerabilities). In general, historical information about many threats is weak, particularly with regard to human threats; thus, experience in this area is important. Some threat data -- especially on physical threats such as fires or floods -- is stronger. Care needs to be taken in using any statistical threat data; the source of the data or the analysis may be inaccurate or incomplete. In general, the greater the likelihood of a threat occurring, the greater the risk.

7.1.3 Interpreting Risk Analysis Results[63]

The risk assessment is used to support two related functions: the acceptance of risk and the selection of cost-effective controls. To accomplish these functions, the risk assessment must produce a meaningful output that reflects what is truly important to the organization. Limiting the risk interpretation activity to the most significant risks is another way that the risk management process can be focused to reduce the overall effort while still yielding useful results.

Risk Analysis Results

Risk analysis results are typically represented quantitatively and/or qualitatively. Quantitative measures may be expressed in terms of reduced expected monetary losses, such as annualized loss expectancies or single occurrences of loss. Qualitative measures are descriptive, expressed in terms such as high, medium, or low, or rankings on a scale of 1 to 10.

If risks are interpreted consistently across an organization, the results can be used to prioritize systems to be secured.

Risk management can *help* a manager select the most appropriate controls; however, it is not a magic wand that instantly eliminates all difficult issues. The quality of the output depends on the quality of the input and the type of analytical methodology used. In some cases, the amount of work required to achieve high-quality input will be too costly. In other cases, achieving high-quality input may be impossible, especially for such variables as the prevalence of a particular threat or the anticipated effectiveness of a proposed safeguard. For all practical purposes, complete information is never available; uncertainty is always present. Despite these drawbacks, risk management provides a very powerful tool for analyzing the risk associated with computer systems.

7.2 Risk Mitigation

Risk mitigation involves the selection and implementation of security controls to reduce risk to a level acceptable to management, within applicable constraints. Although there is flexibility in how risk assessment is conducted, the sequence of identifying boundaries, analyzing input, and producing an output is quite natural. The process of risk mitigation has greater flexibility, and the sequence will differ more, depending on organizational culture and the purpose of the risk management

[63] The NIST Risk Management Framework refers to risk interpretation as risk measurement. The term "interpretation" was chosen to emphasize the wide variety of possible outputs from a risk assessment.

activity. Although these activities are discussed

How Risk Management Works

* There are many possible approaches to safeguard selection. Some involve looping back and reexamining risk analysis data.

Figure 7.2 shows the flow of risk management activities and processes. A major division in risk management (shown by the verticle line) is between risk assessment and risk mitigation. Both are critical parts of the risk management process. Uncertainty is always present.

Figure 7.2

in a specific sequence, they need not be performed in that sequence. In particular, the selection of safeguards and risk acceptance testing are likely to be performed simultaneously.[64]

[64] This is often viewed as a circular, iterative process.

7.2.1 Selecting Safeguards

A primary function of computer security risk management is the identification of appropriate controls. In designing (or reviewing) the security of a system, it may be obvious that some controls should be added (e.g., because they are required by law or because they are clearly cost-effective). It may also be just as obvious that other controls may be too expensive (considering both monetary and nonmonetary factors). For example, it may be immediately apparent to a manager that closing and locking the door to a particular room that contains local area network equipment is a needed control, while posting a guard at the door would be too expensive and not user-friendly.

In every assessment of risk, there will be many areas for which it will not be obvious what kind of controls are appropriate. Even considering only monetary issues, such as whether a control would cost more than the loss it is supposed to prevent, the selection of controls is not simple. However, in selecting appropriate controls, managers need to consider many factors, including:

- organizational policy, legislation, and regulation;
- safety, reliability, and quality requirements;
- system performance requirements;
- timeliness, accuracy, and completeness requirements;
- the life cycle costs of security measures;
- technical requirements; and
- cultural constraints.

What Is a *What If* Analysis?

A *what if* analysis looks at the costs and benefits of various combinations of controls to determine the optimal combination for a particular circumstance. In this simple example (which addresses only one control), suppose that hacker break-ins alert agency computer security personnel to the security risks of using passwords. They may wish to consider replacing the password system with stronger identification and authentication mechanisms, or just strengthening their password procedures. First, the **status quo** is examined. The system in place puts minimal demands upon users and system administrators, but the agency has had three hacker break-ins in the last six months.

What if passwords are strengthened? Personnel may be required to change passwords more frequently or may be required to use a numeral or other nonalphabetic character in their password. There are no direct monetary expenditures, but staff and administrative overhead (e.g., training and replacing forgotten passwords) is increased. Estimates, however, are that this will reduce the number of successful hacker break-ins to three or four per year.

What if stronger identification and authentication technology is used? The agency may wish to implement stronger safeguards in the form of one-time cryptographic-based passwords so that, even if a password were obtained, it would be useless. Direct costs may be estimated at $45,000, and yearly recurring costs at $8,000. An initial training program would be required, at a cost of $17,500. The agency estimates, however, that this would prevent virtually all break-ins.

Computer security personnel use the results of this analysis to make a recommendation to their management officer, who then weighs the costs and benefits, takes into account other constraints (e.g., budget), and selects a solution.

One method of selecting safeguards uses a "what if" analysis. With this method, the effect of adding various safeguards (and, therefore, reducing vulnerabilities) is tested to see what difference each makes with regard to cost, effectiveness, and other relevant factors, such as those listed above. Trade-offs among the factors can be seen. The analysis of trade-offs also supports the acceptance of residual risk, discussed below. This method typically involves multiple iterations of the risk analysis to see how the proposed changes affect the risk analysis result.

Another method is to categorize types of safeguards and recommend implementing them for various levels of risk. For example, stronger controls would be implemented on high-risk systems than on low-risk systems. This method normally does not require multiple iterations of the risk analysis.

As with other aspects of risk management, screening can be used to concentrate on the highest-risk areas. For example once could focus on risks with very severe consequences, such as a very high dollar loss or loss of life or on the threats that are most likely to occur.

7.2.2 Accept Residual Risk

At some point, management needs to decide if the operation of the computer system is acceptable, given the kind and severity of remaining risks. Many managers do not fully understand computer-based risk for several reasons: (1) the type of risk may be different from risks previously associated with the organization or function; (2) the risk may be technical and difficult for a lay person to understand, or (3) the proliferation and decentralization of computing power can make it difficult to identify key assets that may be at risk.

Risk acceptance, like the selection of safeguards, should take into account various factors besides those addressed in the risk assessment. In addition, risk acceptance should take into account the limitations of the risk assessment. (See the section below on uncertainty.) Risk acceptance is linked to the selection of safeguards since, in some cases, risk may have to be accepted because safeguards are too expensive (in either monetary or nonmonetary factors).

Within the federal government, the acceptance of risk is closely linked with the authorization to use a computer system, often called *accreditation*, discussed in Chapters 8 and 9. Accreditation is the acceptance of risk by management resulting in a formal approval for the system to become operational or remain so. As discussed earlier in this chapter, one of the two primary functions of risk management is the interpretation of risk for the purpose of risk acceptance.

7.2.3 Implementing Controls and Monitoring Effectiveness

Merely selecting appropriate safeguards does not reduce risk; those safeguards need to be effectively implemented. Moreover, to continue to be effective, risk management needs to be an ongoing process. This requires a periodic assessment and improvement of safeguards and re-

analysis of risks. Chapter 8 discusses how periodic risk assessment is an integral part of the overall management of a system. (See especially the diagram on page 83.)

The risk management process normally produces security requirements that are used to design, purchase, build, or otherwise obtain safeguards or implement system changes. The integration of risk management into the life cycle process is discussed in Chapter 8.

7.3 Uncertainty Analysis

> While uncertainty is always present it should not invalidate a risk assessment. Data and models, while imperfect, can be good enough for a given purpose.

Risk management often must rely on speculation, best guesses, incomplete data, and many unproven assumptions. The uncertainty analysis attempts to document this so that the risk management results can be used knowledgeably. There are two primary sources of uncertainty in the risk management process: (1) a lack of confidence or precision in the risk management model or methodology and (2) a lack of sufficient information to determine the exact value of the elements of the risk model, such as threat frequency, safeguard effectiveness, or consequences.

The risk management framework presented in this chapter is a generic description of risk management elements and their basic relationships. For a methodology to be useful, it should further refine the relationships and offer some means of screening information. In this process, assumptions may be made that do not accurately reflect the user's environment. This is especially evident in the case of safeguard selection, where the number of relationships among assets, threats, and vulnerabilities can become unwieldy.

The data are another source of uncertainty. Data for the risk analysis normally come from two sources: statistical data and expert analysis. Statistics and expert analysis can sound more authoritative than they really are. There are many potential problems with statistics. For example, the sample may be too small, other parameters affecting the data may not be properly accounted for, or the results may be stated in a misleading manner. In many cases, there may be insufficient data. When expert analysis is used to make projections about future events, it should be recognized that the projection is subjective and is based on assumptions made (but not always explicitly articulated) by the expert.

7.4 Interdependencies

Risk management touches on every control and every chapter in this handbook. It is, however, most closely related to life cycle management and the security planning process. The requirement to perform risk management is often discussed in organizational policy and is an issue for organizational oversight. These issues are discussed in Chapters 5 and 6.

7.5 Cost Considerations

The building blocks of risk management presented in this chapter can be used creatively to develop methodologies that concentrate expensive analysis work where it is most needed. Risk management can become expensive very quickly if an expansive boundary and detailed scope are selected. It is very important to use screening techniques, as discussed in this chapter, to limit the overall effort. The goals of risk management should be kept in mind as a methodology is selected or developed. The methodology should concentrate on areas where identification of risk and the selection of cost-effective safeguards are needed.

The cost of different methodologies can be significant. A "back-of-the-envelope" analysis or high-medium-low ranking can often provide all the information needed. However, especially for the selection of expensive safeguards or the analysis of systems with unknown consequences, more in-depth analysis may be warranted.

References

Caelli, William, Dennis Longley, and Michael Shain. *Information Security Handbook*. New York, NY: Stockton Press, 1991.

Carroll, J.M. *Managing Risk: A Computer-Aided Strategy*. Boston, MA: Butterworths 1984.

Gilbert, Irene. *Guide for Selecting Automated Risk Analysis Tools*. Special Publication 500-174. Gaithersburg, MD: National Institute of Standards and Technology, October 1989.

Jaworski, Lisa. "Tandem Threat Scenarios: A Risk Assessment Approach." *Proceedings of the 16th National Computer Security Conference*, Baltimore, MD: Vol. 1, 1993. pp. 155-164.

Katzke, Stuart. "A Framework for Computer Security Risk Management." *8th Asia Pacific Information Systems Control Conference Proceedings*. EDP Auditors Association, Inc., Singapore, October 12-14, 1992.

Levine, M. "Audit Serve Security Evaluation Criteria." *Audit Vision*. 2(2), 1992. pp. 29-40.

National Bureau of Standards. *Guideline for Automatic Data Processing Risk Analysis*. Federal Information Processing Standard Publication 65. August 1979.

National Institute of Standards and Technology. *Guideline for the Analysis of Local Area Network Security*. Federal Information Processing Standard Publication 191. November 1994.

O'Neill, M., and F. Henninge, Jr., "Understanding ADP System and Network Security Considerations and Risk Analysis." *ISSA Access*. 5(4), 1992. pp. 14-17.

Proceedings, 4th International Computer Security Risk Management Model Builders Workshop. University of Maryland, National Institute of Standards and Technology, College Park, MD, August 6-8, 1991.

Proceedings, 3rd International Computer Security Risk Management Model Builders Workshop, Los Alamos National Laboratory, National Institute of Standards and Technology, National Computer Security Center, Santa Fe, New Mexico, August 21-23, 1990.

Proceedings, 1989 Computer Security Risk Management Model Builders Workshop, AIT Corporation, Communications Security Establishment, National Computer Security Center, National Institute of Standards and Technology, Ottawa, Canada, June 20-22, 1989.

Proceedings, 1988 Computer Security Risk Management Model Builders Workshop, Martin Marietta, National Bureau of Standards, National Computer Security Center, Denver, Colorado, May 24-26, 1988.

Spiegel, L. "Good LAN Security Requires Analysis of Corporate Data." *Infoworld*. 15(52), 1993. p. 49.

Wood, C. "Building Security Into Your System Reduces the Risk of a Breach." *LAN Times*. 10(3), 1993. p. 47.

Wood C., et al., *Computer Security: A Comprehensive Controls Checklist*. New York, NY: John Wiley & Sons, 1987.

Chapter 8

SECURITY AND PLANNING IN THE COMPUTER SYSTEM LIFE CYCLE

Like other aspects of information processing systems, security is most effective and efficient if planned and managed throughout a computer system's life cycle, from initial planning, through design, implementation, and operation, to disposal.[65] Many security-relevant events and analyses occur during a system's life. This chapter explains the relationship among them and how they fit together.[66] It also discusses the important role of security planning in helping to ensure that security issues are addressed comprehensively.

This chapter examines:

- system security plans,
- the components of the computer system life cycle,
- the benefits of integrating security into the computer system life cycle, and
- techniques for addressing security in the life cycle.

8.1 Computer Security Act Issues for Federal Systems

Planning is used to help ensure that security is addressed in a comprehensive manner throughout a system's life cycle. For federal systems, the Computer Security Act of 1987 sets forth a statutory requirement for the preparation of computer security plans for all sensitive systems.[67] The intent and spirit of the Act is to improve computer security in the federal government, not to create paperwork. In keeping with this intent, the Office of Management and Budget (OMB) and NIST have guided agencies toward a planning process that emphasizes good planning and management of computer security within an agency and for each computer system. As emphasized in this

---

[65] A computer system refers to a collection of processes, hardware, and software that perform a function. This includes applications, networks, or support systems.

[66] Although this chapter addresses a life cycle process that starts with system initiation, the process can be initiated at any point in the life cycle.

[67] An organization will typically have many computer security plans. However, it is not necessary that a separate and distinct plan exist for every physical system (e.g., PCs). Plans may address, for example, the computing resources within an operational element, a major application, or a group of similar systems (either technologically or functionally).

chapter, computer *security* management should be a part of computer *systems* management. The benefit of having a distinct computer security plan is to ensure that computer security is not overlooked.

The Act required the submission of plans to NIST and the National Security Agency (NSA) for review and comment, a process which has been completed. Current guidance on implementing the Act requires agencies to obtain independent review of computer security plans. This review may be internal or external, as deemed appropriate by the agency.

> "The purpose of the system security plan is to provide a basic overview of the security and privacy requirements of the subject system and the agency's plan for meeting those requirements. The system security plan may also be viewed as documentation of the structured process of planning adequate, cost-effective security protection for a system."
>
> - OMB Bulletin 90-08

A "typical" plan briefly describes the important security considerations for the system and provides references to more detailed documents, such as system security plans, contingency plans, training programs, accreditation statements, incident handling plans, or audit results. This enables the plan to be used as a management tool without requiring repetition of existing documents. For smaller systems, the plan may include all security documentation. As with other security documents, if a plan addresses specific vulnerabilities or other information that could compromise the system, it should be kept private. It also has to be kept up-to-date.

8.2 Benefits of Integrating Security in the Computer System Life Cycle

Although a computer security plan can be developed for a system at any point in the life cycle, the recommended approach is to draw up the plan at the beginning of the computer system life cycle. Security, like other aspects of a computer system, is best managed if planned for throughout the computer system life cycle. It has long been a tenet of the computer community that it costs ten times more to add a feature in a system *after* it has been designed than to include the feature in the system at the initial design phase. The principal reason for implementing security during a system's development is that it is more difficult to implement it later (as is usually reflected in the higher costs of doing so). It also tends to disrupt ongoing operations.

> Different people can provide security input throughout the life cycle of a system, including the accrediting official, data users, systems users, and system technical staff.

Security also needs to be incorporated into the later phases of the computer system life cycle to help ensure that security keeps up with changes in the system's environment, technology, procedures, and personnel. It also ensures that security is considered in system upgrades, including the purchase of new components or the design of new modules. Adding new security

controls to a system after a security breach, mishap, or audit can lead to haphazard security that can be more expensive and less effective that security that is already integrated into the system. It can also significantly degrade system performance. Of course, it is virtually impossible to anticipate the whole array of problems that may arise during a system's lifetime. Therefore, it is generally useful to update the computer security plan at least at the end of each phase in the life cycle and after each re-accreditation. For many systems, it may be useful to update the plan more often.

Life cycle management also helps document security-relevant decisions, in addition to helping assure management that security is fully considered in all phases. This documentation benefits system management officials as well as oversight and independent audit groups. System management personnel use documentation as a self-check and reminder of why decisions were made so that the impact of changes in the environment can be more easily assessed. Oversight and independent audit groups use the documentation in their reviews to verify that system management has done an adequate job and to highlight areas where security may have been overlooked. This includes examining whether the documentation accurately reflects how the system is actually being operated.

Within the federal government, the Computer Security Act of 1987 and its implementing instructions provide specific requirements for computer security plans. These plans are a form of documentation that helps ensure that security is considered not only during system design and development but also throughout the rest of the life cycle. Plans can also be used to be sure that requirements of Appendix III to OMB Circular A-130, as well as other applicable requirements, have been addressed.

8.3 Overview of the Computer System Life Cycle

There are many models for the computer system life cycle but most contain five basic phases, as pictured in Figure 8.1.

- *Initiation*. During the initiation phase, the need for a system is expressed and the purpose of the system is documented.

- *Development/Acquisition*. During this phase the system is designed, purchased, programmed, developed, or otherwise constructed. This phase often consists of other defined cycles, such as the system development cycle or the acquisition cycle.

- *Implementation*. After initial system testing, the system is installed or fielded.

- *Operation/Maintenance*. During this phase the system performs its work. The system is almost always modified by the addition of hardware and software and by numerous other

events.

- *Disposal*. The computer system is disposed of once the transition to a new computer system is completed.

Each phase can apply to an entire system, a new component or module, or a system upgrade. As with other aspects of systems management, the level of detail and analysis for each activity described here is determined by many factors including size, complexity, system cost, and sensitivity.

> Many different "life cycles" are associated with computer systems, including the system development, acquisition, and information life cycles.

Many people find the concept of a computer system life cycle confusing because many cycles occur within the broad framework of the *entire* computer system life cycle. For example, an organization could develop a system, using a system *development* life cycle. During the system's life, the organization might purchase new components, using the *acquisition* life cycle.

Moreover, the computer system life cycle itself is merely one component of other life cycles. For example, consider the *information life cycle*. Normally information, such as personnel data, is used much longer than the life of one computer system. If an employee works for an organization for thirty years and collects retirement for another twenty, the employee's automated personnel record will probably pass through many different organizational computer systems owned by the company. In addition, parts of the information will also be used in other computer systems, such as those of the Internal Revenue Service and the Social Security Administration.

8.4 Security Activities in the Computer System Life Cycle[68]

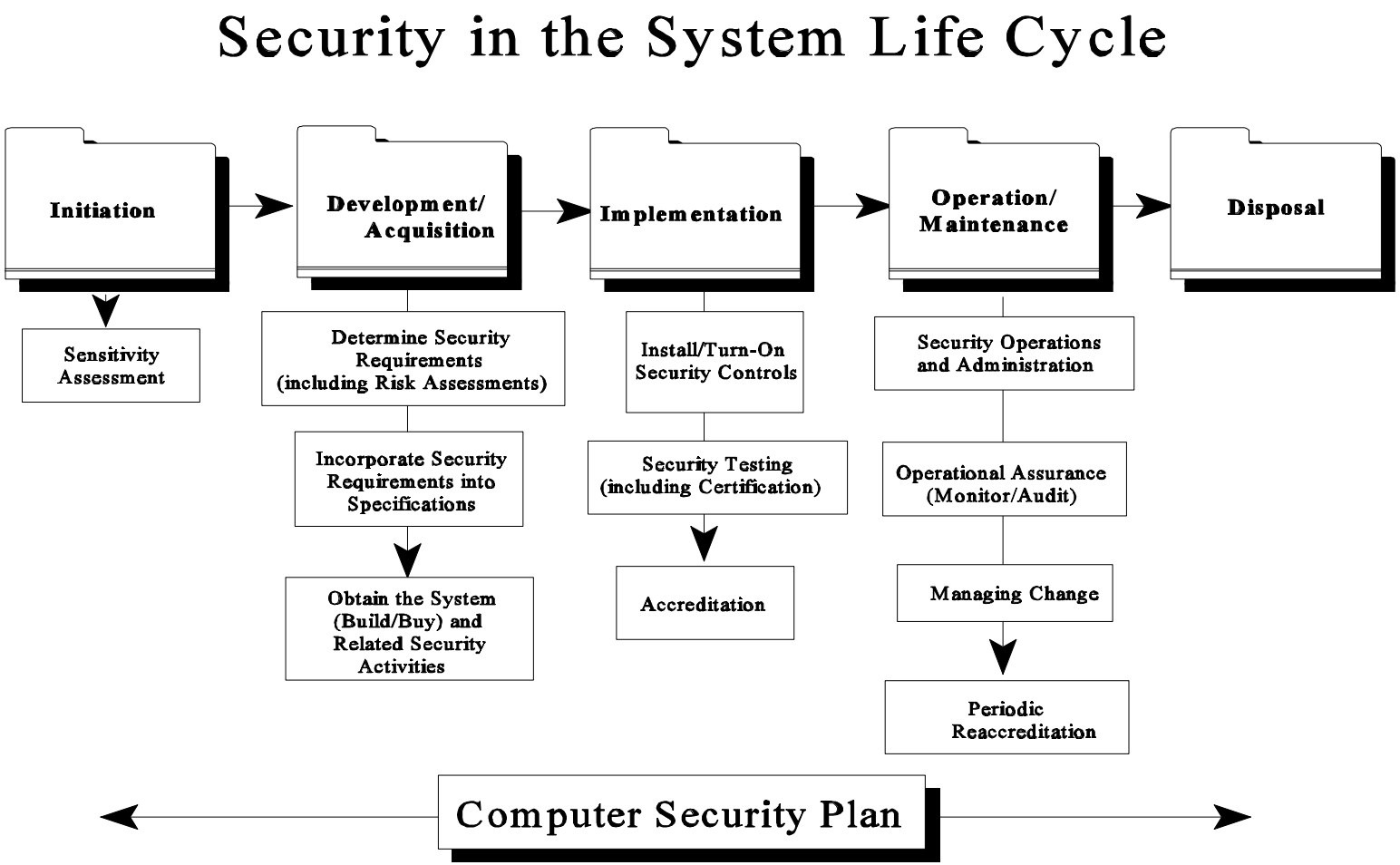
This section reviews the security activities that arise in each stage of the computer system life cycle. (See Figure 8.1.)

8.4.1 Initiation

The conceptual and early design process of a system involves the discovery of a need for a new system or enhancements to an existing system; early ideas as to system characteristics and proposed functionality; brainstorming sessions on architectural, performance, or functional system aspects; and environmental, financial, political, or other constraints. At the same time, the basic *security* aspects of a system should be developed along with the early system design. This can be

[68] For brevity and because of the uniqueness of each system, none of these discussions can include the details of all possible security activities at any particular life cycle phase.

done through a *sensitivity assessment.*



The life cycle process described in this chapter consists of five separate phases. Security issues are present in each.

Figure 8.1

8.4.1.1 Conducting a Sensitivity Assessment

> The definition of *sensitive* is often misconstrued. *Sensitive* is synonymous with *important* or *valuable*. Some data is sensitive because it must be kept confidential. Much more data, however, is sensitive because its integrity or availability must be assured. The Computer Security Act and OMB Circular A-130 clearly state that information is sensitive if its unauthorized disclosure, modification (i.e., loss of integrity), or unavailability would harm the agency. In general, the more important a system is to the mission of the agency, the more sensitive it is.

A *sensitivity assessment* looks at the sensitivity of both the information to be processed and the system itself. The assessment should consider legal implications, organization policy (including federal and agency policy if a federal system), and the functional needs of the system. Sensitivity is normally expressed in terms of integrity, availability, and confidentiality. Such factors as the importance of the system to the organization's mission and the consequences of unauthorized modification, unauthorized disclosure, or unavailability of the system or data need to be examined when assessing sensitivity. To address these types of issues, the people who use or own the system or information should participate in the assessment.

A sensitivity assessment should answer the following questions:

- What information is handled by the system?
- What kind of potential damage could occur through error, unauthorized disclosure or modification, or unavailability of data or the system?
- What laws or regulations affect security (e.g., the Privacy Act or the Fair Trade Practices Act)?
- To what threats is the system or information particularly vulnerable?
- Are there significant environmental considerations (e.g., hazardous location of system)?
- What are the security-relevant characteristics of the user community (e.g., level of technical sophistication and training or security clearances)?
- What internal security standards, regulations, or guidelines apply to this system?

The sensitivity assessment starts an analysis of security that continues throughout the life cycle. The assessment helps determine if the project needs special security oversight, if further analysis is

needed before committing to begin system development (to ensure feasibility at a reasonable cost), or in rare instances, whether the security requirements are so strenuous and costly that system development or acquisition will not be pursued. The sensitivity assessment can be included with the system initiation documentation either as a separate document or as a section of another planning document. The development of security features, procedures, and assurances, described in the next section, builds on the sensitivity assessment.

A sensitivity assessment can also be performed during the planning stages of system upgrades (for either upgrades being procured or developed in house). In this case, the assessment focuses on the affected areas. If the upgrade significantly affects the original assessment, steps can be taken to analyze the impact on the rest of the system. For example, are new controls needed? Will some controls become unnecessary?

8.4.2 Development/Acquisition

For most systems, the development/acquisition phase is more complicated than the initiation phase. Security activities can be divided into three parts:

- determining security features, assurances, and operational practices;
- incorporating these security requirements into design specifications; and
- actually acquiring them.

These divisions apply to systems that are designed and built in house, to systems that are purchased, and to systems developed using a hybrid approach.

During this phase, technical staff and system sponsors should actively work together to ensure that the technical designs reflect the system's security needs. As with development and incorporation of other system requirements, this process requires an open dialogue between technical staff and system sponsors. It is important to address security requirements effectively in synchronization with development of the overall system.

8.4.2.1 Determining Security Requirements

During the first part of the development/ acquisition phase, system planners define the requirements of the system. *Security requirements should be developed at the same time.* These requirements can be expressed as technical features (e.g., access controls), assurances (e.g., background checks for system developers), or operational practices (e.g., awareness and training). System security requirements, like other system requirements, are derived from a number of sources including law, policy, applicable standards and guidelines, functional needs of the system, and cost-benefit trade-offs.

Law. Besides specific laws that place security requirements on information, such as the Privacy Act of 1974, there are laws, court cases, legal opinions, and other similar legal material that may affect security directly or indirectly.

Policy. As discussed in Chapter 5, management officials issue several different types of policy. System security requirements are often derived from issue-specific policy.

Standards and Guidelines. International, national, and organizational standards and guidelines are another source for determining security features, assurances, and operational practices. Standards and guidelines are often written in an "if...then" manner (e.g., if the system is encrypting data, then a particular cryptographic algorithm should be used). Many organizations specify baseline controls for different types of systems, such as administrative, mission- or business-critical, or proprietary. As required, special care should be given to interoperability standards.

Functional Needs of the System. The purpose of security is to support the function of the system, not to undermine it. Therefore, many aspects of the function of the system will produce related security requirements.

Cost-Benefit Analysis. When considering security, cost-benefit analysis is done through risk assessment, which examines the assets, threats, and vulnerabilities of the system in order to determine the most appropriate, cost-effective safeguards (that comply with applicable laws, policy, standards, and the functional needs of the system). Appropriate safeguards are normally those whose anticipated benefits outweigh their costs. Benefits and costs include monetary and nonmonetary issues, such as prevented losses, maintaining an organization's reputation, decreased user friendliness, or increased system administration.

Risk assessment, like cost-benefit analysis, is used to support decision making. It helps managers select cost-effective safeguards. The extent of the risk assessment, like that of other cost-benefit analyses, should be commensurate with the complexity and cost (normally an indicator of complexity) of the system and the expected benefits *of the assessment*. Risk assessment is further discussed in Chapter 7.

Risk assessment can be performed during the requirements analysis phase of a procurement or the design phase of a system development cycle. Risk should also normally be assessed during the development/acquisition phase of a system upgrade. The risk assessment may be performed once or multiple times, depending upon the project's methodology.

Care should be taken in differentiating between *security* risk assessment and *project* risk analysis. Many system development and acquisition projects analyze the risk of failing to successfully complete the project – a different activity from *security* risk assessment.

8.4.2.2 Incorporating Security Requirements Into Specifications

Determining security features, assurances, and operational practices can yield significant security information and often voluminous requirements. This information needs to be validated, updated, and organized into the detailed security protection requirements and specifications used by systems designers or purchasers. Specifications can take on quite different forms, depending on the methodology used for to develop the system, or whether the system, or parts of the system, are being purchased off the shelf.

As specifications are developed, it may be necessary to update initial risk assessments. A safeguard recommended by the risk assessment could be incompatible with other requirements, or a control may be difficult to implement. For example, a security requirement that prohibits dial-in access could prevent employees from checking their e-mail while away from the office.[69]

> Developing testing specifications early can be critical to being able to cost-effectively test security features.

Besides the technical and operational controls of a system, assurance also should be addressed. The degree to which assurance (that the security features and practices can and do work correctly and effectively) is needed should be determined early. Once the desired level of assurance is determined, it is necessary to figure out how the system will be tested or reviewed to determine whether the specifications have been satisfied (to obtain the desired assurance). This applies to both system developments and acquisitions. For example, if rigorous assurance is needed, the ability to test the system or to provide another form of initial and ongoing assurance needs to be designed into the system or otherwise provided for. See Chapter 9 for more information.

8.4.2.3 Obtaining the System and Related Security Activities

During this phase, the system is actually built or bought. If the system is being built, security activities may include developing the system's security aspects, monitoring the development process itself for security problems, responding to changes, and monitoring threat. Threats or vulnerabilities that may arise during the development phase include Trojan horses, incorrect code, poorly functioning development tools, manipulation of code, and malicious insiders.

If the system is being acquired off the shelf, security activities may include monitoring to ensure security is a part of market surveys, contract solicitation documents, and evaluation of proposed systems. Many systems use a combination of development and acquisition. In this case, security activities include both sets.

---

[69] This is an example of a risk-based decision.

As the system is built or bought, choices are made about the system, which can affect security. These choices include selection of specific off-the-shelf products, finalizing an architecture, or selecting a processing site or platform. Additional security analysis will probably be necessary.

> In federal government contracting, it is often useful if personnel with security expertise participate as members of the source selection board to help evaluate the security aspects of proposals.

In addition to obtaining the system, operational practices need to be developed. These refer to human activities that take place around the system such as contingency planning, awareness and training, and preparing documentation. The chapters in the Operational Controls section of this handbook discuss these areas. These need to be developed along with the system, although they are often developed by different individuals. These areas, like technical specifications, should be considered from the beginning of the development and acquisition phase.

8.4.3 Implementation

A separate implementation phase is not always specified in some life cycle planning efforts. (It is often incorporated into the end of development and acquisition or the beginning of operation and maintenance.) However, from a security point of view, a critical security activity, *accreditation*, occurs between development and the start of system operation. The other activities described in this section, turning on the controls and testing, are often incorporated at the end of the development/acquisition phase.

8.4.3.1 Install/Turn-On Controls

While obvious, this activity is often overlooked. When acquired, a system often comes with security features disabled. These need to be enabled and configured. For many systems this is a complex task requiring significant skills. Custom-developed systems may also require similar work.

8.4.3.2 Security Testing

System security testing includes both the testing of the particular parts of the system that have been developed or acquired and the testing of the entire system. Security management, physical facilities, personnel, procedures, the use of commercial or in-house services (such as networking services), and contingency planning are examples of areas that affect the security of the entire system, but may be specified outside of the development or acquisition cycle. Since only items within the development or acquisition cycle will have been tested during system acceptance testing, separate tests or reviews may need to be performed for these additional security elements.

interim period, presumably after security upgrades have been made.

8.4.4 Operation and Maintenance

Many security activities take place during the operational phase of a system's life. In general, these fall into three areas: (1) security operations and administration; (2) operational assurance; and (3) periodic re-analysis of the security. Figure 8.2 diagrams the flow of security activities during the operational phase.

8.4.4.1 Security Operations and Administration

Operation of a system involves many security activities discussed throughout this handbook. Performing backups, holding training classes, managing cryptographic keys, keeping up with user administration and access privileges, and updating security software are some examples.

8.4.4.2 Operational Assurance

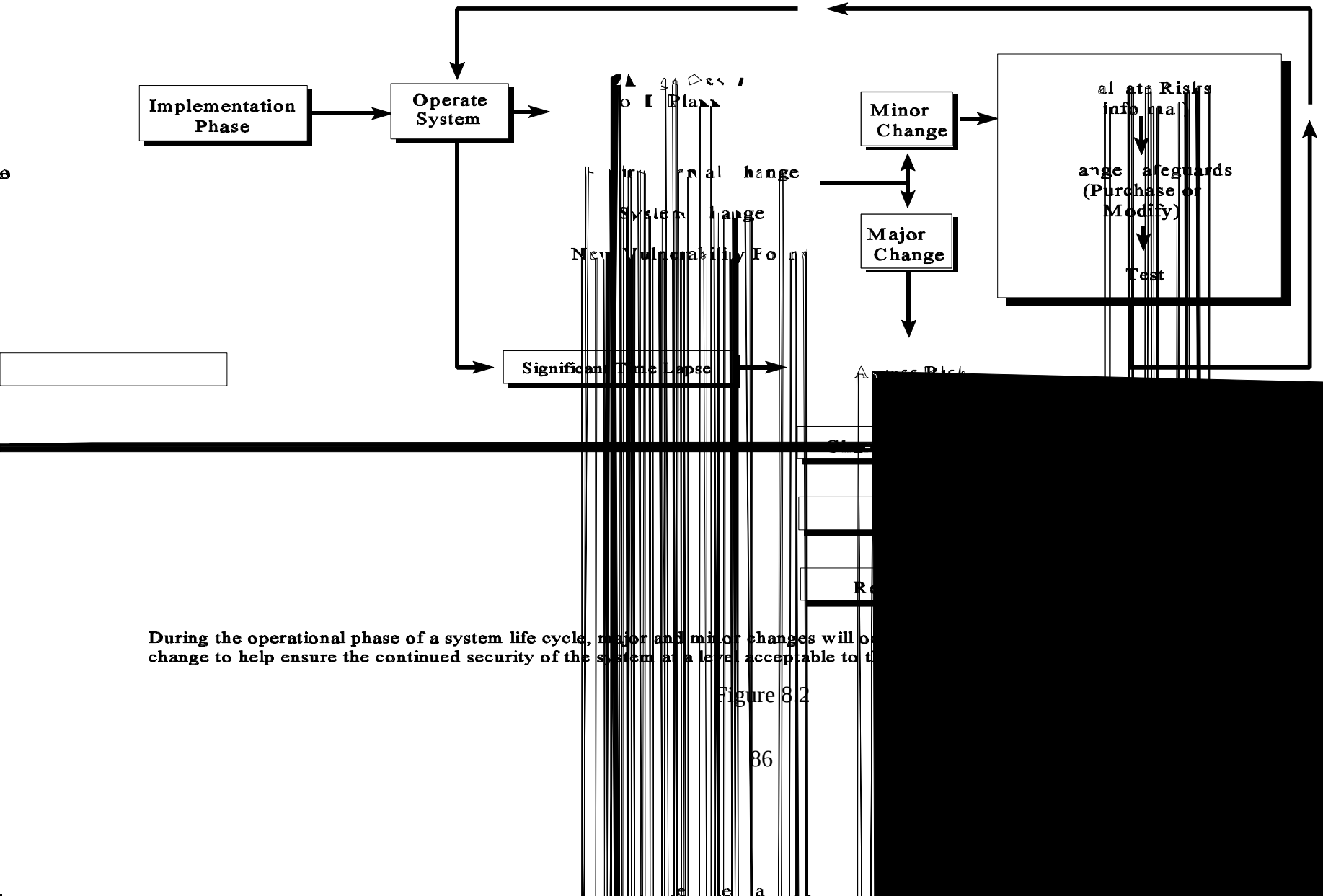
> Operational assurance examines whether a system is operated according to its current security requirements. This includes both the actions of people who operate or use the system and the functioning of technical controls.

Security is *never* perfect when a system is implemented. In addition, system users and operators discover new ways to intentionally or unintentionally bypass or subvert security. Changes in the system or the environment can create new vulnerabilities. Strict adherence to procedures is rare over time, and procedures become outdated. Thinking risk is minimal, users may tend to bypass security measures and procedures.

As shown in Figure 8.2, changes occur. Operational assurance is one way of becoming aware of these changes whether they are new vulnerabilities (or old vulnerabilities that have not been corrected), system changes, or environmental changes. Operational assurance is the process of reviewing an operational system to see that security controls, both automated and manual, are functioning correctly and effectively.

To maintain operational assurance, organizations use two basic methods: *system audits* and *monitoring*. These terms are used loosely within the computer security community and often overlap. A system audit is a one-time or periodic event to evaluate security. Monitoring refers to an ongoing activity that examines either the system or the users. In general, the more "real-time" an activity is, the more it falls into the category of monitoring. (See Chapter 9.)

Operational Phase



During the operational phase of a system life cycle, major and minor changes will o
change to help ensure the continued security of the system at a level acceptable to t

Figure 8.2

8.4.4.3 Managing Change

Computer systems and the environments in which they operate change continually. In response to various events such as user complaints, availability of new features and services, or the discovery of new threats and vulnerabilities, system managers and users modify the system and incorporate new features, new procedures, and software updates.

> Security change management helps develop new security requirements.

The environment in which the system operates also changes. Networking and interconnections tend to increase. A new user group may be added, possibly external groups or anonymous groups. New threats may emerge, such as increases in network intrusions or the spread of personal computer viruses. If the system has a configuration control board or other structure to manage technical system changes, a security specialist can be assigned to the board to make determinations about whether (and if so, how) changes will affect security.

Security should also be considered during system upgrades (and other planned changes) and in determining the impact of unplanned changes. As shown in Figure 8.2, when a change occurs or is planned, a determination is made whether the change is major or minor. A major change, such as reengineering the structure of the system, significantly affects the system. Major changes often involve the purchase of new hardware, software, or services or the development of new software modules.

An organization does not need to have a specific cutoff for major-minor change decisions. A sliding scale between the two can be implemented by using a combination of the following methods:

- *Major change*. A major change requires analysis to determine security requirements. The process described above can be used, although the analysis may focus only on the area(s) in which the change has occurred or will occur. If the original analysis and system changes have been documented throughout the life cycle, the analysis will normally be much easier. Since these changes result in significant system acquisitions, development work, or changes in policy, the system should be reaccredited to ensure that the residual risk is still acceptable.

- *Minor change*. Many of the changes made to a system do not require the extensive analysis performed for major changes, but do require some analysis. Each change can involve a limited risk assessment that weighs the pros (benefits) and cons (costs) and that can even be performed on-the-fly at meetings. Even if the analysis is conducted informally, decisions should still be appropriately documented. This process recognizes that even "small" decisions should be

risk-based.

8.4.4.4 Periodic Reaccreditation

Periodically, it is useful to formally reexamine the security of a system from a wider perspective. The analysis, which leads to reaccreditation, should address such questions as: Is the security still sufficient? Are major changes needed?

The reaccreditation should address high-level security and management concerns as well as the implementation of the security. It is not always necessary to perform a new risk assessment or certification in conjunction with the re-accreditation, but the activities support each other (and both need be performed periodically). The more extensive system changes have been, the more extensive the analyses should be (e.g., a risk assessment or re-certification). A risk assessment is likely to uncover security concerns that result in system changes. After the system has been changed, it may need testing (including certification). Management then reaccredits the system for continued operation if the risk is acceptable.

> It is important to consider legal requirements for records retention when disposing of computer systems. For federal systems, system management officials should consult with their agency office responsible for retaining and archiving federal records.

8.4.5 Disposal

The disposal phase of the computer system life cycle involves the disposition of information, hardware, and software. Information may be moved to another system, archived, discarded, or destroyed. When archiving information, consider the method for retrieving the information in the future. The technology used to create the records may not be readily available in the future.

Hardware and software can be sold, given away, or discarded. There is rarely a need to destroy hardware, except for some storage media containing confidential information that cannot be sanitized without destruction. The disposition of software needs to be in keeping with its license or other agreements with the developer, if applicable. Some licenses are

> **Media Sanitization**
>
> Since electronic information is easy to copy and transmit, information that is sensitive to disclosure often needs to be controlled throughout the computer system life cycle so that managers can ensure its proper disposition. The removal of information from a storage medium (such as a hard disk or tape) is called *sanitization*. Different kinds of sanitization provide different levels of protection. A distinction can be made between clearing information (rendering it unrecoverable by keyboard attack) and purging (rendering information unrecoverable against laboratory attack). There are three general methods of purging media: overwriting, degaussing (for magnetic media only), and destruction.

site-specific or contain other agreements that prevent the software from being transferred.

Measures may also have to be taken for the future use of data that has been encrypted, such as taking appropriate steps to ensure the secure long-term storage of cryptographic keys.

8.5 Interdependencies

Like many management controls, life cycle planning relies upon other controls. Three closely linked control areas are policy, assurance, and risk management.

Policy. The development of system-specific policy is an integral part of determining the security requirements.

Assurance. Good life cycle management provides assurance that security is appropriately considered in system design and operation.

Risk Management. The maintenance of security throughout the operational phase of a system is a process of risk management: analyzing risk, reducing risk, and monitoring safeguards. Risk assessment is a critical element in designing the security of systems and in reaccreditations.

8.6 Cost Considerations

Security is a factor throughout the life cycle of a system. Sometimes security choices are made by default, without anyone analyzing why choices are made; sometimes security choices are made carefully, based on analysis. The first case is likely to result in a system with poor security that is susceptible to many types of loss. In the second case, the cost of life cycle management should be *much smaller* than the losses avoided. The major cost considerations for life cycle management are personnel costs and some delays as the system progresses through the life cycle for completing analyses and reviews and obtaining management approvals.

It is possible to overmanage a system: to spend more time planning, designing, and analyzing risk than is necessary. Planning, by itself, does not further the mission or business of an organization. Therefore, while security life cycle management can yield significant benefits, the effort should be commensurate with the system's size, complexity, and sensitivity and the risks associated with the system. In general, the higher the value of the system, the newer the system's architecture, technologies, and practices, and the worse the impact if the system security fails, the more effort should be spent on life cycle management.

References

Communications Security Establishment. *A Framework for Security Risk Management in*

Information Technology Systems. Canada.

Dykman, Charlene A. ed., and Charles K. Davis, asc. ed. *Control Objectives – Controls in an Information Systems Environment: Objectives, Guidelines, and Audit Procedures*. (fourth edition). Carol Stream, IL: The EDP Auditors Foundation, Inc., April 1992.

Guttman, Barbara. *Computer Security Considerations in Federal Procurements: A Guide for Procurement Initiators, Contracting Officers, and Computer Security Officials*. Special Publication 800-4. Gaithersburg, MD: National Institute of Standards and Technology, March 1992.

