STUDENT EDITION

Cyber Attacks

Protecting NATIONAL Infrastructure

Edward G. Amoroso

B/H

Cyber Attacks

Protecting National Infrastructure

Student Edition

Edward G. Amoroso



AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Butterworth-Heinemann is an Imprint of Elsevier



Acquiring Editor: Pam Chester
Development Editor: David Bevans
Project Manager: Paul Gottehrer
Designer: Alisa Andreola

Butterworth-Heinemann is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA

Copyright © 2013 Elsevier Inc. All rights reserved

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods or professional practices, may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information or methods described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

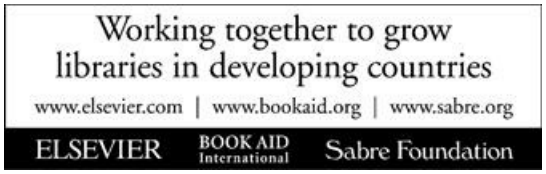
Library of Congress Cataloging-in-Publication Data
Amoroso, Edward G.
Cyber attacks : protecting national infrastructure / Edward Amoroso, John R. Vacca.–Student ed.
p. cm.
Summary: "Ten basic principles that will reduce the risk of cyber attack to national infrastructure in a substantive manner"–Provided by publisher.
ISBN 978-0-12-391855-0 (hardback)
1. Cyberterrorism–United States–Prevention. 2. Computer networks–Security measures. 3. Cyberspace–Security measures. 4. Computer crimes–United States–Prevention. 5. National security–United States. I. Vacca, John R. II. Title.
HV6773.2.A47 2012
363.325'90046780973–dc22
2012000035

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-391855-0

Printed in the United States of America

12 13 14 15 16 10 9 8 7 6 5 4 3 2 1



For information on all BH publications visit our website at www.elsevierdirect.com/security

Preface

Man did not enter into society to become worse than he was before, nor to have fewer rights than he had before, but to have those rights better secured.

Thomas Paine in Common Sense

Before you invest any of your time with this book, please take a moment and look over the following points. They outline my basic philosophy of national infrastructure security. I think that your reaction to these points will give you a pretty good idea of what your reaction will be to the book.

1. Citizens of free nations cannot hope to express or enjoy their freedoms if basic security protections are not provided. Security does not suppress freedom—it makes freedom possible.
2. In virtually every modern nation, computers and networks power critical infrastructure elements. As a result, cyber attackers can use computers and networks to damage or ruin the infrastructures that citizens rely on.
3. Security protections, such as those in security books, were designed for small-scale environments such as enterprise computing environments. These protections do not extrapolate to the protection of massively complex infrastructure.
4. Effective national cyber protections will be driven largely by cooperation and coordination between commercial, industrial, and government organizations. Thus, organizational management issues will be as important to national defense as technical issues.
5. Security is a process of risk reduction, not risk removal. Therefore, concrete steps can and should be taken to reduce, but not remove, the risk of cyber attack to national infrastructure.
6. The current risk of catastrophic cyber attack to national infrastructure must be viewed as extremely high, by any realistic measure. Taking little or no action to reduce this risk would be a foolish national decision.

The chapters of this book are organized around 10 basic principles that *will* reduce the risk of cyber attack to national infrastructure in a substantive manner. They are driven by experiences gained managing the security of one of the largest, most complex infrastructures in the world, by years of learning from various commercial and government organizations, and by years of interaction with students and academic researchers in the security field. They are also driven by personal experiences dealing with a wide range of successful and unsuccessful cyber attacks, including ones directed at infrastructure of considerable value. The implementation of the 10 principles in this book will require national resolve and changes to the way computing and networking elements are designed, built, and operated in the context of national infrastructure. My hope is that the suggestions offered in these pages will make this process easier.

Student Edition

To make it easier to teach these basic principles in the classroom, *Cyber Attacks Student Edition* adds new material developed by John R. Vacca, Editor-in-Chief of *Computer and Information Security Handbook* (Morgan Kaufmann Publishers) aimed specifically at enhancing the student experience, making it appropriate as a core textbook for instructors teaching courses in cyber security, information security, digital security, national security, intelligence studies, technology and infrastructure protection and similar courses.

Cyber Attacks Student Edition features the addition of **case studies** to illustrate actual implementation scenarios discussed in the text. The *Student Edition* also adds a host of **new pedagogical elements** to enhance learning, including chapter outlines, chapter summaries, learning checklists, chapter-by-chapter study questions, and more.

Instructor Support for *Cyber Attacks Student Edition* includes Test Bank, Lecture Slides, Lesson Plans, and Solutions Manual available online at http://textbooks.elsevier.com/web/Manuals.aspx?isbn=9780123918550.