Institute of Internal Auditors Research Foundation. *System Auditability and Control Report*. Altamonte Springs, FL: The Institute of Internal Auditors, 1991.

Murphy, Michael, and Xenia Ley Parker. *Handbook of EDP Auditing*, especially Chapter 2 "The Auditing Profession," and Chapter 3, "The EDP Auditing Profession." Boston, MA: Warren, Gorham & Lamont, 1989.

National Bureau of Standards. *Guideline for Computer Security Certification and Accreditation*. Federal Information Processing Standard Publication 102. September 1983.

National Institute of Standards and Technology. "Disposition of Sensitive Automated Information." Computer Systems Laboratory Bulletin. October 1992.

National Institute of Standards and Technology. "Sensitivity of Information." Computer Systems Laboratory Bulletin. November 1992.

Office of Management and Budget. "Guidance for Preparation of Security Plans for Federal Computer Systems That Contain Sensitive Information." OMB Bulletin 90-08. 1990.

Ruthberg, Zella G, Bonnie T. Fisher and John W. Lainhart IV. *System Development Auditor*. Oxford, England: Elsevier Advanced Technology, 1991.

Ruthberg, Z., et al. *Guide to Auditing for Controls and Security: A System Development Life Cycle Approach*. Special Publication 500-153. Gaithersburg, MD: National Bureau of Standards. April 1988.

Vickers Benzel, T. C. *Developing Trusted Systems Using DOD-STD-2167A*. Oakland, CA: IEEE Computer Society Press, 1990.

Wood, C. "Building Security Into Your System Reduces the Risk of a Breach." *LAN Times*, 10(3), 1993. p 47.

Chapter 9

ASSURANCE

Computer security assurance is the degree of confidence one has that the security measures, both technical and operational, work as intended to protect the system and the information it processes. Assurance is not, however, an absolute guarantee that the measures work as intended. Like the closely related areas of reliability and quality, assurance can be difficult to analyze; however, it is something people expect and obtain (though often without realizing it). For example, people may routinely get product recommendations from colleagues but may not consider such recommendations as providing assurance.

Assurance is a degree of confidence, not a true measure of how secure the system actually is. This distinction is necessary because it is extremely difficult -- and in many cases virtually impossible -- to know exactly how secure a system is.

> Security assurance is the degree of confidence one has that the security controls operate correctly and protect the system as intended.

Assurance is a challenging subject because it is difficult to describe and even more difficult to quantify. Because of this, many people refer to assurance as a "warm fuzzy feeling" that controls work as intended. However, it is possible to apply a more rigorous approach by knowing two things: (1) who needs to be assured and (2) what types of assurance can be obtained. The person who needs to be assured is the management official who is ultimately responsible for the security of the system. Within the federal government, this person is the *authorizing or accrediting official*.[71]

There are many methods and tools for obtaining assurance. For discussion purposes, this chapter categorizes assurance in terms of a general system life cycle. The chapter first discusses planning for assurance and then presents the two categories of assurance methods and tools: (1) design and implementation assurance and (2) operational assurance. Operational assurance is further categorized into audits and monitoring.

The division between design and implementation assurance and operational assurance can be fuzzy. While such issues as configuration management or audits are discussed under operational assurance, they may also be vital during a system's development. The discussion tends to focus more on technical issues during design and implementation assurance and to be a mixture of

[71] Accreditation is a process used primarily within the federal government. It is the process of managerial authorization for processing. Different agencies may use other terms for this approval function. The terms used here are consistent with Federal Information Processing Standard 102, *Guideline for Computer Security Certification and Accreditation*. (See reference section of this chapter.)

management, operational, and technical issues under operational assurance. The reader should keep in mind that the division is somewhat artificial and that there is substantial overlap.

9.1 Accreditation and Assurance

Accreditation is a management official's formal acceptance of the adequacy of a system's security. The best way to view computer security accreditation is as a form of quality control. It forces managers and technical staff to work together to find workable, cost-effective solutions given security needs, technical constraints, operational constraints, and mission or business requirements. The accreditation process obliges managers to make the critical decision regarding the adequacy of security safeguards and, therefore, to recognize and perform their role in securing their systems. In order for the decisions to be sound, they need to be based on reliable information about the implementation of both technical and nontechnical safeguards. These include:

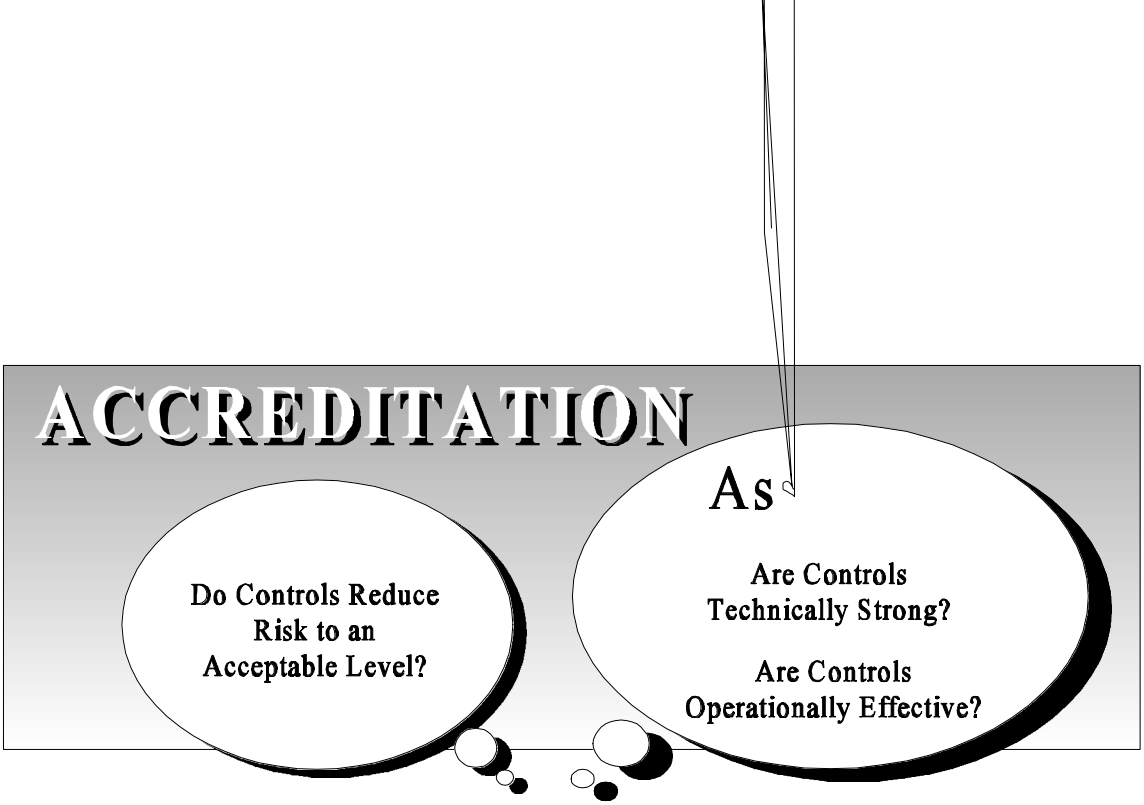
- Technical features (Do they operate as intended?).
- Operational practices (Is the system operated according to stated procedures?).
- Overall security (Are there threats which the technical features and operational practices do not address?).
- Remaining risks (Are they acceptable?).

A computer system should be accredited before the system becomes operational with periodic reaccreditation after major system changes or when significant time has elapsed.[72] Even if a system was not initially accredited, the accreditation process can be initiated at any time. Chapter 8 further discusses accreditation.

9.1.1 Accreditation and Assurance

Assurance is an extremely important -- but not the only -- element in accreditation. As shown in the diagram, assurance addresses whether the technical measures and procedures operate either (1) according to a set of security requirements and specifications or (2) according to general quality principles. Accreditation also addresses whether the system's security requirements are correct and well implemented and whether the level of quality is sufficiently high. These activities are discussed in Chapters 7 and 8.

[72] OMB Circular A-130 requires management security authorization of operation for federal systems.


ACCREDITATION
As
Do Controls Reduce Risk to an Acceptable Level?
Are Controls Technically Strong?
Are Controls Operationally Effective?

the selection of assurance methods should be coordinated with the accrediting official.

In selecting assurance methods, the need for assurance should be weighed against its cost. Assurance can be quite expensive, especially if extensive testing is done. Each method has strengths and weaknesses in terms of cost and what kind of assurance is actually being delivered. A combination of methods can often provide greater assurance, since no method is foolproof, and can be less costly than extensive testing.

The accrediting official is not the only arbiter of assurance. Other officials who use the system should also be consulted. (For example, a Production Manager who relies on a Supply System should provide input to the Supply Manager.) In addition, there may be constraints outside the accrediting official's control that also affect the selection of methods. For instance, some of the methods may unduly restrict competition in acquisitions of federal information processing resources or may be contrary to the organization's privacy policies. Certain assurance methods may be required by organizational policy or directive.

9.2 Planning and Assurance

Assurance planning should begin during the planning phase of the system life cycle, either for new systems or a system upgrades. Planning for assurance when planning for other system requirements makes sense. If a system is going to need extensive testing, it should be built to facilitate such testing.

Planning for assurance helps a manager make decisions about what kind of assurance will be cost-effective. If a manager waits until a system is built or bought to consider assurance, the number of ways to obtain assurance may be much smaller than if the manager had planned for it earlier, and the remaining assurance options may be more expensive.

9.3 Design and Implementation Assurance

Design and implementation assurance should be examined from two points of view: the component and the system. Component assurance looks at the security of a specific product or system component, such as an operating system, application, security add-on, or telecommunications module. System assurance looks at the security of the entire system, including the interaction between products and modules.

Design and implementation assurance addresses whether the features of a system, application, or component meets security requirements and specifications and whether they are they are well designed and well built. Chapter 8 discusses the source for security requirements and specifications. Design and implementation assurance examines system design, development, and installation. Design and implementation assurance is usually associated

with the development/acquisition and implementation phase of the system life cycle; however, it should also be considered throughout the life cycle as the system is modified.

As stated earlier, assurance can address whether the product or system meets a set of security specifications, or it can provide other evidence of quality. This section outlines the major methods for obtaining design and implementation assurance.

9.3.1 Testing and Certification

Testing can address the quality of the system as built, as implemented, or as operated. Thus, it can be performed throughout the development cycle, after system installation, and throughout its operational phase. Some common testing techniques include functional testing (to see if a given function works according to its requirements) or penetration testing (to see if security can be bypassed). These techniques can range from trying several test cases to in-depth studies using metrics, automated tools, or multiple detailed test cases.

Certification is a formal process for testing components or systems against a specified set of security requirements. Certification is normally performed by an independent reviewer, rather than one involved in building the system. Certification is more often cost-effective for complex or high-risk systems. Less formal security testing can be used for lower-risk systems. Certification can be performed at many stages of the system design and implementation process and can take place in a laboratory, operating environment, or both.

9.3.2 NIST Conformance Testing and Validation Suites

NIST produces validation suites and conformance testing to determine if a product (software, hardware, firmware) meets specified standards. These test suites are developed for specific standards and use many methods. Conformance to standards can be important for many reasons, including interoperability or strength of security provided. NIST publishes a list of validated products quarterly.

9.3.3 Use of Advanced or Trusted Development

In the development of both commercial off-the-shelf products and more customized systems, the use of advanced or trusted system architectures, development methodologies, or software engineering techniques can provide assurance. Examples include security design and development reviews, formal modeling, mathematical proofs, ISO 9000 quality techniques, or use of security architecture concepts, such as a trusted computing base (TCB) or reference monitor.

9.3.4 Use of Reliable Architectures

Some system architectures are intrinsically more reliable, such as systems that use fault-tolerance,

redundance, shadowing, or redundant array of inexpensive disks (RAID) features. These examples are primarily associated with system availability.

9.3.5 Use of Reliable Security

One factor in reliable security is the concept of *ease of safe use*, which postulates that a system that is easier to secure will be more likely to be secure. Security features may be more likely to be used when the initial system defaults to the "most secure" option. In addition, a system's security may be deemed more reliable if it does not use very new technology that has not been tested in the "real" world (often called "bleeding-edge" technology). Conversely, a system that uses older, well-tested software may be less likely to contain bugs.

9.3.6 Evaluations

A product evaluation normally includes testing. Evaluations can be performed by many types of organizations, including government agencies, both domestic and foreign; independent organizations, such as trade and professional organizations; other vendors or commercial groups; or individual users or user consortia. Product reviews in trade literature are a form of evaluation, as are more formal reviews made against specific criteria. Important factors for using evaluations are the degree of independence of the evaluating group, whether the evaluation criteria reflect needed security features, the rigor of the testing, the testing environment, the age of the evaluation, the competence of the evaluating organization, and the limitations placed on the evaluations by the evaluating group (e.g., assumptions about the threat or operating environment).

9.3.7 Assurance Documentation

The ability to describe security requirements and how they were met can reflect the degree to which a system or product designer understands applicable security issues. Without a good understanding of the requirements, it is not likely that the designer will be able to meet them.

Assurance documentation can address the security either for a system or for specific components. System-level documentation should describe the system's security requirements and how they have been implemented, including *interrelationships* among applications, the operating system, or networks. System-level documentation addresses more than just the operating system, the security system, and applications; it describes the system as *integrated* and *implemented in a particular environment*. Component documentation will generally be an off-the-shelf product, whereas the system designer or implementer will generally develop system documentation.

9.3.8 Accreditation of Product to Operate in Similar Situation

The accreditation of a product or system to operate in a similar situation can be used to provide

some assurance. However, it is important to realize that an accreditation is environment- and system-specific. Since accreditation balances risk against advantages, the same product may be appropriately accredited for one environment but not for another, even by the same accrediting official.

9.3.9 Self-Certification

A vendor's, integrator's, or system developer's self-certification does not rely on an impartial or independent agent to perform a technical evaluation of a system to see how well it meets a stated security requirement. Even though it is not impartial, it can still provide assurance. The self-certifier's reputation is on the line, and a resulting certification report can be read to determine whether the security requirement was defined and whether a meaningful review was performed.

A hybrid certification is possible where the work is performed under the auspices or review of an independent organization by having that organization analyze the resulting report, perform spot checks, or perform other oversight. This method may be able to combine the lower cost and greater speed of a self-certification with the impartiality of an independent review. The review, however, may not be as thorough as independent evaluation or testing.

9.3.10 Warranties, Integrity Statements, and Liabilities

Warranties are another source of assurance. If a manufacturer, producer, system developer, or integrator is willing to correct errors within certain time frames or by the next release, this should give the system manager a sense of commitment to the product and of the product's quality. An integrity statement is a formal declaration or certification of the product. It can be backed up by a promise to (a) fix the item (warranty) or (b) pay for losses (liability) if the product does not conform to the integrity statement.

9.3.11 Manufacturer's Published Assertions

A manufacturer's or developer's published assertion or formal declaration provides a limited amount of assurance based exclusively on reputation.

9.3.12 Distribution Assurance

It is often important to know that software has arrived unmodified, especially if it is distributed electronically. In such cases, checkbits or digital signatures can provide high assurance that code has not been modified. Anti-virus software can be used to check software that comes from sources with unknown reliability (such as a bulletin board).

9.4 Operational Assurance

Design and implementation assurance addresses the quality of security features built into systems. Operational assurance addresses whether the system's technical features are being bypassed or have vulnerabilities and whether required procedures are being followed. It does not address changes in the system's security requirements, which could be caused by changes to the system and its operating or threat environment. (These changes are addressed in Chapter 8.)

Security tends to degrade during the operational phase of the system life cycle. System users and operators discover new ways to intentionally or unintentionally bypass or subvert security (especially if there is a perception that bypassing security improves functionality). Users and administrators often think that nothing will happen to them or their system, so they shortcut security. Strict adherence to procedures is rare, and they become outdated, and errors in the system's administration commonly occur.

Organizations use two basic methods to maintain operational assurance:

- A *system audit* -- a *one-time* or *periodic* event to evaluate security. An audit can vary widely in scope: it may examine an entire system for the purpose of reaccreditation or it may investigate a single anomalous event.

- *Monitoring* -- an *ongoing* activity that checks on the system, its users, or the environment.

In general, the more "real-time" an activity is, the more it falls into the category of monitoring. This distinction can create some unnecessary linguistic hairsplitting, especially concerning system-generated audit trails. Daily or weekly reviewing of the audit trail (for unauthorized access attempts) is generally monitoring, while an historical review of several months' worth of the trail (tracing the actions of a specific user) is probably an audit.

9.4.1 Audit Methods and Tools

An audit conducted to support operational assurance examines whether the system is meeting stated or implied security requirements including system and organization policies. Some audits also examine whether security requirements are appropriate, but this is outside the scope of operational assurance. (See Chapter 8.) Less formal audits are often called *security reviews*.

Audits can be self-administered or independent (either internal or external).[74] Both types can provide excellent information about technical, procedural, managerial, or other aspects of security. The essential difference between a self-audit and an independent audit is objectivity. Reviews done by system management staff, often called self-audits/ assessments, have an inherent conflict of interest. The system management staff may have little incentive to say that the computer system was poorly designed or is sloppily operated. On the other hand, they may be motivated by a strong desire to improve the security of the system. In addition, they are knowledgeable about the system and may be able to find hidden problems.

> A person who performs an independent audit should be free from personal and external constraints which may impair their independence and should be organizationally independent.

The independent auditor, by contrast, should have no professional stake in the system. Independent audit may be performed by a professional audit staff in accordance with generally accepted auditing standards.

There are many methods and tools, some of which are described here, that can be used to audit a system. Several of them overlap.

9.4.1.1 Automated Tools

Even for small multiuser computer systems, it is a big job to manually review security features. Automated tools make it feasible to review even large computer systems for a variety of security flaws.

There are two types of automated tools: (1) active tools, which find vulnerabilities by trying to exploit them, and (2) passive tests, which only examine the system and infer the existence of problems from the state of the system.

Automated tools can be used to help find a variety of threats and vulnerabilities, such as improper access controls or access control configurations, weak passwords, lack of integrity of the system software, or not using all relevant software updates and patches. These tools are often very successful at finding vulnerabilities and are sometimes used by hackers to break into systems. Not taking advantage of these tools puts system administrators at a disadvantage. Many of the tools

[74] An example of an internal auditor in the federal government is the Inspector General. The General Accounting Office can perform the role of external auditor in the federal government. In the private sector, the corporate audit staff serves the role of internal auditor, while a public accounting firm would be an external auditor.

are simple to use; however, some programs (such as access-control auditing tools for large mainframe systems) require specialized skill to use and interpret.

9.4.1.2 Internal Controls Audit

> The General Accounting Office provides standards and guidance for internal controls audits of federal agencies.

An auditor can review controls in place and determine whether they are effective. The auditor will often analyze both computer and noncomputer-based controls. Techniques used include inquiry, observation, and testing (of both the controls themselves and the data). The audit can also detect illegal acts, errors, irregularities, or a lack of compliance with laws and regulations. Security checklists and penetration testing, discussed below, may be used.

9.4.1.3 Security Checklists

> Warning: Security Checklists that are passed (e.g., with a B+ or better score) are often used mistakenly as proof (instead of an indication) that security is sufficient. Also, managers of systems which "fail" a checklist often focus too much attention on "getting the points," rather than whether the security measures makes sense in the particular environment and are correctly implemented.

Within the government, the computer security plan provides a checklist against which the system can be audited. This plan, discussed in Chapter 8, outlines the major security considerations for a system, including management, operational, and technical issues. One advantage of using a computer security plan is that it reflects the unique security environment of the system, rather than a generic list of controls. Other checklists can be developed, which include national or organizational security policies and practices (often referred to as *baselines*). Lists of "generally accepted security practices" (GSSPs) can also be used. Care needs to be taken so that deviations from the list are not automatically considered wrong, since they may be appropriate for the system's particular environment or technical constraints.

Checklists can also be used to verify that changes to the system have been reviewed from a security point of view. A common audit examines the system's configuration to see if major changes (such as connecting to the Internet) have occurred that have not yet been analyzed from a security point of view.

9.4.1.4 Penetration Testing

Penetration testing can use many methods to attempt a system break-in. In addition to using active automated tools as described above, penetration testing can be done "manually." The most useful type of penetration testing is to use methods that might really be used against the system. For hosts on the Internet, this would certainly include automated tools. For many systems, lax

procedures or a lack of internal controls on applications are common vulnerabilities that penetration testing can target. Another method is "social engineering," which involves getting users or administrators to divulge information about systems, including their passwords.[75]

9.4.2 Monitoring Methods and Tools

Security monitoring is an ongoing activity that looks for vulnerabilities and security problems. Many of the methods are similar to those used for audits, but are done more regularly or, for some automated tools, in real time.

9.4.2.1 Review of System Logs

As discussed in Chapter 8, a periodic review of system-generated logs can detect security problems, including attempts to exceed access authority or gain system access during unusual hours.

9.4.2.2 Automated Tools

Several types of automated tools monitor a system for security problems. Some examples follow:

- *Virus scanners* are a popular means of checking for virus infections. These programs test for the presence of viruses in executable program files.

- *Checksumming* presumes that program files should not change between updates. They work by generating a mathematical value based on the contents of a particular file. When the integrity of the file is to be verified, the checksum is generated on the current file and compared with the previously generated value. If the two values are equal, the integrity of the file is verified. Program checksumming can detect viruses, Trojan horses, accidental changes to files caused by hardware failures, and other changes to files. However, they may be subject to covert replacement by a system intruder. Digital signatures can also be used.

- *Password crackers* check passwords against a dictionary (either a "regular" dictionary or a specialized one with easy-to-guess passwords) and also check if passwords are common permutations of the user ID. Examples of special dictionary entries could be the names of regional sports teams and stars; common permutations could be the user ID spelled backwards.

---

[75] While penetration testing is a very powerful technique, it should preferably be conducted with the knowledge and consent of system management. Unknown penetration attempts can cause a lot of stress among operations personnel, and may create unnecessary disturbances.

- *Integrity verification programs* can be used by such applications to look for evidence of data tampering, errors, and omissions. Techniques include consistency and reasonableness checks and validation during data entry and processing. These techniques can check data elements, as input or as processed, against expected values or ranges of values; analyze transactions for proper flow, sequencing, and authorization; or examine data elements for expected relationships. These programs comprise a very important set of processes because they can be used to convince people that, if they do what they should not do, accidentally or intentionally, they will be caught. Many of these programs rely upon logging of individual user activities.

- *Intrusion detectors* analyze the system audit trail, especially log-ons, connections, operating system calls, and various command parameters, for activity that could represent unauthorized activity. Intrusion detection is covered in Chapters 12 and 18.

- *System performance monitoring* analyzes system performance logs in real time to look for availability problems, including active attacks (such as the 1988 Internet worm) and system and network slowdowns and crashes.

9.4.2.3 Configuration Management

From a security point of view, configuration management provides assurance that the system in operation is the correct version (configuration) of the system and that any changes to be made are reviewed for security implications. Configuration management can be used to help ensure that changes take place in an identifiable and controlled environment and that they do not unintentionally harm any of the system's properties, including its security. Some organizations, particularly those with very large systems (such as the federal government), use a configuration control board for configuration management. When such a board exists, it is helpful to have a computer security expert participate. In any case, it is useful to have computer security officers participate in system management decision making.

Changes to the system can have security implications because they may introduce or remove vulnerabilities and because significant changes may require updating the contingency plan, risk analysis, or accreditation.

9.4.2.4 Trade Literature/Publications/Electronic News

In addition to monitoring the system, it is useful to monitor external sources for information. Such sources as trade literature, both printed and electronic, have information about security vulnerabilities, patches, and other areas that impact security. The Forum of Incident Response Teams (FIRST) has an electronic mailing list that receives information on threats, vulnerabilities,

and patches.[76]

9.5 Interdependencies

Assurance is an issue for every control and safeguard discussed in this handbook. Are user ID and access privileges kept up to date? Has the contingency plan been tested? Can the audit trail be tampered with? One important point to be reemphasized here is that assurance is not only for technical controls, but for operational controls as well. Although the chapter focused on information systems assurance, it is also important to have assurance that management controls are working well. Is the security program effective? Are policies understood and followed? As noted in the introduction to this chapter, the need for assurance is more widespread than people often realize.

Life Cycle. Assurance is closely linked to the planning for security in the system life cycle. Systems can be designed to facilitate various kinds of testing against specified security requirements. By planning for such testing early in the process, costs can be reduced; in some cases, without proper planning, some kinds of assurance cannot be otherwise obtained.

9.6 Cost Considerations

There are many methods of obtaining assurance that security features work as anticipated. Since assurance methods tend to be qualitative rather than quantitative, they will need to be evaluated. Assurance can also be quite expensive, especially if extensive testing is done. It is useful to evaluate the amount of assurance received for the cost to make a best-value decision. In general, personnel costs drive up the cost of assurance. Automated tools are generally limited to addressing specific problems, but they tend to be less expensive.

References

Borsook, P. "Seeking Security." *Byte*. 18(6), 1993. pp. 119-128.

Dykman, Charlene A. ed., and Charles K. Davis, asc. ed. *Control Objectives – Controls in an Information Systems Environment: Objectives, Guidelines, and Audit Procedures*. (fourth edition). Carol Stream, IL: The EDP Auditors Foundation, Inc., April 1992.

Farmer, Dan and Wietse Venema. "Improving the Security of Your Site by Breaking Into It." Available from FTP.WIN.TUE.NL. 1993.

[76]For information on FIRST, send e-mail to FIRST-SEC@FIRST.ORG.

II. Management Controls

Guttman, Barbara. *Computer Security Considerations in Federal Procurements: A Guide for Procurement Initiators, Contracting Officers, and Computer Security Officials*. Special Publication 800-4. Gaithersburg, MD: National Institute of Standards and Technology, March 1992.

Howe, D. "Information System Security Engineering: Cornerstone to the Future." *Proceedings of the 15th National Computer Security Conference*, Vol 1. (Baltimore, MD) Gaithersburg, MD: National Institute of Standards and Technology, 1992. pp. 244-251.

Levine, M. "Audit Serve Security Evaluation Criteria." *Audit Vision*. 2(2). 1992, pp. 29-40.

National Bureau of Standards. *Guideline for Computer Security Certification and Accreditation*. Federal Information Processing Standard Publication 102. September 1983.

National Bureau of Standards. *Guideline for Lifecycle Validation, Verification, and Testing of Computer Software*. Federal Information Processing Standard Publication 101. June 1983.

National Bureau of Standards. *Guideline for Software Verification and Validation Plans*. Federal Information Processing Standard Publication 132. November 1987.

Nuegent, W., J. Gilligan, L. Hoffman, and Z. Ruthberg. *Technology Assessment: Methods for Measuring the Level of Computer Security*. Special Publication 500-133. Gaithersburg, MD: National Bureau of Standards, 1985.

Peng, Wendy W., and Dolores R. Wallace. *Software Error Analysis*. Special Publication 500-209. Gaithersburg, MD: National Institute of Standards and Technology, 1993.

Peterson, P. "Infosecurity and Shrinking Media." *ISSA Access*. 5(2), 1992. pp. 19-22.

Pfleeger, C., S. Pfleeger, and M. Theofanos, "A Methodology for Penetration Testing." *Computers and Security.* 8(7), 1989. pp. 613-620.

Polk, W. Timothy, and Lawrence Bassham. *A Guide to the Selection of Anti-Virus Tools and Techniques*. Special Publication 800-5. Gaithersburg, MD: National Institute of Standards and Technology, December 1992.

Polk, W. Timothy. *Automated Tools for Testing Computer System Vulnerability*. Special Publication 800-6. Gaithersburg, MD: National Institute of Standards and Technology, December 1992.

President's Council on Integrity and Efficiency. *Review of General Controls in Federal Computer Systems*. Washington, DC: President's Council on Integrity and Efficiency, October 1988.

President's Council on Management Improvement and the President's Council on Integrity and Efficiency. *Model Framework for Management Control Over Automated Information System*. Washington, DC: President's Council on Management Improvement, January 1988.

Ruthberg, Zella G, Bonnie T. Fisher and John W. Lainhart IV. *System Development Auditor*. Oxford, England: Elsevier Advanced Technology, 1991.

Ruthburg, Zella, et al. *Guide to Auditing for Controls and Security: A System Development Life Cycle Approach*. Special Publication 500-153. Gaithersburg, MD: National Bureau of Standards, April 1988.

Strategic Defense Initiation Organization. *Trusted Software Methodology*. Vols. 1 and II. SDI-S-SD-91-000007. June 17, 1992.

Wallace, Dolores, and J.C. Cherniasvsky. *Guide to Software Acceptance.* Special Publication 500-180. Gaithersburg, MD: National Institute of Standards and Technology, April 1990.

Wallace, Dolores, and Roger Fugi. *Software Verification and Validation: Its Role in Computer Assurance and Its Relationship with Software Product Management Standards*. Special Publication 500-165. Gaithersburg, MD: National Institute of Standards and Technology, September 1989.

Wallace, Dolores R., Laura M. Ippolito, and D. Richard Kuhn. *High Integrity Software Standards and Guidelines*. Special Publication 500-204. Gaithersburg, MD: National Institute of Standards and Technology, 1992.

Wood, C., et al. *Computer Security: A Comprehensive Controls Checklist*. New York, NY: John Wiley & Sons, 1987.

III. OPERATIONAL CONTROLS

Chapter 10

PERSONNEL/USER ISSUES

Many important issues in computer security involve human users, designers, implementors, and managers. A broad range of security issues relate to how these individuals interact with computers and the access and authorities they need to do their job. No computer system can be secured without properly addressing these security issues.[77]

This chapter examines issues concerning the staffing of positions that interact with computer systems; the administration of users on a system, including considerations for terminating employee access; and special considerations that may arise when contractors or the public have access to systems. Personnel issues are closely linked to logical access controls, discussed in Chapter 17.

10.1 Staffing

The staffing process generally involves at least four steps and can apply equally to general users as well as to application managers, system management personnel, and security personnel. These four steps are: (1) defining the job, normally involving the development of a position description; (2) determining the sensitivity of the position; (3) filling the position, which involves screening applicants and selecting an individual; and (4) training.

10.1.1 Groundbreaking – Position Definition

Early in the process of defining a position, security issues should be identified and dealt with. Once a position has been broadly defined, the responsible supervisor should determine the type of computer access needed for the position. There are two general principles to apply when granting access: *separation of duties* and *least privilege*.

Separation of duties refers to dividing roles and responsibilities so that a single individual cannot subvert a critical process. For example, in financial systems, no single individual should normally be given authority to issue checks. Rather, one person initiates a request for a payment and another authorizes that same payment. In effect, checks and balances need to be designed into both the process as well as the specific, individual positions of personnel who will implement the process. Ensuring that such duties are well defined is the responsibility of management.

Least privilege refers to the security objective of granting users *only those accesses they need to*

[77] A distinction is made between users and personnel, since some users (e.g., contractors and members of the public) may not be considered personnel (i.e., employees).

perform their official duties. Data entry clerks, for example, may not have any need to run analysis reports of their database. However, least privilege does not mean that all users will have extremely little functional access; some employees will have significant access if it is required for their position. However, applying this principle may limit the damage resulting from accidents, errors, or unauthorized use of system resources. It is important to make certain that the implementation of least privilege does not interfere with the ability to have personnel substitute for each other without undue delay. Without careful planning, access control can interfere with contingency plans.

10.1.2 Determining Position Sensitivity

Knowledge of the duties and access levels that a particular position will require is necessary for determining the sensitivity of the position. The responsible management official should correctly identify position sensitivity levels so that appropriate, cost-effective screening can be completed.

Various levels of sensitivity are assigned to positions in the federal government. Determining the appropriate level is based upon such factors as the type and degree of harm (e.g., disclosure of private information, interruption of critical processing, computer fraud) the individual can cause through misuse of the computer system as well as more traditional factors, such as access to classified information and fiduciary responsibilities. Specific agency guidance should be followed on this matter.

It is important to select the appropriate position sensitivity, since controls in excess of the sensitivity of the position wastes resources, while too little may cause unacceptable risks.

10.1.3 Filling the Position -- Screening and Selecting

Once a position's sensitivity has been determined, the position is ready to be staffed. In the federal government, this typically includes publishing a formal vacancy announcement and identifying which applicants meet the position requirements. More sensitive positions typically require *preemployment* background screening; screening after employment has commenced (post-entry-on-duty) may suffice for less sensitive positions.

> In general, it is more effective to use separation of duties and least privilege to limit the sensitivity of the position, rather than relying on screening to reduce the risk to the organization.

Background screening helps determine whether a particular individual is suitable for a given position. For example, in positions with high-level fiduciary responsibility, the screening process will attempt to ascertain the person's trustworthiness and appropriateness for a particular position. In the federal government, the screening process is formalized through a series of background checks conducted through a central investigative office within the

organization or through another organization (e.g., the Office of Personnel Management).

Within the Federal Government, the most basic screening technique involves a check for a criminal history, checking FBI fingerprint records, and other federal indices.[78] More extensive background checks examine other factors, such as a person's work and educational history, personal interview, history of possession or use of illegal substances, and interviews with current and former colleagues, neighbors, and friends. The exact type of screening that takes place depends upon the sensitivity of the position and applicable agency implementing regulations. Screening is not conducted by the prospective employee's manager; rather, agency security and personnel officers should be consulted for agency-specific guidance.

Outside of the Federal Government, employee screening is accomplished in many ways. Policies vary considerably among organizations due to the sensitivity of examining an individual's background and qualifications. Organizational policies and procedures normally try to balance fears of invasiveness and slander against the need to develop confidence in the integrity of employees. One technique may be to place the individual in a less sensitive position initially.

For both the Federal Government and private sector, finding something compromising in a person's background does not necessarily mean they are unsuitable for a particular job. A determination should be made based on the type of job, the type of finding or incident, and other relevant factors. In the federal government, this process is referred to as *adjudication*.

10.1.4 Employee Training and Awareness

Even after a candidate has been hired, the staffing process cannot yet be considered complete – employees still have to be trained to do their job, which includes computer security responsibilities and duties. As discussed in Chapter 13, such security training can be very cost-effective in promoting security.

Some computer security experts argue that employees must receive initial computer security training before they are granted any access to computer systems. Others argue that this must be a risk-based decision, perhaps granting only restricted access (or, perhaps, only access to their PC) until the required training is completed. Both approaches recognize that adequately trained employees are crucial to the effective functioning of computer systems and applications. Organizations may provide introductory training prior to granting any access with follow-up more extensive training. In addition, although training of new users is critical, it is important to recognize that security training and awareness activities should be ongoing during the time an

[78] In the federal government, separate and unique screening procedures are not established for each position. Rather, positions are categorized by general sensitivity and are assigned a corresponding level of background investigation or other checks.

individual is a system user. (See Chapter 13 for a more thorough discussion.)

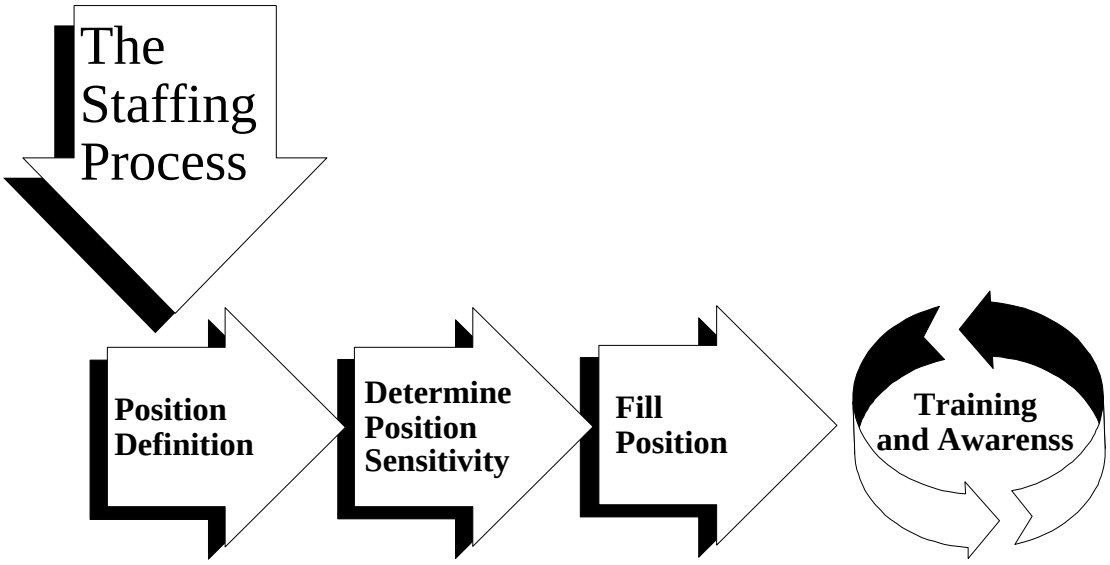


Figure 10.1

10.2 User Administration

Effective administration of users' computer access is essential to maintaining system security. *User account management* focuses on identification, authentication, and access authorizations. This is augmented by the process of *auditing* and otherwise periodically verifying the legitimacy of current accounts and access authorizations. Finally, there are considerations involved in the *timely modification or removal of access* and associated issues for employees who are reassigned, promoted, or terminated, or who retire.

10.2.1 User Account Management

User account management involves (1) the process of requesting, establishing, issuing, and closing user accounts; (2) tracking users and their respective access authorizations; and (3) managing these functions.

User account management typically begins with a request from the user's supervisor to the system manager for a system account. If a user is to have access to a particular application, this request may be sent through the application manager to the system manager. This will ensure that the systems office receives formal approval from the "application manager" for the employee to be given access. The request will normally state the level of access to be granted, perhaps by function or by specifying a particular user profile. (Often when more than one employee is doing the same job, a "profile" of permitted authorizations is created.)

Systems operations staff will normally then use the account request to create an account for the new user. The access levels of the account will be consistent with those requested by the supervisor. This account will normally be assigned selected access authorizations. These are sometimes built directly into applications, and other times rely upon the operating system. "Add-on" access applications are also used. These access levels and authorizations are often tied to specific access levels within an application.

Example of Access Levels Within an Application

Level	Function
1	Create Records
2	Edit *Group A* records
3	Edit *Group B* records
4	Edit *all* records

Next, employees will be given their account information, including the account identifier (e.g., user ID) and a means of authentication (e.g., password or smart card/PIN). One issue that may arise at this stage is whether the user ID is to be tied to the particular *position* an employee holds (e.g., ACC5 for an accountant) or the *individual employee* (e.g., BSMITH for Brenda Smith). Tying user IDs to positions may simplify administrative overhead in some cases; however, it may make auditing more difficult as one tries to trace the actions of a particular individual. It is normally more advantageous to tie the user ID to the individual employee. However, if the user ID is created and tied to a position, procedures will have to be established to change them if employees switch jobs or are otherwise reassigned.

When employees are given their account, it is often convenient to provide initial or refresher training and awareness on computer security issues. Users should be asked to review a set of rules and regulations for system access. To indicate their understanding of these rules, many organizations require employees to sign an "acknowledgment statement," which may also state causes for dismissal or prosecution under the Computer Fraud and Abuse Act and other

applicable state and local laws.[79]

When user accounts are no longer required, the supervisor should inform the application manager and system management office so accounts can be removed in a timely manner. One useful secondary check is to work with the local organization's personnel officer to establish a procedure for routine notification of employee departures to the systems office. Further issues are discussed in the "Termination" section of this chapter.

Sample User Account and Password Acknowledgment Form

I hereby acknowledge personal receipt of the system password(s) associated with the user Ids listed below. I understand that I am responsible for protecting the password(s), will comply with all applicable system security standards, and will not divulge my password(s) to any person. I further understand that I must report to the Information Systems Security Officer any problem I encounter in the use of the password(s) or when I have reason to believe that the private nature of my password(s) has been compromised.

It is essential to realize that *access and authorization administration is a continuing process*. New user accounts are added while others are deleted. Permissions change: sometimes permanently, sometimes temporarily. New applications are added, upgraded, and removed. Tracking this information to keep it up to date is not easy, but is necessary to allow users access to only those functions necessary to accomplish their assigned responsibilities – thereby helping to maintain the principle of *least privilege*. In managing these accounts, there is a need to balance timeliness of service and record keeping. While sound record keeping practices are necessary, delays in processing requests (e.g., change requests) may lead to requests for more access than is really necessary – just to avoid delays should such access ever be required.

Managing this process of user access is also one that, particularly for larger systems, is often decentralized. Regional offices may be granted the authority to create accounts and change user access authorizations or to submit forms requesting that the centralized access control function make the necessary changes. Approval of these changes is important – it may require the approval of the file owner and the supervisor of the employee whose access is being changed.

10.2.2 Audit and Management Reviews

From time to time, it is necessary to review user account management on a system. Within the area of user access issues, such reviews may examine the levels of access each individual has, conformity with the concept of least privilege, whether all accounts are still active, whether management authorizations are up-to-date, whether required training has been completed, and so forth.

[79] Whenever users are asked to sign a document, appropriate review by organizational legal counsel and, if applicable, by employee bargaining units should be accomplished.

These reviews can be conducted on *at least* two levels:[80] (1) on an application-by-application basis or (2) on a systemwide basis. Both kinds of reviews can be conducted by, among others, in-house systems personnel (a self-audit), the organization's internal audit staff, or external auditors. For example, a good practice is for application managers (and data owners, if different) to review all access levels of all application users every month – and sign a formal access approval list, which will provide a written record of the approvals. While it may initially appear that such reviews should be conducted by systems personnel, they usually are not fully effective. System personnel *can* verify that users only have those accesses that their managers have specified. However because access requirements may change over time, it is important to involve the application manager, who is often the only individual in a position to know current access requirements.

Outside audit organizations (e.g., the Inspector General [IG] or the General Accounting Office) may also conduct audits. For example, the IG may direct a more extensive review of permissions. This may involve discussing the need for particular access levels for specific individuals or the number of users with sensitive access. For example, how many employees should really have authorization to the check-printing function? (Auditors will also examine non-computer access by reviewing, for example, who should have physical access to the check printer or blank-check stock.)

10.2.3 Detecting Unauthorized/Illegal Activities

Several mechanisms are used besides auditing[81] and analysis of audit trails to detect unauthorized and illegal acts. (See Chapters 9 and 18.) For example, fraudulent activities may require the regular physical presence of the perpetrator(s). In such cases, the fraud may be detected during the employee's absence. Mandatory vacations for critical systems and applications personnel can help detect such activity (however, this is not a guarantee, for example, if problems are saved for the employees to handle upon their return). It is useful to avoid creating an excessive dependence upon any single individual, since the system will have to function during periods of absence. Particularly within the government, periodic rescreening of personnel is used to identify possible indications of illegal activity (e.g., living a lifestyle in excess of known income level).

10.2.4 Temporary Assignments and In-house Transfers

One significant aspect of managing a system involves keeping user access authorizations up to date. Access authorizations are typically changed under two types of circumstances: (1) change in job role, either temporarily (e.g., while covering for an employee on sick leave) or permanently

---

[80] Note that this is not an either/or distinction.

[81] The term *auditing* is used here in a broad sense to refer to the review and analysis of past events.

(e.g., after an in-house transfer) and (2) termination discussed in the following section.