- **Test Bank**—Compose, customize, and deliver exams using an online assessment package in a free Windows-based authoring tool that makes it easy to build tests using the unique multiple choice and true or false questions created for *Cyber Attacks Student Edition*. What's more, this authoring tool allows you to export customized exams directly to Blackboard, WebCT, eCollege, Angel, and other leading systems. All test bank files are also conveniently offered in Word format.
- **PowerPoint Lecture Slides**—Reinforce key topics with focused PowerPoints, which provide a perfect visual outline with which to augment your lecture. Each individual book chapter has its own dedicated slideshow.
- **Lesson Plans**—Design your course around customized lesson plans. Each individual lesson plan acts as separate syllabi containing content synopses, key terms, content synopses, directions to supplementary websites, and more open-ended critical thinking questions designed to spur class discussion. These lesson plans also delineate and connect chapter-based learning objectives to specific teaching resources, making it easy to catalogue the resources at your disposal.

Acknowledgments

The cyber security experts in the AT&T Chief Security Office, my colleagues across AT&T Labs and the AT&T Chief Technology Office, my colleagues across the entire AT&T business, and my graduate and undergraduate students in the Computer Science Department at the Stevens Institute of Technology have had a profound impact on my thinking and on the contents of this book. In addition, many prominent enterprise customers of AT&T with whom I've had the pleasure of serving, especially those in the United States Federal Government, have been great influencers in the preparation of this material.

I'd also like to extend a great thanks to my wife Lee, daughter Stephanie (17), son Matthew (15), and daughter Alicia (9) for their collective patience with my busy schedule.

TABLE OF CONTENTS

Title

Copyright

Preface

Acknowledgments

1. Introduction

National Cyber Threats, Vulnerabilities, and Attacks

Botnet Threat

National Cyber Security Methodology Components

Deception

Separation

Diversity

Consistency

Depth

Discretion

Collection

Correlation

Awareness

Response

Implementing the Principles Nationally

Protecting the Critical National Infrastructure Against Cyber Attacks

Summary

Chapter Review Questions/Exercises

2. Deception

Scanning Stage

Deliberately Open Ports

Discovery Stage

Deceptive Documents

Exploitation Stage

Procurement Tricks

Exposing Stage

Interfaces Between Humans and Computers

National Deception Program

The Deception Planning Process Against Cyber Attacks

Summary

Chapter Review Questions/Exercises

3. Separation

What Is Separation?

Functional Separation

National Infrastructure Firewalls

DDOS Filtering

SCADA Separation Architecture

Physical Separation

Insider Separation

Asset Separation

Multilevel Security (MLS)

Protecting the Critical National Infrastructure Through Use of Separation

Summary

Chapter Review Questions/Exercises

4. Diversity

Diversity and Worm Propagation

Desktop Computer System Diversity

Diversity Paradox of Cloud Computing

Network Technology Diversity

Physical Diversity

National Diversity Program

Critical Infrastructure Resilience and Diversity Initiative

Summary

Chapter Review Questions/Exercises

5. Commonality

Meaningful Best Practices for Infrastructure Protection

Locally Relevant and Appropriate Security Policy

Culture of Security Protection

Infrastructure Simplification

Certification and Education

Career Path and Reward Structure

Responsible Past Security Practice

National Commonality Program

How Critical National Infrastructure Systems Demonstrate Commonality

Summary

Chapter Review Questions/Exercises

6. Depth

Effectiveness of Depth

Layered Authentication

Layered E-Mail Virus and Spam Protection

Layered Access Controls

Layered Encryption

Layered Intrusion Detection

National Program of Depth

Practical Ways for Achieving Information Assurance in Infrastructure Networked Environments

Summary

Chapter Review Questions/Exercises

7. Discretion

Trusted Computing Base

Security Through Obscurity

Information Sharing

Information Reconnaissance

Obscurity Layers

Organizational Compartments

National Discretion Program

Top-Down and Bottom-Up Sharing of Sensitive Information

Summary

Chapter Review Questions/Exercises

8. Collection

Collecting Network Data

Collecting System Data

Security Information and Event Management

Large-Scale Trending

Tracking a Worm

National Collection Program

Data Collection Efforts: Systems and Assets

Summary

Chapter Review Questions/Exercises

9. Correlation

Conventional Security Correlation Methods

Quality and Reliability Issues in Data Correlation

Correlating Data to Detect a Worm

Correlating Data to Detect a Botnet

Large-Scale Correlation Process

National Correlation Program

Correlation Rules for Critical National Infrastructure Cyber Security

Summary

Chapter Review Questions/Exercises

10. Awareness

Detecting Infrastructure Attacks

Managing Vulnerability Information

Cyber Security Intelligence Reports

Risk Management Process

Security Operations Centers

National Awareness Program

Connecting Current Cyber Security Operation Centers to Enhance Situational Awareness

Summary

Chapter Review Questions/Exercises

11. Response

Pre- Versus Post-Attack Response

Indications and Warning

Incident Response Teams

Forensic Analysis

Law Enforcement Issues

Disaster Recovery

National Response Program

The Critical National Infrastructure Incident Response Framework

Transitioning from NIPP Steady State to Incident Response Management

Summary

Chapter Review Questions/Exercises

APPENDIX A. National Infrastructure Protection Criteria

Deception Requirements

Separation Requirements

Commonality Requirements

Diversity Requirements

Depth Requirements

Response Requirements

Awareness Requirements

Discretion Requirements

Collection Requirements

Correlation Requirements

APPENDIX B. Case Studies

John R. Vacca

Case Study 1: Cyber Storm

Case Study 2: Cyber Attacks on Critical Infrastructures—A Risk to the Nation

Case Study 3: Department of Homeland Security Battle Insider Threats and Maintain National Cyber Security

Case Study 4: Cyber Security Development Life Cycle

Case Study 5

REVIEW. Answers to Review Questions/Exercises, Hands-On Projects, Case Projects, and Optional Team Case Projects by Chapter

Chapter 1: Introduction

Chapter 2: Deception

Chapter 3: Separation

Chapter 4: Diversity

Chapter 5: Commonality

Chapter 6: Depth

Chapter 7: Discretion

Chapter 8: Collection

Chapter 9: Correlation

Chapter 10: Awareness

Chapter 11: Response

Index

1

Introduction

Chapter Outline

National Cyber Threats, Vulnerabilities, and Attacks
Botnet Threat
National Cyber Security Methodology Components
Deception
Separation
Diversity
Consistency
Depth
Discretion
Collection
Correlation
Awareness
Response
Implementing the Principles Nationally
Protecting the Critical National Infrastructure Against Cyber Attacks
Summary
Chapter Review Questions/Exercises

Somewhere in his writings—and I regret having forgotten where—John Von Neumann draws attention to what seemed to him a contrast. He remarked that for simple mechanisms it is often easier to describe how they work than what they do, while for more complicated mechanisms it was usually the other way round.

Edsger W. Dijkstra[1]

National infrastructure refers to the complex, underlying delivery and support systems for all large-scale services considered absolutely essential to a nation. These services include emergency response, law enforcement databases, supervisory control and data acquisition (SCADA) systems, power control networks, military support services, consumer entertainment systems, financial applications, and mobile telecommunications. Some national services are provided directly by government, but most are provided by commercial groups such as Internet service providers, airlines, and banks. In addition, certain services considered essential to one nation might include infrastructure support that is controlled by organizations from another nation. This global interdependency is consistent with the trends referred to collectively by Thomas Friedman as a "flat world."[2]

National infrastructure, especially in the United States, has always been vulnerable to malicious physical attacks such as equipment tampering, cable cuts, facility bombing, and asset theft. The events of September 11, 2001, for example, are the most prominent and recent instance of a massive physical attack directed at national infrastructure. During the past couple of decades, however, vast portions of national infrastructure have become reliant on software, computers, and networks. This reliance typically includes remote access, often over the Internet, to the systems that control national services. Adversaries thus can initiate cyber attacks on infrastructure using worms, viruses, leaks, and the like. These attacks indirectly target national infrastructure through their associated automated controls systems (see Figure 1.1).

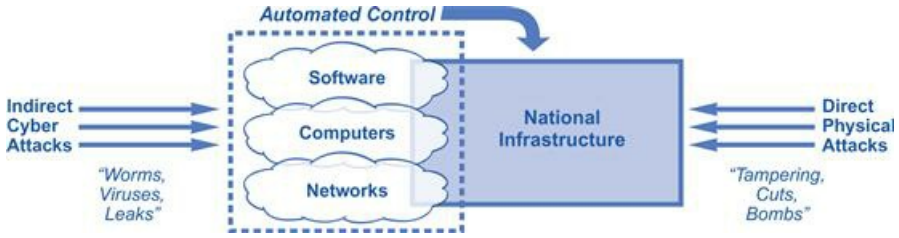


Figure 1.1 National infrastructure cyber and physical attacks.