Users often are required to perform duties outside their normal scope during the absence of others. This requires additional access authorizations. Although necessary, such extra access authorizations should be granted sparingly and monitored carefully, consistent with the need to maintain separation of duties for internal control purposes. Also, they should be removed promptly when no longer required.

Permanent changes are usually necessary when employees change positions within an organization. In this case, the process of granting account authorizations (described in Section 10.2.1) will occur again. At this time, however, is it also important that access authorizations of the prior position be removed. Many instances of "authorization creep" have occurred with employees continuing to maintain access rights for previously held positions within an organization. This practice is inconsistent with the principle of least privilege.

10.2.5 Termination

Termination of a user's system access generally can be characterized as either "friendly" or "unfriendly." Friendly termination may occur when an employee is voluntarily transferred, resigns to accept a better position, or retires. Unfriendly termination may include situations when the user is being fired for cause, "RIFed,"[82] or involuntarily transferred. Fortunately, the former situation is more common, but security issues have to be addressed in both situations.

10.2.5.1 Friendly Termination

Friendly termination refers to the removal of an employee from the organization when there is no reason to believe that the termination is other than mutually acceptable. Since terminations can be expected regularly, this is usually accomplished by implementing a standard set of procedures for outgoing or transferring employees. These are part of the standard employee "out-processing," and are put in place, for example, to ensure that system accounts are removed in a timely manner. Out-processing often involves a sign-out form initialed by each functional manager with an interest in the separation. This normally includes the group(s) managing access controls, the control of keys, the briefing on the responsibilities for confidentiality and privacy, the library, the property clerk, and several other functions not necessarily related to information security.

In addition, other issues should be examined as well. The continued availability of data, for example, must often be assured. In both the manual and the electronic worlds, this may involve documenting procedures or filing schemes, such as how documents are stored on the hard disk, and how are they backed up. Employees should be instructed whether or not to "clean up" their

[82] *RIF* is a term used within the government as shorthand for "reduction in force."

PC before leaving. If cryptography is used to protect data, the availability of cryptographic keys to management personnel must be ensured. Authentication tokens must be collected.

Confidentiality of data can also be an issue. For example, do employees know what information they are allowed to share with their immediate organizational colleagues? Does this differ from the information they may share with the public? These and other organizational-specific issues should be addressed throughout an organization to ensure continued access to data and to provide continued confidentiality and integrity during personnel transitions. (Many of these issues should be addressed on an ongoing basis, not just during personnel transitions.) The training and awareness program normally should address such issues.

10.2.5.2 Unfriendly Termination

Unfriendly termination involves the removal of an employee under involuntary or adverse conditions. This may include termination for cause, RIF, involuntary transfer, resignation for "personality conflicts," and situations with pending grievances. The tension in such terminations may multiply and complicate security issues. Additionally, all of the issues involved in friendly terminations are still present, but addressing them may be considerably more difficult.

The greatest threat from unfriendly terminations is likely to come from those personnel who are capable of changing code or modifying the system or applications. For example, systems personnel are ideally positioned to wreak considerable havoc on systems operations. Without appropriate safeguards, personnel with such access can place logic bombs (e.g., a hidden program to erase a disk) in code that will not even execute until after the employee's departure. Backup copies can be destroyed. There are even examples where code has been "held hostage." But other employees, such as general users, can also cause damage. Errors can be input purposefully, documentation can be misfiled, and other "random" errors can be made. Correcting these situations can be extremely resource intensive.

Given the potential for adverse consequences, security specialists routinely recommend that system access be terminated as quickly as possible in such situations. If employees are to be fired, system access should be removed at the same time (or just before) the employees are notified of their dismissal. When an employee notifies an organization of a resignation and it can be reasonably expected that it is on unfriendly terms, system access should be immediately terminated. During the "notice" period, it may be necessary to assign the individual to a restricted area and function. This may be particularly true for employees capable of changing programs or modifying the system or applications. In other cases, physical removal from their offices (and, of course, logical removal, *when logical access controls exist*) may suffice.

10.3 Contractor Access Considerations

Many federal agencies as well as private organizations use contractors and consultants to assist with computer processing. Contractors are often used for shorter periods of time than regular employees. This factor may change the cost-effectiveness of conducting screening. The often higher turnover among contractor personnel generates additional costs for security programs in terms of user administration.

10.4 Public Access Considerations

Many federal agencies have begun to design, develop, and implement public access systems for electronic dissemination of information to the public. Some systems provide electronic interaction by allowing the public to send information to the government (e.g., electronic tax filing) as well as to receive it. When systems are made available for access by the public (or a large or significant subset thereof), additional security issues arise due to: (1) increased threats against public access systems and (2) the difficulty of security administration.

> OMB Circular A-130, Appendix III "Security of Federal Automated Information" and NIST *CSL Bulletin* "Security Issues in Public Access Systems" both recommend segregating information made directly accessible to the public from official records.

While many computer systems have been victims of hacker attacks, public access systems are well known and have published phone numbers and network access IDs. In addition, a successful attack could result in a lot of publicity. For these reasons, public access systems are subject to a greater threat from hacker attacks on the confidentiality, availability, and integrity of information processed by a system. In general, it is safe to say that when a system is made available for public access, the risk to the system increases – and often the constraints on its use are tightened.

Besides increased risk of hackers, public access systems can be subject to insider malice. For example, an unscrupulous user, such as a disgruntled employee, may try to introduce errors into data files intended for distribution in order to embarrass or discredit the organization. Attacks on public access systems could have a substantial impact on the organization's reputation and the level of public confidence due to the high visibility of public access systems. Other security problems may arise from unintentional actions by untrained users.

In systems without public access, there are procedures for enrolling users that often involve some user training and frequently require the signing of forms acknowledging user responsibilities. In addition, user profiles can be created and sophisticated audit mechanisms can be developed to detect unusual activity by a user. In public access systems, users are often anonymous. This can complicate system security administration.

In most systems without public access, users are typically a mix of known employees or contractors. In this case, imperfectly implemented access control schemes may be tolerated. However, when opening up a system to public access, additional precautions may be necessary because of the increased threats.

10.5 Interdependencies

User issues are tied to topics throughout this handbook.

Training and Awareness discussed in Chapter 13 is a critical part of addressing the user issues of computer security.

Identification and Authentication and *Access Controls* in a computer system can only prevent people from doing what the computer is instructed they are not allowed to do, as stipulated by *Policy*. The recognition by computer security experts that much more harm comes from people doing what they are allowed to do, but should not do, points to the importance of considering user issues in the computer security picture, and the importance of *Auditing*.

Policy, particularly its compliance component, is closely linked to personnel issues. A deterrent effect arises among users when they are aware that their misconduct, intentional or unintentional, will be detected.

These controls also depend on manager's (1) selecting the right type and level of access for their employees and (2) informing system managers of which employees need accounts and what type and level of access they require, and (3) promptly informing system managers of changes to access requirements. Otherwise, accounts and accesses can be granted to or maintained for people who should not have them.

10.6 Cost Considerations

There are many security costs under the category of user issues. Among these are:

Screening -- Costs of initial background screening and periodic updates, as appropriate.[83]

Training and Awareness -- Costs of training needs assessments, training materials, course fees, and so forth, as discussed separately in Chapter 13.

User Administration -- Costs of managing identification and authentication which, particularly for

[83] When analyzing the costs of screening, it is important to realize that screening is often conducted to meet requirements wholly unrelated to computer security.

large distributed systems, may be rather significant.

Access Administration -- Particularly beyond the initial account set-up, are ongoing costs of maintaining user accesses currently and completely.

Auditing -- Although such costs can be reduced somewhat when using automated tools, consistent, resource-intensive human review is still often necessary to detect and resolve security anomalies.

References

Fites, P., and M. Kratz. *Information Systems Security: A Practitioner's Reference*. New York, NY: Van Nostrand Reinhold, 1993. (See especially Chapter 6.)

National Institute of Standards and Technology. "Security Issues in Public Access Systems." Computer Systems Laboratory Bulletin. May 1993.

North, S. "To Catch a `Crimoid.'" *Beyond Computing.* 1(1), 1992. pp. 55-56.

Pankau, E. "The Consummate Investigator." *Security Management*. 37(2), 1993. pp. 37-41.

Schou, C., W. Machonachy, F. Lynn McNulty, and A. Chantker. "Information Security Professionalism for the 1990s." *Computer Security Journal*. 9(1), 1992. pp. 27-38.

Wagner, M. "Possibilities Are Endless, and Frightening." *Open Systems Today*. November 8 (136), 1993. pp. 16-17.

Wood, C. "Be Prepared Before You Fire." *Infosecurity News*. 5(2), 1994. pp. 51-54.

Wood, C. "Duress, Terminations and Information Security." *Computers and Security*. 12(6), 1993. pp. 527-535.

Chapter 11

PREPARING FOR CONTINGENCIES AND DISASTERS

A *computer security contingency* is an event with the potential to disrupt computer operations, thereby disrupting critical mission and business functions. Such an event could be a power outage, hardware failure, fire, or storm. If the event is very destructive, it is often called a disaster.[84]

To avert potential contingencies and disasters or minimize the damage they cause organizations can take steps early to control the event. Generally called *contingency planning*,[85] this activity is closely related to incident handling, which primarily addresses malicious technical threats such as hackers and viruses.[86]

> Contingency planning directly supports an organization's goal of continued operations. Organizations practice contingency planning because it makes good business sense.

Contingency planning involves more than planning for a move offsite after a disaster destroys a data center. It also addresses how to keep an organization's critical functions operating in the event of disruptions, both large and small. This broader perspective on contingency planning is based on the distribution of computer support throughout an organization.

This chapter presents the contingency planning process in six steps:[87]

1. *Identifying the mission- or business-critical functions*.
2. *Identifying the resources that support the critical functions.*
3. *Anticipating potential contingencies or disasters*.
4. *Selecting contingency planning strategies*.

---

[84] There is no distinct dividing line between disasters and other contingencies.

[85] Other names include disaster recovery, business continuity, continuity of operations, or business resumption planning.

[86] Some organizations include incident handling as a subset of contingency planning. The relationship is further discussed in Chapter 12, Incident Handling.

[87] Some organizations and methodologies may use a different order, nomenclature, number, or combination of steps. The specific steps can be modified, as long as the basic functions are addressed.

5. *Implementing the contingency strategies*.

6. *Testing and revising the strategy*.

11.1 Step 1: Identifying the Mission- or Business-Critical Functions

Protecting the continuity of an organization's mission or business is very difficult if it is not clearly identified. Managers need to understand the organization from a point of view that usually extends beyond the area they control. The definition of an organization's critical mission or business functions is often called a *business plan*.

> This chapter refers to an organization as having critical *mission* or *business* functions. In government organizations, the focus is normally on performing a mission, such as providing citizen benefits. In private organizations, the focus is normally on conducting a business, such as manufacturing widgets.

Since the development of a business plan will be used to support contingency planning, it is necessary not only to identify critical missions and businesses, but also to *set priorities for* them. A fully redundant capability for each function is prohibitively expensive for most organizations. In the event of a disaster, certain functions will not be performed. If appropriate priorities have been set (and approved by senior management), it could mean the difference in the organization's ability to survive a disaster.

11.2 Step 2: Identifying the Resources That Support Critical Functions

After identifying critical missions and business functions, it is necessary to identify the supporting resources, the time frames in which each resource is used (e.g., is the resource needed constantly or only at the end of the month?), and the effect on the mission or business of the unavailability of the resource. In identifying resources, a traditional problem has been that different managers oversee different resources. They may not realize how resources interact to support the organization's mission or business. Many of these resources are *not* computer resources. Contingency planning should address all the resources needed to perform a function, regardless whether they directly relate to a computer.[88]

> In many cases, the longer an organization is without a resource, the more critical the situation becomes. For example, the longer a garbage collection strike lasts, the more critical the situation becomes.

[88] However, since this is a computer security handbook, the descriptions here focus on the computer-related resources. The logistics of coordinating contingency planning for computer-related and other resources is an important consideration.

The analysis of needed resources should be conducted by those who understand how the function is performed and the dependencies of various resources on other resources and other critical relationships. This will allow an organization to *assign priorities* to resources since not all elements of all resources are crucial to the critical functions.

11.2.1 Human Resources

People are perhaps an organization's most obvious resource. Some functions require the effort of specific individuals, some require specialized expertise, and some only require individuals who can be trained to perform a specific task. Within the information technology field, human resources include both operators (such as technicians or system programmers) and users (such as data entry clerks or information analysts).

Resources That Support Critical Functions

Human Resources
Processing Capability
Computer-Based Services
Data and Applications
Physical Infrastructure
Documents and Papers

11.2.2 Processing Capability

Traditionally contingency planning has focused on processing power (i.e., if the data center is down, how can applications dependent on it continue to be processed?). Although the need for data center backup remains vital, today's other processing alternatives are also important. Local area networks (LANs), minicomputers, workstations, and personal computers in all forms of centralized and distributed processing may be performing critical tasks.

11.2.3 Automated Applications and Data

Computer systems run applications that process data. Without current electronic versions of both applications and data, computerized processing may not be possible. If the processing is being performed on alternate hardware, the applications must be compatible with the alternate hardware, operating systems and other software (including version and configuration), and

Contingency Planning Teams

To understand what resources are needed from each of the six resource categories and to understand how the resources support critical functions, it is often necessary to establish a contingency planning team. A typical team contains representatives from various organizational elements, and is often headed by a contingency planning coordinator. It has representatives from the following three groups:

1. business-oriented groups , such as representatives from functional areas;
2. facilities management; and
3. technology management.

Various other groups are called on as needed including financial management, personnel, training, safety, computer security, physical security, and public affairs.

numerous other technical factors. Because of the complexity, it is normally necessary to periodically verify compatibility. (See Step 6, Testing and Revising.)

11.2.4 Computer-Based Services

An organization uses many different kinds of computer-based services to perform its functions. The two most important are normally communications services and information services. Communications can be further categorized as data and voice communications; however, in many organizations these are managed by the same service. Information services include any source of information outside of the organization. Many of these sources are becoming automated, including on-line government and private databases, news services, and bulletin boards.

11.2.5 Physical Infrastructure

For people to work effectively, they need a safe working environment and appropriate equipment and utilities. This can include office space, heating, cooling, venting, power, water, sewage, other utilities, desks, telephones, fax machines, personal computers, terminals, courier services, file cabinets, and many other items. In addition, computers also need space and utilities, such as electricity. Electronic and paper media used to store applications and data also have physical requirements.

11.2.6 Documents and Papers

Many functions rely on vital records and various documents, papers, or forms. These records could be important because of a legal need (such as being able to produce a signed copy of a loan) or because they are the only record of the information. Records can be maintained on paper, microfiche, microfilm, magnetic media, or optical disk.

11.3 Step 3: Anticipating Potential Contingencies or Disasters

Although it is impossible to think of *all* the things that can go wrong, the next step is to identify a likely range of problems. The development of scenarios will help an organization develop a plan to address the wide range of things that can go wrong.

Scenarios should include small and large contingencies. While some general classes of contingency scenarios are obvious, imagination and creativity, as well as research, can point to other possible, but less obvious, contingencies. The contingency scenarios should address each of the resources described above. The following are *examples* of some of the types of questions that contingency scenarios may address:

Human Resources: Can people get to work? Are key personnel willing to cross a picket line? Are there critical skills and knowledge possessed by one person? Can people easily get to an alternative site?

Processing Capability: Are the computers harmed? What happens if some of the computers are inoperable, but not all?

Automated Applications and Data: Has data integrity been affected? Is an application sabotaged? Can an application run on a different processing platform?

Computer-Based Services: Can the computers communicate? To where? Can people communicate? Are information services down? For how long?

Examples of Some Less Obvious Contingencies

1. A computer center in the basement of a building had a minor problem with rats. Exterminators killed the rats, but the bodies were not retrieved because they were hidden under the raised flooring and in the pipe conduits. Employees could only enter the data center with gas masks because of the decomposing rats.

2. After the World Trade Center explosion when people reentered the building, they turned on their computer systems to check for problems. Dust and smoke damaged many systems when they were turned on. If the systems had been cleaned *first*, there would not have been significant damage.

Infrastructure: Do people have a place to sit? Do they have equipment to do their jobs? Can they occupy the building?

Documents/Paper: Can needed records be found? Are they readable?

11.4 Step 4: Selecting Contingency Planning Strategies

The next step is to plan how to recover needed resources. In evaluating alternatives, it is necessary to consider what controls are in place to prevent and minimize contingencies. Since no set of controls can cost-effectively prevent all contingencies, it is necessary to coordinate prevention and recovery efforts.

A contingency planning strategy normally consists of three parts: emergency response, recovery, and resumption.[89] *Emergency response* encompasses the initial actions taken to protect lives and limit damage. *Recovery* refers to the steps that are taken to continue support for critical functions. *Resumption* is the return to normal operations. The relationship between recovery and resumption is important. The longer it takes to resume normal operations, the longer the

[89] Some organizations divide a contingency strategy into emergency response, backup operations, and recovery. The different terminology can be confusing (especially the use of conflicting definitions of *recovery*), although the basic functions performed are the same.

organization will have to operate in the recovery mode.

The selection of a strategy needs to be based on practical considerations, including feasibility and cost. The different categories of resources should each be considered. Risk assessment can be used to help estimate the cost of options to decide on an optimal strategy. For example, is it more expensive to purchase and maintain a generator or to move processing to an alternate site, considering the likelihood of losing electrical power for various lengths of time? Are the consequences of a loss of computer-related resources sufficiently high to warrant the cost of various recovery strategies? The risk assessment should focus on areas where it is not clear which strategy is the best.

In developing contingency planning strategies, there are many factors to consider in addressing each of the resources that support critical functions. Some examples are presented in the sidebars.

Example 1: If the system administrator for a LAN has to be out of the office for a long time (due to illness or an accident), arrangements are made for the system administrator of another LAN to perform the duties. Anticipating this, the absent administrator should have taken steps beforehand to keep documentation current. This strategy is inexpensive, but service will probably be significantly reduced on both LANs which may prompt the manager of the loaned administrator to partially renege on the agreement.

Example 2: An organization depends on an on-line information service provided by a commercial vendor. The organization is no longer able to obtain the information manually (e.g., from a reference book) within acceptable time limits and there are no other comparable services. In this case, the organization relies on the contingency plan of the service provider. The organization pays a premium to obtain priority service in case the service provider has to operate at reduced capacity.

Example #3: A large mainframe data center has a contract with a hot site vendor, has a contract with the telecommunications carrier to reroute communications to the hot site, has plans to move people, and stores up-to-date copies of data, applications and needed paper records off-site. The contingency plan is expensive, but management has decided that the expense is fully justified.

Example #4. An organization distributes its processing among two major sites, each of which includes small to medium processors (personal computers and minicomputers). If one site is lost, the other can carry the critical load until more equipment is purchased. Routing of data and voice communications can be performed transparently to redirect traffic. Backup copies are stored at the other site. This plan requires tight control over the architectures used and types of applications that are developed to ensure compatibility. In addition, personnel at both sites must be cross-trained to perform all functions.

11.4.1 Human Resources

To ensure an organization has access to workers with the right skills and knowledge, training and documentation of knowledge are needed. During a major contingency, people will be under significant stress and may panic. If the contingency is a regional disaster, their first concerns will probably be their family and property. In addition, many people will be either unwilling or unable to come to work. Additional hiring or temporary services can be used. The use of additional personnel may introduce security vulnerabilities.

Contingency planning, especially for emergency response, normally places the highest emphasis on the protection of human life.

11.4.2 Processing Capability

Strategies for processing capability are normally grouped into five categories: hot site; cold site; redundancy; reciprocal agreements; and hybrids. These terms originated with recovery strategies for data centers but can be applied to other platforms.

1. *Hot site* – A building already equipped with processing capability and other services.

2. *Cold site* – A building for housing processors that can be easily adapted for use.

3. *Redundant site* – A site equipped and configured exactly like the primary site. (Some organizations plan on having reduced processing capability after a disaster and use partial redundancy. The stocking of spare personal computers or LAN servers also provides some redundancy.)

4. *Reciprocal agreement* – An agreement that allows two organizations to back each other up. (While this approach often sounds desirable, contingency planning experts note that this alternative has the greatest chance of failure due to problems keeping agreements and plans up-to-date as systems and personnel change.)

5. *Hybrids* – Any combinations of the above such as using having a hot site as a backup in case a redundant or reciprocal agreement site is damaged by a separate contingency.

Recovery may include several stages, perhaps marked by increasing availability of processing capability. Resumption planning may include contracts or the ability to place contracts to replace equipment.

11.4.3 Automated Applications and Data

Normally, the primary contingency strategy for applications and data is *regular backup* and secure *offsite storage*. Important decisions to be addressed include how often the backup is performed, how often it is stored off-site, and how it is transported (to storage, to an alternate processing site, or to support the resumption of normal operations).

> The need for computer security does not go away when an organization is processing in a contingency mode. In some cases, the need may increase due to sharing processing facilities, concentrating resources in fewer sites, or using additional contractors and consultants. Security should be an important consideration when selecting contingency strategies.

11.4.4 Computer-Based Services

Service providers may offer contingency services. Voice communications carriers often can reroute calls (transparently to the user) to a new location. Data communications carriers can also reroute traffic. Hot sites are usually capable of receiving data and voice communications. If one service provider is down, it may be possible to use another. However, the type of communications carrier lost, either local or long distance, is important. Local voice service may be carried on cellular. Local data communications, especially for large volumes, is normally more difficult. In addition, resuming normal operations may require another rerouting of communications services.

11.4.5 Physical Infrastructure

Hot sites and cold sites may also offer office space in addition to processing capability support. Other types of contractual arrangements can be made for office space, security services, furniture, and more in the event of a contingency. If the contingency plan calls for moving offsite, procedures need to be developed to ensure a smooth transition back to the primary operating facility or to a new facility. Protection of the physical infrastructure is normally an important part of the emergency response plan, such as use of fire extinguishers or protecting equipment from water damage.

11.4.6 Documents and Papers

The primary contingency strategy is usually backup onto magnetic, optical, microfiche, paper, or other medium and offsite storage. Paper documents are generally harder to backup than electronic ones. A supply of forms and other needed papers can be stored offsite.

11.5 Step 5: Implementing the Contingency Strategies

Once the contingency planning strategies have been selected, it is necessary to make appropriate preparations, document the strategies, and train employees. Many of these tasks are ongoing.

11.5.1 Implementation

Much preparation is needed to implement the strategies for protecting critical functions and their supporting resources. For example, one common preparation is to establish procedures for backing up files and applications. Another is to establish contracts and agreements, *if* the contingency strategy calls for them. Existing service contracts may need to be renegotiated to add contingency services. Another preparation may be to purchase equipment, especially to support a redundant capability.

It is important to keep preparations, including documentation, up-to-date. Computer systems change rapidly and so should backup services and redundant equipment. Contracts and agreements may also need to reflect the changes. If additional equipment is needed, it must be maintained and periodically replaced when it is no longer dependable or no longer fits the organization's architecture.

Backing up data files and applications is a critical part of virtually every contingency plan. Backups are used, for example, to restore files after a personal computer virus corrupts the files or after a hurricane destroys a data processing center.

Preparation should also include formally designating people who are responsible for various tasks in the event of a contingency. These people are often referred to as the contingency response team. This team is often composed of people who were a part of the contingency planning team.

There are many important implementation issues for an organization. Two of the most important are 1) how many plans should be developed? and 2) who prepares each plan? Both of these questions revolve around the organization's overall strategy for contingency planning. The answers should be documented in organization policy and procedures.

How Many Plans?

Some organizations have just one plan for the entire organization, and others have a plan for every distinct computer system, application, or other resource. Other approaches recommend a plan for each business or mission function, with separate plans, as needed, for critical resources.

Relationship Between Contingency Plans and Computer Security Plans

For small or less complex systems, the contingency plan may be a part of the computer security plan. For larger or more complex systems, the computer security plan could contain a brief synopsis of the contingency plan, which would be a separate document.

The answer to the question, therefore, depends upon the unique circumstances for each organization. But it is critical to coordinate between resource managers and functional managers who are responsible for the mission or business.

Who Prepares the Plan?

If an organization decides on a centralized approach to contingency planning, it may be best to name a *contingency planning coordinator*. The coordinator prepares the plans in cooperation with various functional and resource managers. Some organizations place responsibility directly with the functional and resource managers.

11.5.2 Documenting

The contingency plan needs to be written, kept up-to-date as the system and other factors change, and stored in a safe place. A written plan is critical during a contingency, especially if the person who developed the plan is unavailable. It should clearly state in simple language the sequence of tasks to be performed in the event of a contingency so that someone with minimal knowledge could immediately begin to execute the plan. It is generally helpful to store up-to-date copies of the contingency plan in several locations, including any off-site locations, such as alternate processing sites or backup data storage facilities.

11.5.3 Training

All personnel should be trained in their contingency-related duties. New personnel should be trained as they join the organization, refresher training may be needed, and personnel will need to practice their skills.

Training is particularly important for effective employee response during emergencies. There is no time to check a manual to determine correct procedures if there is a fire. Depending on the nature of the emergency, there may or may not be time to protect equipment and other assets. Practice is necessary in order to react correctly, especially when human safety is involved.

11.6 Step 6: Testing and Revising

A contingency plan should be tested periodically because there will undoubtedly be flaws in the plan and in its implementation. The plan will become dated as time passes and as the resources used to support critical functions change. Responsibility for keeping the contingency plan current should be specifically assigned. The extent and frequency of testing will vary between organizations and among systems. There are several types of testing, including reviews, analyses, and simulations of disasters.

> Contingency plan maintenance can be incorporated into procedures for change management so that upgrades to hardware and software are reflected in the plan.

A *review* can be a simple test to check the accuracy of contingency plan documentation. For instance, a reviewer could check if individuals listed are still in the organization and still have the responsibilities that caused them to be included in the plan. This test can check home and work telephone numbers, organizational codes, and building and room numbers. The review can determine if files can be restored from backup tapes or if employees know emergency procedures.

An *analysis* may be performed on the entire plan or portions of it, such as emergency response procedures. It is beneficial if the analysis is performed by someone who did *not* help develop the contingency plan but has a good working knowledge of the critical function and supporting resources. The analyst(s) may mentally follow the strategies in the contingency plan, looking for flaws in the logic or process used by the plan's developers. The analyst may also interview functional managers, resource managers, and their staff to uncover missing or unworkable pieces of the plan.

> The results of a "test" often implies a grade assigned for a specific level of performance, or simply pass or fail. However, in the case of contingency planning, a test should be used to improve the plan. If organizations do not use this approach, flaws in the plan may remain hidden and uncorrected.

Organizations may also arrange *disaster simulations*. These tests provide valuable information about flaws in the contingency plan and provide practice for a real emergency. While they can be expensive, these tests can also provide critical information that can be used to ensure the continuity of important functions. In general, the more critical the functions and the resources addressed in the contingency plan, the more cost-beneficial it is to perform a disaster simulation.

11.7 Interdependencies

Since all controls help to prevent contingencies, there is an interdependency with all of the controls in the handbook.

Risk Management provides a tool for analyzing the security costs and benefits of various contingency planning options. In addition, a risk management effort can be used to help identify critical resources needed to support the organization and the likely threat to those resources. It is not necessary, however, to perform a risk assessment prior to contingency planning, since the identification of critical resources can be performed during the contingency planning process itself.

Physical and Environmental Controls help prevent contingencies. Although many of the other controls, such as logical access controls, also prevent contingencies, the major threats that a contingency plan addresses are physical and environmental threats, such as fires, loss of power, plumbing breaks, or natural disasters.

Incident Handling can be viewed as a subset of contingency planning. It is the emergency response capability for various technical threats. Incident handling can also help an organization prevent future incidents.

Support and Operations in most organizations includes the periodic backing up of files. It also

includes the prevention and recovery from more common contingencies, such as a disk failure or corrupted data files.

Policy is needed to create and document the organization's approach to contingency planning. The policy should explicitly assign responsibilities.

11.8 Cost Considerations

The cost of developing and implementing contingency planning strategies can be significant, especially if the strategy includes contracts for backup services or duplicate equipment. There are too many options to discuss cost considerations for each type.

One contingency cost that is often overlooked is the cost of testing a plan. Testing provides many benefits and should be performed, although some of the less expensive methods (such as a review) may be sufficient for less critical resources.

References

Alexander, M. ed. "Guarding Against Computer Calamity." *Infosecurity News*. 4(6), 1993. pp. 26-37.

Coleman, R. "Six Steps to Disaster Recovery." *Security Management*. 37(2), 1993. pp. 61-62.

Dykman, C., and C. Davis, eds. *Control Objectives - Controls in an Information Systems Environment: Objectives, Guidelines, and Audit Procedures*, fourth edition. Carol Stream, IL: The EDP Auditors Foundation, Inc., 1992 (especially Chapter 3.5).

Fites, P., and M. Kratz, *Information Systems Security: A Practitioner's Reference*. New York, NY: Van Nostrand Reinhold, 1993 (esp. Chapter 4, pp. 95-112).

FitzGerald, J. "Risk Ranking Contingency Plan Alternatives." *Information Executive*. 3(4), 1990. pp. 61-63.

Helsing, C. "Business Impact Assessment." *ISSA Access*. 5(3), 1992, pp. 10-12.

Isaac, I. *Guide on Selecting ADP Backup Process Alternatives*. Special Publication 500-124. Gaithersburg, MD: National Bureau of Standards, November 1985.

Kabak, I., and T. Beam, "On the Frequency and Scope of Backups." *Information Executive*, 4(2), 1991. pp. 58-62.

Kay, R. "What's Hot at Hotsites?" *Infosecurity News*. 4(5), 1993. pp. 48-52.

Lainhart, J., and M. Donahue. *Computerized Information Systems (CIS) Audit Manual: A Guideline to CIS Auditing in Governmental Organizations*. Carol Stream, IL: The EDP Auditors Foundation Inc., 1992.

National Bureau of Standards. *Guidelines for ADP Contingency Planning*. Federal Information Processing Standard 87. 1981.

Rhode, R., and J. Haskett. "Disaster Recovery Planning for Academic Computing Centers." *Communications of the ACM*. 33(6), 1990. pp. 652-657.

Chapter 12

COMPUTER SECURITY INCIDENT HANDLING

Computer systems are subject to a wide range of mishaps – from corrupted data files, to viruses, to natural disasters. Some of these mishaps can be fixed through standard operating procedures. For example, frequently occurring events (e.g., a mistakenly deleted file) can usually be readily repaired (e.g., by restoration from the backup file). More severe mishaps, such as outages caused by natural disasters, are normally addressed in an organization's contingency plan. Other damaging events result from *deliberate malicious technical activity* (e.g., the creation of viruses or system hacking).

A computer security incident can result from a computer virus, other malicious code, or a system intruder, either an insider or an outsider. It is used in this chapter to broadly refer to those incidents resulting from deliberate malicious technical activity.[90] It can more generally refer to those incidents that, without technically expert response, could result in severe damage.[91] This definition of a computer security incident is somewhat flexible and may vary by organization and computing environment.

> Malicious code include viruses as well as Trojan horses and worms. A virus is a code segment that replicates by attaching copies of itself to existing executables. A Trojan horse is a program that performs a desired task, but also includes unexpected functions. A worm is a self-replicating program.

Although the threats that hackers and malicious code pose to systems and networks are well known, the occurrence of such harmful events remains unpredictable. Security incidents on larger networks (e.g., the Internet), such as break-ins and service disruptions, have harmed various organizations' computing capabilities. When initially confronted with such incidents, most organizations respond in an *ad hoc* manner. However recurrence of similar incidents often makes it cost-beneficial to develop a standing capability for quick discovery of and response to such events. This is especially true, since incidents can often "spread" when left unchecked thus increasing damage and seriously harming an organization.

Incident handling is closely related to contingency planning as well as support and operations. An incident handling capability may be viewed as a component of contingency planning, because it provides the ability to react quickly and efficiently to disruptions in normal processing. Broadly speaking, contingency planning addresses events with the potential to interrupt system operations. Incident handling can be considered that portion of contingency planning that responds to

---

[90] Organizations may wish to expand this to include, for example, incidents of theft.

[91] Indeed, damage may result, despite the best efforts to the contrary.

malicious technical threats.

This chapter describes how organizations can address computer security incidents (in the context of their larger computer security program) by developing a *computer security incident handling capability*.[92]

Many organizations handle incidents as part of their user support capability (discussed in Chapter 14) or as a part of general system support.

12.1 Benefits of an Incident Handling Capability

The primary benefits of an incident handling capability are *containing* and *repairing* damage from incidents, and *preventing* future damage. In addition, there are less obvious side benefits related to establishing an incident handling capability.

12.1.1 Containing and Repairing Damage From Incidents

When left unchecked, malicious software can significantly harm an organization's computing, depending on the technology and its connectivity. An incident handling capability provides a way for users to report incidents[93] and the appropriate response and assistance to be provided to aid in recovery. Technical capabilities (e.g., trained personnel and virus identification software) are prepositioned, ready to be used as necessary. Moreover, the organization will have already made important contacts with other supportive sources (e.g., legal, technical, and managerial) to aid in containment and recovery efforts.

> Some organizations suffer repeated outbreaks of viruses because the viruses are never completely eradicated. For example suppose two LANs, Personnel and Budget, are connected, and a virus has spread within each. The administrators of each LAN detect the virus and decide to eliminate it on their LAN. The Personnel LAN administrator first eradicates the virus, but since the Budget LAN is not yet virus-free, the Personnel LAN is reinfected. Somewhat later, the Budget LAN administrator eradicates the virus. However, the virus reinfects the Budget LAN from the Personnel LAN. Both administrators may think all is well, but both are reinfected. An incident handling capability allows organizations to address recovery and containment of such incidents in a skilled, coordinated manner.

Without an incident handling capability, certain responses – although well intentioned – can actually make matters worse. In some cases, individuals have unknowingly infected anti-virus software with viruses and then spread them to

[92] See NIST Special Publication 800-3, *Establishing an Incident Response Capability*, November 1991.

[93] A good incident handling capability is closely linked to an organization's training and awareness program. It will have educated users about such incidents and what to do when they occur. This can increase the likelihood that incidents will be reported early, thus helping to minimize damage.

other systems. When viruses spread to local area networks (LANs), most or all of the connected computers can be infected within hours. Moreover, *uncoordinated* efforts to rid LANs of viruses can prevent their eradication.

Many organizations use large LANs internally and also connect to public networks, such as the Internet. By doing so, organizations increase their exposure to threats from intruder activity, especially if the organization has a high profile (e.g., perhaps it is involved in a controversial program). An incident handling capability can provide enormous benefits by responding quickly to suspicious activity and coordinating incident handling with responsible offices and individuals, as necessary. Intruder activity, whether hackers or malicious code, can often affect many systems located at many different network sites; thus, handling the incidents can be logistically complex and can require information from outside the organization. By planning ahead, such contacts can be preestablished and the speed of response improved, thereby containing and minimizing damage. Other organizations may have already dealt with similar situations and may have very useful guidance to offer in speeding recovery and minimizing damage.

12.1.2 Preventing Future Damage

An incident handling capability also assists an organization in preventing (or at least minimizing) damage from future incidents. Incidents can be studied internally to gain a better understanding of the organizations's threats and vulnerabilities so more effective safeguards can be implemented. Additionally, through outside contacts (established by the incident handling capability) early warnings of threats and vulnerabilities can be provided. Mechanisms will already be in place to warn users of these risks.

The incident handling capability allows an organization to learn from the incidents that it has experienced. Data about past incidents (and the corrective measures taken) can be collected. The data can be analyzed for patterns – for example, which viruses are most prevalent, which corrective actions are most successful, and which systems and information are being targeted by hackers. Vulnerabilities can also be identified in this process – for example, whether damage is occurring to systems when a new software package or patch is used. Knowledge about the types of threats that are occurring and the presence of vulnerabilities can aid in identifying security solutions. This information will also prove useful in creating a more effective training and awareness program, and thus help reduce the potential for losses. The incident handling capability assists the training and awareness program by providing information to users as to (1) measures that can help avoid incidents (e.g., virus scanning) and (2) what should be done in case an incident does occur.

> The sharing of incident data among organizations can help at both the national and the international levels to prevent and respond to breaches of security in a timely, coordinated manner.

Of course, the organization's attempts to prevent future losses does not occur in a vacuum. With a sound incident handling

capability, contacts will have been established with counterparts outside the organization. This allows for *early warning* of threats and vulnerabilities that the organization may have not yet experienced. Early preventative measures (generally more cost-effective than repairing damage) can then be taken to reduce future losses. Data is also shared outside the organization to allow others to learn from the organization's experiences.

12.1.3 Side Benefits

Finally, establishing an incident handling capability helps an organization in perhaps unanticipated ways. Three are discussed here.

Uses of Threat and Vulnerability Data. Incident handling can greatly enhance the risk assessment process. An incident handling capability will allow organizations to collect threat data that may be useful in their risk assessment and safeguard selection processes (e.g., in designing new systems). Incidents can be logged and analyzed to determine whether there is a recurring problem (or if other patterns are present, as are sometimes seen in hacker attacks), which would not be noticed if each incident were only viewed in isolation. Statistics on the numbers and types of incidents in the organization can be used in the risk assessment process as an indication of vulnerabilities and threats.[94]

Enhancing Internal Communications and Organization Preparedness. Organizations often find that an incident handling capability enhances internal communications and the readiness of the organization to respond to any type of incident, not just computer security incidents. Internal communications will be improved; management will be better organized to receive communications; and contacts within public affairs, legal staff, law enforcement, and other groups will have been preestablished. The structure set up for reporting incidents can also be used for other purposes.

Enhancing the Training and Awareness Program. The organization's training process can also benefit from incident handling experiences. Based on incidents reported, training personnel will have a better understanding of users' knowledge of security issues. Trainers can use actual incidents to vividly illustrate the importance of computer security. Training that is based on current threats and controls recommended by incident handling staff provides users with information more specifically directed to their current needs – thereby reducing the risks to the organization from incidents.

[94] It is important, however, *not* to assume that since only *n* reports were made, that *n* is the total number of incidents; it is not likely that all incidents will be reported.

12.2 Characteristics of a Successful Incident Handling Capability

A successful incident handling capability has several core characteristics:

- an understanding of the constituency it will serve;
- an educated constituency;
- a means for centralized communications;
- expertise in the requisite technologies; and
- links to other groups to assist in incident handling (as needed).

12.2.1 Defining the Constituency to Be Served

The constituency includes computer users and program managers. Like any other customer-vendor relationship, the constituency will tend to take advantage of the capability if the services rendered are valuable.

> The focus of a computer security incident handling capability may be external as well as internal. An incident that affects an organization may also affect its trading partners, contractors, or clients. In addition, an organization's computer security incident handling capability may be able to help other organizations and, therefore, help protect the community as a whole.

The constituency is not always the entire organization. For example, an organization may use several types of computers and networks but may decide that its incident handling capability is cost-justified only for its personal computer users. In doing so, the organization may have determined that computer viruses pose a much larger risk than other malicious technical threats on other platforms. Or, a large organization composed of several sites may decide that current computer security efforts at some sites do not require an incident handling capability, whereas other sites do (perhaps because of the criticality of processing).

12.2.2 Educated Constituency

Users need to know about, accept, and trust the incident handling capability or it will not be used. Through training and awareness programs, users can become knowledgeable about the existence of the capability and how to recognize and report incidents. Users trust

> Managers need to know details about incidents, including who discovered them and how, so that they can prevent similar incidents in the future. However users will not be forthcoming if they fear reprisal or that they will become scapegoats. Organizations may need to offer incentives to employees for reporting incidents and offer guarantees against reprisal or other adverse actions. It may also be useful to consider anonymous reporting.

in the value of the service will build with reliable performance.

12.2.3 Centralized Reporting and Communications

Successful incident handling requires that users be able to report incidents to the incident handling team in a convenient, straightforward fashion; this is referred to as *centralized reporting*. A successful incident handling capability depends on timely reporting. If it is difficult or time consuming to report incidents, the incident handling capability may not be fully used. Usually, some form of a hotline, backed up by pagers, works well.

Centralized communications is very useful for accessing or distributing information relevant to the incident handling effort. For example, if users are linked together via a network, the incident handling capability can then use the network to send out timely announcements and other information. Users can take advantage of the network to retrieve security information stored on servers and communicate with the incident response team via e-mail.

12.2.4 Technical Platform and Communications Expertise

The technical staff members who comprise the incident handling capability need specific knowledge, skills, and abilities. Desirable qualifications for technical staff members may include the ability to:

- work expertly with some or all of the constituency's core technology;
- work in a group environment;
- communicate effectively with different types of users, who will range from system administrators to unskilled users to management to law-enforcement officials;
- be on-call 24 hours as needed; and
- travel on short notice (of course, this depends upon the physical location of the constituency to be served).

12.2.5 Liaison With Other Organizations

Due to increasing computer connectiv94 -0grludor ectiv94l on networrs canaeffeco mayt Thre foes, a oOrganizatio'sr

supporting organizations.

Especially important to incident handling are contacts with investigative agencies, such as federal (e.g., the FBI), state, and local law enforcement. Laws that affect computer crime vary among localities and states, and some actions may be state (but not federal) crimes. It is important for teams to be familiar with current laws and to have established contacts within law enforcement and investigative agencies.

The Forum of Incident Response and Security Teams

The 1988 Internet worm incident highlighted the need for better methods for responding to and sharing information about incidents. It was also clear that any single team or "hot line" would simply be overwhelmed. Out of this was born the concept of a coalition of response teams – each with its own constituency, but working together to share information, provide alerts, and support each other in the response to incidents. The Forum of Incident Response and Security Teams (FIRST) includes teams from government, industry, computer manufacturers, and academia. NIST serves as the secretariat of FIRST.

Incidents can also garner much media attention and can reflect quite negatively on an organization's image. An incident handling capability may need to work closely with the organization's public affairs office, which is trained in dealing with the news media. In presenting information to the press, it is important that (1) attackers are not given information that would place the organization at greater risk and (2) potential legal evidence is properly protected.

12.3 Technical Support for Incident Handling

Incident handling will be greatly enhanced by technical mechanisms that enable the dissemination of information quickly and conveniently.

12.3.1 Communications for Centralized Reporting of Incidents

The technical ability to report incidents is of primary importance, since without knowledge of an incident, response is precluded. Fortunately, such technical mechanisms are already in place in many organizations.

For rapid response to constituency problems, a simple telephone "hotline" is practical and convenient. Some agencies may already have a number used for emergencies or for obtaining help with other problems; it may be practical (and cost-effective) to also use this number for incident handling. It may be necessary to provide 24-hour coverage for the hotline. This can be done by staffing the answering center, by providing an answering service for nonoffice hours, or by using a combination of an answering machine and personal pagers.

If additional mechanisms for contacting the incident handling team can be provided, it may increase access and thus benefit incident handling efforts. A centralized e-mail address that forwards mail to staff members would permit the constituency to conveniently exchange information with the team. Providing a fax number to users may also be helpful.

> One way to establish a centralized reporting and incident response capability, while minimizing expenditures, is to use an existing Help Desk. Many agencies already have central Help Desks for fielding calls about commonly used applications, troubleshooting system problems, and providing help in detecting and eradicating computer viruses. By expanding the capabilities of the Help Desk and publicizing its telephone number (or e-mail address), an agency may be able to significantly improve its ability to handle many different types of incidents at minimal cost.