A seemingly obvious approach to dealing with this national cyber threat would involve the use of well-known computer security techniques. After all, computer security has matured substantially in the past couple of decades, and considerable expertise now exists on how to protect software, computers, and networks. In such a national scheme, safeguards such as firewalls, intrusion detection systems, antivirus software, passwords, scanners, audit trails, and encryption would be directly embedded into infrastructure, just as they are currently in small-scale environments. These national security systems would be connected to a centralized threat management system, and incident response would follow a familiar sort of enterprise process. Furthermore, to ensure security policy compliance, one would expect the usual programs of end-user awareness, security training, and third-party audit to be directed toward the people building and operating national infrastructure. Virtually every national infrastructure protection initiative proposed to date has followed this seemingly straightforward path.[3]

While well-known computer security techniques will certainly be useful for national infrastructure, most practical experience to date suggests that this conventional approach will not be sufficient. A primary reason is the size, scale, and scope inherent in complex national infrastructure. For example, where an enterprise might involve manageably sized assets, national infrastructure will require unusually powerful computing support with the ability to handle enormous volumes of data. Such volumes will easily exceed the storage and processing capacity of typical enterprise security tools such as a commercial threat management system. Unfortunately, this incompatibility conflicts with current initiatives in government and industry to reduce costs through the use of common commercial off-the-shelf products.

National infrastructure databases far exceed the size of even the largest commercial databases.

In addition, whereas enterprise systems can rely on manual intervention by a local expert during a

security disaster, large-scale national infrastructure generally requires a carefully orchestrated response by teams of security experts using predetermined processes. These teams of experts will often work in different groups, organizations, or even countries. In the worst cases, they will cooperate only if forced by government, often sharing just the minimum amount of information to avoid legal consequences. An additional problem is that the complexity associated with national infrastructure leads to the bizarre situation where response teams often have partial or incorrect understanding about how the underlying systems work. For these reasons, seemingly convenient attempts to apply existing small-scale security processes to large-scale infrastructure attacks will ultimately fail (see Figure 1.2).

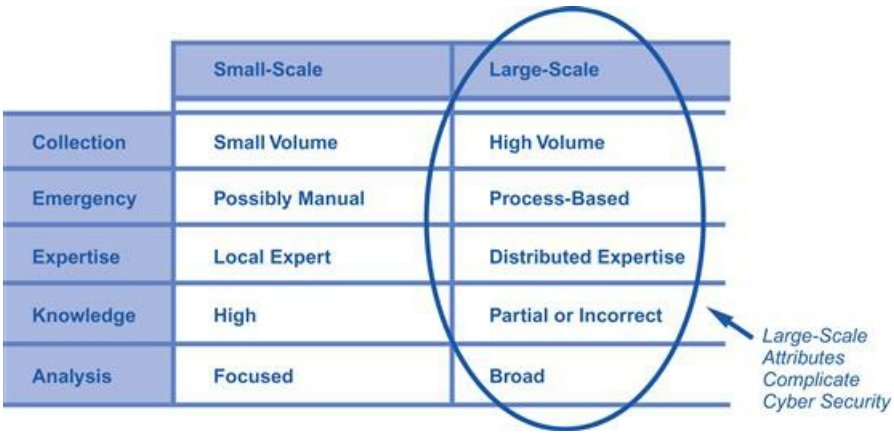


Figure 1.2 Differences between small- and large-scale cyber security.

As a result, a brand-new type of national infrastructure protection methodology is required—one that combines the best elements of existing computer and network security techniques with the unique and difficult challenges associated with complex, large-scale national services. This book offers just such a protection methodology for national infrastructure. It is based on a quarter century of practical experience designing, building, and operating cyber security systems for government, commercial, and consumer infrastructure. It is represented as a series of protection principles that can be applied to new or existing systems. Because of the unique needs of national infrastructure, especially its massive size, scale, and scope, some aspects of the methodology will be unfamiliar to the computer security community. In fact, certain elements of the approach, such as our favorable view of "security through obscurity," might appear in direct conflict with conventional views of how computers and networks should be protected.

National Cyber Threats, Vulnerabilities, and Attacks

Conventional computer security is based on the oft-repeated taxonomy of security threats which includes confidentiality, integrity, availability, and theft. In the broadest sense, all four diverse threat types will have applicability in national infrastructure. For example, protections are required equally to deal with sensitive information leaks (confidentiality), worms affecting the operation of some critical application (integrity), botnets knocking out an important system (availability), or citizens having their identities compromised (theft). Certainly, the availability threat to national services must be viewed as particularly important, given the nature of the threat and its relation to national assets. One should thus expect particular attention to availability threats to national infrastructure. Nevertheless, it makes sense to acknowledge that all four types of security threats in the conventional taxonomy of computer security must be addressed in any national infrastructure protection methodology.

Any of the most common security concern—confidentiality, integrity, availability, and theft—threaten our national infrastructure.

Vulnerabilities are more difficult to associate with any taxonomy. Obviously, national infrastructure must address well-known problems such as improperly configured equipment, poorly designed local area networks, unpatched system software, exploitable bugs in application code, and locally disgruntled employees. The problem is that the most fundamental vulnerability in national infrastructure involves the staggering complexity inherent in the underlying systems. This complexity is so pervasive that many times security incidents uncover aspects of computing functionality that were previously unknown to anyone, including sometimes the system designers. Furthermore, in certain cases, the optimal security solution involves simplifying and cleaning up poorly conceived infrastructure. This is bad news, because most large organizations are inept at simplifying much of anything.

The best one can do for a comprehensive view of the vulnerabilities associated with national infrastructure is to address their relative exploitation points. This can be done with an abstract national infrastructure cyber security model that includes three types of malicious adversaries: *external adversary* (hackers on the Internet), *internal adversary* (trusted insiders), and *supplier adversary* (vendors and partners). Using this model, three exploitation points emerge for national infrastructure: *remote access* (Internet and telework), *system administration and normal usage* (management and use of software, computers, and networks), and *supply chain* (procurement and outsourcing) (see Figure 1.3).

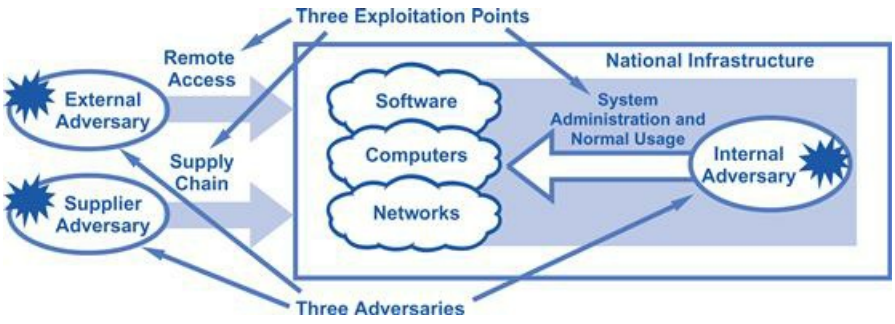


Figure 1.3 Adversaries and exploitation points in national infrastructure.

These three exploitation points and three types of adversaries can be associated with a variety of possible motivations for initiating either a full or test attack on national infrastructure.

Five Possible Motivations for an Infrastructure Attack

- *Country-sponsored warfare*—National infrastructure attacks sponsored and funded by enemy countries must be considered the most significant potential motivation, because the intensity of adversary capability and willingness to attack is potentially unlimited.
- *Terrorist attack*—The terrorist motivation is also signifi cant, especially because groups driven by terror can easily obtain sufficient capability and funding to perform significant attacks on infrastructure.
- *Commercially motivated attack*—When one company chooses to utilize cyber attacks to gain a commercial advantage, it becomes a national infrastructure incident if the target company is a purveyor of some national asset.
- *Financially driven criminal attack*—Identify theft is the most common example of a fi nancially driven attack by criminal groups, but other cases exist, such as companies being extorted to avoid a cyber incident.
- *Hacking*—One must not forget that many types of attacks are still driven by the motivation of hackers, who are often just mischievous youths trying to learn or to build a reputation within the hacking community. This is much less a sinister motivation, and national leaders should try to identify better ways to tap this boundless capability and energy.

Each of the three exploitation points might be utilized in a cyber attack on national infrastructure. For example, a supplier might use a poorly designed supply chain to insert Trojan horse code into a software component that controls some national asset, or a hacker on the Internet might take advantage of some unprotected Internet access point to break into a vulnerable service. Similarly, an insider might use trusted access for either system administration or normal system usage to create an attack. The potential also exists for an external adversary to gain valuable insider access through patient, measured means, such as gaining employment in an infrastructure-supporting organization and then becoming trusted through a long process of work performance. In each case, the possibility exists that a limited type of engagement might be performed as part of a planned test or exercise. This seems especially likely if the attack is country or terrorist sponsored, because it is consistent with past practice.

When to issue a vulnerability risk advisory and when to keep the risk confidential must be determined on a case-by-case basis, depending on the threat.