12.3.2 Rapid Communications Facilities

Some form of rapid communications is essential for quickly communicating with the constituency as well as with management officials and outside organizations. The team may need to send out security advisories or collect information quickly, thus some convenient form of communications, such as electronic mail, is generally highly desirable. With electronic mail, the team can easily direct information to various subgroups within the constituency, such as system managers or network managers, and broadcast general alerts to the entire constituency as needed. When connectivity already exists, e-mail has low overhead and is easy to use. (However, it is possible for the e-mail system itself to be attacked, as was the case with the 1988 Internet worm.)

Although there are substitutes for e-mail, they tend to increase response time. An electronic bulletin board system (BBS) can work well for distributing information, especially if it provides a convenient user interface that encourages its use. A BBS connected to a network is more convenient to access than one requiring a terminal and modem; however, the latter may be the only alternative for organizations without sufficient network connectivity. In addition, telephones, physical bulletin boards, and flyers can be used.

12.3.3 Secure Communications Facilities

Incidents can range from the trivial to those involving national security. Often when exchanging information about incidents, using encrypted communications may be advisable. This will help prevent the unintended distribution of incident-related information. Encryption technology is available for voice, fax, and e-mail communications.

12.4 Interdependencies

An incident handling capability generally depends upon other safeguards presented in this handbook. The most obvious is the strong link to other components of the contingency plan. The following paragraphs detail the most important of these interdependencies.

Contingency Planning. As discussed in the introduction to this chapter, an incident handling capability can be viewed as the component of contingency planning that deals with responding to technical threats, such as viruses or hackers. Close coordination is necessary with other contingency planning efforts, particularly when planning for contingency processing in the event of a serious unavailability of system resources.

Support and Operations. Incident handling is also closely linked to support and operations, especially user support and backups. For example, for purposes of efficiency and cost savings, the incident handling capability is often co-operated with a user "help desk." Also, backups of system resources may need to be used when recovering from an incident.

Training and Awareness. The training and awareness program can benefit from lessons learned during incident handling. Incident handling staff will be able to help assess the level of user awareness about current threats and vulnerabilities. Staff members may be able to help train system administrators, system operators, and other users and systems personnel. Knowledge of security precautions (resulting from such training) helps reduce future incidents. It is also important that users are trained what to report and how to report it.

Risk Management. The risk analysis process will benefit from statistics and logs showing the numbers and types of incidents that have occurred and the types of controls that are effective in preventing incidents. This information can be used to help select appropriate security controls and practices.

12.5 Cost Considerations

There are a number of start-up costs and funding issues to consider when planning an incident handling capability. Because the success of an incident handling capability relies so heavily on users' perceptions of its worth and whether they use it, it is very important that the capability be able to meet users' requirements. Two important funding issues are:

Personnel. An incident handling capability plan might call for at least one manager and one or more technical staff members (or their equivalent) to accomplish program objectives. Depending on the scope of the effort, however, full-time staff members may not be required. In some situations, some staff may be needed part-time or on an on-call basis. Staff may be performing incident handling duties as an adjunct responsibility to their normal assignments.

Education and Training. Incident handling staff will need to keep current with computer system and security developments. Budget allowances need to be made, therefore, for attending conferences, security seminars, and other continuing-education events. If an organization is located in more than one geographic areas, funds will probably be needed for travel to other sites for handling incidents.

References

Brand, Russell L. *Coping With the Threat of Computer Security Incidents: A Primer from Prevention Through Recovery*. July 1989.

Fedeli, Alan. "Organizing a Corporate Anti-Virus Effort." *Proceedings of the Third Annual Computer VIRUS Clinic*, Nationwide Computer Corp. March 1990.

Holbrook, P., and J. Reynolds, eds. *Site Security Handbook*. RFC 1244 prepared for the Internet Engineering Task Force, 1991. FTP from csrc.nist.gov:/put/secplcy/rfc1244.txt.

National Institute of Standards and Technology. "Establishing a Computer Security Incident Response Capability." Computer Systems Laboratory Bulletin. Gaithersburg, MD. February 1992.

Padgett, K. *Establishing and Operating an Incident Response Team*. Los Alamos, NM: Los Alamos National Laboratory, 1992.

Pethia, Rich, and Kenneth van Wyk. *Computer Emergency Response - An International Problem*. 1990.

Quarterman, John. *The Matrix - Computer Networks and Conferencing Systems Worldwide.* Digital Press, 1990.

Scherlis, William, S. Squires, and R. Pethia. *Computer Emergency Response*. 1989.

Schultz, E., D. Brown, and T. Longstaff. *Responding to Computer Security Incidents: Guidelines for Incident Handling*. University of California Technical Report UCRL-104689, 1990.

Proceedings of the Third Invitational Workshop on Computer Security Incident Response. August 1991.

Wack, John. *Establishing an Incident Response Capability*. Special Publication 800-3. Gaithersburg, MD: National Institute of Standards and Technology. November 1991.

Chapter 13

AWARENESS, TRAINING, AND EDUCATION

People, who are all fallible, are usually recognized as one of the weakest links in securing systems. The purpose of computer security awareness, training, and education is to enhance security by:

- improving awareness of the need to protect system resources;
- developing skills and knowledge so computer users can perform their jobs more securely; and
- building in-depth knowledge, as needed, to design, implement, or operate security programs for organizations and systems.

Making computer system users aware of their security responsibilities and teaching them correct practices helps users change their behavior.[95] It also supports *individual accountability*, which is one of the most important ways to improve computer security. Without knowing the necessary security measures (and to how to use them), users cannot be truly accountable for their actions. The importance of this training is emphasized in the Computer Security Act, which requires training for those involved with the management, use, and operation of federal computer systems.

This chapter first discusses the two overriding benefits of awareness, training, and education, namely: (1) improving employee behavior and (2) increasing the ability to hold employees accountable for their actions. Next, awareness, training, and education are discussed separately, with techniques used for each. Finally, the chapter presents one approach for developing a computer security awareness and training program.[96]

13.1 Behavior

People are a crucial factor in ensuring the security of computer systems and valuable information resources. Human actions account for a far greater degree of computer-related loss than all other sources combined. Of such losses, the actions of an organization's insiders normally cause far more harm than the actions of outsiders. (Chapter 4 discusses the major sources of computer-related loss.)

---

[95] One often-cited goal of training is changing people's attitudes. This chapter views changing attitudes as just one step toward changing behavior.

[96] This chapter does not discuss the specific contents of training programs. See the references for details of suggested course contents.

The major causes of loss due to an organization's own employees are: errors and omissions, fraud, and actions by disgruntled employees. One principal purpose of security awareness, training, and education is to reduce errors and omissions. However, it can also reduce fraud and unauthorized activity by disgruntled employees by increasing employees' knowledge of their accountability and the penalties associated with such actions.

Management sets the example for behavior within an organization. If employees know that management does not care about security, no training class teaching the importance of security and imparting valuable skills can be truly effective. This "tone from the top" has myriad effects an organization's security program.

13.2 Accountability

> One of the keys to a successful computer security program is security awareness and training. If employees are not informed of applicable organizational policies and procedures, they cannot be expected to act effectively to secure computer resources.

Both the *dissemination* and the *enforcement* of policy are critical issues that are implemented and strengthened through training programs. Employees cannot be expected to follow policies and procedures of which they are unaware. In addition, enforcing penalties may be difficult if users can claim ignorance when caught doing something wrong.

Training employees may also be necessary to show that a standard of *due care* has been taken in protecting information. Simply issuing policy, with no follow-up to implement that policy, may not suffice.

Many organizations use *acknowledgment statements* which state that employees have read and understand computer security requirements. (An example is provided in Chapter 10.)

13.3 Awareness

> Security *awareness* programs: (1) set the stage for *training* by changing organizational attitudes to realize the importance of security and the adverse consequences of its failure; and (2) remind users of the procedures to be followed.

Awareness stimulates and motivates those being trained to care about security and to remind them of important security practices. Explaining what happens to an organization, its mission, customers, and employees if security fails motivates people to take security seriously.

Awareness can take on different forms for particular audiences. Appropriate awareness for management officials might stress management's pivotal role in establishing organizational

attitudes toward security. Appropriate awareness for other groups, such as system programmers or information analysts, should address the need for security as it relates to their job. In today's systems environment, almost everyone in an organization may have access to system resources – and therefore may have the potential to cause harm.

Comparative Framework

	AWARENESS	TRAINING	EDUCATION
Attribute:	"What"	"How"	"Why"
Level:	Information	Knowledge	Insight
Objective:	Recognition	Skill	Understanding
Teaching Method:	Media - Videos -Newsletters -Posters, etc.	Practical Instruction - Lecture - Case study workshop - Hands-on practice	Theoretical Instruction - Discussion Seminar - Background reading
Test Measure:	True/False Multiple Choice (identify learning)	Problem Solving (apply learning)	Eassay (interpret learning)
Impact Timeframe:	Short-term	Intermediate	Long-term

Figure 13.1 compares some of the differences in awareness, training, and education.

Awareness is used to reinforce the fact that security supports the mission of the organization by protecting valuable resources. If employees view security as just bothersome rules and procedures, they are more likely to ignore them. In addition, they may not make needed suggestions about improving security nor recognize and report security threats and vulnerabilities.

Awareness also is used to remind people of basic security practices, such as logging off a computer system or locking doors.

Techniques. A security awareness program can use many teaching methods, including video

tapes, newsletters, posters, bulletin boards, flyers, demonstrations, briefings, short reminder notices at log-on, talks, or lectures. Awareness is often incorporated into basic security training and can use any method that can change employees' attitudes.

Effective security awareness programs need to be designed with the recognition that people tend to practice a *tuning out* process (also known as *acclimation*). For example, after a while, a security poster, no matter how well designed, will be ignored; it will, in effect, simply blend into the environment. For this reason, awareness techniques should be creative and frequently changed.

> Employees often regard computer security as an obstacle to productivity. A common feeling is that they are paid to produce, not to protect. To help motivate employees, awareness should emphasize how security, from a broader perspective, contributes to productivity. The consequences of poor security should be explained, while avoiding the fear and intimidation that employees often associate with security.

13.4 Training

The purpose of training is to teach people the skills that will enable them to perform their jobs more securely. This includes teaching people *what* they should do and *how* they should (or can) do it. Training can address many levels, from basic security practices to more advanced or specialized skills. It can be specific to one computer system or generic enough to address all systems.

Training is most effective when targeted to a specific audience. This enables the training to focus on security-related job skills and knowledge that people need performing their duties. Two types of audiences are general users and those who require specialized or advanced skills.

General Users. Most users need to understand good computer security practices, such as:

- protecting the physical area and equipment (e.g., locking doors, caring for floppy diskettes);
- protecting passwords (if used) or other authentication data or tokens (e.g., never divulge PINs); and
- reporting security violations or incidents (e.g., whom to call if a virus is suspected).

In addition, general users should be taught the organization's policies for protecting information and computer systems and the roles and responsibilities of various organizational units with which they may have to interact.

In teaching general users, care should be taken not to overburden them with unneeded details. These people are the target of multiple training programs, such as those addressing safety, sexual harassment, and AIDS in the workplace. The training should be made useful by addressing security issues that *directly* affect the users. The goal is to improve basic security practices, *not* to make everyone literate in all the jargon or philosophy of security.

Specialized or Advanced Training. Many groups need more advanced or more specialized training than just basic security practices. For example, managers may need to understand security consequences and costs so they can factor security into their decisions, or system administrators may need to know how to implement and use specific access control products.

There are many different ways to identify individuals or groups who need specialized or advanced training. One method is to look at job categories, such as executives, functional managers, or technology providers. Another method is to look at job functions, such as system design, system operation, or system use. A third method is to look at the specific technology and products used, especially for advanced training for user groups and training for a new system. This is further discussed in the section 13.6 of this chapter.

One group that has been targeted for specialized training is executives and functional managers. The training for management personnel is specialized (rather than advanced) because managers do *not* (as a general rule) need to understand the technical details of security. However, they do need to understand how to organize, direct, and evaluate security measures and programs. They also need to understand risk acceptance.

Techniques. A security training program normally includes training classes, either strictly devoted to security or as added special sections or modules within existing training classes. Training may be computer- or lecture-based (or both), and may include hands-on practice and case studies. Training, like awareness, also happens on the job.

13.5 Education

Security education is more in-depth than security training and is targeted for security professionals and those whose jobs require *expertise* in security.

Techniques. Security education is normally outside the scope of most organization awareness and training programs. It is more appropriately a part of *employee career development*. Security education is obtained through college or graduate classes or through specialized training programs. Because of this, most computer security programs focus primarily on awareness and

training, as does the remainder of this chapter.[97]

13.6 Implementation[98]

An effective computer security awareness and training (CSAT) program requires proper planning, implementation, maintenance, and periodic evaluation. The following seven steps constitute *one approach* for developing a CSAT program.[99]

Step 1: Identify Program Scope, Goals, and Objectives.

Step 2: Identify Training Staff.

Step 3: Identify Target Audiences.

Step 4: Motivate Management and Employees.

Step 5: Administer the Program.

Step 6: Maintain the Program.

Step 7: Evaluate the Program.

13.6.1 Identify Program Scope, Goals, and Objectives

The first step in developing a CSAT program is to determine the program's scope, goals, and objectives. The scope of the CSAT program should provide training to all types of people who interact with computer systems. The scope of the program can be an entire organization or a subunit. Since users need training which relates directly to their use of

> The Computer Security Act of 1987 requires federal agencies to "provide for the mandatory periodic training in computer security awareness and accepted computer practices of all employees who are involved with the management, use, or operation of each federal computer system within or under the supervision of that agency." The scope and goals of federal computer security awareness and training programs must implement this broad mandate. (Other federal requirements for computer security training are contained in OMB Circular A-130, Appendix III, and OPM regulations.)

[97] Unfortunately, college and graduate security courses are not widely available. In addition, the courses may only address general security.

[98] This section is based on material prepared by the Department of Energy's Office of Information Management for its unclassified security program.

[99] This approach is presented to familiarize the reader with some of the important implementation issues. It is not the only approach to implementing an awareness and training program.

particular systems, a large organizationwide program may need to be supplemented by more specific programs. In addition, the organization should specifically address whether the program applies to employees only or also to other users of organizational systems.

Generally, the overall goal of a CSAT program is to sustain an appropriate level of protection for computer resources by increasing employee awareness of their computer security responsibilities and the ways to fulfill them. More specific goals may need to be established. Objectives should be defined to meet the organization's specific goals.

13.6.2 Identify Training Staff

There are many possible candidates for conducting the training including internal training departments, computer security staff, or contract services. Regardless of who is chosen, it is important that trainers have sufficient knowledge of computer security issues, principles, and techniques. It is also vital that they know how to communicate information and ideas effectively.

13.6.3 Identify Target Audiences

Not everyone needs the same degree or type of computer security information to do their jobs. A CSAT program that distinguishes between groups of people, presents only the information needed by the particular audience, and omits irrelevant information will have the best results. Segmenting audiences (e.g., by their function or familiarity with the system) can also improve the effectiveness of a CSAT program. For larger organizations, some individuals will fit into more than one group. For smaller organizations, segmenting may not be needed. The following methods are some examples of ways to do this.

Segment according to level of awareness. Individuals may be separated into groups according to their current level of awareness. This may require research to determine how well employees follow computer security procedures or understand how computer security fits into their jobs.

Segment according to general job task or function. Individuals may be grouped as data providers, data processors, or data users.

Segment according to specific job category. Many organizations assign individuals to job categories. Since each job category generally has different job responsibilities, training for each will be different. Examples of job categories could be general management, technology management, applications development, or security.

Segment according to level of computer knowledge. Computer experts may be expected to find a program containing highly technical information more valuable than one covering the management issues in computer security. Similarly, a computer novice would benefit more from a training program that presents introductory fundamentals.

Segment according to types of technology or systems used. Security techniques used for each off-the-shelf product or application system will usually vary. The users of major applications will normally require training specific to that application.

13.6.4 Motivate Management and Employees

To successfully implement an awareness and training program, it is important to gain the *support* of management and employees. Consideration should be given to using motivational techniques to show management and employees how their participation in the CSAT program will benefit the organization.

Management. Motivating management normally relies upon increasing awareness. Management needs to be aware of the losses that computer security can reduce and the role of training in computer security. Management commitment is necessary because of the resources used in developing and implementing the program and also because the program affects their staff.

Employees. Motivation of managers alone is not enough. Employees often need to be convinced of the merits of computer security and how it relates to their jobs. Without appropriate training, many employees will not fully comprehend the value of the system resources with which they work.

> Employees and managers should be solicited to provide input to the CSAT program. Individuals are more likely to support a program when they have actively participated in its development.

Some awareness techniques were discussed above. Regardless of the techniques that are used, employees should feel that their cooperation will have a beneficial impact on the organization's future (and, consequently, their own).

13.6.5 Administer the Program

There are several important considerations for administering the CSAT program.

Visibility. The visibility of a CSAT program plays a key role in its success. Efforts to achieve high visibility should begin during the early stages of CSAT program development. However, care should be give not to promise what cannot be delivered.

Training Methods. The methods used in the CSAT program should be consistent with the material presented and tailored to the audience's needs. Some training and

> The Federal Information Systems Security Educators' Association and NIST Computer Security Program Managers' Forum provide two means for federal government computer security program managers and training officers to share training ideas and materials.

awareness methods and techniques are listed above (in the *Techniques* sections). Computer security awareness and training can be added to existing courses and presentations or taught separately. On-the-job training should also be considered.

Training Topics. There are more topics in computer security than can be taught in any one course. Topics should be selected based on the audience's requirements.

Training Materials. In general, higher-quality training materials are more favorably received and are more expensive. Costs, however, can be minimized since training materials can often be obtained from other organizations. The cost of modifying materials is normally less than developing training materials from scratch.

Training Presentation. Consideration should be given to the frequency of training (e.g., annually or as needed), the length of training presentations (e.g., 20 minutes for general presentations, one hour for updates or one week for an off-site class), and the style of training presentation (e.g., formal presentation, informal discussion, computer-based training, humorous).

13.6.6 Maintain the Program

Computer technology is an ever-changing field. Efforts should be made to keep abreast of changes in computer technology and security requirements. A training program that meets an organization's needs today may become ineffective when the organization starts to use a new application or changes its environment, such as by connecting to the Internet. Likewise, an awareness program can become obsolete if laws or organization policies change. For example, the awareness program should make employees aware of a new policy on e-mail usage. Employees may discount the CSAT program, and by association the importance of computer security, if the program does not provide current information.

13.6.7 Evaluate the Program

It is often difficult to measure the effectiveness of an awareness or training program. Nevertheless, an evaluation should attempt to ascertain how much information is retained, to what extent computer security procedures are being followed, and general attitudes toward computer security. The results of such an evaluation should help identify and correct problems. Some evaluation methods (which can be used in conjunction with one another) are:

- Use student evaluations.
- Observe how well employees follow recommended security procedures.
- Test employees on material covered.

- Monitor the number and kind of computer security incidents reported before and after the program is implemented.[100]

13.7 Interdependencies

Training can, and in most cases should, be used to support every control in the handbook. All controls are more effective if designers, implementers, and users are thoroughly trained.

Policy. Training is a critical means of informing employees of the contents of and reasons for the organization's policies.

Security Program Management. Federal agencies need to ensure that appropriate computer security awareness and training is provided, as required under the Computer Security Act of 1987. A security program should ensure that an organization is meeting all applicable laws and regulations.

Personnel/User Issues. Awareness, training, and education are often included with other personnel/user issues. Training is often required before access is granted to a computer system.

13.8 Cost Considerations

The major cost considerations in awareness, training, and education programs are:

- the cost of preparing and updating materials, including the time of the preparer;
- the cost of those providing the instruction;
- employee time attending courses and lectures or watching videos; and
- the cost of outside courses and consultants (both of which may including travel expenses), including course maintenance.

References

Alexander, M. ed. "Multimedia Means Greater Awareness." *Infosecurity News*. 4(6), 1993. pp. 90-94.

[100] The number of incidents will not necessarily go down. For example, virus-related losses may decrease when users know the proper procedures to avoid infection. On the other hand, reports of incidents may go up as users employ virus scanners and find more viruses. In addition, users will now know that virus incidents should be reported and to whom the reports should be sent.

Burns, G.M. "A Recipe for a Decentralized Security Awareness Program." *ISSA Access*. Vol. 3, Issue 2, 2nd Quarter 1990. pp. 12-54.

Code of Federal Regulations. 5 CFR 930. Computer Security Training Regulation.

Flanders, D. "Security Awareness - A 70% Solution." Fourth Workshop on Computer Security Incident Handling, August 1992.

Isaacson, G. "Security Awareness: Making It Work." *ISSA Access*. 3(4), 1990. pp. 22-24.

National Aeronautics and Space Administration. *Guidelines for Development of Computer Security Awareness and Training (CSAT) Programs*. Washington, DC. NASA Guide 2410.1. March 1990.

Maconachy, V. "Computer Security Education, Training, and Awareness: Turning a Philosophical Orientation Into Practical Reality." *Proceedings of the 12th National Computer Security Conference*. National Institute of Standards and Technology and National Computer Security Center. Washington, DC. October 1989.

Maconachy, V. "Panel: Federal Information Systems Security Educators' Association (FISSEA)." *Proceeding of the 15th National Computer Security Conference*. National Institute of Standards and Technology and National Computer Security Center. Baltimore, MD. October 1992.

Suchinsky, A. "Determining Your Training Needs." *Proceedings of the 13th National Computer Security Conference*. National Institute of Standards and Technology and National Computer Security Center. Washington, DC. October 1990.

Todd, M.A. and Guitian C. *"Computer Security Training Guidelines*." Special Publication 500-172. Gaithersburg, MD: National Institute of Standards and Technology. November 1989.

U.S. Department of Energy. *Computer Security Awareness and Training Guideline* (Vol. 1). Washington, DC. DOE/MA-0320. February 1988.

Wells, R.O. "Security Awareness for the Non-Believers." *ISSA Access.* Vol. 3, Issue 2, 2nd Quarter 1990. pp. 10-61.

Chapter 14

SECURITY CONSIDERATIONS IN COMPUTER SUPPORT AND OPERATIONS

Computer support and operations refers to everything done to run a computer system. This includes both system administration and tasks external to the system that support its operation (e.g., maintaining documentation). It does not include system planning or design. The support and operation of any computer system, from a three-person local area network to a worldwide application serving thousands of users, is critical to maintaining the security of a system. Support and operations are routine activities that enable computer systems to function correctly. These include fixing software or hardware problems, loading and maintaining software, and helping users resolve problems.

System management and administration staff generally perform support and operations tasks although sometimes users do. Larger systems may have full-time operators, system programmers, and support staff performing these tasks. Smaller systems may have a part-time administrator.

The failure to consider security as part of the support and operations of computer systems is, for many organizations, their Achilles heel. Computer security system literature includes many examples of how organizations undermined their often expensive security measures because of poor documentation, old user accounts, conflicting software, or poor control of maintenance accounts. Also, an organization's policies and procedures often fail to address many of these important issues.

The important security considerations within some of the major categories of support and operations are:

- user support,
- software support,
- configuration management,
- backups,
- media controls,
- documentation, and
- maintenance.

The primary goal of computer support and operations is the continued and correct operation of a computer system. One of the goals of computer security is the availability and integrity of systems. These goals are very closely linked.

Some special considerations are noted for larger or smaller systems.[101]

This chapter addresses the support and operations activities directly related to security. Every control discussed in this handbook relies, in one way or another, on computer system support and operations. This chapter, however, focuses on areas *not covered in other chapters*. For example, operations personnel normally create user accounts on the system. This topic is covered in the Identification and Authentication chapter, so it is not discussed here. Similarly, the input from support and operations staff to the security awareness and training program is covered in the Security Awareness, Training, and Education chapter.

14.1 User Support

In many organizations, user support takes place through a Help Desk. Help Desks can support an entire organization, a subunit, a specific system, or a combination of these. For smaller systems, the system administrator normally provides direct user support. Experienced users provide informal user support on most systems.

An important security consideration for user support personnel is being able to recognize which problems (brought to their attention by users) are security-related. For example, users' inability to log onto a computer system may result from the disabling of their accounts due to too many failed access attempts. This could indicate the presence of hackers trying to guess users' passwords.

> User support should be closely linked to the organization's incident handling capability. In many cases, the same personnel perform these functions.

In general, system support and operations staff need to be able to identify security problems, respond appropriately, and inform appropriate individuals. A wide range of possible security problems exist. Some will be internal to custom applications, while others apply to off-the-shelf products. Additionally, problems can be software- or hardware-based.

The more responsive and knowledgeable system support and operation staff personnel are, the less user support will be provided informally. The support other users provide is important, but they may not be aware of the "whole picture."

> Small systems are especially susceptible to viruses, while networks are particularly susceptible to hacker attacks, which can be targeted at multiple systems. System support personnel should be able to recognize attacks and know how to respond.

[101] In general, larger systems include mainframes, large minicomputers, and WANs. Smaller systems include PCs and LANs.

14.2 Software Support

Software is the heart of an organization's computer operations, whatever the size and complexity of the system. Therefore, it is essential that software function correctly and be protected from corruption. There are many elements of software support.

One is *controlling what software is used on a system*. If users or systems personnel can load and execute any software on a system, the system is more vulnerable to viruses, to unexpected software interactions, and to software that may subvert or bypass security controls. One method of controlling software is to inspect or test software before it is loaded (e.g., to determine compatibility with custom applications or identify other unforeseen interactions). This can apply to new software packages, to upgrades, to off-the-shelf products, or to custom software, as deemed appropriate. In addition to controlling the loading and execution of new software, organizations should also give care to the configuration and use of powerful system utilities. System utilities can compromise the integrity of operating systems and logical access controls.

A second element in software support can be to ensure that *software has not been modified without proper authorization*. This involves the protection of software and backup copies. This can be done with a combination of logical and physical access controls.

Many organizations also include a program to ensure that software is properly licensed, as required. For example, an organization may audit systems for illegal copies of copyrighted software. This problem is primarily associated with PCs and LANs, but can apply to any type of system.

Viruses take advantage of the weak software controls in personal computers. Also, there are powerful utilities available for PCs that can restore deleted files, find hidden files, and interface directly with PC hardware, bypassing the operating system. Some organizations use personal computers without floppy drives in order to have better control over the system.

There are several widely available utilities that look for security problems in both networks and the systems attached to them. Some utilities look for and try to exploit security vulnerabilities. (This type of software is further discussed in Chapter 9.)

14.3 Configuration Management

Closely related to software support is *configuration management* – the process of keeping track of changes to the system and, if needed, approving them.[102] Configuration management normally addresses hardware, software, networking, and other changes; it can be formal or informal. The primary security goal of configuration management is ensuring that changes to the system do not

[102] This chapter only addresses configuration management during the operational phase. Configuration management can have extremely important security consequences during the development phase of a system.

unintentionally or unknowingly diminish security. Some of the methods discussed under software support, such as inspecting and testing software changes, can be used. Chapter 9 discusses other methods.

Note that the security goal is to know what changes occur, not to prevent security from being changed. There may be circumstances when security will be reduced. However, the decrease in security should be the result of a decision based on all appropriate factors.

> For networked systems, configuration management should include external connections. Is the computer system connected? To what other systems? In turn, to what systems are these systems and organizations connected?

A second security goal of configuration management is ensuring that changes to the system are reflected in other documentation, such as the contingency plan. If the change is major, it may be necessary to reanalyze some or all of the security of the system. This is discussed in Chapter 8.

14.4 Backups

Support and operations personnel and sometimes users back up software and data. This function is critical to contingency planning. Frequency of backups will depend upon how often data changes and how important those changes are. Program managers should be consulted to determine what backup schedule is appropriate. Also, as a safety measure, it is useful to test that backup copies are actually usable. Finally, backups should be stored securely, as appropriate (discussed below).

> Users of smaller systems are often responsible for their own backups. However, in reality they do not always perform backups regularly. Some organizations, therefore, task support personnel with making backups periodically for smaller systems, either automatically (through server software) or manually (by visiting each machine).

14.5 Media Controls

Media controls include a variety of measures to provide physical and environmental protection and accountability for tapes, diskettes, printouts, and other media. From a security perspective, media controls should be designed to prevent the loss of confidentiality, integrity, or availability of information, including data or software, when stored outside the system. This can include storage of information before it is input to the system and after it is output.

The extent of media control depends upon many factors, including the type of data, the quantity of media, and the nature of the user environment. Physical and environmental protection is used to prevent unauthorized individuals from accessing the media. It also protects against such

factors as heat, cold, or harmful magnetic fields. When necessary, logging the use of individual media (e.g., a tape cartridge) provides detailed accountability – to hold authorized people responsible for their actions.

14.5.1 Marking

Controlling media may require some form of physical labeling. The labels can be used to identify media with special handling instructions, to locate needed information, or to log media (e.g., with serial/control numbers or bar codes) to support accountability. Identification is often by colored labels on diskettes or tapes or banner pages on printouts.

If labeling is used for special handling instructions, it is critical that people be appropriately trained. The marking of PC input and output is generally the responsibility of the *user*, not the system support staff. Marking backup diskettes can help prevent them from being accidentally overwritten.

> Typical markings for media could include: Privacy Act Information, Company Proprietary, or Joe's Backup Tape. In each case, the individuals handling the media must know the applicable handling instructions. For example, at the Acme Patent Research Firm, proprietary information may not leave the building except under the care of a security officer. Also, Joe's Backup Tape should be easy to find in case something happens to Joe's system.

14.5.2 Logging

The logging of media is used to support accountability. Logs can include control numbers (or other tracking data), the times and dates of transfers, names and signatures of individuals involved, and other relevant information. Periodic spot checks or audits may be conducted to determine that no controlled items have been lost and that all are in the custody of individuals named in control logs. Automated media tracking systems may be helpful for maintaining inventories of tape and disk libraries.

14.5.3 Integrity Verification

When electronically stored information is read into a computer system, it may be necessary to determine whether it has been read correctly or subject to any modification. The integrity of electronic information can be verified using error detection and correction or, if intentional modifications are a threat, cryptographic-based technologies. (See Chapter 19.)

14.5.4 Physical Access Protection

Media can be stolen, destroyed, replaced with a look-alike copy, or lost. Physical access controls, which can limit these problems, include locked doors, desks, file cabinets, or safes.

If the media requires protection at all times, it may be necessary to actually output data to the

media in a secure location (e.g., printing to a printer in a locked room instead of to a general-purpose printer in a common area).

Physical protection of media should be extended to backup copies stored offsite. They generally should be accorded an equivalent level of protection to media containing the same information stored onsite. (Equivalent protection does not mean that the security measures need to be exactly the same. The controls at the off-site location are quite likely to be different from the controls at the regular site.) Physical access is discussed in Chapter 15.

14.5.5 Environmental Protection

Magnetic media, such as diskettes or magnetic tape, require environmental protection, since they are sensitive to temperature, liquids, magnetism, smoke, and dust. Other media (e.g., paper and optical storage) may have different sensitivities to environmental factors.

14.5.6 Transmittal

Media control may be transferred both within the organization and to outside elements. Possibilities for securing such transmittal include sealed and marked envelopes, authorized messenger or courier, or U.S. certified or registered mail.

14.5.7 Disposition

When media is disposed of, it may be important to ensure that information is not improperly disclosed. This applies both to media that is *external* to a computer system (such as a diskette) and to media *inside* a computer system, such as a hard disk. The process of removing information from media is called *sanitization.*

> Many people throw away old diskettes, believing that erasing the files on the diskette has made the data unretrievable. In reality, however, erasing a file simply removes the pointer to that file. The pointer tells the computer where the file is physically stored. Without this pointer, the files will not appear on a directory listing. This does *not* mean that the file was removed. Commonly available utility programs can often retrieve information that is presumed deleted.

Three techniques are commonly used for media sanitization: overwriting, degaussing, and destruction. *Overwriting* is an effective method for clearing data from magnetic media. As the name implies, overwriting uses a program to write (1s, 0s, or a combination) onto the media. Common practice is to overwrite the media three times. Overwriting should not be confused with merely deleting the pointer to a file (which typically happens when a *delete* command is used). Overwriting requires that the media be in working order. *Degaussing* is a method to magnetically erase data from magnetic media. Two types of degausser exist: strong permanent magnets and electric degaussers. The final method of sanitization is *destruction* of the media by shredding or burning.

14.6 Documentation

Documentation of all aspects of computer support and operations is important to ensure continuity and consistency. Formalizing operational practices and procedures with sufficient detail helps to eliminate security lapses and oversights, gives new personnel sufficiently detailed instructions, and provides a quality assurance function to help ensure that operations will be performed correctly and efficiently.

The security of a system also needs to be documented. This includes many types of documentation, such as security plans, contingency plans, risk analyses, and security policies and procedures. Much of this information, particularly risk and threat analyses, has to be protected against unauthorized disclosure. Security documentation also needs to be both current and accessible. Accessibility should take special factors into account (such as the need to find the contingency plan during a disaster).

Security documentation should be designed to fulfill the needs of the different types of people who use it. For this reason, many organizations separate documentation into *policy* and *procedures*. A *security procedures manual* should be written to inform various system users how to do their jobs securely. A security procedures manual for systems operations and support staff may address a wide variety of technical and operational concerns in considerable detail.

14.7 Maintenance

System maintenance requires either physical or logical access to the system. Support and operations staff, hardware or software vendors, or third-party service providers may maintain a system. Maintenance may be performed on site, or it may be necessary to move equipment to a repair site. Maintenance may also be performed remotely via communications connections. If someone who does not normally have access to the system performs maintenance, then a security vulnerability is introduced.

In some circumstances, it may be necessary to take additional precautions, such as conducting background investigations of service personnel. Supervision of maintenance personnel may prevent some problems, such as "snooping around" the physical area. However, once someone has access to the system, it is very difficult for supervision to prevent damage done through the maintenance process.

Many computer systems provide *maintenance accounts*. These special log-in accounts are normally preconfigured at the factory with pre-set, widely known passwords. *It is critical to change these passwords or*

> One of the most common methods hackers use to break into systems is through maintenance accounts that still have factory-set or easily guessed passwords.

otherwise disable the accounts until they are needed. Procedures should be developed to ensure that only authorized maintenance personnel can use these accounts. If the account is to be used remotely, authentication of the maintenance provider can be performed using call-back confirmation. This helps ensure that remote diagnostic activities actually originate from an established phone number at the vendor's site. Other techniques can also help, including encryption and decryption of diagnostic communications; strong identification and authentication techniques, such as tokens; and remote disconnect verification.

Larger systems may have *diagnostic ports*. In addition, manufacturers of larger systems and third-party providers may offer more diagnostic and support services. It is critical to ensure that these ports are only used by authorized personnel and cannot be accessed by hackers.

14.8 Interdependencies

There are support and operations components in most of the controls discussed in this handbook.

Personnel. Most support and operations staff have special access to the system. Some organizations conduct background checks on individuals filling these positions to screen out possibly untrustworthy individuals.

Incident Handling. Support and operations may include an organization's incident handling staff. Even if they are separate organizations, they need to work together to recognize and respond to incidents.

Contingency Planning. Support and operations normally provides technical input to contingency planning and carries out the activities of making backups, updating documentation, and practicing responding to contingencies.

Security Awareness, Training, and Education. Support and operations staff should be trained in security procedures and should be aware of the importance of security. In addition, they provide technical expertise needed to teach users how to secure their systems.

Physical and Environmental. Support and operations staff often control the immediate physical area around the computer system.

Technical Controls. The technical controls are installed, maintained, and used by support and operations staff. They create the user accounts, add users to access control lists, review audit logs for unusual activity, control bulk encryption over telecommunications links, and perform the countless operational tasks needed to use technical controls effectively. In addition, support and operations staff provide needed input to the selection of controls based on their knowledge of system capabilities and operational constraints.

Assurance. Support and operations staff ensure that changes to a system do not introduce security vulnerabilities by using assurance methods to evaluate or test the changes and their effect on the system. Operational assurance is normally performed by support and operations staff.

14.9 Cost Considerations

The cost of ensuring adequate security in day-to-day support and operations is largely dependent upon the size and characteristics of the operating environment and the nature of the processing being performed. If sufficient support personnel are already available, it is important that they be trained in the security aspects of their assigned jobs; it is usually not necessary to hire additional support and operations security specialists. Training, both initial and ongoing, is a cost of successfully incorporating security measures into support and operations activities.

Another cost is that associated with creating and updating documentation to ensure that security concerns are appropriately reflected in support and operations policies, procedures, and duties.

References

Bicknell, Paul. "Data Security for Personal Computers." *Proceedings of the 15th National Computer Security Conference*. Vol. I. National Institute of Standards and Technology and National Computer Security Center. Baltimore, MD. October 1992.

Caelli, William, Dennis Longley, and Michael Shain. *Information Security Handbook*. New York, NY: Stockton Press, 1991.

Carnahan, Lisa J. "A Local Area Network Security Architecture." *Proceedings of the 15th National Computer Security Conference*. Vol. I. National Institute of Standards and Technology and National Computer Security Center. Baltimore, MD. 1992.

Carroll, J.M. *Managing Risk: A Computer-Aided Strategy*. Boston, MA: Butterworths, 1984.

Chapman, D. Brent. "Network (In)Security Through IP Packet Filtering." *Proceedings of the 3rd USENIX UNIX Security Symposium*, 1992.

Curry, David A. *UNIX System Security: A Guide for Users and System Administrators*. Reading, MA: Addison-Wesley Publishing Co., Inc., 1992.

Garfinkel, Simson, and Gene Spafford. *Practical UNIX Security*. Sebastopol, CA: O'Reilly & Associates, 1991.

Holbrook, Paul, and Joyce Reynolds, eds. *Site Security Handbook*. Available by anonymous ftp

from nic.ddn.mil (in rfc directory).

Internet Security for System & Network Administrators. Computer Emergency Response Team Security Seminars, CERT Coordination Center, 1993.

Murray, W.H. "Security Considerations for Personal Computers." *Tutorial: Computer and Network Security*. Oakland, CA: IEEE Computer Society Press, 1986.

Parker, Donna B. *Managers Guide to Computer Security*. Reston, VA: Reston Publishing, Inc., 1981.

Pfleeger, Charles P. *Security in Computing*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989.

Chapter 15

PHYSICAL AND ENVIRONMENTAL SECURITY

The term *physical and environmental security*, as used in this chapter, refers to measures taken to protect systems, buildings, and related supporting infrastructure against threats associated with their physical environment.[103] Physical and environmental security controls include the following three broad areas:

> Physical and environmental security controls are implemented to protect the facility housing system resources, the system resources themselves, and the facilities used to support their operation.

1. The physical facility is usually the building, other structure, or vehicle housing the system and network components. Systems can be characterized, based upon their operating location, as static, mobile, or portable. Static systems are installed in structures at fixed locations. Mobile systems are installed in vehicles that perform the function of a structure, but not at a fixed location. Portable systems are not installed in fixed operating locations. They may be operated in wide variety of locations, including buildings or vehicles, or in the open. The physical characteristics of these structures and vehicles determine the level of such physical threats as fire, roof leaks, or unauthorized access.

2. The facility's general geographic operating location determines the characteristics of *natural threats*, which include earthquakes and flooding; *man-made threats* such as burglary, civil disorders, or interception of transmissions and emanations; and *damaging nearby activities*, including toxic chemical spills, explosions, fires, and electromagnetic interference from emitters, such as radars.

3. Supporting facilities are those services (both technical and human) that underpin the operation of the system. The system's operation usually depends on supporting facilities such as electric power, heating and air conditioning, and telecommunications. The failure or substandard performance of these facilities may interrupt operation of the system and may cause physical damage to system hardware or stored data.

This chapter first discusses the benefits of physical security measures, and then presents an overview of common physical and environmental security controls. Physical and environmental security measures result in many benefits, such as protecting employees. This chapter focuses on the protection of computer systems from the following:

---

[103] This chapter draws upon work by Robert V. Jacobson, International Security Technology, Inc., funded by the Tennessee Valley Authority.

Interruptions in Providing Computer Services. An external threat may interrupt the scheduled operation of a system. The magnitude of the losses depends on the duration and timing of the service interruption and the characteristics of the operations end users perform.

Physical Damage. If a system's hardware is damaged or destroyed, it usually has to be repaired or replaced. Data may be destroyed as an act of sabotage by a physical attack on data storage media (e.g., rendering the data unreadable or only partly readable). If data stored by a system for operational use is destroyed or corrupted, the data needs to be restored from back-up copies or from the original sources before the system can be used. The magnitude of loss from physical damage depends on the cost to repair or replace the damaged hardware *and* data, as well as costs arising from service interruptions.

Unauthorized Disclosure of Information. The physical characteristics of the facility housing a system may permit an intruder to gain access both to media external to system hardware (such as diskettes, tapes and printouts) and to media within system components (such as fixed disks), transmission lines or display screens. All may result in loss of disclosure-sensitive information.

Loss of Control over System Integrity. If an intruder gains access to the central processing unit, it is usually possible to reboot the system and *bypass* logical access controls. This can lead to information disclosure, fraud, replacement of system and application software, introduction of a Trojan horse, and more. Moreover, if such access is gained, it may be very difficult to determine what has been modified, lost, or corrupted.

Physical Theft. System hardware may be stolen. The magnitude of the loss is determined by the costs to replace the stolen hardware and restore data stored on stolen media. Theft may also result in service interruptions.

This chapter discusses seven major areas of physical and environmental security controls:

- physical access controls,
- fire safety,
- supporting utilities,
- structural collapse,
- plumbing leaks,
- interception of data, and
- mobile and portable systems.

15.1 Physical Access Controls

Physical access controls restrict the entry and exit of personnel (and often equipment and media) from an area, such as an office building, suite, data center, or room containing a LAN server.

Life Safety

It is important to understand that the objectives of physical access controls may be in conflict with those of *life safety*. Simply stated, life safety focuses on providing easy exit from a facility, particularly in an emergency, while physical security strives to control entry. In general, life safety must be given first consideration, but it is usually possible to achieve an effective balance between the two goals.

For example, it is often possible to equip emergency exit doors with a time delay. When one pushes on the panic bar, a loud alarm sounds, and the door is released after a brief delay. The expectation is that people will be deterred from using such exits improperly, but will not be significantly endangered during an emergency evacuation.

The controls over physical access to the elements of a system can include controlled areas, barriers that isolate each area, entry points in the barriers, and screening measures at each of the entry points. In addition, staff members who work in a restricted area serve an important role in providing physical security, as they can be trained to challenge people they do not recognize.

Physical access controls should address not only the area containing system hardware, but also locations of wiring used to connect elements of the system, the electric power service, the air conditioning and heating plant, telephone and data lines, backup media and source documents, and any other elements required system's operation. This means that all the areas in the building(s) that contain system elements must be identified.

There are many types of physical access controls, including badges, memory cards, guards, keys, true-floor-to-true-ceiling wall construction, fences, and locks.

It is also important to review the effectiveness of physical access controls in each area, both during normal business hours, and at other times – particularly when an area may be unoccupied. Effectiveness depends on both the characteristics of the control devices used (e.g., keycard-controlled doors) and the implementation and operation. Statements to the effect that "only authorized persons may enter this area" are not particularly effective. Organizations should determine whether intruders can easily defeat the controls, the extent to which strangers are challenged, and the effectiveness of other control procedures. Factors like these modify the effectiveness of physical controls.

The feasibility of surreptitious entry also needs to be considered. For example, it may be possible to go over the top of a partition that stops at the underside of a suspended ceiling or to cut a hole

in a plasterboard partition in a location hidden by furniture. If a door is controlled by a combination lock, it may be possible to observe an authorized person entering the lock combination. If keycards are not carefully controlled, an intruder may be able to steal a card left on a desk or use a card passed back by an accomplice.

Corrective actions can address any of the factors listed above. Adding an additional barrier reduces the risk to the areas behind the barrier. Enhancing the screening at an entry point can reduce the number of penetrations. For example, a guard may provide a higher level of screening than a keycard-controlled door, or an anti-passback feature can be added. Reorganizing traffic patterns, work flow, and work areas may reduce the number of people who need access to a restricted area. Physical modifications to barriers can reduce the vulnerability to surreptitious entry. Intrusion detectors, such as closed-circuit television cameras, motion detectors, and other devices, can detect intruders in unoccupied spaces.

15.2 Fire Safety Factors

Building fires are a particularly important security threat because of the potential for complete destruction of both hardware and data, the risk to human life, and the pervasiveness of the damage. Smoke, corrosive gases, and high humidity from a localized fire can damage systems throughout an entire building. Consequently, it is important to evaluate the fire safety of buildings that house systems. Following are important factors in determining the risks from fire.

Ignition Sources. Fires begin because something supplies enough heat to cause other materials to burn. Typical ignition sources are failures of electric devices and wiring, carelessly discarded cigarettes, improper storage of materials subject to spontaneous combustion, improper operation of heating devices, and, of course, arson.

Types of Building Construction

There are four basic kinds of building construction: (a) light frame, (b) heavy timber, (c) incombustible, and (d) fire resistant. Note that the term *fireproof* is not used because no structure can resist a fire indefinitely. Most houses are light frame, and cannot survive more than about thirty minutes in a fire. Heavy timber means that the basic structural elements have a minimum thickness of four inches. When such structures burn, the char that forms tends to insulate the interior of the timber and the structure may survive for an hour or more depending on the details. Incombustible means that the structure members will not burn. This almost always means that the members are steel. Note, however, that steel loses it strength at high temperatures, at which point the structure collapses. Fire resistant means that the structural members are incombustible and are insulated. Typically, the insulation is either concrete that encases steel members, or is a mineral wool that is sprayed onto the members. Of course, the heavier the insulation, the longer the structure will resist a fire.

Note that a building constructed of reinforced concrete can still be destroyed in a fire if there is sufficient fuel present and fire fighting is ineffective. The prolonged heat of a fire can cause differential expansion of the concrete which causes *spalling*. Portions of the concrete split off, exposing the reinforcing, and the interior of the concrete is subject to additional spalling. Furthermore, as heated floor slabs expand outward, they deform supporting columns. Thus, a reinforced concrete parking garage with open exterior walls and a relatively low fire load has a low fire risk, but a similar archival record storage facility with closed exterior walls and a high fire load has a higher risk even though the basic building material is incombustible.

Fuel Sources. If a fire is to grow, it must have a supply of fuel, material that will burn to support its growth, and an adequate supply of oxygen. Once a fire becomes established, it depends on the combustible materials in the building (referred to as the fire load) to support its further growth. The more fuel per square meter, the more intense the fire will be.

Building Operation. If a building is well maintained and operated so as to minimize the accumulation of fuel (such as maintaining the integrity of fire barriers), the fire risk will be minimized.

Building Occupancy. Some occupancies are inherently more dangerous than others because of an above-average number of potential ignition sources. For example, a chemical warehouse may contain an above-average fuel load.

Fire Detection. The more quickly a fire is detected, all other things being equal, the more easily it can be extinguished, minimizing damage. It is also important to accurately pinpoint the location of the fire.

Fire Extinguishment. A fire will burn until it consumes all of the fuel in the building or until it is extinguished. Fire extinguishment may be automatic, as with an automatic sprinkler system or a HALON discharge system, or it may be performed by people using portable extinguishers, cooling the fire site with a stream of water, by limiting the supply of oxygen with a blanket of foam or powder, or by breaking the combustion chemical reaction chain.

When properly installed, maintained, and provided with an adequate supply of water, automatic sprinkler systems are highly effective in protecting buildings and their contents.[104] Nonetheless, one often hears uninformed persons speak of the *water damage* done by sprinkler systems as a disadvantage. *Fires that trigger sprinkler systems* cause the water damage.[105] In short, sprinkler systems reduce fire damage, protect

Halons have been identified as harmful to the Earth's protective ozone layer. So, under an international agreement (known as the Montreal Protocol), production of halons ended January 1, 1994. In September 1992, the General Services Administration issued a moratorium on halon use by federal agencies.

[104] As discussed in this section, many variables affect fire safety and should be taken into account in selecting a fire extinguishment system. While automatic sprinklers can be very effective, selection of a fire extinguishment system for a particular building should take into account the particular fire risk factors. Other factors may include rate changes from either a fire insurance carrier or a business interruption insurance carrier. Professional advice is required.

[105] Occurrences of accidental discharge are extremely rare, and, in a fire, only the sprinkler heads in the immediate area of the fire open and discharge water.

the lives of building occupants, and limit the fire damage to the building itself. All these factors contribute to more rapid recovery of systems following a fire.

Each of these factors is important when estimating the occurrence rate of fires and the amount of damage that will result. The objective of a fire-safety program is to optimize these factors to minimize the risk of fire.

15.3 Failure of Supporting Utilities

Systems and the people who operate them need to have a reasonably well-controlled operating environment. Consequently, failures of heating and air-conditioning systems will usually cause a service interruption and may damage hardware. These utilities are composed of many elements, each of which must function properly.

For example, the typical air-conditioning system consists of (1) air handlers that cool and humidify room air, (2) circulating pumps that send chilled water to the air handlers, (3) chillers that extract heat from the water, and (4) cooling towers that discharge the heat to the outside air. Each of these elements has a mean-time-between-failures (MTBF) and a mean-time-to-repair (MTTR). Using the MTBF and MTTR values for each of the elements of a system, one can estimate the occurrence rate of system failures and the range of resulting service interruptions.

This same line of reasoning applies to electric power distribution, heating plants, water, sewage, and other utilities required for system operation or staff comfort. By identifying the failure modes of each utility and estimating the MTBF and MTTR, necessary failure threat parameters can be developed to calculate the resulting risk. The risk of utility failure can be reduced by substituting units with lower MTBF values. MTTR can be reduced by stocking spare parts on site and training maintenance personnel. And the outages resulting from a given MTBF can be reduced by installing redundant units under the assumption that failures are distributed randomly in time. Each of these strategies can be evaluated by comparing the reduction in risk with the cost to achieve it.

15.4 Structural Collapse

A building may be subjected to a load greater than it can support. Most commonly this is a result of an earthquake, a snow load on the roof beyond design criteria, an explosion that displaces or cuts structural members, or a fire that weakens structural members. Even if the structure is not completely demolished, the authorities may decide to ban its further use, sometimes even banning entry to remove materials. This threat applies primarily to high-rise buildings and those with large interior spaces without supporting columns.

15.5 Plumbing Leaks

While plumbing leaks do not occur every day, they can be seriously disruptive. The building's plumbing drawings can help locate plumbing lines that might endanger system hardware. These lines include hot and cold water, chilled water supply and return lines, steam lines, automatic sprinkler lines, fire hose standpipes, and drains. If a building includes a laboratory or manufacturing spaces, there may be other lines that conduct water, corrosive or toxic chemicals, or gases.

As a rule, analysis often shows that the cost to relocate threatening lines is difficult to justify. However, the location of shutoff valves and procedures that should be followed in the event of a failure must be specified. Operating and security personnel should have this information immediately available for use in an emergency. In some cases, it may be possible to relocate system hardware, particularly distributed LAN hardware.

15.6 Interception of Data

Depending on the type of data a system processes, there may be a significant risk if the data is intercepted. There are three routes of data interception: direct observation, interception of data transmission, and electromagnetic interception.

Direct Observation. System terminal and workstation display screens may be observed by unauthorized persons. In most cases, it is relatively easy to relocate the display to eliminate the exposure.

Interception of Data Transmissions. If an interceptor can gain access to data transmission lines, it may be feasible to tap into the lines and read the data being transmitted. Network monitoring tools can be used to capture data packets. Of course, the interceptor cannot control what is transmitted, and so may not be able to immediately observe data of interest. However, over a period of time there may be a serious level of disclosure. Local area networks typically broadcast messages.[106] Consequently, all traffic, including passwords, could be retrieved. Interceptors could also transmit spurious data on tapped lines, either for purposes of disruption or for fraud.

Electromagnetic Interception. Systems routinely radiate electromagnetic energy that can be detected with special-purpose radio receivers. Successful interception will depend on the signal strength at the receiver location; the greater the separation between the system and the receiver, the lower the success rate. TEMPEST shielding, of either equipment or rooms, can be used to minimize the spread of electromagnetic signals. The signal-to-noise ratio at the receiver,

[106] An insider may be able to easily collect data by configuring their ethernet network interface to receive all network traffic, rather than just network traffic intended for this node. This is called the *promiscuous* mode.

determined in part by the number of competing emitters will also affect the success rate. The more workstations of the same type in the same location performing "random" activity, the more difficult it is to intercept a given workstation's radiation. On the other hand, the trend toward wireless (i.e., deliberate radiation) LAN connections may increase the likelihood of successful interception.

15.7 Mobile and Portable Systems

The analysis and management of risk usually has to be modified if a system is installed in a vehicle or is portable, such as a laptop computer. The system in a vehicle will share the risks of the vehicle, including accidents and theft, as well as regional and local risks.

Portable and mobile systems share an increased risk of theft and physical damage. In addition, portable systems can be "misplaced" or left unattended by careless users. Secure storage of laptop computers is often required when they are not in use.

> Encryption of data files on stored media may also be a cost-effective precaution against disclosure of confidential information if a laptop computer is lost or stolen.

If a mobile or portable system uses particularly valuable or important data, it may be appropriate to either store its data on a medium that can be removed from the system when it is unattended or to encrypt the data. In any case, the issue of how custody of mobile and portable computers are to be controlled should be addressed. Depending on the sensitivity of the system and its application, it may be appropriate to require briefings of users and signed briefing acknowledgments. (See Chapter 10 for an example.)

15.8 Approach to Implementation

Like other security measures, physical and environmental security controls are selected because they are cost-beneficial. This does not mean that a user must conduct a detailed cost-benefit analysis for the selection of every control. There are four general ways to justify the selection of controls:

1. *They are required by law or regulation*. Fire exit doors with panic bars and exit lights are examples of security measures required by law or regulation. Presumably, the regulatory authority has considered the costs and benefits and has determined that it is in the public interest to require the security measure. A lawfully conducted organization has no option but to implement all required security measures.

2. *The cost is insignificant, but the benefit is material*. A good example of this is a facility with a key-locked low-traffic door to a restricted access. The cost of keeping the door

locked is minimal, but there is a significant benefit. Once a significant benefit/minimal cost security measure has been identified, no further analysis is required to justify its implementation.

3. The security measure addresses a potentially "fatal" security exposure but has a reasonable cost. Backing up system software and data is an example of this justification . For most systems, the cost of making regular backup copies is modest (compared to the costs of operating the system), the organization would not be able to function if the stored data were lost, and the cost impact of the failure would be material. In such cases, it would not be necessary to develop any further cost justification for the backup of software and data. However, this justification depends on what constitutes a *modest* cost, and it does not identify the optimum backup schedule. Broadly speaking, a cost that does not require budgeting of additional funds would qualify.

4. The security measure is estimated to be cost-beneficial. If the cost of a potential security measure is significant, and it cannot be justified by any of the first three reasons listed above, then its cost (both implementation and ongoing operation) and its benefit (reduction in future expected losses) need to be analyzed to determine if it is cost-beneficial. In this context, *cost-beneficial* means that the reduction in expected loss is significantly greater than the cost of implementing the security measure.

Arriving at the fourth justification requires a detailed analysis. Simple rules of thumb do not apply. Consider, for example, the threat of electric power failure and the security measures that can protect against such an event. The threat parameters, rate of occurrence, and range of outage durations depend on the location of the system, the details of its connection to the local electric power utility, the details of the internal power distribution system, and the character of other activities in the building that use electric power. The system's potential losses from service interruption depends on the details of the functions it performs. Two systems that are otherwise identical can support functions that have quite different degrees of urgency. Thus, two systems may have the same electric power failure threat and vulnerability parameters, yet entirely different loss potential parameters.

Furthermore, a number of different security measures are available to address electric power failures. These measures differ in both cost and performance. For example, the cost of an uninterruptible power supply (UPS) depends on the size of the electric load it can support, the number of minutes it can support the load, and the speed with which it assumes the load when the primary power source fails. An on-site power generator could also be installed either in place of a UPS (accepting the fact that a power failure will cause a brief service interruption) or in order to provide long-term backup to a UPS system. Design decisions include the magnitude of the load the generator will support, the size of the on-site fuel supply, and the details of the facilities to switch the load from the primary source or the UPS to the on-site generator.

This example shows systems with a wide range of risks and a wide range of available security measures (including, of course, no action), each with its own cost factors and performance parameters.

15.9 Interdependencies

Physical and environmental security measures rely on and support the proper functioning of many of the other areas discussed in this handbook. Among the most important are the following:

Logical Access Controls. Physical security controls augment technical means for controlling access to information and processing. Even if the most advanced and best-implemented logical access controls are in place, if physical security measures are inadequate, logical access controls may be circumvented by directly accessing the hardware and storage media. For example, a computer system may be rebooted using different software.

Contingency Planning. A large portion of the contingency planning process involves the failure of physical and environmental controls. Having sound controls, therefore, can help minimize losses from such contingencies.

Identification and Authentication (I&A). Many physical access control systems require that people be identified and authenticated. Automated physical security access controls can use the same types of I&A as other computer systems. In addition, it is possible to use the same tokens (e.g., badges) as those used for other computer-based I&A.

Other. Physical and environmental controls are also closely linked to the activities of the local guard force, fire house, life safety office, and medical office. These organizations should be consulted for their expertise in planning controls for the systems environment.

15.10 Cost Considerations

Costs associated with physical security measures range greatly. Useful generalizations about costs, therefore, are difficult make. Some measures, such as keeping a door locked, may be a trivial expense. Other features, such as fire-detection and -suppression systems, can be far more costly. Cost considerations should include operation. For example, adding controlled-entry doors requires persons using the door to stop and unlock it. Locks also require physical key management and accounting (and rekeying when keys are lost or stolen). Often these effects will be inconsequential, but they should be fully considered. As with other security measures, the objective is to select those that are cost-beneficial.

References

Alexander, M., ed. "Secure Your Computers and Lock Your Doors." *Infosecurity News*. 4(6), 1993. pp. 80-85.

Archer, R. "Testing: Following Strict Criteria." *Security Dealer*. 15(5), 1993. pp. 32-35.

Breese, H., ed. *The Handbook of Property Conservation*. Norwood, MA: Factory Mutual Engineering Corp.

Chanaud, R. "Keeping Conversations Confidential." *Security Management*. 37(3), 1993. pp. 43-48.

Miehl, F. "The Ins and Outs of Door Locks." *Security Management*. 37(2), 1993. pp. 48-53.

National Bureau of Standards. *Guidelines for ADP Physical Security and Risk Management*. Federal Information Processing Standard Publication 31. June 1974.

Peterson, P. "Infosecurity and Shrinking Media." *ISSA Access*. 5(2), 1992. pp. 19-22.

Roenne, G. "Devising a Strategy Keyed to Locks." *Security Management*. 38(4), 1994. pp. 55-56.

Zimmerman, J. "Using Smart Cards - A Smart Move." *Security Management*. 36(1), 1992. pp. 32-36.

IV. TECHNICAL CONTROLS

Chapter 16

IDENTIFICATION AND AUTHENTICATION

For most systems, identification and authentication (I&A) is the first line of defense. I&A is a technical measure that prevents unauthorized people (or unauthorized processes) from entering a computer system.

I&A is a critical building block of computer security since it is the basis for most types of access control and for establishing user accountability.[107] Access control often requires that the system be able to identify and differentiate among users. For example, access control is often based on *least privilege*, which refers to the granting to users of only those accesses required to perform their duties. User accountability requires the linking of activities on a computer system to specific individuals and, therefore, requires the system to identify users.

Identification is the means by which a user *provides* a claimed identity to the system. *Authentication*[108] is the means of establishing the *validity* of this claim.

This chapter discusses the basic means of identification and authentication, the current technology used to provide I&A, and some important implementation issues.

> A typical user identification could be JSMITH (for Jane Smith). This information can be known by system administrators and other system users. A typical user authentication could be Jane Smith's password, which is kept secret. This way system administrators can set up Jane's access and see her activity on the audit trail, and system users can send her e-mail, but no one can pretend to be Jane.

Computer systems recognize people based on the authentication data the systems *receive.* Authentication presents several challenges: collecting authentication data, transmitting the data securely, and knowing whether the person who was originally authenticated is *still* the person using the computer system. For example, a user may walk away from a terminal while still logged on, and another person may start using it.

There are three means of authenticating a user's identity *which can be used alone or in combination*:

- something the individual *knows* (a secret– e.g., a password, Personal Identification Number (PIN), or cryptographic key);

[107] Not all types of access control require identification and authentication.

[108] Computers also use authentication to verify that a message or file has not been altered and to verify that a message originated with a certain person. This chapter only addresses user authentication. The other forms of authentication are addressed in the Chapter 19.

- something the individual *possesses* (a token – e.g., an ATM card or a smart card); and
- something the individual *is* (a biometric – e.g., such characteristics as a voice pattern, handwriting dynamics, or a fingerprint).

> For most applications, trade-offs will have to be made among security, ease of use, and ease of administration, especially in modern networked environments.

While it may appear that any of these means could provide strong authentication, there are problems associated with each. If people wanted to pretend to be someone else on a computer system, they can guess or learn that individual's password; they can also steal or fabricate tokens. Each method also has drawbacks for legitimate users and system administrators: users forget passwords and may lose tokens, and administrative overhead for keeping track of I&A data and tokens can be substantial. Biometric systems have significant technical, user acceptance, and cost problems as well.

This section explains current I&A technologies and their benefits and drawbacks as they relate to the three means of authentication. Although some of the technologies make use of cryptography because it can significantly strengthen authentication, the explanations of cryptography appear in Chapter 19, rather than in this chapter.

16.1 I&A Based on Something the User Knows

The most common form of I&A is a user ID coupled with a password. This technique is based solely on something the user knows. There are other techniques besides *conventional* passwords that are based on knowledge, such as knowledge of a cryptographic key.

16.1.1 Passwords

In general, password systems work by requiring the user to enter a user ID and password (or passphrase or personal identification number). The system compares the password to a previously stored password for that user ID. If there is a match, the user is authenticated and granted access.

Benefits of Passwords. Passwords have been successfully providing security for computer systems for a long time. They are integrated into many operating systems, and users and system administrators are familiar with them. When properly managed in a controlled environment, they can provide effective security.

Problems With Passwords. The security of a password system is dependent upon keeping passwords secret. Unfortunately, there are many ways that the secret may be divulged. All of the

problems discussed below can be significantly mitigated by improving password security, as discussed in the sidebar. However, there is no fix for the problem of electronic monitoring, except to use more advanced authentication (e.g., based on cryptographic techniques or tokens).

1. Guessing or finding passwords. If users select their own passwords, they tend to make them easy to remember. That often makes them easy to guess. The names of people's children, pets, or favorite sports teams are common examples. On the other hand, assigned passwords may be difficult to remember, so users are more likely to write them down. Many computer systems are shipped with administrative accounts that have preset passwords. Because these passwords are standard, they are easily "guessed." Although security practitioners have been warning about this problem for years, many system administrators still do not change default passwords. Another method of learning passwords is to observe someone entering a password or PIN. The observation can be done by someone in the same room or by someone some distance away using binoculars. This is often referred to as *shoulder surfing*.

2. Giving passwords away. Users may share their passwords. They may give their password to a co-worker in order to share files. In addition, people can be tricked into divulging their passwords. This process is referred to as *social engineering*.

Improving Password Security

Password generators. If users are not allowed to generate their own passwords, they cannot pick easy-to-guess passwords. Some generators create only pronounceable nonwords to help users remember them. However, users tend to write down hard-to-remember passwords.

Limits on log-in attempts. Many operating systems can be configured to lock a user ID after a set number of failed log-in attempts. This helps to prevent guessing of passwords.

Password attributes. Users can be instructed, or the system can force them, to select passwords (1) with a certain minimum length, (2) with special characters, (3) that are unrelated to their user ID, or (4) to pick passwords which are not in an on-line dictionary. This makes passwords more difficult to guess (but more likely to be written down).

Changing passwords. Periodic changing of passwords can reduce the damage done by stolen passwords and can make brute-force attempts to break into systems more difficult. Too frequent changes, however, can be irritating to users.

Technical protection of the password file. Access control and one-way encryption can be used to protect the password file itself.

Note: Many of these techniques are discussed in FIPS 112, *Password Usage* and FIPS 181, *Automated Password Generator*.

3. Electronic monitoring. When passwords are transmitted to a computer system, they can be electronically monitored. This can happen on the network used to transmit the password or on the computer system itself. Simple encryption of a password that will be used again does not solve this problem because encrypting the same password will create the same ciphertext; the ciphertext becomes the password.

4. Accessing the password file. If the password file is not protected by strong access controls, the file can be downloaded. Password files are often protected with one-way encryption[109] so that plain-text passwords are not available to system administrators or hackers (if they successfully bypass access controls). Even if the file is encrypted, brute force can be used to learn passwords if the file is downloaded (e.g., by encrypting English words and comparing them to the file).

Passwords Used as Access Control. Some mainframe operating systems and many PC applications use passwords as a means of restricting access to specific resources within a system. Instead of using mechanisms such as access control lists (see Chapter 17), access is granted by entering a password. The result is a proliferation of passwords that can reduce the overall security of a system. While the use of passwords as a means of access control is common, it is an approach that is often less than optimal and not cost-effective.

16.1.2 Cryptographic Keys

Although the authentication derived from the knowledge of a cryptographic key may be based entirely on something the user knows, it is necessary for the user to also possess (or have access to) something that can perform the cryptographic computations, such as a PC or a smart card. For this reason, the protocols used are discussed in the Smart Tokens section of this chapter. However, it is possible to implement these types of protocols without using a smart token. Additional discussion is also provided under the Single Log-in section.

16.2 I&A Based on Something the User Possesses

Although some techniques are based solely on something the user possesses, most of the techniques described in this section are combined with something the user knows. This combination can provide significantly stronger security than either something the user knows or possesses alone.[110]

Objects that a user possesses for the purpose of I&A are called *tokens*. This section divides tokens into two categories: *memory tokens* and *smart tokens*.

---

[109] One-way encryption algorithms only provide for the encryption of data. The resulting ciphertext cannot be decrypted. When passwords are entered into the system, they are one-way encrypted, and the result is compared with the stored ciphertext. (See the Chapter 19.)

[110] For the purpose of understanding how possession-based I&A works, it is not necessary to distinguish whether possession of a token in various systems is identification or authentication.

16.2.1 Memory Tokens

Memory tokens store, but do not process, information. Special reader/writer devices control the writing and reading of data to and from the tokens. The most common type of memory token is a magnetic striped card, in which a thin stripe of magnetic material is affixed to the surface of a card (e.g., as on the back of credit cards). A common application of memory tokens for authentication to computer systems is the automatic teller machine (ATM) card. This uses a combination of something the user possesses (the card) with something the user knows (the PIN).

Some computer systems authentication technologies are based solely on possession of a token, but they are less common. Token-only systems are more likely to be used in other applications, such as for physical access. (See Chapter 15.)

Benefits of Memory Token Systems. Memory tokens when used with PINs provide significantly more security than passwords. In addition, memory cards are inexpensive to produce. For a hacker or other would-be masquerader to pretend to be someone else, the hacker must have both a valid token *and* the corresponding PIN. This is much more difficult than obtaining a valid password and user ID combination (especially since most user IDs are common knowledge).

Another benefit of tokens is that they can be used in support of log generation without the need for the employee to key in a user ID for each transaction or other logged event since the token can be scanned repeatedly. If the token is required for physical entry and exit, then people will be forced to remove the token when they leave the computer. This can help maintain authentication.

Problems With Memory Token Systems. Although sophisticated technical attacks are possible against memory token systems, most of the problems associated with them relate to their cost, administration, token loss, user dissatisfaction, and the compromise of PINs. Most of the techniques for increasing the security of memory token systems relate to the protection of PINs. Many of the techniques discussed in the sidebar on Improving Password Security apply to PINs.

> *1. Requires special reader*. The need for a special reader increases the cost of using memory tokens. The readers used for memory tokens must include both the physical unit that reads the card and a processor that determines whether the card and/or the PIN entered with the card is valid. If the PIN or token is validated by a processor that is not physically located with the reader, then the authentication data is vulnerable to electronic monitoring (although cryptography can be used to solve this problem).

2. Token loss. A lost token may prevent the user from being able to log in until a replacement is provided. This can increase administrative overhead costs.

The lost token could be found by someone who wants to break into the system, or could be stolen or forged. If the token is also used with a PIN, any of the methods described above in password problems can be used to obtain the PIN. Common methods are finding the PIN taped to the card or observing the PIN being entered by the legitimate user. In addition, any information stored on the magnetic stripe that has not been encrypted can be read.

> Attacks on memory-card systems have sometimes been quite creative. One group stole an ATM machine that they installed at a local shopping mall. The machine collected valid account numbers and corresponding PINs, which the thieves used to forge cards. The forged cards were then used to withdraw money from legitimate ATMs.

3. User Dissatisfaction. In general, users want computers to be easy to use. Many users find it inconvenient to carry and present a token. However, their dissatisfaction may be reduced if they see the need for increased security.

16.2.2 Smart Tokens

A smart token expands the functionality of a memory token by incorporating one or more integrated circuits into the token itself. When used for authentication, a smart token is another example of authentication based on something a user possesses (i.e., the token itself). A smart token typically requires a user also to provide something the user knows (i.e., a PIN or password) in order to "unlock" the smart token for use.

There are many different types of smart tokens. In general, smart tokens can be divided three different ways based on physical characteristics, interface, and protocols used. These three divisions are not mutually exclusive.

Physical Characteristics. Smart tokens can be divided into two groups: smart cards and other types of tokens. A smart card looks like a credit card, but incorporates an embedded microprocessor. Smart cards are defined by an International Standards Organization (ISO) standard. Smart tokens that are not smart cards can look like calculators, keys, or other small portable objects.

Interface. Smart tokens have either a manual or an electronic interface. Manual or human interface tokens have displays and/or keypads to allow humans to communicate with the card. Smart tokens with electronic interfaces must be read by special reader/writers. Smart cards, described above, have an electronic interface. Smart tokens that look like calculators usually have a manual interface.

Protocol. There are many possible protocols a smart token can use for authentication. In general, they can be divided into three categories: static password exchange, dynamic password generators, and challenge-response.

- *Static* tokens work similarly to memory tokens, except that the users authenticate themselves *to the token* and then the token authenticates the user to the computer.

- A token that uses a *dynamic password generator* protocol creates a unique value, for example, an eight-digit number, that changes periodically (e.g., every minute). If the token has a manual interface, the user simply reads the current value and then types it into the computer system for authentication. If the token has an electronic interface, the transfer is done automatically. If the correct value is provided, the log-in is permitted, and the user is granted access to the system.

- Tokens that use a *challenge-response* protocol work by having the computer generate a challenge, such as a random string of numbers. The smart token then generates a response based on the challenge. This is sent back to the computer, which authenticates the user based on the response. The challenge-response protocol is based on cryptography. Challenge-response tokens can use either electronic or manual interfaces.

There are other types of protocols, some more sophisticated and some less so. The three types described above are the most common.

Benefits of Smart Tokens

Smart tokens offer great flexibility and can be used to solve many authentication problems. The benefits of smart tokens vary, depending on the type used. In general, they provide greater security than memory cards. Smart tokens can solve the problem of electronic monitoring even if the authentication is done across an open network by using *one-time passwords*.

1. One-time passwords. Smart tokens that use either dynamic password generation or challenge-response protocols can create one-time passwords. Electronic monitoring is not a problem with one-time passwords because each time the user is authenticated to the computer, a different "password" is used. (A hacker could learn the one-time password through electronic monitoring, but would be of no value.)

2. Reduced risk of forgery. Generally, the memory on a smart token is not readable unless the PIN is entered. In addition, the tokens are more complex and, therefore, more difficult to forge.

3. Multi-application. Smart tokens with electronic interfaces, such as smart cards, provide a way for users to access many computers using many networks with only one log-in. This is

further discussed in the Single Log-in section of this chapter. In addition, a single smart card can be used for multiple functions, such as physical access or as a debit card.

Problems with Smart Tokens

Like memory tokens, most of the problems associated with smart tokens relate to their cost, the administration of the system, and user dissatisfaction. Smart tokens are generally less vulnerable to the compromise of PINs because authentication usually takes place on the card. (It is possible, of course, for someone to watch a PIN being entered and steal that card.) Smart tokens cost more than memory cards because they are more complex, particularly challenge-response calculators.

1. Need reader/writers or human intervention. Smart tokens can use either an electronic or a human interface. An electronic interface requires a reader, which creates additional expense. Human interfaces require more actions from the user. This is especially true for challenge-response tokens with a manual interface, which require the user to type the challenge into the smart token and the response into the computer. This can increase user dissatisfaction.

Electronic reader/writers can take many forms, such as a slot in a PC or a separate external device. Most human interfaces consist of a keypad and display.

2. Substantial Administration. Smart tokens, like passwords and memory tokens, require strong administration. For tokens that use cryptography, this includes key management. (See Chapter 19.)

16.3 I&A Based on Something the User Is

Biometric authentication technologies use the unique characteristics (or attributes) of an individual to authenticate that person's identity. These include physiological attributes (such as fingerprints, hand geometry, or retina patterns) or behavioral attributes (such as voice patterns and hand-written signatures). Biometric authentication technologies based upon these attributes have been developed for computer log-in applications.

Biometric authentication is technically complex and expensive, and user acceptance can be difficult. However, advances continue to be made to make the technology more reliable, less costly, and more user-friendly.

Biometric systems can provide an increased level of security for computer systems, but the technology is still less mature than that of memory tokens or smart tokens. Imperfections in biometric authentication devices arise from technical difficulties in measuring and profiling physical attributes as well as from the somewhat variable nature of physical attributes. These may change, depending on various conditions. For example, a person's speech pattern may change under stressful conditions or when suffering from a sore throat or cold.

Biometric authentication generally operates in the following manner:

Before any authentication attempts, a user is "enrolled" by creating a reference profile (or template) based on the desired physical attribute. The resulting template is associated with the identity of the user and stored for later use.

When attempting authentication, the user's biometric attribute is measured. The previously stored reference profile of the biometric attribute is compared with the measured profile of the attribute taken from the user. The result of the comparison is then used to either accept or reject the user.

Due to their relatively high cost, biometric systems are typically used with other authentication means in environments requiring high security.

16.4 Implementing I&A Systems

Some of the important implementation issues for I&A systems include administration, maintaining authentication, and single log-in.

16.4.1 Administration

Administration of authentication data is a critical element for all types of authentication systems. The administrative overhead associated with I&A can be significant. I&A systems need to create, distribute, and store authentication data0 -28uoatsctions indlT* -0.09e asreating a sctions indsuesgBistribute,n of teuoatometric systems andlT* -0.09e asreating a d store g a sfile os

of these issues are discussed in Chapter 10 under User Administration.

In addition, I&A administrative tasks should address lost or stolen passwords or tokens. It is often necessary to monitor systems to look for stolen or shared accounts.

> One method of looking for improperly used accounts is for the computer to inform users when they last logged on. This allows users to check if someone else used their account.

Authentication data needs to be stored securely, as discussed with regard to accessing password files. The value of authentication data lies in the data's confidentiality, integrity, and availability. If confidentiality is compromised, someone may be able to use the information to masquerade as a legitimate user. If system administrators can read the authentication file, they can masquerade as another user. Many systems use encryption to hide the authentication data from the system administrators.[111] If integrity is compromised, authentication data can be added or the system can be disrupted. If availability is compromised, the system cannot authenticate users, and the users may not be able to work.

16.4.2 Maintaining Authentication

So far, this chapter has discussed initial authentication only. It is also possible for someone to use a legitimate user's account after log-in.[112] Many computer systems handle this problem by logging a user out or locking their display or session after a certain period of inactivity. However, these methods can affect productivity and can make the computer less user-friendly.

16.4.3 Single Log-in

From an efficiency viewpoint, it is desirable for users to authenticate themselves only once and then to be able to access a wide variety of applications and data available on local and remote systems, even if those systems require users to authenticate themselves. This is known as *single log-in*.[113] If the access is within the same host computer, then the use of a modern access control system (such as an access control list) should allow for a single log-in. If the access is across multiple platforms, then the issue is more complicated, as discussed below. There are three main

---

[111] Masquerading *by* system administrators cannot be prevented entirely. However, controls can be set up so that improper actions by the system administrator can be detected in audit records.

[112] After a user signs on, the computer treats all commands originating from the user's physical device (such as a PC or terminal) as being from that user.

[113] Single log-in is somewhat of a misnomer. It is currently not feasible to have one sign-on for every computer system a user might wish to access. The types of single log-in described apply mainly to groups of systems (e.g., within an organization or a consortium).

techniques that can provide single log-in across multiple computers: host-to-host authentication, authentication servers, and user-to-host authentication.

Host-to-Host Authentication. Under a host-to-host authentication approach, users authenticate themselves once to a host computer. That computer then authenticates itself to other computers and vouches for the specific user. Host-to-host authentication can be done by passing an identification, a password, or by a challenge-response mechanism or other one-time password scheme. Under this approach, it is necessary for the computers to recognize each other and to trust each other.

Authentication Servers. When using authentication server, the users authenticate themselves to a special host computer (the authentication server). This computer then authenticates the user to other host computers the user wants to access. Under this approach, it is necessary for the computers to trust the authentication server. (The authentication server need not be a separate computer, although in some environments this may be a cost-effective way to increase the security of the server.) Authentication servers can be distributed geographically or logically, as needed, to reduce workload.

> Kerberos and SPX are examples of network authentication server protocols. They both use cryptography to authenticate users to computers on networks.

User-to-Host. A user-to-host authentication approach requires the user to log-in to each host computer. However, a smart token (such as a smart card) can contain all authentication data and perform that service for the user. To users, it looks as though they were only authenticated once.

16.5 Interdependencies

There are many interdependencies among I&A and other controls. Several of them have been discussed in the chapter.

Logical Access Controls. Access controls are needed to protect the authentication database. I&A is often the basis for access controls. Dial-back modems and firewalls, discussed in Chapter 17, can help prevent hackers from trying to log-in.

Audit. I&A is necessary if an audit log is going to be used for individual accountability.

Cryptography. Cryptography provides two basic services to I&A: it protects the confidentiality of authentication data, and it provides protocols for proving knowledge and/or possession of a token without having to transmit data that could be replayed to gain access to a computer system.

16.6 Cost Considerations

In general, passwords are the least expensive authentication technique and generally the least secure. They are already embedded in many systems. Memory tokens are less expensive than smart tokens, but have less functionality. Smart tokens with a human interface do not require readers, but are more inconvenient to use. Biometrics tend to be the most expensive.

For I&A systems, the cost of administration is often underestimated. Just because a system comes with a password system does not mean that using it is free. For example, there is significant overhead to administering the I&A system.

References

Alexander, M., ed. "Keeping the Bad Guys Off-Line." *Infosecurity News.* 4(6), 1993. pp. 54-65.

American Bankers Association. *American National Standard for Financial Institution Sign-On Authentication for Wholesale Financial Transactions*. ANSI X9.26-1990. Washington, DC, February 28, 1990.

CCITT Recommendation X.509. The Directory - Authentication Framework. November 1988 (Developed in collaboration, and technically aligned, with ISO 9594-8).

Department of Defense. Password Management Guideline. CSC-STD-002-85. April 12, 1985.

Feldmeier, David C., and Philip R. Kam. "UNIX Password Security - Ten Years Later." *Crypto '89 Abstracts*. Santa Barbara, CA: Crypto '89 Conference, August 20-24, 1989.

Haykin, Martha E., and Robert B. J. Warnar. *Smart Card Technology: New Methods for Computer Access Control.* Special Publication 500-157. Gaithersburg, MD: National Institute of Standards and Technology, September 1988.

Kay, R. "Whatever Happened to Biometrics?" *Infosecurity News*. 4(5), 1993. pp. 60-62.

National Bureau of Standards. *Password Usage.* Federal Information Processing Standard Publication 112. May 30, 1985.

National Institute of Standards and Technology. *Automated Password Generator*. Federal Information Processing Standard Publication 181. October, 1993.

National Institute of Standards and Technology. *Guideline for the Use of Advanced Authentication Technology Alternatives*. Federal Information Processing Standard Publication

190. October, 1994.

Salamone, S. "Internetwork Security: Unsafe at Any Node?" *Data Communications.* 22(12), 1993. pp. 61-68.

Sherman, R. "Biometric Futures." *Computers and Security*. 11(2), 1992. pp. 128-133.

Smid, Miles, James Dray, and Robert B. J. Warnar. "A Token-Based Access Control System for Computer Networks." *Proceedings of the 12th National Commuter Security Conference*. National Institute of Standards and Technology, October 1989.

Steiner, J.O., C. Neuman, and J. Schiller. "Kerberos: An Authentication Service for Open Network Systems." *Proceedings Winter USENIX*. Dallas, Texas, February 1988. pp. 191-202.

Troy, Eugene F. *Security for Dial-Up Lines*. Special Publication 500-137, Gaithersburg, MD: National Bureau of Standards, May 1986.

Chapter 17

LOGICAL ACCESS CONTROL

> Logical access controls provide a technical means of controlling what information users can utilize, the programs they can run, and the modifications they can make.

On many multiuser systems, requirements for using (and prohibitions against the use of) various computer resources[114] vary considerably. Typically, for example, some information must be accessible to all users,[115] some may be needed by several groups or departments, and some should be accessed by only a few individuals. While it is obvious that users must have access to the information they need to do their jobs, it may also be required to deny access to non-job-related information. It may also be important to control the *kind of access* that is afforded (e.g., the ability for the average user to execute, but not change, system programs). These types of access restrictions enforce policy and help ensure that unauthorized actions are not taken.

> The term *access* is often confused with *authorization* and *authentication*.
>
> *Access* is the *ability* to do something with a computer resource. This usually refers to a technical ability (e.g., read, create, modify, or delete a file, execute a program, or use an external connection).
>
> *Authorization* is the *permission* to use a computer resource. Permission is granted, directly or indirectly, by the application or system owner.
>
> *Authentication* is proving (to some reasonable degree) that users are who they claim to be.

Access is the ability to do something with a computer resource (e.g., use, change, or view). *Access control* is the means by which the ability is explicitly enabled or restricted in some way (usually through physical and system-based controls). Computer-based access controls are called *logical access controls*. Logical access controls can prescribe not only who or what (e.g., in the case of a process) is to have access to a specific system resource but also the type of access that is permitted. These controls may be built into the operating system, may be incorporated into applications programs or major utilities (e.g., database management systems or communications systems), or may be implemented through add-on security packages. Logical access controls may be implemented internally to the computer system being protected or may be implemented in external devices.

---

[114] The term *computer resources* includes information as well as system resources, such as programs, subroutines, and hardware (e.g., modems, communications lines).

[115] *Users* need not be actual human users. They could include, for example, a program or another computer requesting use of a system resource.

Logical access controls can help protect:

- operating systems and other system software from unauthorized modification or manipulation (and thereby help ensure the system's integrity and availability);
- the integrity and availability of information by restricting the number of users and processes with access; and
- confidential information from being disclosed to unauthorized individuals.

Controlling access is normally thought of as applying to human users (e.g., will technical access be provided for user JSMITH to the file "payroll.dat") but access can be provided to other computer systems. Also, access controls are often incorrectly thought of as only applying to *files*. However, they also protect other system resources such as the ability to place an outgoing long-distance phone call though a system modem (as well as, perhaps, the information that can be sent over such a call). Access controls can also apply to specific functions within an application and to specific fields of a file.

This chapter first discusses basic criteria that can be used to decide whether a particular user should be granted access to a particular system resource. It then reviews the use of these criteria by those who set policy (usually system-specific policy), commonly used *technical mechanisms* for implementing logical access control, and issues related to administration of access controls.

17.1 Access Criteria

In deciding whether to permit someone to use a system resource logical access controls examine whether *the user is authorized for the type of access requested.* (Note that this inquiry is usually distinct from the question of whether the user is authorized to use the system *at all*, which is usually addressed in an identification and authentication process.)

When determining what kind of technical access to allow to specific data, programs, devices, and resources, it is important to consider who will have access and what kind of access they will be allowed. It may be desirable for everyone in the organization to have access to some information on the system, such as the data displayed on an organization's daily calendar of nonconfidential meetings. The program that formats and displays the calendar, however, might be modifiable by only a very few system administrators, while the operating system controlling that program might be directly accessible by still fewer.

The system uses various criteria to determine if a request for access will be granted. They are typically used in some combination. Many of the advantages and complexities involved in implementing and managing access control are related to the different kinds of user accesses supported.

17.1.1 Identity

It is probably fair to say that the majority of access controls are based upon the identity of the user (either human or process), which is usually obtained through identification and authentication (I&A). (See Chapter 16.) The identity is usually unique, to support individual accountability, but can be a group identification or can even be anonymous. For example, public information dissemination systems may serve a large group called "researchers" in which the individual researchers are not known.

17.1.2 Roles

Access to information may also be controlled by the job assignment or function (i.e., the *role*) of the user who is seeking access. Examples of roles include data entry clerk, purchase officer, project leader, programmer, and technical editor. Access rights are grouped by role name, and the use of resources is restricted to individuals authorized to assume the associated role. An individual may be authorized for more than one role, but may be required to act in only a single role at a time. Changing roles may require logging out and then in again, or entering a role-changing command. Note that use of roles is *not* the same as shared-use accounts. An individual may be assigned a standard set of rights of a shipping department data entry clerk, for example, but the account would still be tied to that individual's identity to allow for auditing. (See Chapter 18.)

Many systems already support a small number of special-purpose roles, such as System Administrator or Operator. For example, an individual who is logged on in the role of a System Administrator can perform operations that would be denied to the same individual acting in the role of an ordinary user.

Recently, the use of roles has been expanded beyond system tasks to application-oriented activities. For example, a user in a company could have an Order Taking role, and would be able to collect and enter customer billing information, check on availability of particular items, request shipment of items, and issue invoices. In addition, there could be an Accounts Receivable role, which would receive payments and credit them to particular invoices. A Shipping role, could then be responsible for shipping products and updating the inventory. To provide additional security, constraints could be imposed so a single user would never be simultaneously authorized to assume all three roles. Constraints of this kind are sometimes referred to as *separation of duty constraints*.