At each exploitation point, the vulnerability being used might be a well-known problem previously reported in an authoritative public advisory, or it could be a proprietary issue kept hidden by a local organization. It is entirely appropriate for a recognized authority to make a detailed public vulnerability advisory if the benefits

of notifying the good guys outweigh the risks of alerting the bad guys. This cost–benefit result usually occurs when many organizations can directly benefit from the information and can thus take immediate action. When the reported vulnerability is unique and isolated, however, then reporting the details might be irresponsible, especially if the notification process does not enable a more timely fix. This is a key issue, because many government authorities continue to consider new rules for mandatory reporting. If the information being demanded is not properly protected, then the reporting process might result in more harm than good.

Botnet Threat

Perhaps the most insidious type of attack that exists today is the *botnet*.[4] In short, a botnet involves remote control of a collection of compromised end-user machines, usually broadband-connected PCs. The controlled end-user machines, which are referred to as *bots*, are programmed to attack some target that is designated by the botnet controller. The attack is tough to stop because end-user machines are typically administered in an ineffective manner. Furthermore, once the attack begins, it occurs from sources potentially scattered across geographic, political, and service provider boundaries. Perhaps worse, bots are programmed to take commands from multiple controller systems, so any attempts to destroy a given controller result in the bots simply homing to another one.

The Five Entities That Comprise a Botnet Attack

- *Botnet operator*—This is the individual, group, or country that creates the botnet, including its setup and operation.When the botnet is used for financial gain, it is the operator who will benefit. Law enforcement and cyber security initiatives have found it very difficult to identify the operators. The press, in particular, has done a poor job reporting on the presumed identity of botnet operators, often suggesting sponsorship by some country when little supporting evidence exists.
- *Botnet controller*—This is the set of servers that command and control the operation of a botnet. Usually these servers have been maliciously compromised for this purpose. Many times, the real owner of a server that has been compromised will not even realize what has occurred. The type of activity directed by a controller includes all recruitment, setup, communication, and attack activity. Typical botnets include a handful of controllers, usually distributed across the globe in a non-obvious manner.
- *Collection of bots*—These are the end-user, broadband-connected PCs infected with botnet malware. They are usually owned and operated by normal citizens, who become unwitting and unknowing dupes in a botnet attack. When a botnet includes a concentration of PCs in a given region, observers often incorrectly attribute the attack to that region. The use of smart mobile devices in a botnet will grow as upstream capacity and device processing power increase.
- *Botnet software drop*—Most botnets include servers designed to store software that might be useful for the botnets during their lifecycle. Military personnel might refer to this as an arsenal. Like controllers, botnet software drop points are usually servers compromised for this purpose, often unknown to the normal server operator.
- *Botnet target*—This is the location that is targeted in the attack. Usually, it is a website, but it can really be any device, system, or network that is visible to the bots. In most cases, botnets target prominent and often controversial websites, simply because they are visible via the Internet and generally have a great deal at stake in terms of their availability. This increases gain and leverage for the attacker. Logically, however, botnets can target anything visible.

The way a botnet works is that the controller is set up to communicate with the bots via some designated protocol, most often Internet Relay Chat (IRC). This is done via malware inserted into the end-user PCs that

comprise the bots. A great challenge in this regard is that home PCs and laptops are so poorly administered. Amazingly, over time, the day-to-day system and security administration task for home computers has gravitated to the end user. This obligation results in both a poor user experience and general dissatisfaction with the security task. For example, when a typical computer buyer brings a new machine home, it has probably been preloaded with security software by the retailer. From this point onward, however, that home buyer is then tasked with all responsibility for protecting the machine. This includes keeping firewall, intrusion detection, antivirus, and antispam software up to date, as well as ensuring that all software patches are current. When these tasks are not well attended, the result is a more vulnerable machine that is easily turned into a bot. (Sadly, even if a machine is properly managed, expert bot software designers might find a way to install the malware anyway.)

Home PC users may never know they are being used for a botnet scheme.

Once a group of PCs has been compromised into bots, attacks can thus be launched by the controller via a command to the bots, which would then do as they are instructed. This might not occur instantaneously with the infection; in fact, experience suggests that many botnets lay dormant for a great deal of time. Nevertheless, all sorts of attacks are possible in a botnet arrangement, including the now-familiar *distributed denial of service attack* (DDOS). In such a case, the bots create more inbound traffic than the target gateway can handle. For example, if some theoretical gateway allows for 1 Gbps of inbound traffic, and the botnet creates an inbound stream larger than 1 Gbps, then a logjam results at the inbound gateway, and a denial of service condition occurs (see Figure 1.4).

A DDOS attack is like a cyber traffic jam.



Figure 1.4 Sample DDOS attack from a botnet.