The use of roles can be a very effective way of providing access control. The process of defining roles should be based on a thorough analysis of how an organization operates and should include input from a wide spectrum of users in an organization.

17.1.3 Location

Access to particular system resources may also be based upon physical or logical location. For example, in a prison, all users in areas to which prisoners are physically permitted may be limited to read-only access. Changing or deleting is limited to areas to which prisoners are denied

physical access. The same authorized users (e.g., prison guards) would operate under significantly different logical access controls, depending upon their physical location. Similarly, users can be restricted based upon network addresses (e.g., users from sites within a given organization may be permitted greater access than those from outside).

17.1.4 Time

Time-of-day or day-of-week restrictions are common limitations on access. For example, use of confidential personnel files may be allowed only during normal working hours – and maybe denied before 8:00 a.m. and after 6:00 p.m. and all day during weekends and holidays.

17.1.5 Transaction

Another approach to access control can be used by organizations handling transactions (e.g., account inquiries). Phone calls may first be answered by a computer that requests that callers key in their account number and perhaps a PIN. Some routine transactions can then be made directly, but more complex ones may require human intervention. In such cases, the computer, which already knows the account number, can grant a clerk, for example, access to a particular account *for the duration of the transaction*. When completed, the access authorization is terminated. This means that users have no choice in which accounts they have access to, and can reduce the potential for mischief. It also eliminates employee browsing of accounts (e.g., those of celebrities or their neighbors) and can thereby heighten privacy.

17.1.6 Service Constraints

Service constraints refer to those restrictions that depend upon the parameters that may arise during use of the application or that are preestablished by the resource owner/manager. For example, a particular software package may only be licensed by the organization for five users at a time. Access would be denied for a sixth user, even if the user were otherwise authorized to use the application. Another type of service constraint is based upon application content or numerical thresholds. For example, an ATM machine may restrict transfers of money between accounts to certain dollar limits or may limit maximum ATM withdrawals to $500 per day. Access may also be selectively permitted based on the type of service requested. For example, users of computers on a network may be permitted to exchange electronic mail but may not be allowed to log in to each others' computers.

17.1.7 Common Access Modes

In addition to considering criteria for *when* access should occur, it is also necessary to consider the *types* of access, or *access modes*. The concept of access modes is fundamental to access control. Common access modes, which can be used in both operating or application systems,

include the following:[116]

> *Read* access provides users with the capability to view information in a system resource (such as a file, certain records, certain fields, or some combination thereof), but not to *alter* it, such as delete from, add to, or modify in any way. One must assume that information can be copied and printed if it can be read (although perhaps only manually, such as by using a print screen function and retyping the information into another file).
>
> *Write* access allows users to add to, modify, or delete information in system resources (e.g., files, records, programs). Normally user have read access to anything they have write access to.
>
> *Execute* privilege allows users to run programs.
>
> *Delete* access allows users to erase system resources (e.g., files, records, fields, programs).[117] Note that if users have write access but not delete access, they could overwrite the field or file with gibberish or otherwise inaccurate information and, in effect, delete the information.

Other specialized access modes (more often found in applications) include:

> *Create* access allows users to create new files, records, or fields.
>
> *Search* access allows users to list the files in a directory.

Of course, these criteria can be used in conjunction with one another. For example, an organization may give authorized individuals write access to an application at any time from within the office but only read access during normal working hours if they dial-in.

Depending upon the technical mechanisms available to implement logical access control, a wide variety of access permissions and restrictions are possible. No discussion can present all possibilities.

17.2 Policy: The Impetus for Access Controls

Logical access controls are a technical means of implementing *policy decisions*. Policy is made by

---

[116] These access modes are described generically; exact definitions and capabilities will vary from implementation to implementation. Readers are advised to consult their system and application documentation.

[117] "Deleting" information does not necessarily physically remove the data from the storage media. This can have serious implications for information that must be kept confidential. See "Disposition of Sensitive Automated Information," CSL Bulletin, NIST, October 1992.

a management official responsible for a particular system, application, subsystem, or group of systems. The development of an access control policy may not be an easy endeavor. It requires balancing the often-competing interests of security, operational requirements, and user-friendliness. In addition, technical constraints have to be considered.

This chapter discusses issues relating to the technical implementation of logical access controls – not the actual policy decisions as to who *should* have what type of access. These decisions are typically included in system-specific policy, as discussed in Chapters 5 and 10.

Once these policy decisions have been made, they will be *implemented* (or *enforced*) through logical access controls. In doing so, it is important to realize that the capabilities of various types of technical mechanisms (for logical access control) vary greatly.[118]

A few *simple* examples of *specific policy issues* are provided below; it is important to recognize, however, that *comprehensive system-specific policy* is significantly more complex.

1. The director of an organization's personnel office could decide that all clerks can update all files, to increase the efficiency of the office. Or the director could decide that clerks can only view and update specific files, to help prevent information browsing.

2. In a disbursing office, a single individual is usually prohibited from both requesting and authorizing that a particular payment be made. This is a *policy decision* taken to reduce the likelihood of embezzlement and fraud.

3. Decisions may also be made regarding access to the system itself. In the government, for example, the senior information resources management official may decide that agency systems that process information protected by the Privacy Act may not be used to process public-access database applications.

17.3 Technical Implementation Mechanisms

Many mechanisms have been developed to provide internal and external access controls, and they vary significantly in terms of precision, sophistication, and cost. These methods are not mutually exclusive and are often employed in combination. Managers need to analyze their organization's protection requirements to select the most appropriate, cost-effective logical access controls.

17.3.1 Internal Access Controls

Internal access controls are a logical means of separating what defined users (or user groups) can or cannot do with system resources. Five methods of internal access control are discussed in this section: passwords, encryption, access control lists, constrained user interfaces, and labels.

[118] Some policies may not be technically implementable; appropriate technical controls may simply not exist.

17.3.1.1 Passwords

Passwords are most often associated with user authentication. (See Chapter 16.) However, they are also used to protect data and applications on many systems, including PCs. For instance, an accounting application may require a password to access certain financial data or to invoke a restricted application (or function of an application).[119]

> The use of passwords as a means of access control can result in a proliferation of passwords that can reduce overall security.

Password-based access control is often inexpensive because it is already included in a large variety of applications. However, users may find it difficult to remember additional application passwords, which, if written down or poorly chosen, can lead to their compromise. Password-based access controls for PC applications are often easy to circumvent if the user has access to the operating system (and knowledge of what to do). As discussed in Chapter 16, there are other disadvantages to using passwords.

17.3.1.2 Encryption

Another mechanism that can be used for logical access control is encryption. Encrypted information can only be decrypted by those possessing the appropriate cryptographic key. This is especially useful if strong physical access controls cannot be provided, such as for laptops or floppy diskettes. Thus, for example, if information is encrypted on a laptop computer, and the laptop is stolen, the information cannot be accessed. While encryption can provide strong access control, it is accompanied by the need for strong key management. Use of encryption may also affect availability. For example, lost or stolen keys or read/write errors may prevent the decryption of the information. (See the cryptography chapter.)

17.3.1.3 Access Control Lists

Access Control Lists (ACLs) refer to a register of: (1) users (including groups, machines, processes) who have been given permission to use a particular system resource, and (2) the types of access they have been permitted.

ACLs vary considerably in their capability and flexibility. Some only allow specifications for certain pre-set groups (e.g., owner, group, and world) while more advanced ACLs allow much more flexibility, such as *user-defined* groups. Also, more advanced ACLs can be used to explicitly *deny* access to a particular individual or group. With more advanced ACLs, access can be at the discretion of the policymaker (and implemented by the security administrator) or

[119] Note that this password is normally *in addition* to the one supplied initially to log onto the system.

individual user, depending upon how the controls are technically implemented.

Elementary ACLs. Elementary ACLs (e.g., "permission bits") are a widely available means of providing access control on multiuser systems. In this scheme, a short, predefined list of the access rights to files or other system resources is maintained.

Elementary ACLs are typically based on the concepts of *owner*, *group*, and *world*. For each of these, a set of access modes (typically chosen from read, write, execute, and delete) is specified by the owner (or custodian) of the resource. The owner is usually its creator, though in some cases, ownership of resources may be automatically assigned to project administrators, regardless of the identity of the creator. File owners often have all privileges for their resources.

Example of Elementary ACL for the file "payroll":

Owner: PAYMANAGER
Access: Read, Write, Execute, Delete

Group: COMPENSATION-OFFICE
Access: Read, Write, Execute, Delete

"World"
Access: None

In addition to the privileges assigned to the owner, each resource is associated with *a named group of users*. Users who are members of the group can be granted modes of access distinct from nonmembers, who belong to the rest of the "world" that includes all of the system's users. User groups may be arranged according to departments, projects, or other ways appropriate for the particular organization. For example, groups may be established for members of the Personnel and Accounting departments. The system administrator is normally responsible for technically maintaining and changing the membership of a group, based upon input from the owners/custodians of the particular resources to which the groups may be granted access.

As the name implies, however, the technology is not particularly flexible. It may not be possible to explicitly deny access to an individual who is a member of the file's group. Also, it may not be possible for two groups to easily share information (without exposing it to the "world"), since the list is predefined to only include one group. If two groups wish to share information, an owner may make the file

Since one would presume that no one would have access without being granted access, why would it be desirable to explicitly deny access? Consider a situation in which a group name has already been established for 50 employees. If it were desired to exclude five of the individuals from that group, it would be easier for the access control administrator to simply grant access to that group and take it away from the five rather than grant access to 45 people. Or, consider the case of a complex application in which many groups of users are defined. It may be desired, for some reason, to prohibit Ms. *X* from generating a particular report (perhaps she is under investigation). In a situation in which group names are used (and perhaps modified by others), this explicit denial may be a safety check to restrict Ms. *X*'s access – in case someone were to redefine a group (with access to the report generation function) to include Ms. *X*. She would still be denied access.

available to be read by "world." This may disclose information that should be restricted. Unfortunately, elementary ACLs have no mechanism to easily permit such sharing.

Advanced ACLs. Like elementary ACLs, advanced ACLs provide a form of access control based upon a logical registry. They do, however, provide *finer precision* in control.

Advanced ACLs can be very useful in many complex information sharing situations. They provide a great deal of flexibility in implementing system-specific policy and allow for customization to meet the security requirements of functional managers. Their flexibility also makes them more of a challenge to manage. The rules for determining access in the face of apparently conflicting ACL entries are not uniform across all implementations and can be confusing to security administrators. When such systems are introduced, they should be coupled with training to ensure their correct use.

Example of Advanced ACL for the file "payroll"

PAYMGR:	R,	W,	E,	D	
J. Anderson:		R,	W,	E,	-
L. Carnahan:		-,	-,	-,	-
B. Guttman:	R,	W,	E,	-	
E. Roback:	R,	W,	E,	-	
H. Smith:	R,	-,	-,	-	
PAY-OFFICE:		R,	-,	-,	-
WORLD:	-,	-,	-,	-	

17.3.1.4 Constrained User Interfaces

Often used in conjunction with ACLs are *constrained user interfaces*, which restrict users' access to specific functions by never allowing them to request the use of information, functions, or other specific system resources for which they do not have access. Three major types exist: (1) *menus*, (2) *database views*, and (3) *physically constrained user interfaces*.

Constrained user interfaces can provide a form of access control that closely models how an organization operates. Many systems allow administrators to restrict users' ability to use the operating system or application system directly. Users can only execute commands that are provided by the administrator, typically in the form of a *menu*. Another means of restricting users is through restricted *shells* which limit the system commands the user can invoke. The use of menus and shells can often make the system easier to use and can help reduce errors.

Menu-driven systems are a common constrained user interface, where different users are provided different menus on the same system.

Database views is a mechanism for restricting user access to data contained in a database. It may be necessary to allow a user to access a database, but that user may not need access to all the data in the database (e.g., not all fields of a record nor all records in the database). Views can be used to enforce complex access requirements that are often needed in database situations, such as those based on the content of a field. For example, consider the situation where clerks maintain

personnel records in a database. Clerks are assigned a range of clients based upon last name (e.g., A-C, D-G). Instead of granting a user access to all records, the view can grant the user access to the record based upon the first letter of the last name field.

Physically constrained user interfaces can also limit a user's abilities. A common example is an ATM machine, which provides only a limited number of physical buttons to select options; no alphabetic keyboard is usually present.

17.3.1.5 Security Labels

A security label is a designation assigned to a resource (such as a file). Labels can be used for a variety of purposes, including controlling access, specifying protective measures, or indicating additional handling instructions. In many implementations, once this designator has been set, it cannot be changed (except perhaps under carefully controlled conditions that are subject to auditing).

Data Categorization

One tool that is used to increase the ease of security labelling is categorizing data by similar protection requirements. For example, a label could be developed for "organization proprietary data." This label would mark information that can be disclosed only to the organization's employees. Another label, "public data" could be used to mark information that is available to anyone.

When used for access control, labels are also assigned to *user sessions*. Users are permitted to initiate sessions with specific labels only. For example, a file bearing the label "Organization Proprietary Information" would not be accessible (readable) except during user sessions with the corresponding label. Moreover, only a restricted set of users would be able to initiate such sessions. The labels of the session and those of the files accessed during the session are used, in turn, to label output from the session. This ensures that information is uniformly protected throughout its life on the system.

For systems with stringent security requirements (such as those processing national security information), labels may be useful in access control.

Labels are a very strong form of access control; however, they are often inflexible and can be expensive to administer. Unlike permission bits or access control lists, labels cannot ordinarily be changed. Since labels are permanently linked to specific information, data cannot be disclosed by a user copying information and changing the access to that file so that the information is more accessible than the original owner intended. By removing users' ability to arbitrarily designate the accessibility of files they own, opportunities for certain kinds of human errors and malicious software problems are eliminated. In the example above, it would not be possible to copy Organization Proprietary Information into a file with a different label. This prevents inappropriate disclosure, but can interfere with legitimate extraction of some information.

Labels are well suited for consistently and uniformly enforcing access restrictions, although their administration and inflexibility can be a significant deterrent to their use.

17.3.2 External Access Controls

External access controls are a means of controlling interactions between the system and outside people, systems, and services. External access controls use a wide variety of methods, often including a separate physical device (e.g., a computer) that is between the system being protected and a network.

> One of the most common PPDs is the *dial-back modem*. A typical dial-back modem sequence follows: a user calls the dial-back modem and enters a password. The modem hangs up on the user and performs a table lookup for the password provided. If the password is found, the modem places a return call to the user (at a previously specified number) to initiate the session. The return call itself also helps to protect against the use of lost or compromised accounts. This is, however, not always the case. Malicious hackers can use such advance functions as call forwarding to reroute calls.

17.3.2.1 Port Protection Devices

Fitted to a communications port of a host computer, a port protection device (PPD) authorizes access to the port itself, prior to and independent of the computer's own access control functions. A PPD can be a separate device in the communications stream,[120] or it may be incorporated into a communications device (e.g., a modem). PPDs typically require a separate authenticator, such as a password, in order to access the communications port.

17.3.2.2 Secure Gateways/Firewalls

Often called *firewalls*, secure gateways block or filter access between two networks, often between a private[121] network and a larger, more public network such as the Internet, which attract malicious hackers. Secure gateways allow internal users to connect to external networks and at the same time prevent malicious hackers from compromising the internal systems.[122]

Some secure gateways are set up to allow all traffic to pass through except for specific traffic

---

[120] Typically PPDs are found only in serial communications streams.

[121] *Private network* is somewhat of a misnomer. *Private* does not mean that the organization's network is totally inaccessible to outsiders or prohibits use of the outside network from insiders (or the network would be disconnected). It also does not mean that all the information on the network requires confidentiality protection. It does mean that a network (or part of a network) is, in some way, separated from another network.

[122] Questions frequently arise as to whether secure gateways help prevent the spread of viruses. In general, having a gateway scan transmitted files for viruses requires more system overhead than is practical, especially since the scanning would have to handle many different file formats. However, secure gateways may reduce the spread of network worms.

which has known or suspected vulnerabilities or security problems, such as remote log-in services. Other secure gateways are set up to disallow all traffic except for specific types, such as e-mail. Some secure gateways can make access-control decisions based on the location of the requester. There are several technical approaches and mechanisms used to support secure gateways.

Because gateways provide security by restricting services or traffic, they can affect a system's usage. For this reason, firewall experts always emphasize the need for policy, so that appropriate officials decide how the organization will balance operational needs and security.

Types of Secure Gateways

There are many types of secure gateways. Some of the most common are packet filtering (or screening) routers, proxy hosts, bastion hosts, dual-homed gateways, and screened-host gateways.

In addition to reducing the risks from malicious hackers, secure gateways have several other benefits. They can reduce internal system security overhead, since they allow an organization to concentrate security efforts on a limited number of machines. (This is similar to putting a guard on the first floor of a building instead of needing a guard on every floor.)

A second benefit is the centralization of services. A secure gateway can be used to provide a central management point for various services, such as advanced authentication (discussed in Chapter 16), e-mail, or public dissemination of information. Having a central management point can reduce system overhead and improve service.

17.3.2.3 Host-Based Authentication

Host-based authentication grants access based upon the *identity of the host* originating the request, instead of the identity of the user making the request. Many network applications in use today use host-based authentication to determine whether access is allowed. Under certain circumstances it is fairly easy to masquerade as the legitimate host, especially if the masquerading host is physically located close to the host being impersonated. Security measures to protect against misuse of some host-based authentication systems are available (e.g., Secure RPC[123] uses DES to provide a more secure identification of the client host).

An example of host-based authentication is the Network File System (NFS) which allows a server to make file systems/directories available to specific machines.

17.4 Administration of Access Controls

[123] RPC, or Remote Procedure Call, is the service used to implement NFS.

One of the most complex and challenging aspects of access control, administration involves implementing, monitoring, modifying, testing, and terminating user accesses on the system. These can be demanding tasks, even though they typically do not include making the actual decisions as to the type of access each user may have.[124] Decisions regarding accesses should be guided by organizational policy, employee job descriptions and tasks, information sensitivity, user "need-to-know" determinations, and many other factors.

There are three basic approaches to administering access controls: centralized, decentralized, or a combination of these. Each has relative advantages and disadvantages. Which is most appropriate in a given situation will depend upon the particular organization and its circumstances.

System and Security Administration

The administration of systems and security requires access to advanced functions (such as setting up a user account). The individuals who technically set up and modify who has access to what are very powerful users on the system; they are often called system or security administrators. On some systems, these users are referred to as having *privileged accounts*.

The type of access of these accounts varies considerably. Some administrator privileges, for example, may allow an individual to administer only one application or subsystem, while a higher level of privileges may allow for oversight and establishment of subsystem administrators.

Normally, users who are security administrators have two accounts: one for regular use and one for security use. This can help protect the security account from compromise. Furthermore, additional I&A precautions, such as ensuring that administrator passwords are robust and changed regularly, are important to minimize opportunities for unauthorized individuals to gain access to these functions.

17.4.1 Centralized Administration

Using centralized administration, one office or individual is responsible for configuring access controls. As users' information processing needs change, their accesses can be modified only through the central office, usually after requests have been approved by the appropriate official. This allows very strict control over information, because the ability to make changes resides with very few individuals. Each user's account can be centrally monitored, and closing all accesses for any user can be easily accomplished if that individual leaves the organization. Since relatively few individuals oversee the process, consistent and uniform procedures and criteria are usually not difficult to enforce. However, when changes are needed quickly, going through a central administration office can be frustrating and time-consuming.

[124] As discussed in the policy section earlier in this chapter, those decisions are usually the responsibility of the applicable application manager or cognizant management official. See also the discussion of system-specific policy in Chapters 5 and 10.

17.4.2 Decentralized Administration

In decentralized administration, access is directly controlled by the owners or creators of the files, often the functional manager. This keeps control in the hands of those most accountable for the information, most familiar with it and its uses, and best able to judge who needs what kind of access. This may lead, however, to a lack of consistency among owners/creators as to procedures and criteria for granting user accesses and capabilities. Also, when requests are not processed centrally, it may be much more difficult to form a systemwide composite view of all user accesses on the system at any given time. Different application or data owners may inadvertently implement combinations of accesses that introduce conflicts of interest or that are in some other way not in the organization's best interest.[125] It may also be difficult to ensure that all accesses are properly terminated when an employee transfers internally or leaves an organization.

17.4.3 Hybrid Approach

A hybrid approach combines centralized and decentralized administration. One typical arrangement is that central administration is responsible for the broadest and most basic accesses, and the owners/creators of files control types of accesses or changes in users' abilities for the files under their control. The main disadvantage to a hybrid approach is adequately defining which accesses should be assignable locally and which should be assignable centrally.

17.5 Coordinating Access Controls

It is vital that access controls protecting a system work together. At a minimum, three basic types of access controls should be considered: physical, operating system, and application. In general, access controls within an application are the most specific. However, for application access controls to be fully effective they need to be supported by operating system access controls. Otherwise access can be made to application resources without going through the application.[126] Operating system and application access controls need to be supported by physical access controls.

17.6 Interdependencies

Logical access controls are closely related to many other controls. Several of them have been discussed in the chapter.

---

[125] Without necessary review mechanisms, central administration does not *a priori* preclude this.

[126] For example, logical access controls within an application block User A from viewing File F. However, if operating systems access controls do not also block User A from viewing File F, User A can use a utility program (or another application) to view the file.

Policy and Personnel. The most fundamental interdependencies of logical access control are with policy and personnel. Logical access controls are the technical implementation of system-specific and organizational policy, which stipulates *who* should be able to access what kinds of information, applications, and functions. These decisions are normally based on the principles of separation of duties and least privilege.

Audit Trails. As discussed earlier, logical access controls can be difficult to implement correctly. Also, it is sometimes *not possible* to make logical access control as precise, or fine-grained, as would be ideal for an organization. In such situations, users may either deliberately or inadvertently abuse their access. For example, access controls cannot prevent a user from modifying data the user is authorized to modify, even if the modification is incorrect. Auditing provides a way to identify abuse of access permissions. It also provides a means to review the actions of system or security administrators.

Identification and Authentication. In most logical access control scenarios, the identity of the user must be established before an access control decision can be made. The access control process then associates the permissible forms of accesses with that identity. This means that access control can only be as effective as the I&A process employed for the system.

Physical Access Control. Most systems can be compromised if someone can physically access the machine (i.e., CPU or other major components) by, for example, restarting the system with different software. Logical access controls are, therefore, dependent on physical access controls (with the exception of encryption, which can depend solely on the strength of the algorithm and the secrecy of the key).

17.7 Cost Considerations

Incorporating logical access controls into a computer system involves the purchase or use of access control mechanisms, their implementation, and changes in user behavior.

Direct Costs. Among the direct costs associated with the use of logical access controls are the purchase and support of hardware, operating systems, and applications that provide the controls, and any add-on security packages. The most significant personnel cost in relation to logical access control is usually for administration (e.g., initially determining, assigning, and keeping access rights up to date). Label-based access control is available in a limited number of commercial products, but at greater cost and with less variety of selection. Role-based systems are becoming more available, but there are significant costs involved in customizing these systems for a particular organization. Training users to understand and use an access control system is another necessary cost.

Indirect Costs. The primary indirect cost associated with introducing logical access controls into

a computer system is the effect on user productivity. There may be additional overhead involved in having individual users properly determine (when under their control) the protection attributes of information. Another indirect cost that may arise results from users not being able to immediately access information necessary to accomplish their jobs because the permissions were incorrectly assigned (or have changed). This situation is familiar to most organizations that put strong emphasis on logical access controls.

References

Abrams, M.D., et al. *A Generalized Framework for Access Control: An Informal Description*. McLean, VA: Mitre Corporation, 1990.

Baldwin, R.W. "Naming and Grouping Privileges to Simplify Security Management in Large Databases." *1990 IEEE Symposium on Security and Privacy Proceedings*. Oakland, CA: IEEE Computer Society Press, May 1990. pp. 116-132.

Caelli, William, Dennis Longley, and Michael Shain. *Information Security Handbook*. New York, NY: Stockton Press, 1991.

Cheswick, William, and Steven Bellovin. *Firewalls and Internet Security*. Reading, MA: Addison-Wesley Publishing Company, 1994.

Curry, D. *Improving the Security of Your UNIX System, ITSTD-721-FR-90-21*. Menlo Park, CA: SRI International, 1990.

Dinkel, Charles. *Secure Data Network System Access Control Documents*. NISTIR 90-4259. Gaithersburg, MD: National Institute of Standards and Technology, 1990.

Fites, P., and M. Kratz. *Information Systems Security: A Practitioner's Reference*. New York, NY: Van Nostrand Reinhold, 1993. Especially Chapters 1, 9, and 12.

Garfinkel, S., and Spafford, G. "UNIX Security Checklist." *Practical UNIX Security*. Sebastopol, CA: O'Riley & Associates. Inc., 1991. pp. 401-413.

Gasser, Morrie. *Building a Secure Computer System*. New York, NY: Van Nostrand Reinhold, 1988.

Haykin, M., and R. Warner. *Smart Card Technology: New Methods for Computer Access Control*. Spec Pub 500-157. Gaithersburg, MD: National Institute of Standards and Technology, 1988.

Landwehr, C., C. Heitmeyer, and J. McLean. "A Security Model for Military Message Systems." *ACM Transactions on Computer Systems*, Vol. 2, No. 3, August 1984.

National Bureau of Standards. *Guidelines for Security of Computer Applications*. Federal Information Processing Standard Publication 73. June 1980.

Pfleeger, Charles. *Security in Computing*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989.

President's Council on Integrity and Efficiency. *Review of General Controls in Federal Computer Systems*. Washington, DC: President's Council on Integrity and Efficiency, October 1988.

S. Salamone, "Internetwork Security: Unsafe at Any Node?" *Data Communications*. 22(12), 1993. pp. 61-68.

Sandhu, R. "Transaction Control Expressions for Separation of Duty." *Fourth Annual Computer Security Applications Conference Proceedings*. Orlando, FL, December 1988, pp. 282-286.

Thomsen, D.J. "Role-based Application Design and Enforcement." *Fourth IFIP Workshop on Database Security Proceedings.* International Federation for Information Processing, Halifax, England, September 1990.

T. Whiting. "Understanding VAX/VMS Security." *Computers and Security*. 11(8), 1992. pp. 695-698.

Chapter 18

AUDIT TRAILS

Audit trails maintain a record of system activity both by system and application processes and by user activity of systems and applications.[127] In conjunction with appropriate tools and procedures, audit trails can assist in detecting security violations, performance problems, and flaws in applications.[128]

The Difference Between Audit Trails and Auditing

An *audit trail* is a series of records of computer events, about an operating system, an application, or user activities. A computer system may have several audit trails, each devoted to a particular type of activity.

Auditing is the 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 *auditability* of the computer system. Auditing is discussed in the assurance chapter.

Audit trails may be used as either a support for regular system operations or a kind of insurance policy or as both of these. As insurance, audit trails are maintained but are not used unless needed, such as after a system outage. As a support for operations, audit trails are used to help system administrators ensure that the system or resources have not been harmed by hackers, insiders, or technical problems.

This chapter focuses on audit trails as a technical control, rather than the process of security auditing, which is a review and analysis of the security of a system as discussed in Chapter 9. This chapter discusses the benefits and objectives of audit trails, the types of audit trails, and some common implementation issues.

18.1 Benefits and Objectives

An *event* is any action that happens on a computer system. Examples include logging into a system, executing a program, and opening a file.

Audit trails can provide a means to help accomplish *several* security-related objectives, including individual accountability,

[127] Some security experts make a distinction between an *audit trail* and an *audit log* as follows: a *log* is a record of events made by a particular software package, and an *audit trail* is an entire history of an event, possibly using several logs. However, common usage within the security community does not make use of this definition. Therefore, this document does not distinguish between trails and logs.

[128] The type and amount of detail recorded by audit trails vary by both the technical capability of the logging application and the managerial decisions. Therefore, when we state that "audit trails can...," the reader should be aware that capabilities vary widely.

reconstruction of events, intrusion detection, and problem analysis.

18.1.1 Individual Accountability

Audit trails are a technical mechanism that help managers maintain individual accountability. By advising users that they are personally accountable for their actions, which are tracked by an audit trail that logs user activities, managers can help promote proper user behavior.[129] Users are less likely to attempt to circumvent security policy if they know that their actions will be recorded in an audit log.

For example, audit trails 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 upon 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.

Audit trails work in concert with logical access controls, which restrict use of system resources. Granting users access to particular resources usually means that they need that access to accomplish their job. Authorized access, of course, can be misused, which is where audit trail analysis is useful. While users cannot be prevented from using resources to which they have legitimate access authorization, audit trail analysis is used to examine their actions. For example, consider a personnel office in which users have access to those personnel records for which they are responsible. Audit trails can reveal that an individual is printing far more records than the average user, which could indicate the selling of personal data. Another example may be an engineer who is using a computer for the design of a new product. Audit trail analysis could reveal that an outgoing modem was used extensively by the engineer the week before quitting. This could be used to investigate whether proprietary data files were sent to an unauthorized party.

18.1.2 Reconstruction of Events

Audit trails can also be used to reconstruct events after a problem has occurred. Damage can be more easily assessed by reviewing audit trails of system activity to pinpoint how, when, and why normal operations ceased. Audit trail analysis can often distinguish between operator-induced errors (during which the system may have performed exactly as instructed) or system-created errors (e.g., arising from a poorly tested piece of replacement code). If, for example, a system fails or the integrity of a file (either program or data) is questioned, an analysis of the audit trail

[129] For a fuller discussion of changing employee behavior, see Chapter 13.

can reconstruct the series of steps taken by the system, the users, and the application. Knowledge of the conditions that existed at the time of, for example, a system crash, can be useful in avoiding future outages. Additionally, if a technical problem occurs (e.g., the corruption of a data file) audit trails can aid in the recovery process (e.g., by using the record of changes made to reconstruct the file).

18.1.3 Intrusion Detection

> *Intrusion detection* refers to the process of identifying attempts to penetrate a system and gain unauthorized access.

If audit trails have been designed and implemented to record appropriate information, they can assist in intrusion detection. Although normally thought of as a real-time effort, intrusions can be detected *in real time*, by examining audit records as they are created (or through the use of other kinds of warning flags/notices), or *after the fact* (e.g., by examining audit records in a batch process).

Real-time intrusion detection is primarily aimed at outsiders attempting to gain unauthorized access to the system. It may also be used to detect changes in the system's performance indicative of, for example, a virus or worm attack.[130] There may be difficulties in implementing real-time auditing, including unacceptable system performance.

After-the-fact identification may indicate that unauthorized access was attempted (or was successful). Attention can then be given to damage assessment or reviewing controls that were attacked.

18.1.4 Problem Analysis

Audit trails may also be used as on-line tools to help identify problems other than intrusions as they occur. This is often referred to as *real-time auditing* or monitoring. If a system or application is deemed to be critical to an organization's business or mission, real-time auditing may be implemented to monitor the status of these processes (although, as noted above, there can be difficulties with real-time analysis). An analysis of the audit trails may be able to verify that the *system* operated normally (i.e., that an error may have resulted from operator error, as opposed to a system-originated error). Such use of audit trails may be complemented by system performance logs. For example, a significant increase in the use of system resources (e.g., disk file space or outgoing modem use) *could* indicate a security problem.

[130] Viruses and worms of forms of malicious code. A virus is a code segment that replicates by attaching copies of itself to existing executables. A worm is a self-replicating program.

18.2 Audit Trails and Logs

A system can maintain several different audit trails concurrently. There are typically two kinds of audit records, (1) an event-oriented log and (2) a record of every keystroke, often called keystroke monitoring. Event-based logs usually contain records describing *system* events, *application* events, or *user* events.

An audit trail should include sufficient information to establish what events occurred and who (or what) caused them. In general, an event record should specify when the event occurred, the user ID associated with the event, the program or command used to initiate the event, and the result. Date and time can help determine if the user was a masquerader or the actual person specified.

18.2.1 Keystroke Monitoring[131]

Keystroke monitoring is the process used to view or record both the keystrokes entered by a computer user and the computer's response during an interactive session. Keystroke monitoring is usually considered a special case of audit trails. Examples of keystroke monitoring would include viewing characters as they are typed by users, reading users' electronic mail, and viewing other recorded information typed by users.

Some forms of routine system maintenance may record user keystrokes. This could constitute keystroke monitoring if the keystrokes are preserved along with the user identification so that an administrator could determine the keystrokes entered by specific users. Keystroke monitoring is conducted in an effort to protect systems and data from intruders who access the systems without authority or in excess of their assigned authority. Monitoring keystrokes typed by intruders can help administrators assess and repair damage caused by intruders.

18.2.2 Audit Events

System audit records are generally used to monitor and fine-tune system performance. *Application audit trails* may be used to discern flaws in applications, or violations of security policy committed within an application. *User audits records* are generally used to hold individuals accountable for their actions. An analysis of user audit records may expose a variety

---

[131] The Department of Justice has advised that an ambiguity in U.S. law makes it unclear whether keystroke monitoring is considered equivalent to an unauthorized telephone wiretap. The ambiguity results from the fact that current laws were written years before such concerns as keystroke monitoring or system intruders became prevalent. Additionally, no legal precedent has been set to determine whether keystroke monitoring is legal or illegal. System administrators conducting such monitoring might be subject to criminal and civil liabilities. The Department of Justice advises system administrators to protect themselves by giving notice to system users if keystroke monitoring is being conducted. Notice should include agency/organization policy statements, training on the subject, and a banner notice on each system being monitored. [NIST, *CSL Bulletin*, March 1993]

of security violations, which might range from simple browsing to attempts to plant Trojan horses or gain unauthorized privileges.

Sample System Log File Showing Authentication Messages

```
Jan 27 17:14:04 host1 login: ROOT LOGIN console
Jan 27 17:15:04 host1 shutdown: reboot by root
Jan 27 17:18:38 host1 login: ROOT LOGIN console
Jan 27 17:19:37 host1 reboot: rebooted by root
Jan 28 09:46:53 host1 su: 'su root' succeeded for user1 on /dev/ttyp0
Jan 28 09:47:35 host1 shutdown: reboot by user1
Jan 28 09:53:24 host1 su: 'su root' succeeded for user1 on /dev/ttyp1
Feb 12 08:53:22 host1 su: 'su root' succeeded for user1 on /dev/ttyp1
Feb 17 08:57:50 host1 date: set by user1
Feb 17 13:22:52 host1 su: 'su root' succeeded for user1 on /dev/ttyp0
```

The system itself enforces certain aspects of policy (particularly *system-specific* policy) such as access to files and access to the system itself. Monitoring the alteration of systems configuration files that implement the policy is important. If special accesses (e.g., security administrator access) have to be used to alter configuration files, the system should generate audit records whenever these accesses are used.

Application-Level Audit Record for a Mail Delivery System

```
Apr  9 11:20:22 host1 AA06370: from=<user2@host2>, size=3355, class=0
Apr  9 11:20:23 host1 AA06370: to=<user1@host1>, delay=00:00:02,
stat=Sent
Apr  9 11:59:51 host1 AA06436: from=<user4@host3>, size=1424, class=0
Apr  9 11:59:52 host1 AA06436: to=<user1@host1>, delay=00:00:02,
stat=Sent
Apr  9 12:43:52 host1 AA06441: from=<user2@host2>, size=2077, class=0
Apr  9 12:43:53 host1 AA06441: to=<user1@host1>, delay=00:00:01,
stat=Sent
```

Sometimes a finer level of detail than system audit trails is required. *Application audit trails* can provide this greater level of recorded detail. If an application is critical, it can be desirable to record not only who invoked the application, but certain details specific to each use. For example, consider an e-mail application. It may be desirable to record who sent mail, as well as to whom they sent mail and the length of messages. Another example would be that of a database application. It may be useful to record who accessed what database as well as the individual rows or columns of a table that were read (or changed or deleted), instead of just recording the execution of the database program.

User Log Showing a Chronological List of Commands Executed by Users

```
rcp         user1     ttyp0     0.02 secs Fri Apr  8 16:02
ls          user1     ttyp0     0.14 secs Fri Apr  8 16:01
clear       user1     ttyp0     0.05 secs Fri Apr  8 16:01
rpcinfo     user1     ttyp0     0.20 secs Fri Apr  8 16:01
nroff       user2     ttyp2     0.75 secs Fri Apr  8 16:00
sh          user2     ttyp2     0.02 secs Fri Apr  8 16:00
mv          user2     ttyp2     0.02 secs Fri Apr  8 16:00
sh          user2     ttyp2     0.03 secs Fri Apr  8 16:00
col         user2     ttyp2     0.09 secs Fri Apr  8 16:00
man         user2     ttyp2     0.14 secs Fri Apr  8 15:57
```

A *user audit trail* monitors and logs user activity in a system or application by recording events initiated by the user (e.g., access of a file, record or field, use of a modem).

Flexibility is a critical feature of audit trails. Ideally (from a security point of view), a system administrator would have the ability to monitor all system and user activity, but could choose to log only certain functions at the system level, and within certain applications. The decision of how much to log and how much to review should be a function of application/data sensitivity and should be decided by each functional manager/application owner with guidance from the system administrator and the computer security manager/officer, weighing the costs and benefits of the logging.[132]

18.2.2.1 System-Level Audit Trails

A system audit trail should be able to identify failed log-on attempts, especially if the system does not limit the number of failed log-on attempts. Unfortunately, some system-level audit trails cannot detect attempted log-ons, and therefore, cannot log them for later review. These audit trails can only monitor and log successful log-ons and subsequent activity. To effectively detect intrusion, a record of failed log-on attempts is required.

If a system-level audit capability exists, the audit trail should capture, at a minimum, any attempt to log on (successful or unsuccessful), the log-on ID, date and time of each log-on attempt, date and time of each log-off, the devices used, and the function(s) performed once logged on (e.g., the applications that the user tried, successfully or unsuccessfully, to invoke). System-level logging also typically includes information that is not specifically security-related, such as system operations, cost-accounting charges, and network performance.

[132] In general, audit logging can have privacy implications. Users should be aware of applicable privacy laws, regulations, and policies that may apply in such situations.

18.2.2.2 Application-Level Audit Trails

System-level audit trails may not be able to track and log events *within* applications, or may not be able to provide the level of detail needed by application or data owners, the system administrator, or the computer security manager. In general, application-level audit trails monitor and log user activities, including data files opened and closed, specific actions, such as reading, editing, and deleting records or fields, and printing reports. Some applications may be sensitive enough from a data availability, confidentiality, and/or integrity perspective that a "before" and "after" picture of each modified record (or the data element(s) changed within a record) should be captured by the audit trail.

18.2.2.3 User Audit Trails

User audit trails can usually log:

- all commands directly initiated by the user;
- all identification and authentication attempts; and
- files and resources accessed.

It is most useful if options and parameters are also recorded from commands. It is much more useful to know that a user tried to delete a log file (e.g., to hide unauthorized actions) than to know the user merely issued the delete command, possibly for a personal data file.

18.3 Implementation Issues

Audit trail data requires protection, since the data should be available for use when needed and is not useful if it is not accurate. *Also, the best planned and implemented audit trail is of limited value without timely review of the logged data.* Audit trails may be reviewed periodically, as needed (often triggered by occurrence of a security event), automatically in realtime, or in some combination of these. System managers and administrators, with

Audit Logs for Physical Access

Physical access control systems (e.g., a card/key entry system or an alarm system) use software and audit trails similar to general-purpose computers. The following are *examples* of criteria that may be used in selecting which events to log:

The date and time the access was attempted or made should be logged, as should the gate or door through which the access was attempted or made, and the individual (or user ID) making the attempt to access the gate or door.

Invalid attempts should be monitored and logged by noncomputer audit trails just as they are for computer-system audit trails. Management should be made aware if someone attempts to gain access during unauthorized hours.

Logged information should also include attempts to add, modify, or delete physical access privileges (e.g., granting a new employee access to the building or granting transferred employees access to their new office [and, of course, deleting their old access, as applicable]).

As with system and application audit trails, auditing of noncomputer functions can be implemented to send messages to security personnel indicating valid or invalid attempts to gain access to controlled spaces. In order not to desensitize a guard or monitor, all access should not result in messages being sent to a screen. Only exceptions, such as failed access attempts, should be highlighted to those monitoring access.

guidance from computer security personnel, should determine how long audit trail data will be maintained – either on the system or in archive files.

Following are examples of implementation issues that may have to be addressed when using audit trails.

18.3.1 Protecting Audit Trail Data

Access to on-line audit logs should be strictly controlled. Computer security managers and system administrators or managers should have access for review purposes; however, security and/or administration personnel who maintain logical access functions may have no need for access to audit logs.

It is particularly important to ensure the *integrity* of audit trail data against modification. One way to do this is to use digital signatures. (See Chapter 19.) Another way is to use write-once devices. The audit trail files needs to be protected since, for example, intruders may try to "cover their tracks" by modifying audit trail records. Audit trail records should be protected by strong access controls to help prevent unauthorized access. The integrity of audit trail information may be particularly important when legal issues arise, such as when audit trails are used as legal evidence. (This may, for example, require daily printing and signing of the logs.) Questions of such legal issues should be directed to the cognizant legal counsel.

The confidentiality of audit trail information may also be protected, for example, if the audit trail is recording information about users that may be disclosure-sensitive such as transaction data containing personal information (e.g., "before" and "after" records of modification to income tax data). Strong access controls and encryption can be particularly effective in preserving confidentiality.

18.3.2 Review of Audit Trails

Audit trails can be used to review what occurred after an event, for periodic reviews, and for real-time analysis. Reviewers should know what to look for to be effective in spotting unusual activity. They need to understand what normal activity looks like. Audit trail review can be easier if the audit trail function can be queried by user ID, terminal ID, application name, date and time, or some other set of parameters to run reports of selected information.

Audit Trail Review After an Event. Following a known system or application software problem, a known violation of existing requirements by a user, or some unexplained system or user problem, the appropriate system-level or application-level administrator should review the audit trails. Review by the application/data owner would normally involve a separate report, based upon audit trail data, to determine if their resources are being misused.

Periodic Review of Audit Trail Data. Application owners, data owners, system administrators, data processing function managers, and computer security managers should determine how much review of audit trail records is necessary, based on the importance of identifying unauthorized activities. This determination should have a direct correlation to the frequency of periodic reviews of audit trail data.

Real-Time Audit Analysis. Traditionally, audit trails are analyzed in a batch mode at regular intervals (e.g., daily). Audit records are archived during that interval for later analysis. Audit analysis tools can also be used in a real-time, or near real-time fashion. Such intrusion detection tools are based on audit reduction, attack signature, and variance techniques. Manual review of audit records in real time is almost never feasible on large multiuser systems due to the volume of records generated. However, it might be possible to view all records associated with a particular user or application, and view them in real time.[133]

18.3.3 Tools for Audit Trail Analysis

Many types of tools have been developed to help to reduce the amount of information contained in audit records, as well as to distill useful information from the raw data. Especially on larger systems, audit trail software can create very large files, which can be extremely difficult to analyze manually. The use of automated tools is likely to be the difference between unused audit trail data and a robust program. Some of the types of tools include:

Audit reduction tools are preprocessors designed to reduce the volume of audit records to facilitate manual review. Before a security review, these tools can remove many audit records known to have little security significance. (This alone may cut in half the number of records in the audit trail.) These tools generally remove records generated by specified classes of events, such as records generated by nightly backups might be removed.

Trends/variance-detection tools look for anomalies in user or system behavior. It is possible to construct more sophisticated processors that monitor usage trends and detect major variations. For example, if a user typically logs in at 9 a.m., but appears at 4:30 a.m. one morning, this may indicate a security problem that may need to be investigated.

Attack signature-detection tools look for an *attack signature*, which is a specific sequence of events indicative of an unauthorized access attempt. A simple example would be repeated failed log-in attempts.

[133] This is similar to keystroke monitoring, though, and may be legally restricted.

18.4 Interdependencies

The ability to audit supports many of the controls presented in this handbook. The following paragraphs describe some of the most important interdependencies.

Policy. The most fundamental interdependency of audit trails is with policy. Policy dictates who is authorized access to what system resources. Therefore it specifies, directly or indirectly, what violations of policy should be identified through audit trails.

Assurance. System auditing is an important aspect of operational assurance. The data recorded into an audit trail is used to support a system audit. The analysis of audit trail data and the process of auditing systems are closely linked; in some cases, they may even be the same thing. In most cases, the analysis of audit trail data is a critical part of maintaining operational assurance.

Identification and Authentication. Audit trails are tools often used to help hold users accountable for their actions. To be held accountable, the users must be known to the system (usually accomplished through the identification and authentication process). However, as mentioned earlier, audit trails record events and associate them with the *perceived* user (i.e., the user ID). If a user is impersonated, the audit trail will establish events but *not* the identity of the user.

Logical Access Control. Logical access controls restrict the use of system resources to authorized users. Audit trails complement this activity in two ways. First, they may be used to identify breakdowns in logical access controls or to verify that access control restrictions are behaving as expected, for example, if a particular user is erroneously included in a group permitted access to a file. Second, audit trails are used *to audit use of resources by those who have legitimate access*. Additionally, to protect audit trail files, access controls are used to ensure that audit trails are not modified.

Contingency Planning. Audit trails assist in contingency planning by leaving a record of activities performed on the system or within a specific application. In the event of a technical malfunction, this log can be used to help reconstruct the state of the system (or specific files).

Incident Response. If a security incident occurs, such as hacking, audit records and other intrusion detection methods can be used to help determine the extent of the incident. For example, was just one file browsed, or was a Trojan horse planted to collect passwords?

Cryptography. Digital signatures can be used to protect audit trails from undetected modification. (This does not prevent deletion or modification of the audit trail, but will provide an alert that the audit trail has been altered.) Digital signatures can also be used in conjunction with adding secure time stamps to audit records. Encryption can be used if confidentiality of audit trail information is important.

18.5 Cost Considerations

Audit trails involve many costs. First, some system overhead is incurred recording the audit trail. Additional system overhead will be incurred storing and processing the records. The more detailed the records, the more overhead is required. Another cost involves human and machine time required to do the analysis. This can be minimized by using tools to perform most of the analysis. Many simple analyzers can be constructed quickly (and cheaply) from system utilities, but they are limited to audit reduction and identifying particularly sensitive events. More complex tools that identify trends or sequences of events are slowly becoming available as off-the-shelf software. (If complex tools are not available for a system, development may be prohibitively expensive. Some intrusion detection systems, for example, have taken years to develop.)

The final cost of audit trails is the cost of investigating anomalous events. If the system is identifying too many events as suspicious, administrators may spend undue time reconstructing events and questioning personnel.

References

Fites, P., and M. Kratz. *Information Systems Security: A Practitioner's Reference*. New York: Van Nostrand Reinhold, 1993, (especially Chapter 12, pp. 331 - 350).

Kim, G., and E. Spafford, "Monitoring File System Integrity on UNIX Platforms." *Infosecurity News*. 4(4), 1993. pp. 21-22.

Lunt, T. "Automated Audit Trail Analysis for Intrusion Detection," *Computer Audit Update*, April 1992. pp. 2-8.

National Computer Security Center. *A Guide to Understanding Audit in Trusted Systems*. NCSC-TG-001, Version-2. Ft. Meade, MD, 1988.

National Institute of Standards and Technology. "Guidance on the Legality of Keystroke Monitoring." *CSL Bulletin*. March 1993.

Phillips, P. W. "New Approach Identifies Malicious System Activity." *Signal*. 46(7), 1992. pp. 65-66.

Ruthberg, Z., et al. *Guide to Auditing for Controls and Security: A System Development Life Cycle Approach*. Special Publication 500-153. Gaithersburg, MD: National Bureau of Standards, 1988.

Stoll, Clifford. *The Cuckoo's Egg*. New York, NY: Doubleday, 1989.

Chapter 19

CRYPTOGRAPHY

Cryptography is a branch of mathematics based on the transformation of data. It provides an important tool for protecting information and is used in many aspects of computer security. For example, cryptography can help provide data confidentiality, integrity, electronic signatures, and advanced user authentication. Although modern cryptography relies upon advanced mathematics, users can reap its benefits without understanding its mathematical underpinnings.

This chapter describes cryptography as a tool for satisfying a wide spectrum of computer security needs and requirements. It describes fundamental aspects of the basic cryptographic technologies and some specific ways cryptography can be applied to improve security. The chapter also explores some of the important issues that should be considered when incorporating cryptography into computer systems.

> Cryptography is traditionally associated only with keeping data secret. However, modern cryptography can be used to provide many security services, such as electronic signatures and ensuring that data has not been modified.

19.1 Basic Cryptographic Technologies

Cryptography relies upon two basic components: an *algorithm* (or cryptographic methodology) and a *key*. In modern cryptographic systems, algorithms are complex mathematical formulae and keys are strings of bits. For two parties to communicate, they must use the same algorithm (or algorithms that are designed to work together). In some cases, they must also use the same key. Many cryptographic keys must be kept secret; sometimes algorithms are also kept secret.

> There are two basic types of cryptography: "secret key" and "public key."

There are two basic types of cryptography: *secret key systems* (also called symmetric systems) and *public key systems* (also called asymmetric systems). Table 19.1 compares some of the distinct features of secret and public key systems. Both types of systems offer advantages and disadvantages. Often, the two are combined to form a *hybrid system* to exploit the strengths of each type. To determine which type of cryptography best meets its needs, an organization first has to identify its security requirements and operating environment.

DISTINCT FEATURES	SECRET KEY CRYPTOGRAPHY	PUBLIC KEY CRYPTOGRAPHY
NUMBER OF KEYS	Single key.	Pair of keys.
TYPES OF KEYS	Key is secret.	One key is private, and one key is public.
PROTECTION OF KEYS	Disclosure and modification.	Disclosure and modification for private keys and modification for public keys.
RELATIVE SPEEDS	Faster.	Slower.

Table 19.1

19.1.1 Secret Key Cryptography

In secret key cryptography, two (or more) parties share the same key, and that key is used to encrypt and decrypt data. As the name implies, secret key cryptography relies on keeping the key secret. If the key is compromised, the security offered by cryptography is severely reduced or eliminated. Secret key cryptography assumes that the parties who share a key rely upon each other not to disclose the key and protect it against modification.

Secret key cryptography has been in use for centuries. Early forms merely transposed the written characters to hide the message.

The best known secret key system is the *Data Encryption Standard* (DES), published by NIST as Federal Information Processing Standard (FIPS) 46-2. Although the adequacy of DES has at times been questioned, these claims remain unsubstantiated, and DES remains strong. It is the most widely accepted, publicly available cryptographic system today. The American National Standards Institute (ANSI) has adopted DES as the basis for encryption, integrity, access control, and key management standards.

The *Escrowed Encryption Standard*, published as FIPS 185, also makes use of a secret key system. (See the discussion of Key Escrow Encryption in this chapter.)

19.1.2 Public Key Cryptography

Public key cryptography is a modern invention and requires the use of advanced mathematics.

Whereas secret key cryptography uses a single key shared by two (or more) parties, public key cryptography uses a pair of keys for *each* party. One of the keys of the pair is "public" and the other is "private." The public key can be made known to other parties; the private key must be kept confidential and must be known only to its owner. Both keys, however, need to be protected against modification.

Public key cryptography is particularly useful when the parties wishing to communicate cannot rely upon each other or do not share a common key. There are several public key cryptographic systems. One of the first public key systems is RSA, which can provide many different security services. The Digital Signature Standard (DSS), described later in the chapter, is another example of a public key system.

19.1.3 Hybrid Cryptographic Systems

Public and secret key cryptography have relative advantages and disadvantages. Although public key cryptography does not require users to share a common key, secret key cryptography is much faster: equivalent implementations of secret key cryptography can run 1,000 to 10,000 times faster than public key cryptography.

Secret key systems are often used for bulk data encryption and public key systems for automated key distribution.

To maximize the advantages and minimize the disadvantages of both secret and public key cryptography, a computer system can use both types in a complementary manner, with each performing different functions. Typically, the speed advantage of secret key cryptography means that it is used for encrypting data. Public key cryptography is used for applications that are less demanding to a computer system's resources, such as encrypting the keys used by secret key cryptography (for distribution) or to sign messages.

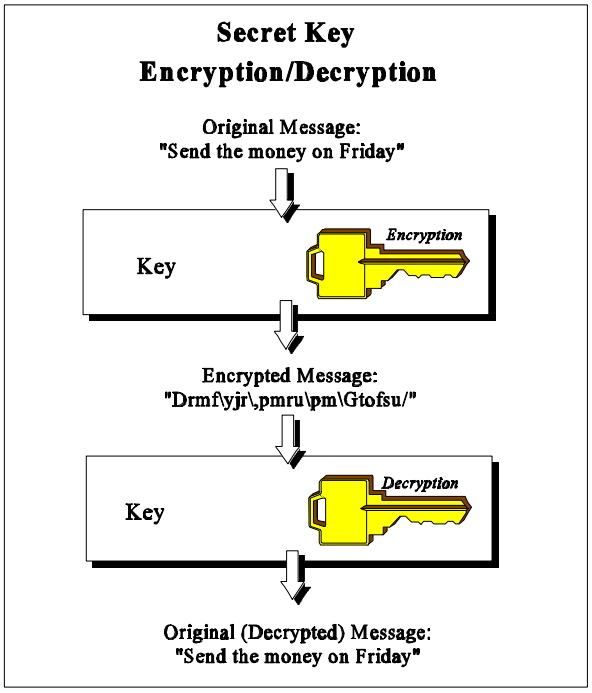
19.1.4 Key Escrow

Because cryptography can provide extremely strong encryption, it can thwart the government's efforts to lawfully perform electronic surveillance. For example, if strong cryptography is used to encrypt a phone conversation, a court-authorized wiretap will not be effective. To meet the needs of the government *and* to provide privacy, the federal government has adopted voluntary key escrow cryptography. This technology allows the use of strong encryption, but also allows the government when legally authorized to obtain decryption keys held by escrow agents. NIST has published the *Escrowed Encryption Standard* as FIPS 185. Under the Federal Government's

voluntary key escrow initiative, the decryption keys are split into parts and given to separate escrow authorities. Access to one part of the key does *not* help decrypt the data; both keys must be obtained.

19.2 Uses of Cryptography

Cryptography is used to protect data *both* inside and outside the boundaries of a computer system. Outside the computer system, cryptography is sometimes the *only* way to protect data. While in a computer system, data is normally protected with logical and physical access controls (perhaps supplemented by cryptography). However, when in transit across communications lines or resident on someone else's computer, data cannot be protected by the originator's[134] logical or physical access controls. Cryptography provides a solution by protecting data even when the data is no longer in the control of the originator.



19.2.1 Data Encryption

One of the best ways to obtain cost-effective data confidentiality is through the use of encryption. Encryption transforms intelligible data, called *plaintext*,[135] into an unintelligible form, called *ciphertext*. This process is reversed through the process of decryption. Once data is encrypted, the ciphertext does not have to be protected against disclosure. However, if ciphertext is modified, it will not decrypt correctly.

Both secret key and public key cryptography can be used for data encryption although not all public key algorithms provide for data encryption.

To use a secret key algorithm, data is encrypted using a key. The same key must be used to

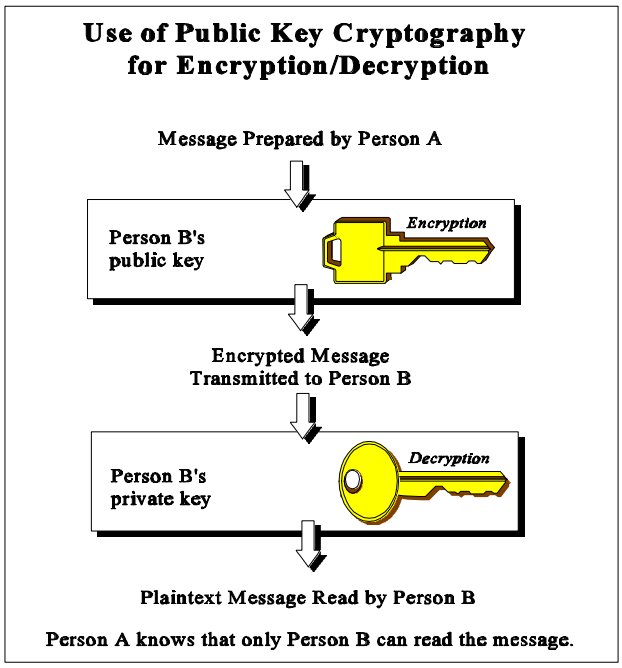
[134] The originator does not have to be the original creator of the data. It can also be a guardian or custodian of the data.

[135] Plaintext can be intelligible to a human (e.g., a novel) or to a machine (e.g., executable code).

decrypt the data.

When public key cryptography is used for encryption, any party may use any other party's public key to encrypt a message; however, only the party with the corresponding private key can decrypt, and thus read, the message.

Since secret key encryption is typically much faster, it is normally used for encrypting larger amounts of data.



19.2.2 Integrity

In computer systems, it is not always possible for humans to scan information to determine if data has been erased, added, or modified. Even if scanning were possible, the individual may have no way of knowing what the correct data should be. For example, "do" may be changed to "do not," or $1,000 may be changed to $10,000. It is therefore desirable to have an automated means of detecting *both* intentional and unintentional modifications of data.

While error detecting codes have long been used in communications protocols (e.g., parity bits), these are more effective in detecting (and correcting) unintentional modifications. They can be defeated by adversaries. Cryptography can effectively detect both intentional and unintentional modification; however, cryptography does not protect files from being modified. Both secret key and public key cryptography can be used to ensure integrity. Although newer public key methods may offer more flexibility than the older secret key method, secret key integrity verification systems have been successfully integrated into many applications.

When secret key cryptography is used, a message authentication code (MAC) is calculated from and appended to the data. To verify that the data has not been modified at a later time, any party with access to the correct secret key can recalculate the MAC. The new MAC is compared with the original MAC, and if they are identical, the verifier has confidence that the data has not been modified by an unauthorized party. FIPS 113, *Computer Data Authentication*, specifies a standard technique for calculating a MAC for integrity verification.

Public key cryptography verifies integrity by using of public key signatures and secure hashes. A secure hash algorithm is used to create a message digest. The message digest, called a hash, is a

short form of the message that changes if the message is modified. The hash is then signed with a private key. Anyone can recalculate the hash and use the corresponding public key to verify the integrity of the message.[136]

19.2.3 Electronic Signatures

Today's computer systems store and process increasing numbers of paper-based documents in electronic form. Having documents in electronic form permits rapid processing and transmission and improves overall efficiency. However, approval of a paper document has traditionally been indicated by a written signature. What is needed, therefore, is the electronic equivalent of a written signature that can be recognized as having the same legal status as a written signature. In addition to the integrity protections, discussed above, cryptography can provide a means of linking a document with a particular person, as is done with a written signature. Electronic signatures can use either secret key or public key cryptography; however, public key methods are generally easier to use.

What Is an Electronic Signature?

An electronic signature is a cryptographic mechanism that performs a similar function to a written signature. It is used to verify the origin and contents of a message. For example, a recipient of data (e.g., an e-mail message) can verify who signed the data and that the data was not modified after being signed. This also means that the originator (e.g., sender of an e-mail message) cannot falsely deny having signed the data.

Cryptographic signatures provide extremely strong proof that a message has not been altered and was signed by a specific key.[137] However, there are other mechanisms besides cryptographic-based electronic signatures that perform a *similar* function. These mechanisms provide some assurance of the origin of a message, some verification of the message's integrity, or both.[138]

---

[136] Sometimes a secure hash is used for integrity verification. However, this can be defeated if the hash is not stored in a secure location, since it may be possible for someone to change the message and then replace the old hash with a new one based on the modified message.

[137] Electronic signatures rely on the secrecy of the keys and the link or binding between the owner of the key and the key itself. If a key is compromised (by theft, coercion, or trickery), then the electronic originator of a message may not be the same as the owner of the key. Although the binding of cryptographic keys to actual people is a significant problem, it does not necessarily make electronic signatures less secure than written signatures. Trickery and coercion are problems for written signatures as well. In addition, written signatures are easily forged.

[138] The strength of these mechanisms relative to electronic signatures varies depending on the specific implementation; however, in general, electronic signatures are stronger and more flexible. These mechanisms may be used in conjunction with electronic signatures or separately, depending upon the system's specific needs and limitations.

- Examination of the transmission path of a message. When messages are sent across a network, such as the Internet, the message source and the physical path of the message are recorded as a part of the message. These can be examined electronically or manually to help ascertain the origin of a message.

- Use of a value-added network provider. If two or more parties are communicating via a third party network, the network provider may be able to provide assurance that messages originate from a given source and have not been modified.

- Acknowledgment statements. The recipient of an electronic message may confirm the message's origin and contents by sending back an acknowledgement statement.

- Use of audit trails. Audit trails can track the sending of messages and their contents for later reference.

Simply taking a digital picture of a written signature does not provide adequate security. Such a *digitized* written signature could easily be copied from one electronic document to another with no way to determine whether it is legitimate. Electronic signatures, on the other hand, are unique to the message being signed and will not verify if they are copied to another document.

19.2.3.1 Secret Key Electronic Signatures

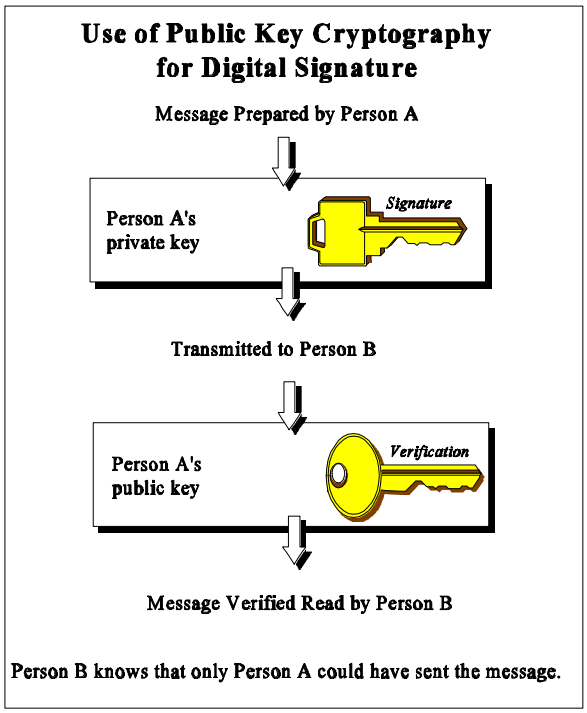
Systems incorporating message authentication technology have been approved for use by the federal government as a replacement for written signatures on electronic documents.

An electronic signature can be implemented using secret key message authentication codes (MACs). For example, if two parties share a secret key, and one party receives data with a MAC that is correctly verified using the shared key, that party may assume that the other party signed the data. This assumes, however, that the two parties trust each other. Thus, through the use of a MAC, in addition to data integrity, a form of electronic signature is obtained. Using additional controls, such as key notarization and key attributes, it is possible to provide an electronic signature even if the two parties do not trust each other.

19.2.3.2 Public Key Electronic Signatures

Another type of electronic signature called a *digital signature* is implemented using public key cryptography. Data is electronically signed by applying the originator's private key to the data. (The exact mathematical process for doing this is not important for this discussion.) To increase the speed of the process, the private key is applied to a shorter form of the data, called a "hash" or "message digest," rather than to the entire set of data. The resulting digital signature can be stored or transmitted along with the data. The signature can be verified by any party using the public key of the signer. This feature is very useful, for example, when distributing signed copies

of virus-free software. Any recipient can verify that the program remains virus-free. If the signature verifies properly, then the verifier has confidence that the data was not modified after being signed and that the owner of the public key was the signer.



NIST has published standards for a digital signature and a secure hash for use by the federal government in FIPS 186, *Digital Signature Standard* and FIPS 180, *Secure Hash Standard*.

19.2.4 User Authentication

Cryptography can increase security in user authentication techniques. As discussed in Chapter 16, cryptography is the basis for several advanced authentication methods. Instead of communicating passwords over an open network, authentication can be performed by demonstrating knowledge of a cryptographic key. Using these methods, a one-time password, which is not susceptible to eavesdropping, can be used. User authentication can use either secret or public key cryptography.

19.3 Implementation Issues

This section explores several important issues that should be considered when using (e.g., designing, implementing, integrating) cryptography in a computer system.

19.3.1 Selecting Design and Implementation Standards

NIST and other organizations have developed numerous standards for designing, implementing, and using cryptography and for integrating it into automated systems. By using these standards, organizations can reduce costs and protect their investments in technology. Standards provide solutions that have been accepted by a wide community and that have been reviewed by experts in relevant areas. Standards help ensure interoperability among different vendors' equipment, thus allowing an

> Applicable security standards provide a common level of security and interoperability among users.

organization to select from among various products in order to find cost-effective equipment.

Managers and users of computer systems will have to select among various standards when deciding to use cryptography. Their selection should be based on cost-effectiveness analysis, trends in the standard's acceptance, and interoperability requirements. In addition, each standard should be carefully analyzed to determine if it is applicable to the organization and the desired application. For example, the Data Encryption Standard and the Escrowed Encryption Standard are both applicable to certain applications involving communications of data over commercial modems. Some federal standards are mandatory for federal computer systems, including DES (FIPS 46-2) and the DSS (FIPS 181).

19.3.2 Deciding on Hardware vs. Software Implementations

The trade-offs among security, cost, simplicity, efficiency, and ease of implementation need to be studied by managers acquiring various security products meeting a standard. Cryptography can be implemented in either hardware or software. Each has its related costs and benefits.

In general, software is less expensive and slower than hardware, although for large applications, hardware may be less expensive. In addition, software may be less secure, since it is more easily modified or bypassed than equivalent hardware products. Tamper resistance is usually considered better in hardware.

In many cases, cryptography is implemented in a hardware device (e.g., electronic chip, ROM-protected processor) but is controlled by software. This software requires integrity protection to ensure that the hardware device is provided with correct information (i.e., controls, data) and is not bypassed. Thus, a hybrid solution is generally provided, even when the basic cryptography is implemented in hardware. Effective security requires the correct management of the entire hybrid solution.

19.3.3 Managing Keys

The proper management of cryptographic keys is essential to the effective use of cryptography for security. Ultimately, the security of information protected by cryptography directly depends upon the protection afforded to keys.

All keys need to be protected against modification, and secret keys and private keys need protection against unauthorized disclosure. Key management involves the procedures and protocols, both manual and automated, used throughout the entire life cycle of the keys. This includes the generation, distribution, storage, entry, use, destruction, and archiving of cryptographic keys.

With secret key cryptography, the secret key(s) should be securely distributed (i.e., safeguarded

against unauthorized replacement, modification, and disclosure) to the parties wishing to communicate. Depending upon the number and location of users, this task may not be trivial. Automated techniques for generating and distributing cryptographic keys can ease overhead costs of key management, but some resources have to be devoted to this task. FIPS 171, *Key Management Using ANSI X9.17*, provides key management solutions for a variety of operational environments.

Public key cryptography users also have to satisfy certain key management requirements. For example6 0827 a private-public key pair is associated with (i.e., generated or held by) a specific user, it is necessary to *bind* the public part of the key pair to the user.[139]

In a small community of users, public keys and their "owners" can be strongly bound by simply exchanging public keys (e.g., putting them on a CD-ROM or other media). However, conducting electronic business on a larger scale6 potentially involving geographically and organizationally distributed users, necessitates a means for obtaining public keys electronically with a high degree of confide827 in their integrity and binding to individuals. The support for the binding between a key and its owner is generally referred to as a *public key infrastructure*.

Users also need to be able enter the community of key holders, generate keys (or have them generated on their behalf), disseminate public keys, revoke keys (in case, for example6 of compromise of the private key), and change keys. In addition, it may be necessary to build in time/date stamping and to archive keys for verification of old signatures.

19.3.4 Security of Cryptographic Modules

Cryptography is typically implemented in a *module* of software, firmware, hardware, or some combination thereof. This module contains the cryptographic algorithm(s), certain control parameters, and temporary storage facilities for the key(s) being used by the algorithm(s). The proper functioning of the cryptography requires the secure design, implementation, and use of the cryptographic module. This 082ludes protecting the module against tampering.

FIPS 140-1, *Security Requirements for Cryptographic Modules*, specifies the physical and logical security requirements for cryptographic modules. The standard defines four security levels for cryptographic modules, with each level providing a significant increase in security over the preceding level. The four levels allow for cost-effective solutions that are appropriate for different degrees of data sensitivity and different application environments. The user can select the best module for any given application or system, avoiding the cost of unnecessary security features.

[139] In some cases, the key may be bound to a position or an organization, rather than to an individual user.

19.3.5 Applying Cryptography to Networks

The use of cryptography within networking applications often requires special considerations. In these applications, the suitability of a cryptographic module may depend on its capability for handling special requirements imposed by locally attached communications equipment or by the network protocols and software.

Encrypted information, MACs, or digital signatures may require transparent communications protocols or equipment to avoid being misinterpreted by the communications equipment or software as control information. It may be necessary to format the encrypted information, MAC, or digital signature to ensure that it does not confuse the communications equipment or software. It is essential that cryptography satisfy the requirements imposed by the communications equipment and does not interfere with the proper and efficient operation of the network.

Data is encrypted on a network using either link or end-to-end encryption. In general, *link encryption* is performed by service providers, such as a data communications provider. Link encryption encrypts all of the data along a communications path (e.g., a satellite link, telephone circuit, or T1 line). Since link encryption also encrypts routing data, communications nodes need to decrypt the data to continue routing. *End-to-end encryption* is generally performed by the end-user organization. Although data remains encrypted when being passed through a network, routing information remains visible. It is possible to combine both types of encryption.

19.3.6 Complying with Export Rules

The U.S. Government controls the export of cryptographic implementations. The rules governing export can be quite complex, since they consider multiple factors. In addition, cryptography is a rapidly changing field, and rules may change from time to time. Questions concerning the export of a particular implementation should be addressed to appropriate legal counsel.

19.4 Interdependencies

There are many interdependencies among cryptography and other security controls highlighted in this handbook. Cryptography both depends on other security safeguards and assists in providing them.

Physical Security. Physical protection of a cryptographic module is required to prevent – or at least detect – physical replacement or modification of the cryptographic system and the keys within it. In many environments (e.g., open offices, portable computers), the cryptographic module itself has to provide the desired levels of physical security. In other environments (e.g., closed communications facilities, steel-encased Cash-Issuing Terminals), a cryptographic module may be safely employed within a secured facility.

User Authentication. Cryptography can be used both to protect passwords that are stored in computer systems and to protect passwords that are communicated between computers. Furthermore, cryptographic-based authentication techniques may be used in conjunction with, or in place of, password-based techniques to provide stronger authentication of users.

Logical Access Control. In many cases, cryptographic software may be embedded within a host system, and it may not be feasible to provide extensive physical protection to the host system. In these cases, logical access control may provide a means of isolating the cryptographic software from other parts of the host system and for protecting the cryptographic software from tampering and the keys from replacement or disclosure. The use of such controls should provide the equivalent of physical protection.

Audit Trails. Cryptography may play a useful role in audit trails. For example, audit records may need to be signed. Cryptography may also be needed to protect audit records stored on computer systems from disclosure or modification. Audit trails are also used to help support electronic signatures.

Assurance. Assurance that a cryptographic module is properly and securely implemented is essential to the effective use of cryptography. NIST maintains validation programs for several of its standards for cryptography. Vendors can have their products validated for conformance to the standard through a rigorous set of tests. Such testing provides increased assurance that a module meets stated standards, and system designers, integrators, and users can have greater confidence that validated products conform to accepted standards.

NIST maintains validation programs for several of its cryptographic standards.

A cryptographic system should be monitored and periodically audited to ensure that it is satisfying its security objectives. All parameters associated with correct operation of the cryptographic system should be reviewed, and operation of the system itself should be periodically tested and the results audited. Certain information, such as secret keys or private keys in public key systems, should not be subject to audit. However, nonsecret or nonprivate keys could be used in a simulated audit procedure.

19.5 Cost Considerations

Using cryptography to protect information has both direct and indirect costs. Cost is determined in part by product availability; a wide variety of products exist for implementing cryptography in integrated circuits, add-on boards or adapters, and stand-alone units.

19.5.1 Direct Costs

The direct costs of cryptography include:

- Acquiring or implementing the cryptographic module and integrating it into the computer system. The medium (i.e., hardware, software, firmware, or combination) and various other issues such as level of security, logical and physical configuration, and special processing requirements will have an impact on cost.

- Managing the cryptography and, in particular, managing the cryptographic keys, which includes key generation, distribution, archiving, and disposition, as well as security measures to protect the keys, as appropriate.

19.5.2 Indirect Costs

The indirect costs of cryptography include:

- A decrease in system or network performance, resulting from the additional overhead of applying cryptographic protection to stored or communicated data.

- Changes in the way users interact with the system, resulting from more stringent security enforcement. However, cryptography can be made nearly transparent to the users so that the impact is minimal.

References

Alexander, M., ed. "Protecting Data With Secret Codes," *Infosecurity News*. 4(6), 1993. pp. 72-78.

American Bankers Association. *American National Standard for Financial Institution Key Management (Wholesale).* ANSI X9.17-1985. Washington, DC., 1985.

Denning, P., and D. Denning, "The Clipper and Capstone Encryption Systems." *American Scientist.* 81(4), 1993. pp. 319-323.

Diffie, W., and M. Hellman. "New Directions in Cryptography." *IEEE Transactions on Information Theory*. Vol. IT-22, No. 6, November 1976. pp. 644-654.

Duncan, R. "Encryption ABCs." *Infosecurity News.* 5(2), 1994. pp. 36-41.

International Organization for Standardization. *Information Processing Systems - Open Systems*

Interconnection Reference Model - Part 2: Security Architecture. ISO 7498/2. 1988.

Meyer, C.H., and S. M. Matyas. *Cryptography: A New Dimension in Computer Data Security*. New York, NY: John Wiley & Sons, 1982.

Nechvatal, James. *Public-Key Cryptography*. Special Publication 800-2. Gaithersburg, MD: National Institute of Standards and Technology, April 1991.

National Bureau of Standards. *Computer Data Authentication*. Federal Information Processing Standard Publication 113. May 30, 1985.

National Institute of Standards and Technology. "Advanced Authentication Technology." *Computer Systems Laboratory Bulletin*. November 1991.

National Institute of Standards and Technology. *Data Encryption Standard*. Federal Information Processing Standard Publication 46-2. December 30, 1993.

National Institute of Standards and Technology. "Digital Signature Standard." *Computer Systems Laboratory Bulletin*. January 1993.

National Institute of Standards and Technology. *Digital Signature Standard*. Federal Information Processing Standard Publication 186. May 1994.

National Institute of Standards and Technology. *Escrowed Encryption Standard*. Federal Information Processing Standard Publication 185. 1994.

National Institute of Standards and Technology. *Key Management Using ANSI X9.17*. Federal Information Processing Standard Publication 171. April 27, 1992.

National Institute of Standards and Technology. *Secure Hash Standard*. Federal Information Processing Standard Publication 180. May 11, 1993.

National Institute of Standards and Technology. *Security Requirements for Cryptographic Modules*. Federal Information Processing Standard Publication 140-1. January 11, 1994.

Rivest, R., A. Shamir, and L. Adleman. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems." *Communications of the ACM.*, Vol. 21, No. 2, 1978. pp. 120-126.

Saltman, Roy G., ed. *Good Security Practices for Electronic Commerce, Including Electronic Data interchange*. Special Publication 800-9. Gaithersburg, MD: National Institute of Standards and Technology. December 1993.

Schneier, B. "A Taxonomy of Encryption Algorithms." *Computer Security Journal*. 9(1), 1193. pp. 39-60.

Schneier, B. "Four Crypto Standards." *Infosecurity News*. 4(2), 1993. pp. 38-39.

Schneier, B. *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. New York, NY: John Wiley & Sons, Inc., 1994.

U.S. Congress, Office of Technology Assessment. "Security Safeguards and Practices." *Defending Secrets, Sharing Data: New Locks and Keys for Electronic Information*. Washington, DC: 1987, pp. 54-72.

V. EXAMPLE

Chapter 20

ASSESSING AND MITIGATING THE RISKS TO A HYPOTHETICAL COMPUTER SYSTEM

This chapter illustrates how a hypothetical government agency (HGA) deals with computer security issues in its operating environment.[140] It follows the evolution of HGA's initiation of an assessment of the threats to its computer security system all the way through to HGA's recommendations for mitigating those risks. In the real world, many solutions exist for computer security problems. No single solution can solve similar security problems in all environments. Likewise, the solutions presented in this example may not be appropriate for all environments.

> This example can be used to help understand how security issues are examined, how some potential solutions are analyzed, how their cost and benefits are weighed, and ultimately how management accepts responsibility for risks.

This case study is provided for illustrative purposes only, and should not be construed as guidance or specific recommendations to solving specific security issues. Because a comprehensive example attempting to illustrate all handbook topics would be inordinately long, this example necessarily simplifies the issues presented and omits many details. For instance, to highlight the similarities and differences among controls in the different processing environments, it addresses some of the major types of processing platforms linked together in a distributed system: personal computers, local-area networks, wide-area networks, and mainframes; it does not show how to secure these platforms.

This section also highlights the importance of management's acceptance of a particular level of risk—this will, of course, vary from organization to organization. It is management's prerogative to decide what level of risk is appropriate, given operating and budget environments and other applicable factors.

20.1 Initiating the Risk Assessment

HGA has information systems that comprise and are intertwined with several different kinds of assets valuable enough to merit protection. HGA's systems play a key role in transferring U.S. Government funds to individuals in the form of paychecks; hence, financial resources are among the assets associated with HGA's systems. The system components owned and operated by HGA

[140] While this chapter draws upon many actual systems, details and characteristics were changed and merged. Although the chapter is arranged around an agency, the case study could also apply to a large division or office within an agency.

are also assets, as are personnel information, contracting and procurement documents, draft regulations, internal correspondence, and a variety of other day-to-day business documents, memos, and reports. HGA's assets include intangible elements as well, such as reputation of the agency and the confidence of its employees that personal information will be handled properly and that the wages will be paid on time.

A recent change in the directorship of HGA has brought in a new management team. Among the new Chief Information Officer's first actions was appointing a Computer Security Program Manager who immediately initiated a comprehensive risk analysis to assess the soundness of HGA's computer security program in protecting the agency's assets and its compliance with federal directives. This analysis drew upon prior risk assessments, threat studies, and applicable internal control reports. The Computer Security Program Manager also established a timetable for periodic reassessments.

Since the wide-area network and mainframe used by HGA are owned and operated by other organizations, they were not treated in the risk assessment as HGA's assets. And although HGA's personnel, buildings, and facilities are essential assets, the Computer Security Program Manager considered them to be outside the scope of the risk analysis.

After examining HGA's computer system, the risk assessment team identified specific threats to HGA's assets, reviewed HGA's and national safeguards against those threats, identified the vulnerabilities of those policies, and recommended specific actions for mitigating the remaining risks to HGA's computer security. The following sections provide highlights from the risk assessment. The assessment addressed many other issues at the programmatic and system levels. However, this chapter focuses on security issues related to the time and attendance application. (Other issues are discussed in Chapter 6.)

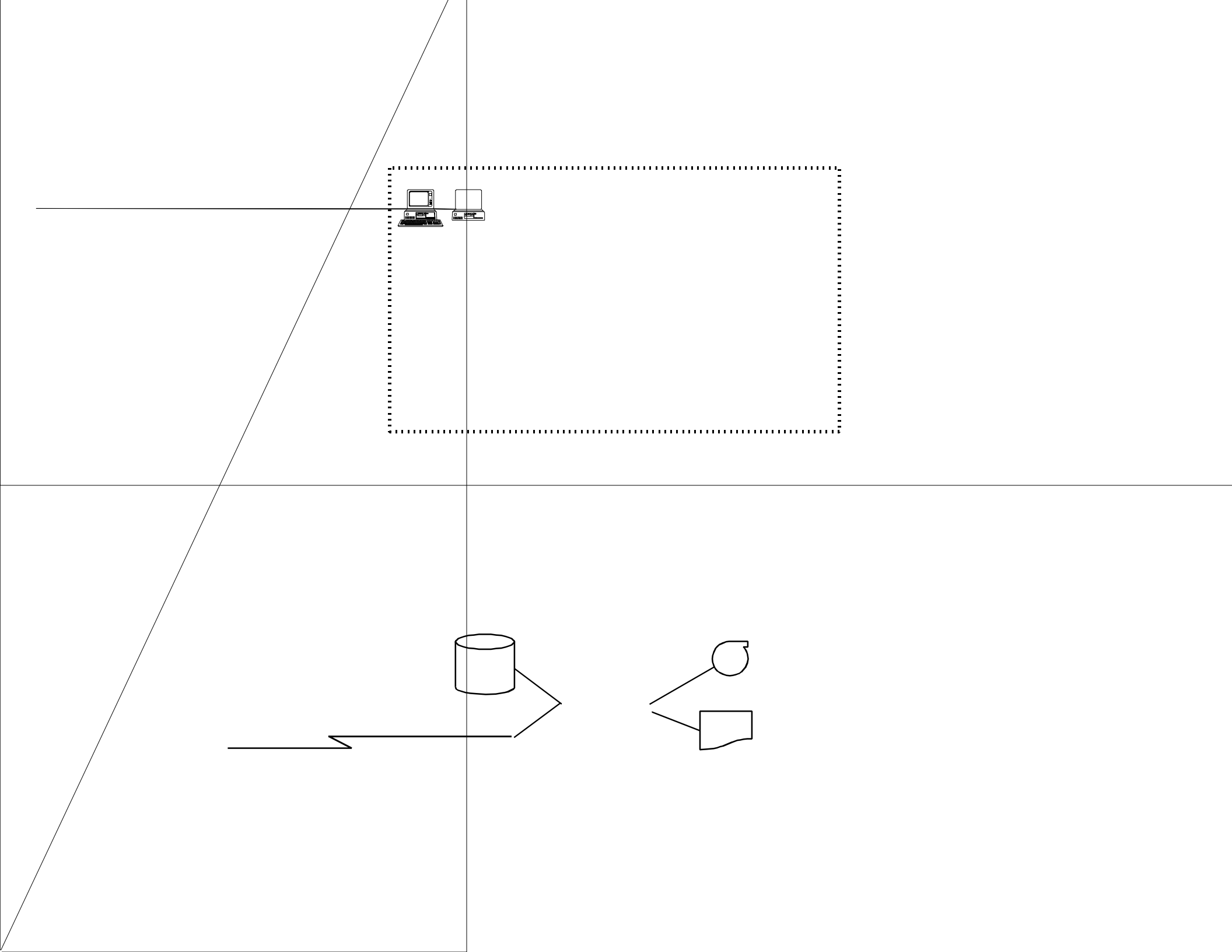
20.2 HGA's Computer System

HGA relies on the distributed computer systems and networks shown in Figure 20.1. They consist of a collection of components, some of which are systems in their own right. Some belong to HGA, but others are owned and operated by other organizations. This section describes these components, their role in the overall distributed system architecture, and how they are used by HGA.

20.2.1 System Architecture

Most of HGA's staff (a mix of clerical, technical, and managerial staff) are provided with personal computers (PCs) located in their offices. Each PC includes hard-disk and floppy-disk drives.

The PCs are connected to a local area network (LAN) so that users can exchange and share



information. The central component of the LAN is a *LAN server*, a more powerful computer that acts as an intermediary between PCs on the network and provides a large volume of disk storage for shared information, including shared application programs. The server provides logical access controls on potentially sharable information via elementary access control lists. These access controls can be used to limit user access to various files and programs stored on the server. Some programs stored on the server can be retrieved via the LAN and executed on a PC; others can only be executed on the server.

To initiate a session on the network or execute programs on the server, users at a PC must log into the server and provide a user identifier and password known to the server. Then they may use files to which they have access.

One of the applications supported by the server is *electronic mail* (e-mail), which can be used by all PC users. Other programs that run on the server can only be executed by a limited set of PC users.

Several printers, distributed throughout HGA's building complex, are connected to the LAN. Users at PCs may direct printouts to whichever printer is most convenient for their use.

Since HGA must frequently communicate with industry, the LAN also provides a connection to the Internet via a *router*. The router is a network interface device that translates between the protocols and addresses associated with the LAN and the Internet. The router also performs *network packet filtering*, a form of network access control, and has recently been configured to disallow non–e-mail (e.g., file transfer, remote log-in) between LAN and Internet computers.

The LAN server also has connections to several other devices.

- A *modem pool* is provided so that HGA's employees on travel can "dial up" via the public switched (telephone) network and read or send e-mail. To initiate a dial-up session, a user must successfully log in. During dial-up sessions, the LAN server provides access only to e-mail facilities; no other functions can be invoked.

- A *special console* is provided for the server administrators who configure the server, establish and delete user accounts, and have other special privileges needed for administrative and maintenance functions. These functions can only be invoked from the *administrator console*; that is, they cannot be invoked from a PC on the network or from a dial-up session.

- A *connection to a government agency X.25-based wide-area network* (WAN) is provided so that information can be transferred to or from other agency systems. One of the other hosts on the WAN is a large multiagency mainframe system. This mainframe is used to collect and process information from a large number of

agencies while providing a range of access controls.

20.2.2 System Operational Authority/Ownership

The system components contained within the large dashed rectangle shown in Figure 20.1 are managed and operated by an organization within HGA known as the Computer Operations Group (COG). This group includes the PCs, LAN, server, console, printers, modem pool, and router. The WAN is owned and operated by a large commercial telecommunications company that provides WAN services under a government contract. The mainframe is owned and operated by a federal agency that acts as a service provider for HGA and other agencies connected to the WAN.

20.2.3 System Applications

PCs on HGA's LAN are used for word processing, data manipulation, and other common applications, including spreadsheet and project management tools. Many of these tasks are concerned with data that are sensitive with respect to confidentiality or integrity. Some of these documents and data also need to be available in a timely manner.

The mainframe also provides storage and retrieval services for other databases belonging to individual agencies. For example, several agencies, including HGA, store their personnel databases on the mainframe; these databases contain dates of service, leave balances, salary and W-2 information, and so forth.

In addition to their time and attendance application, HGA's PCs and the LAN server are used to manipulate other kinds of information that may be sensitive with respect to confidentiality or integrity, including personnel-related correspondence and draft contracting documents.