Any serious present study of cyber security must acknowledge the unique threat posed by botnets. Virtually any Internet-connected system is vulnerable to major outages from a botnet-originated DDOS attack. The physics of the situation are especially depressing; that is, a botnet that might steal 500 Kbps of upstream capacity from each bot (which would generally allow for concurrent normal computing and networking) would only need three bots to collapse a target T1 connection. Following this logic, only 16,000 bots would be required theoretically to fill up a 10-Gbps connection. Because most of the thousands of

botnets that have been observed on the Internet are at least this size, the threat is obvious; however, many recent and prominent botnets such as Storm and Conficker are much larger, comprising as many as several million bots, so the threat to national infrastructure is severe and immediate.

National Cyber Security Methodology Components

Our proposed methodology for protecting national infrastructure is presented as a series of ten basic design and operation principles. The implication is that, by using these principles as a guide for either improving existing infrastructure components or building new ones, the security result will be desirable, including a reduced risk from botnets. The methodology addresses all four types of security threats to national infrastructure; it also deals with all three types of adversaries to national infrastructure, as well as the three exploitation points detailed in the infrastructure model. The list of principles in the methodology serves as a guide to the remainder of this chapter, as well as an outline for the remaining chapters of the book:

- *Chapter 2: Deception*—The openly advertised use of deception creates uncertainty for adversaries because they will not know if a discovered problem is real or a trap. The more common hidden use of deception allows for real-time behavioral analysis if an intruder is caught in a trap. Programs of national infrastructure protection must include the appropriate use of deception, especially to reduce the malicious partner and supplier risk.
- *Chapter 3: Separation*—Network separation is currently accomplished using firewalls, but programs of national infrastructure protection will require three specific changes. Specifically, national infrastructure must include network-based firewalls on high-capacity backbones to throttle DDOS attacks, internal firewalls to segregate infrastructure and reduce the risk of sabotage, and better tailoring of firewall features for specific applications such as SCADA protocols.[5]
- *Chapter 4: Diversity*—Maintaining diversity in the products, services, and technologies supporting national infrastructure reduces the chances that one common weakness can be exploited to produce a cascading attack. A massive program of coordinated procurement and supplier management is required to achieve a desired level of national diversity across all assets. This will be tough, because it conflicts with most cost-motivated information technology procurement initiatives designed to minimize diversity in infrastructure.
- *Chapter 5: Commonality*—The consistent use of security best practices in the administration of national infrastructure ensures that no infrastructure component is either poorly managed or left completely unguarded. National programs of standards selection and audit validation, especially with an emphasis on uniform programs of simplification, are thus required. This can certainly include citizen end users, but one should never rely on high levels of security compliance in the broad population.
- *Chapter 6: Depth*—The use of defense in depth in national infrastructure ensures that no critical asset is reliant on a single security layer; thus, if any layer should fail, an additional layer is always present to mitigate an attack. Analysis is required at the national level to ensure that all critical assets are protected by at least two layers, preferably more.
- *Chapter 7: Discretion*—The use of personal discretion in the sharing of information about national assets is a practical technique that many computer security experts find difficult to accept because it conflicts with popular views on "security through obscurity." Nevertheless, large-scale infrastructure protection cannot be done properly unless a national culture of discretion and secrecy is nurtured. It

goes without saying that such discretion should never be put in place to obscure illegal or unethical practices.

- *Chapter 8: Collection*—The collection of audit log information is a necessary component of an infrastructure security scheme, but it introduces privacy, size, and scale issues not seen in smaller computer and network settings. National infrastructure protection will require a data collection approach that is acceptable to the citizenry and provides the requisite level of detail for security analysis.
- *Chapter 9: Correlation*—Correlation is the most fundamental of all analysis techniques for cyber security, but modern attack methods such as botnets greatly complicate its use for attack-related indicators. National-level correlation must be performed using all available sources and the best available technology and algorithms. Correlating information around a botnet attack is one of the more challenging present tasks in cyber security.
- *Chapter 10: Awareness*—Maintaining situational awareness is more important in large-scale infrastructure protection than in traditional computer and network security because it helps to coordinate the real-time aspect of multiple infrastructure components. A program of national situational awareness must be in place to ensure proper management decision-making for national assets.
- *Chapter 11: Response*—Incident response for national infrastructure protection is especially difficult because it generally involves complex dependencies and interactions between disparate government and commercial groups. It is best accomplished at the national level when it focuses on early indications, rather than on incidents that have already begun to damage national assets.

The balance of this chapter will introduce each principle, with discussion on its current use in computer and network security, as well as its expected benefits for national infrastructure protection.