20.3 Threats to HGA's Assets

Different assets of HGA are subject to different kinds of threats. Some threats are considered less likely than others, and the potential impact of different threats may vary greatly. The likelihood of threats is generally difficult to estimate accurately. Both HGA and the risk assessment's authors have attempted to the extent possible to base these estimates on historical data, but have also tried to anticipate new trends stimulated by emerging technologies (e.g., external networks).

20.3.1 Payroll Fraud

As for most large organizations that control financial assets, attempts at fraud and embezzlement are likely to occur. Historically, attempts at payroll fraud have almost always come from within HGA or the other agencies that operate systems on which HGA depends. Although HGA has thwarted many of these attempts, and some have involved relatively small sums of money, it

considers preventing financial fraud to be a *critical* computer security priority, particularly in light of the potential financial losses and the risks of damage to its reputation with Congress, the public, and other federal agencies.

Attempts to defraud HGA have included the following:

- Submitting fraudulent time sheets for hours or days not worked, or for pay periods following termination or transfer of employment. The former may take the form of overreporting compensatory or overtime hours worked, or underreporting vacation or sick leave taken. Alternatively, attempts have been made to modify time sheet data after being entered and approved for submission to payroll.

- Falsifying or modifying dates or data on which one's "years of service" computations are based, thereby becoming eligible for retirement earlier than allowed, or increasing one's pension amount.

- Creating employee records and time sheets for fictitious personnel, and attempting to obtain their paychecks, particularly after arranging for direct deposit.

20.3.2 Payroll Errors

Of greater likelihood, but of perhaps lesser potential impact on HGA, are errors in the entry of time and attendance data; failure to enter information describing new employees, terminations, and transfers in a timely manner; accidental corruption or loss of time and attendance data; or errors in interagency coordination and processing of personnel transfers.

Errors of these kinds can cause financial difficulties for employees and accounting problems for HGA. If an employee's vacation or sick leave balance became negative erroneously during the last pay period of the year, the employee's last paycheck would be automatically reduced. An individual who transfers between HGA and another agency may risk receiving duplicate paychecks or no paychecks for the pay periods immediately following the transfer. Errors of this sort that occur near the end of the year can lead to errors in W-2 forms and subsequent difficulties with the tax collection agencies.

20.3.3 Interruption of Operations

HGA's building facilities and physical plant are several decades old and are frequently under repair or renovation. As a result, power, air conditioning, and LAN or WAN connectivity for the server are typically interrupted several times a year for periods of up to one work day. For example, on several occasions, construction workers have inadvertently severed power or network cables. Fires, floods, storms, and other natural disasters can also interrupt computer operations, as can equipment malfunctions.

Another threat of small likelihood, but significant potential impact, is that of a malicious or disgruntled employee or outsider seeking to disrupt time-critical processing (e.g., payroll) by deleting necessary inputs or system accounts, misconfiguring access controls, planting computer viruses, or stealing or sabotaging computers or related equipment. Such interruptions, depending upon when they occur, can prevent time and attendance data from getting processed and transferred to the mainframe before the payroll processing deadline.

20.3.4 Disclosure or Brokerage of Information

Other kinds of threats may be stimulated by the growing market for information about an organization's employees or internal activities. Individuals who have legitimate work-related reasons for access to the master employee database may attempt to disclose such information to other employees or contractors or to sell it to private investigators, employment recruiters, the press, or other organizations. HGA considers such threats to be moderately likely and of low to high potential impact, depending on the type of information involved.

20.3.5 Network-Related Threats

Most of the human threats of concern to HGA originate from insiders. Nevertheless, HGA also recognizes the need to protect its assets from outsiders. Such attacks may serve many different purposes and pose a broad spectrum of risks, including unauthorized disclosure or modification of information, unauthorized use of services and assets, or unauthorized denial of services.

As shown in Figure 20.1, HGA's systems are connected to the three external networks: (1) the Internet, (2) the Interagency WAN, and (3) the public-switched (telephone) network. Although these networks are a source of security risks, connectivity with them is essential to HGA's mission and to the productivity of its employees; connectivity cannot be terminated simply because of security risks.

In each of the past few years before establishing its current set of network safeguards, HGA had detected several attempts by outsiders to penetrate its systems. Most, but not all of these, have come from the Internet, and those that succeeded did so by learning or guessing user account passwords. In two cases, the attacker deleted or corrupted significant amounts of data, most of which were later restored from backup files. In most cases, HGA could detect no ill effects of the attack, but concluded that the attacker may have browsed through some files. HGA also conceded that its systems did not have audit logging capabilities sufficient to track an attacker's activities. Hence, for most of these attacks, HGA could not accurately gauge the extent of penetration.

In one case, an attacker made use of a bug in an e-mail utility and succeeded in acquiring System Administrator privileges on the server—a significant breach. HGA found no evidence that the attacker attempted to exploit these privileges before being discovered two days later. When the

attack was detected, COG immediately contacted the HGA's Incident Handling Team, and was told that a bug fix had been distributed by the server vendor several months earlier. To its embarrassment, COG discovered that it had already received the fix, which it then promptly installed. It now believes that no subsequent attacks of the same nature have succeeded.

Although HGA has no evidence that it has been significantly harmed to date by attacks via external networks, it believes that these attacks have great potential to inflict damage. HGA's management considers itself lucky that such attacks have not harmed HGA's reputation and the confidence of the citizens its serves. It also believes the likelihood of such attacks via external networks will increase in the future.

20.3.6 Other Threats

HGA's systems also are exposed to several other threats that, for reasons of space, cannot be fully enumerated here. Examples of threats and HGA's assessment of their probabilities and impacts include those listed in Table 20.1.

20.4 Current Security Measures

HGA has numerous policies and procedures for protecting its assets against the above threats. These are articulated in HGA's *Computer Security Manual*, which implements and synthesizes the requirements of many federal directives, such as Appendix III to OMB Circular A-130, the Computer Security Act of 1987, and the Privacy Act. The manual also includes policies for automated financial systems, such as those based on OMB Circulars A-123 and A-127, as well as the Federal Managers' Financial Integrity Act.

Several examples of those policies follow, as they apply generally to the use and administration of HGA's computer system and specifically to security issues related to time and attendance, payroll, and continuity of operations.

20.4.1 General Use and Administration of HGA's Computer System

HGA's Computer Operations Group (COG) is responsible for controlling, administering, and maintaining the computer resources owned and operated by HGA. These functions are depicted in Figure 20.1 enclosed in the large, dashed rectangle. Only individuals holding the job title System Administrator are authorized to establish log-in IDs and passwords on multiuser HGA systems (e.g., the LAN server). Only HGA's employees and contract personnel may use the system, and only after receiving written authorization from the department supervisor (or, in the case of contractors, the contracting officer) to whom these individuals report.

COG issues copies of all relevant security policies and procedures to new users. Before activating

a system account for a new users, COG requires that they (1) attend a security awareness and training course or complete an interactive computer-aided-instruction training session and (2) sign an acknowledgment form indicating that they understand their security responsibilities.

Authorized users are assigned a secret log-in ID and password, which they must not share with anyone else. They are expected to comply with all of HGA's password selection and security procedures (e.g., periodically changing passwords). Users who fail to do so are subject to a range of penalties.

Examples of Threats to HGA Systems

Potential Threat	Probability	Impact
Accidental Loss/Release of Disclosure-Sensitive Information	Medium	Low/Medium
Accidental Destruction of Information	High	Medium
Loss of Information due to Virus Contamination	Medium	Medium
Misuse of System Resources	Low	Low
Theft	High	Medium
Unauthorized Access to Telecommunications Resources[*]	Medium	Medium
Natural Disaster	Low	High

[*] HGA operates a PBX system, which may be vulnerable to (1) hacker disruptions of PBX availability and, consequently, agency operations, (2) unauthorized access to outgoing phone lines for long-distance services, (3) unauthorized access to stored voice-mail messages, and (4) surreptitious access to otherwise private conversations/data transmissions.

Table 20.1

Users creating data that are sensitive with respect to disclosure or modification are expected to make effective use of the automated access control mechanisms available on HGA computers to reduce the risk of exposure to unauthorized individuals. (Appropriate training and education are in place to help users do this.) In general, access to disclosure-sensitive information is to be granted only to individuals whose jobs require it.

V. Example

20.4.2 Protection Against Payroll Fraud and Errors: Time and Attendance Application

The time and attendance application plays a major role in protecting against payroll fraud and errors. Since the time and attendance application is a component of a larger automated payroll process, many of its functional and security requirements have been derived from both governmentwide and HGA-specific policies related to payroll and leave. For example, HGA must protect personal information in accordance with the Privacy Act. Depending on the specific type of information, it should normally be viewable only by the individual concerned, the individual's supervisors, and personnel and payroll department employees. Such information should also be timely and accurate.

Each week, employees must sign and submit a time sheet that identifies the number of hours they have worked and the amount of leave they have taken. The Time and Attendance Clerk enters the data for a given group of employees and runs an application on the LAN server to verify the data's validity and to ensure that only authorized users with access to the Time and Attendance Clerk's functions can enter time and attendance data. The application performs these security checks by using the LAN server's access control and identification and authentication (I&A) mechanisms. The application compares the data with a limited database of employee information to detect incorrect employee identifiers, implausible numbers of hours worked, and so forth. After correcting any detected errors, the clerk runs another application that formats the time and attendance data into a report, flagging exception/out-of-bound conditions (e.g., negative leave balances).

Department supervisors are responsible for reviewing the correctness of the time sheets of the employees under their supervision and indicating their approval by initialing the time sheets. If they detect significant irregularities and indications of fraud in such data, they must report their findings to the Payroll Office before submitting the time sheets for processing. In keeping with the principle of separation of duty, all data on time sheets and corrections on the sheets that may affect pay, leave, retirement, or other benefits of an individual must be reviewed for validity by at least two authorized individuals (other than the affected individual).

Protection Against Unauthorized Execution

Only users with access to Time and Attendance Supervisor functions may approve and submit time and attendance data — or subsequent corrections thereof — to the mainframe. Supervisors may not approve their own time and attendance data.

Only the System Administrator has been granted access to assign a special access control privilege to server programs. As a result, the server's operating system is designed to prevent a bogus time and attendance application created by any other user from communicating with the WAN and, hence, with the mainframe.

The time and attendance application is supposed to be configured so that the clerk and supervisor functions can only be carried out from specific PCs attached to the LAN and only during normal working hours. Administrators are not authorized to exercise functions of the time and attendance application apart from those concerned with configuring the accounts, passwords, and access permissions for clerks and supervisors. Administrators are expressly prohibited by policy from entering, modifying, or submitting time and attendance data via the time and attendance application or other mechanisms.[141]

Protection against unauthorized execution of the time and attendance application depends on I&A and access controls. While the time and attendance application is accessible from any PC, unlike most programs run by PC users, it does not execute directly on the PC's processor. Instead, it executes on the server, while the PC behaves as a terminal, relaying the user's keystrokes to the server and displaying text and graphics sent from the server. The reason for this approach is that common PC systems do not provide I&A and access controls and, therefore, cannot protect against unauthorized time and attendance program execution. *Any* individual who has access to the PC could run any program stored there.

Another possible approach is for the time and attendance program to perform I&A and access control on its own by requesting and validating a password before beginning each time and attendance session. This approach, however, can be defeated easily by a moderately skilled programming attack, and was judged inadequate by HGA during the application's early design phase.

Recall that the server is a more powerful computer equipped with a multiuser operating system that includes password-based I&A and access controls. Designing the time and attendance application program so that it executes on the server under the control of the server's operating system provides a more effective safeguard against unauthorized execution than executing it on the user's PC.

Protection Against Payroll Errors

The frequency of data entry errors is reduced by having Time and Attendance clerks enter each time sheet into the time and attendance application twice. If the two copies are identical, both are considered error free, and the record is accepted for subsequent review and approval by a supervisor. If the copies are not identical, the discrepancies are displayed, and for each discrepancy, the clerk determines which copy is correct. The clerk then incorporates the corrections into one of the copies, which is then accepted for further processing. If the clerk

[141] Technically, Systems Administrators may still have the ability to do so. This highlights the importance of adequate managerial reviews, auditing, and personnel background checks.

makes the same data-entry error twice, then the two copies will match, and one will be accepted as correct, even though it is erroneous. To reduce this risk, the time and attendance application could be configured to require that the two copies be entered by different clerks.

In addition, each department has one or more Time and Attendance Supervisors who are authorized to review these reports for accuracy and to approve them by running another server program that is part of the time and attendance application. The data are then subjected to a collection of "sanity checks" to detect entries whose values are outside expected ranges. Potential anomalies are displayed to the supervisor prior to allowing approval; if errors are identified, the data are returned to a clerk for additional examination and corrections.

When a supervisor approves the time and attendance data, this application logs into the interagency mainframe via the WAN and transfers the data to a payroll database on the mainframe. The mainframe later prints paychecks or, using a pool of modems that can send data over phone lines, it may transfer the funds electronically into employee-designated bank accounts. Withheld taxes and contributions are also transferred electronically in this manner.

The Director of Personnel is responsible for ensuring that forms describing significant payroll-related personnel actions are provided to the Payroll Office at least one week before the payroll processing date for the first affected pay period. These actions include hiring, terminations, transfers, leaves of absences and returns from such, and pay raises.

The Manager of the Payroll Office is responsible for establishing and maintaining controls adequate to ensure that the amounts of pay, leave, and other benefits reported on pay stubs and recorded in permanent records and those distributed electronically are accurate and consistent with time and attendance data and with other information provided by the Personnel Department. In particular, paychecks must never be provided to anyone who is not a bona fide, active-status employee of HGA. Moreover, the pay of any employee who terminates employment, who transfers, or who goes on leave without pay must be suspended as of the effective date of such action; that is, extra paychecks or excess pay must not be dispersed.

Protection Against Accidental Corruption or Loss of Payroll Data

The same mechanisms used to protect against fraudulent modification are used to protect against accidental corruption of time and attendance data — namely, the access-control features of the server and mainframe operating systems.

COG's nightly backups of the server's disks protect against loss of time and attendance data. To a limited extent, HGA also relies on mainframe administrative personnel to back up time and attendance data stored on the mainframe, even though HGA has no direct control over these individuals. As additional protection against loss of data at the mainframe, HGA retains copies of all time and attendance data on line on the server for at least one year, at which time the data are

archived and kept for three years. The server's access controls for the on-line files are automatically set to read-only access by the time and attendance application at the time of submission to the mainframe. The integrity of time and attendance data will be protected by digital signatures as they are implemented.

The WAN's communications protocols also protect against loss of data during transmission from the server to the mainframe (e.g., error checking). In addition, the mainframe payroll application includes a program that is automatically run 24 hours before paychecks and pay stubs are printed. This program produces a report identifying agencies from whom time and attendance data for the current pay period were expected but not received. Payroll department staff are responsible for reviewing the reports and immediately notifying agencies that need to submit or resubmit time and attendance data. If time and attendance input or other related information is not available on a timely basis, pay, leave, and other benefits are temporarily calculated based on information estimated from prior pay periods.

20.4.3 Protection Against Interruption of Operations

HGA's policies regarding continuity of operations are derived from requirements stated in OMB Circular A-130. HGA requires various organizations within it to develop contingency plans, test them annually, and establish appropriate administrative and operational procedures for supporting them. The plans must identify the facilities, equipment, supplies, procedures, and personnel needed to ensure reasonable continuity of operations under a broad range of adverse circumstances.

COG Contingency Planning

COG is responsible for developing and maintaining a contingency plan that sets forth the procedures and facilities to be used when physical plant failures, natural disasters, or major equipment malfunctions occur sufficient to disrupt the normal use of HGA's PCs, LAN, server, router, printers, and other associated equipment.

The plan prioritizes applications that rely on these resources, indicating those that should be suspended if available automated functions or capacities are temporarily degraded. COG personnel have identified system software and hardware components that are compatible with those used by two nearby agencies. HGA has signed an agreement with those agencies, whereby they have committed to reserving spare computational and storage capacities sufficient to support HGA's system-based operations for a few days during an emergency.

No communication devices or network interfaces may be connected to HGA's systems without written approval of the COG Manager. The COG staff is responsible for installing all known security-related software patches in a timely manner and for maintaining spare or redundant PCs, servers, storage devices, and LAN interfaces to ensure that at least 100 people can simultaneously

perform word processing tasks at all times.

To protect against accidental corruption or loss of data, COG personnel back up the LAN server's disks onto magnetic tape every night and transport the tapes weekly to a sister agency for storage. HGA's policies also stipulate that all PC users are responsible for backing up weekly any significant data stored on their PC's local hard disks. For the past several years, COG has issued a yearly memorandum reminding PC users of this responsibility. COG also strongly encourages them to store significant data on the LAN server instead of on their PC's hard disk so that such data will be backed up automatically during COG's LAN server backups.

To prevent more limited computer equipment malfunctions from interrupting routine business operations, COG maintains an inventory of approximately ten fully equipped spare PC's, a spare LAN server, and several spare disk drives for the server. COG also keeps thousands of feet of LAN cable on hand. If a segment of the LAN cable that runs through the ceilings and walls of HGA's buildings fails or is accidentally severed, COG technicians will run temporary LAN cabling along the floors of hallways and offices, typically restoring service within a few hours for as long as needed until the cable failure is located and repaired.

To protect against PC virus contamination, HGA authorizes only System Administrators approved by the COG Manager to install licensed, copyrighted PC software packages that appear on the COG-approved list. PC software applications are generally installed only on the server. (These stipulations are part of an HGA assurance strategy that relies on the quality of the engineering practices of vendors to provide software that is adequately robust and trustworthy.) Only the COG Manager is authorized to add packages to the approved list. COG procedures also stipulate that every month System Administrators should run virus-detection and other security-configuration validation utilities on the server and, on a spot-check basis, on a number of PCs. If they find a virus, they must immediately notify the agency team that handles computer security incidents.

COG is also responsible for reviewing audit logs generated by the server, identifying audit records indicative of security violations, and reporting such indications to the Incident-Handling Team. The COG Manager assigns these duties to specific members of the staff and ensures that they are implemented as intended.

The COG Manager is responsible for assessing adverse circumstances and for providing recommendations to HGA's Director. Based on these and other sources of input, the Director will determine whether the circumstances are dire enough to merit activating various sets of procedures called for in the contingency plan.

Division Contingency Planning

HGA's divisions also must develop and maintain their own contingency plans. The plans must

identify critical business functions, the system resources and applications on which they depend, and the maximum acceptable periods of interruption that these functions can tolerate without significant reduction in HGA's ability to fulfill its mission. The head of each division is responsible for ensuring that the division's contingency plan and associated support activities are adequate.

For each major application used by multiple divisions, a chief of a single division must be designated as the *application owner*. The designated official (supported by his or her staff) is responsible for addressing that application in the contingency plan and for coordinating with other divisions that use the application.

If a division relies exclusively on computer resources maintained by COG (e.g., the LAN), it need not duplicate COG's contingency plan, but is responsible for reviewing the adequacy of that plan. If COG's plan does not adequately address the division's needs, the division must communicate its concerns to the COG Director. In either situation, the division must make known the criticality of its applications to the COG. If the division relies on computer resources or services that are *not* provided by COG, the division is responsible for (1) developing its own contingency plan or (2) ensuring that the contingency plans of other organizations (e.g., the WAN service provider) provide adequate protection against service disruptions.

20.4.4 Protection Against Disclosure or Brokerage of Information

HGA's protection against information disclosure is based on a need-to-know policy and on personnel hiring and screening practices. The need-to-know policy states that time and attendance information should be made accessible only to HGA employees and contractors whose assigned professional responsibilities require it. Such information must be protected against access from all other individuals, including other HGA employees. Appropriate hiring and screening practices can lessen the risk that an untrustworthy individual will be assigned such responsibilities.

The need-to-know policy is supported by a collection of physical, procedural, and automated safeguards, including the following:

- Time and attendance paper documents are must be stored securely when not in use, particularly during evenings and on weekends. Approved storage containers include locked file cabinets and desk drawers—to which only the owner has the keys. While storage in a container is preferable, it is also permissible to leave time and attendance documents on top of a desk or other exposed surface in a locked office (with the realization that the guard force has keys to the office). (This is a judgment left to local discretion.) Similar rules apply to disclosure-sensitive information stored on floppy disks and other removable magnetic media.

- Every HGA PC is equipped with a key lock that, when locked, disables the PC.

When information is stored on a PC's local hard disk, the user to whom that PC was assigned is expected to (1) lock the PC at the conclusion of each work day and (2) lock the office in which the PC is located.

- The LAN server operating system's access controls provide extensive features for controlling access to files. These include group-oriented controls that allow teams of users to be assigned to named groups by the System Administrator. Group members are then allowed access to sensitive files not accessible to nonmembers. Each user can be assigned to several groups according to need to know. (The reliable functioning of these controls is assumed, perhaps incorrectly, by HGA.)

- All PC users undergo security awareness training when first provided accounts on the LAN server. Among other things, the training stresses the necessity of protecting passwords. It also instructs users to log off the server before going home at night or before leaving the PC unattended for periods exceeding an hour.

20.4.5 Protection Against Network-Related Threats

HGA's current set of external network safeguards has only been in place for a few months. The basic approach is to tightly restrict the kinds of external network interactions that can occur by funneling all traffic to and from external networks through two interfaces that filter out unauthorized kinds of interactions. As indicated in Figure 20.1, the two interfaces are the network router and the LAN server. The only kinds of interactions that these interfaces allow are (1) e-mail and (2) data transfers from the server to the mainframe controlled by a few special applications (e.g., the time and attendance application).

Figure 20.1 shows that the network router is the only direct interface between the LAN and the Internet. The router is a dedicated special-purpose computer that translates between the protocols and addresses associated with the LAN and the Internet. Internet protocols, unlike those used on the WAN, specify that packets of information coming from or going to the Internet must carry an indicator of the kind of service that is being requested or used to process the information. This makes it possible for the router to distinguish e-mail packets from other kinds of packets—for example, those associated with a remote log-in request.[142] The router has been configured by COG to discard all packets coming from or going to the Internet, except those associated with e-mail. COG personnel believe that the router effectively eliminates Internet-based attacks on HGA user accounts because it disallows all remote log-in sessions, even those accompanied by a legitimate password.

[142] Although not discussed in this example, recognize that technical "spoofing" can occur.

The LAN server enforces a similar type of restriction for dial-in access via the public-switched network. The access controls provided by the server's operating system have been configured so that during dial-in sessions, only the e-mail utility can be executed. (HGA policy, enforced by periodic checks, prohibits installation of modems on PCs, so that access must be through the LAN server.) In addition, the server's access controls have been configured so that its WAN interface device is accessible only to programs that possess a special access-control privilege. Only the System Administrator can assign this privilege to server programs, and only a handful of special-purpose applications, like the time and attendance application, have been assigned this privilege.

20.4.6 Protection Against Risks from Non–HGA Computer Systems

HGA relies on systems and components that it cannot control directly because they are owned by other organizations. HGA has developed a policy to avoid undue risk in such situations. The policy states that system components controlled and operated by organizations other than HGA may not be used to process, store, or transmit HGA information without obtaining explicit permission from the application owner and the COG Manager. Permission to use such system components may not be granted without written commitment from the controlling organization that HGA's information will be safeguarded commensurate with its value, as designated by HGA. This policy is somewhat mitigated by the fact that HGA has developed an issue-specific policy on the use of the Internet, which allows for its use for e-mail with outside organizations and access to other resources (but not for transmission of HGA's proprietary data).

20.5 Vulnerabilities Reported by the Risk Assessment Team

The risk assessment team found that many of the risks to which HGA is exposed stem from (1) the failure of individuals to comply with established policies and procedures or (2) the use of automated mechanisms whose assurance is questionable because of the ways they have been developed, tested, implemented, used, or maintained. The team also identified specific vulnerabilities in HGA's policies and procedures for protecting against payroll fraud and errors, interruption of operations, disclosure and brokering of confidential information, and unauthorized access to data by outsiders.

20.5.1 Vulnerabilities Related to Payroll Fraud

Falsified Time Sheets

The primary safeguards against falsified time sheets are review and approval by supervisory personnel, who are not permitted to approve their own time and attendance data. The risk assessment has concluded that, while imperfect, these safeguards are adequate. The related requirement that a clerk and a supervisor must cooperate closely in creating time and attendance

data and submitting the data to the mainframe also safeguards against other kinds of illicit manipulation of time and attendance data by clerks or supervisors acting independently.

Unauthorized Access

When a PC user enters a password to the server during I&A, the password is sent to the server by broadcasting it over the LAN "in the clear." This allows the password to be intercepted easily by any other PC connected to the LAN. In fact, so-called "password sniffer" programs that capture passwords in this way are widely available. Similarly, a malicious program planted on a PC could also intercept passwords before transmitting them to the server. An unauthorized individual who obtained the captured passwords could then run the time and attendance application in place of a clerk or supervisor. Users might also store passwords in a log-on script file.

Bogus Time and Attendance Applications

The server's access controls are probably adequate for protection against bogus time and attendance applications that run on the server. However, the server's operating system and access controls have only been in widespread use for a few years and contain a number of security-related bugs. And the server's access controls are ineffective if not properly configured, and the administration of the server's security features in the past has been notably lax.

Unauthorized Modification of Time and Attendance Data

Protection against unauthorized modification of time and attendance data requires a variety of safeguards because each system component on which the data are stored or transmitted is a potential source of vulnerabilities.

First, the time and attendance data are entered on the server by a clerk. On occasion, the clerk may begin data entry late in the afternoon, and complete it the following morning, storing it in a temporary file between the two sessions. One way to avoid unauthorized modification is to store the data on a diskette and lock it up overnight. After being entered, the data will be stored in another temporary file until reviewed and approved by a supervisor. These files, now stored on the system, must be protected against tampering. As before, the server's access controls, if reliable and properly configured, can provide such protection (as can digital signatures, as discussed later) in conjunction with proper auditing.

Second, when the Supervisor approves a batch of time and attendance data, the time and attendance application sends the data over the WAN to the mainframe. The WAN is a collection of communications equipment and special-purpose computers called "switches" that act as relays, routing information through the network from source to destination. Each switch is a potential site at which the time and attendance data may be fraudulently modified. For example, an HGA PC user might be able to intercept time and attendance data and modify the data enroute to the

payroll application on the mainframe. Opportunities include tampering with incomplete time and attendance input files while stored on the server, interception and tampering during WAN transit, or tampering on arrival to the mainframe prior to processing by the payroll application.

Third, on arrival at the mainframe, the time and attendance data are held in a temporary file on the mainframe until the payroll application is run. Consequently, the mainframe's I&A and access controls must provide a critical element of protection against unauthorized modification of the data.

According to the risk assessment, the server's access controls, with prior caveats, probably provide acceptable protection against unauthorized modification of data stored on the server. The assessment concluded that a WAN-based attack involving collusion between an employee of HGA and an employee of the WAN service provider, although unlikely, should not be dismissed entirely, especially since HGA has only cursory information about the service provider's personnel security practices and no contractual authority over how it operates the WAN.

The greatest source of vulnerabilities, however, is the mainframe. Although its operating system's access controls are mature and powerful, it uses password-based I&A. This is of particular concern, because it serves a large number of federal agencies via WAN connections. A number of these agencies are known to have poor security programs. As a result, one such agency's systems could be penetrated (e.g., from the Internet) and then used in attacks on the mainframe via the WAN. In fact, time and attendance data awaiting processing on the mainframe would probably not be as attractive a target to an attacker as other kinds of data or, indeed, disabling the system, rendering it unavailable. For example, an attacker might be able to modify the employee data base so that it disbursed paychecks or pensions checks to fictitious employees. Disclosure-sensitive law enforcement databases might also be attractive targets.

The access control on the mainframe is strong and provides good protection against intruders breaking into a second application after they have broken into a first. However, previous audits have shown that the difficulties of system administration may present some opportunities for intruders to defeat access controls.

20.5.2 Vulnerabilities Related to Payroll Errors

HGA's management has established procedures for ensuring the timely submission and interagency coordination of paperwork associated with personnel status changes. However, an unacceptably large number of troublesome payroll errors during the past several years has been traced to the late submission of personnel paperwork. The risk assessment documented the adequacy of HGA's safeguards, but criticized the managers for not providing sufficient incentives for compliance.

V. Example

20.5.3 Vulnerabilities Related to Continuity of Operations

COG Contingency Planning

The risk assessment commended HGA for many aspects of COG's contingency plan, but pointed out that many COG personnel were completely unaware of the responsibilities the plan assigned to them. The assessment also noted that although HGA's policies require annual testing of contingency plans, the capability to resume HGA's computer-processing activities at another cooperating agency has never been verified and may turn out to be illusory.

Division Contingency Planning

assessment focused primarily, but not exclusively, on protecting the mainframe.

The risk assessment concluded that significant, avoidable information brokering vulnerabilities were present—particularly due to HGA's lack of compliance with its own policies and procedures. Time and attendance documents were typically not stored securely after hours, and few PCs containing time and attendance information were routinely locked. Worse yet, few were routinely powered down, and many were left logged into the LAN server overnight. These practices make it easy for an HGA employee wandering the halls after hours to browse or copy time and attendance information on another employee's desk, PC hard disk, or LAN server directories.

The risk assessment pointed out that information sent to or retrieved from the server is subject to eavesdropping by other PCs on the LAN. The LAN hardware transmits information by broadcasting it to all connection points on the LAN cable. Moreover, information sent to or retrieved from the server is transmitted in the clear—that is, without encryption. Given the widespread availability of LAN "sniffer" programs, LAN eavesdropping is trivial for a prospective information broker and, hence, is likely to occur.

Last, the assessment noted that HGA's employee master database is stored on the mainframe, where it might be a target for information brokering by employees of the agency that owns the mainframe. It might also be a target for information brokering, fraudulent modification, or other illicit acts by any outsider who penetrates the mainframe via another host on the WAN.

20.5.5 Network-Related Vulnerabilities

The risk assessment concurred with the general approach taken by HGA, but identified several vulnerabilities. It reiterated previous concerns about the lack of assurance associated with the server's access controls and pointed out that these play a critical role in HGA's approach. The assessment noted that the e-mail utility allows a user to include a copy of *any* otherwise accessible file in an outgoing mail message. If an attacker dialed in to the server and succeeded in logging in as an HGA employee, the attacker could use the mail utility to export copies of all the files accessible to that employee. In fact, copies could be mailed to any host on the Internet.

The assessment also noted that the WAN service provider may rely on microwave stations or satellites as relay points, thereby exposing HGA's information to eavesdropping. Similarly, any information, including passwords and mail messages, transmitted during a dial-in session is subject to eavesdropping.

20.6 Recommendations for Mitigating the Identified Vulnerabilities

The discussions in the following subsections were chosen to illustrate a *broad sampling*[143] of handbook topics. Risk management and security program management themes are integral throughout, with particular emphasis given to the selection of risk-driven safeguards.

20.6.1 Mitigating Payroll Fraud Vulnerabilities

To remove the vulnerabilities related to payroll fraud, the risk assessment team recommended[144] the use of stronger authentication mechanisms based on smart tokens to generate one-time passwords that cannot be used by an interloper for subsequent sessions. Such mechanisms would make it very difficult for outsiders (e.g., from the Internet) who penetrate systems on the WAN to use them to attack the mainframe. The authors noted, however, that the mainframe serves many different agencies, and HGA has no authority over the way the mainframe is configured and operated. Thus, the costs and procedural difficulties of implementing such controls would be substantial. The assessment team also recommended improving the server's administrative procedures and the speed with which security-related bug fixes distributed by the vendor are installed on the server.

After input from COG security specialists and application owners, HGA's managers accepted most of the risk assessment team's recommendations. They decided that since the residual risks from the falsification of time sheets were acceptably low, no changes in procedures were necessary. However, they judged the risks of payroll fraud due to the interceptability of LAN server passwords to be unacceptably high, and thus directed COG to investigate the costs and procedures associated with using one-time passwords for Time and Attendance Clerks and supervisor sessions on the server. Other users performing less sensitive tasks on the LAN would continue to use password-based authentication.

While the immaturity of the LAN server's access controls was judged a significant source of risk, COG was only able to identify one other PC LAN product that would be significantly better in this respect. Unfortunately, this product was considerably less friendly to users and application developers, and incompatible with other applications used by HGA. The negative impact of changing PC LAN products was judged too high for the potential incremental gain in security benefits. Consequently, HGA decided to accept the risks accompanying use of the current product, but directed COG to improve its monitoring of the server's access control configuration

---

[143] Some of the controls, such as auditing and access controls, play an important role in many areas. The limited nature of this example, however, prevents a broader discussion.

[144] Note that, for the sake of brevity, the process of evaluating the cost-effectiveness of various security controls is not specifically discussed.

and its responsiveness to vendor security reports and bug fixes.

HGA concurred that risks of fraud due to unauthorized modification of time and attendance data at or in transit to the mainframe should not be accepted unless no practical solutions could be identified. After discussions with the mainframe's owning agency, HGA concluded that the owning agency was unlikely to adopt the advanced authentication techniques advocated in the risk assessment. COG, however, proposed an alternative approach that did not require a major resource commitment on the part of the mainframe owner.

The alternative approach would employ digital signatures based on public key cryptographic techniques to detect unauthorized modification of time and attendance data. The data would be *digitally signed* by the supervisor using a private key prior to transmission to the mainframe. When the payroll application program was run on the mainframe, it would use the corresponding public key to validate the correspondence between the time and attendance data and the signature. Any modification of the data during transmission over the WAN or while in temporary storage at the mainframe would result in a mismatch between the signature and the data. If the payroll application detected a mismatch, it would reject the data; HGA personnel would then be notified and asked to review, sign, and send the data again. If the data and signature matched, the payroll application would process the time and attendance data normally.

HGA's decision to use advanced authentication for time and attendance Clerks and Supervisors can be combined with digital signatures by using smart tokens. Smart tokens are programmable devices, so they can be loaded with private keys and instructions for computing digital signatures without burdening the user. When supervisors approve a batch of time and attendance data, the time and attendance application on the server would instruct the supervisor to insert their token in the token reader/writer device attached to the supervisors' PC. The application would then send a special "hash" (summary) of the time and attendance data to the token via the PC. The token would generate a digital signature using its embedded secret key, and then transfer the signature back to the server, again via the PC. The time and attendance application running on the server would append the signature to the data before sending the data to the mainframe and, ultimately, the payroll application.

Although this approach did not address the broader problems posed by the mainframe's I&A vulnerabilities, it does provide a reliable means of detecting time and attendance data tampering. In addition, it protects against bogus time and attendance submissions from systems connected to the WAN because individuals who lack a time and attendance supervisor's smart token will be unable to generate valid signatures. (Note, however, that the use of digital signatures does require increased administration, particularly in the area of key management.) In summary, digital signatures mitigate risks from a number of different kinds of threats.

HGA's management concluded that digitally signing time and attendance data was a practical, cost-effective way of mitigating risks, and directed COG to pursue its implementation. (They also

noted that it would be useful as the agency moved to use of digital signatures in other applications.) This is an example of developing and providing a solution in an environment over which no single entity has overall authority.

20.6.2 Mitigating Payroll Error Vulnerabilities

After reviewing the risk assessment, HGA's management concluded that the agency's current safeguards against payroll errors and against accidental corruption and loss of time and attendance data were adequate. However, the managers also concurred with the risk assessment's conclusions about the necessity for establishing incentives for complying (and penalties for not complying) with these safeguards. They thus tasked the Director of Personnel to ensure greater compliance with paperwork-handling procedures and to provide quarterly compliance audit reports. They noted that the digital signature mechanism HGA plans to use for fraud protection can also provide protection against payroll errors due to accidental corruption.

20.6.3 Mitigating Vulnerabilities Related to the Continuity of Operations

The assessment recommended that COG institute a program of periodic internal training and awareness sessions for COG personnel having contingency plan responsibilities.The assessment urged that COG undertake a rehearsal during the next three months in which selected parts of the plan would be exercised. The rehearsal should include attempting to initiate some aspect of processing activities at one of the designated alternative sites. HGA's management agreed that additional contingency plan training was needed for COG personnel and committed itself to its first plan rehearsal within three months.

After a short investigation, HGA divisions owning applications that depend on the WAN concluded that WAN outages, although inconvenient, would not have a major impact on HGA. This is because the few time-sensitive applications that required WAN-based communication with the mainframe were originally designed to work with magnetic tape instead of the WAN, and could still operate in that mode; hence courier-delivered magnetic tapes could be used as an alternative input medium in case of a WAN outage. The divisions responsible for contingency planning for these applications agreed to incorporate into their contingency plans both descriptions of these procedures and other improvements.

With respect to mainframe outages, HGA determined that it could not easily make arrangements for a suitable alternative site. HGA also obtained and examined a copy of the mainframe facility's own contingency plan. After detailed study, including review by an outside consultant, HGA concluded that the plan had major deficiencies and posed significant risks because of HGA's reliance on it for payroll and other services. This was brought to the attention of the Director of HGA, who, in a formal memorandum to the head of the mainframe's owning agency, called for (1) a high-level interagency review of the plan by all agencies that rely on the mainframe, and (2) corrective action to remedy any deficiencies found.

HGA's management agreed to improve adherence to its virus-prevention procedures. It agreed (from the point of view of the entire agency) that information stored on PC hard disks is frequently lost. It estimated, however, that the labor hours lost as a result would amount to less than a person year—which HGA management does *not* consider to be unacceptable. After reviewing options for reducing this risk, HGA concluded that it would be cheaper to accept the associated loss than to commit significant resources in an attempt to avoid it. COG volunteered, however, to set up an automated program on the LAN server that e-mails backup reminders to all PC users once each quarter. In addition, COG agreed to provide regular backup services for about 5 percent of HGA's PCs; these will be chosen by HGA's management based on the information stored on their hard disks.

20.6.4 Mitigating Threats of Information Disclosure/Brokering

HGA concurred with the risk assessment's conclusions about its exposure to information-brokering risks, and adopted most of the associated recommendations.

The assessment recommended that HGA improve its security awareness training (e.g., via mandatory refresher courses) and that it institute some form of compliance audits. The training should be sure to stress the penalties for noncompliance. It also suggested installing "screen lock" software on PCs that automatically lock a PC after a specified period of idle time in which no keystrokes have been entered; unlocking the screen requires that the user enter a password or reboot the system.

The assessment recommended that HGA modify its information-handling policies so that employees would be required to store some kinds of disclosure-sensitive information only on PC local hard disks (or floppies), but not on the server. This would eliminate or reduce risks of LAN eavesdropping. It was also recommended that an activity log be installed on the server (and regularly reviewed). Moreover, it would avoid unnecessary reliance on the server's access-control features, which are of uncertain assurance. The assessment noted, however, that this strategy conflicts with the desire to store most information on the server's disks so that it is backed up routinely by COG personnel. (This could be offset by assigning responsibility for someone other than the PC owner to make backup copies.) Since the security habits of HGA's PC users have generally been poor, the assessment also recommended use of hard-disk encryption utilities to protect disclosure-sensitive information on unattended PCs from browsing by unauthorized individuals. Also, ways to encrypt information on the server's disks would be studied.

The assessment recommended that HGA conduct a thorough review of the mainframe's safeguards in these respects, and that it regularly review the mainframe audit log, using a query package, with particular attention to records that describe user accesses to HGA's employee master database.

20.6.5 Mitigating Network-Related Threats

The assessment recommended that HGA:

- require stronger I&A for dial-in access or, alternatively, that a restricted version of the mail utility be provided for dial-in, which would prevent a user from including files in outgoing mail messages;
- replace its current modem pool with encrypting modems, and provide each dial-in user with such a modem; and
- work with the mainframe agency to install a similar encryption capability for server-to-mainframe communications over the WAN.

As with previous risk assessment recommendations, HGA's management tasked COG to analyze the costs, benefits, and impacts of addressing the vulnerabilities identified in the risk assessment. HGA eventually adopted some of the risk assessment's recommendations, while declining others. In addition, HGA decided that its policy on handling time and attendance information needed to be clarified, strengthened, and elaborated, with the belief that implementing such a policy would help reduce risks of Internet and dial-in eavesdropping. Thus, HGA developed and issued a revised policy, stating that users are individually responsible for ensuring that they do not transmit disclosure-sensitive information outside of HGA's facilities via e-mail or other means. It also prohibited them from examining or transmitting e-mail containing such information during dial-in sessions and developed and promulgated penalties for noncompliance.

20.7 Summary

This chapter has illustrated how many of the concepts described in previous chapters might be applied in a federal agency. An integrated example concerning a Hypothetical Government Agency (HGA) has been discussed and used as the basis for examining a number of these concepts. HGA's distributed system architecture and its uses were described. The time and attendance application was considered in some detail.

For context, some national and agency-level policies were referenced. Detailed operational policies and procedures for computer systems were discussed and related to these high-level policies. HGA assets and threats were identified, and a detailed survey of selected safeguards, vulnerabilities, and risk mitigation actions were presented. The safeguards included a wide variety of procedural and automated techniques, and were used to illustrate issues of assurance, compliance, security program oversight, and inter-agency coordination.

As illustrated, effective computer security requires clear direction from upper management.

Upper management must assign security responsibilities to organizational elements and individuals and must formulate or elaborate the security policies that become the foundation for the organization's security program. These policies must be based on an understanding of the organization's mission priorities and the assets and business operations necessary to fulfill them. They must also be based on a pragmatic assessment of the threats against these assets and operations. A critical element is assessment of threat likelihoods. These are most accurate when derived from historical data, but must also anticipate trends stimulated by emerging technologies.

A good security program relies on an integrated, cost-effective collection of physical, procedural, and automated controls. Cost-effectiveness requires targeting these controls at the threats that pose the highest risks while accepting other residual risks. The difficulty of applying controls properly and in a consistent manner over time has been the downfall of many security programs. This chapter has provided numerous examples in which major security vulnerabilities arose from a lack of assurance or compliance. Hence, periodic compliance audits, examinations of the effectiveness of controls, and reassessments of threats are essential to the success of any organization's security program.

Cross Reference and Index

Interdependencies Cross Reference

The following is a cross reference of the interdependencies sections. Note that the references only include specific controls. Some controls were referenced in groups, such as technical controls and occasionally interdependencies were noted for all controls.

Control	Chapters Where It Is Cited
Policy	Program Management Life Cycle Personnel/User Contingency Awareness and Training Logical Access Audit
Program Management	Policy Awareness and Training
Risk Management	Life Cycle Contingency Incident
Life Cycle	Program Management Assurance
Assurance	Life Cycle Support and Operations Audit Cryptography
Personnel	Training and Awareness Support and Operations Access
Training and Awareness	Personnel/User Incident Support and Operations
Contingency	Incident

	Support and Operations Physical and Environmental Audit
Incident	Contingency Support and Operations Audit
Physical and Environment	Contingency Support and Operations Logical Access Cryptography
Support and Operations	Contingency Incident
Identification and Authentication	Personnel/User Physical and Environmental Logical Access Audit Cryptography
Access Controls	Policy Personnel/User Physical and Environmental Identification and Authentication Audit Cryptography
Audit	Identification and Authentication Logical Access Cryptography
Cryptography	Identification and Authentication

General Index