Deception

The principle of *deception* involves the deliberate introduction of misleading functionality or misinformation into national infrastructure for the purpose of tricking an adversary. The idea is that an adversary would be presented with a view of national infrastructure functionality that might include services or interface components that are present for the sole purpose of fakery. Computer scientists refer to this functionality as a *honey pot*, but the use of deception for national infrastructure could go far beyond this conventional view. Specifically, deception can be used to protect against certain types of cyber attacks that no other security method will handle. Law enforcement agencies have been using deception effectively for many years, often catching cyber stalkers and criminals by spoofing the reported identity of an end point. Even in the presence of such obvious success, however, the cyber security community has yet to embrace deception as a mainstream protection measure.

Deception is an oft-used tool by law enforcement agencies to catch cyber stalkers and predators.

Deception in computing typically involves a layer of cleverly designed trap functionality strategically embedded into the internal and external interfaces for services. Stated more simply, deception involves fake functionality embedded into real interfaces. An example might be a deliberately planted trap link on a website that would lead potential intruders into an environment designed to highlight adversary behavior. When the deception is open and not secret, it might introduce uncertainty for adversaries in the exploitation of real vulnerabilities, because the adversary might suspect that the discovered entry point is a trap. When it is hidden and stealth, which is the more common situation, it serves as the basis for real-time forensic analysis of adversary behavior. In either case, the result is a public interface that includes real services, deliberate honey pot traps, and the inevitable exploitable vulnerabilities that unfortunately will be present in all nontrivial interfaces (see Figure 1.5).

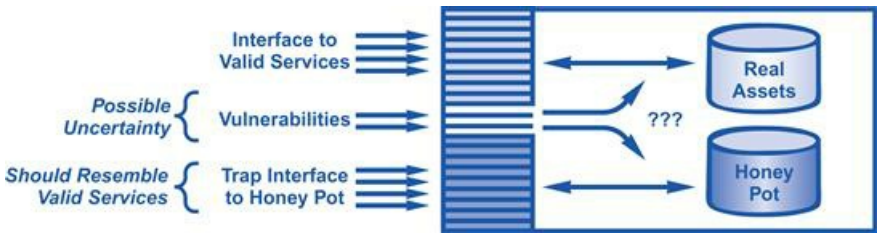


Figure 1.5 Components of an interface with deception.

Only relatively minor tests of honey pot technology have been reported to date, usually in the context of a research effort. Almost no reports are available on the day-to-day use of deception as a structural component of a real enterprise security program. In fact, the vast majority of security programs for companies, government agencies, and national infrastructure would include no such functionality. Academic computer scientists have shown little interest in this type of security, as evidenced by the relatively thin body of literature on the subject. This lack of interest might stem from the discomfort associated with using computing to

mislead. Another explanation might be the relative ineffectiveness of deception against the botnet threat, which is clearly the most important security issue on the Internet today. Regardless of the cause, this tendency to avoid the use of deception is unfortunate, because many cyber attacks, such as subtle break-ins by trusted insiders and Trojan horses being maliciously inserted by suppliers into delivered software, cannot be easily remedied by any other means.

Deception is less effective against botnets than other types of attack methods.

The most direct benefit of deception is that it enables forensic analysis of intruder activity. By using a honey pot, unique insights into attack methods can be gained by watching what is occurring in real time. Such deception obviously works best in a hidden, stealth mode, unknown to the intruder, because if the intruder realizes that some vulnerable exploitation point is a fake, then no exploitation will occur. Honey pot pioneers Cliff Stoll, Bill Cheswick, and Lance Spitzner have provided a majority of the reported experience in real-time forensics using honey pots. They have all suggested that the most difficult task involves creating believability in the trap. It is worth noting that connecting a honey pot to real assets is a terrible idea.

Do not connect honey pots to real assets!

An additional potential benefit of deception is that it can introduce the clever idea that some discovered vulnerability might instead be a deliberately placed trap. Obviously, such an approach is only effective if the use of deception is not hidden; that is, the adversary must know that deception is an approved and accepted technique used for protection. It should therefore be obvious that the major advantage here is that an accidental vulnerability, one that might previously have been an open door for an intruder, will suddenly look like a possible trap. A further profound notion, perhaps for open discussion, is whether just the *implied statement* that deception might be present (perhaps without real justification) would actually reduce risk. Suppliers, for example, might be less willing to take the risk of Trojan horse insertion if the procuring organization advertises an open research and development program of detailed software test and inspection against this type of attack.

Separation

The principle of *separation* involves enforcement of access policy restrictions on the users and resources in a computing environment. Access policy restrictions result in separation domains, which are arguably the most common security architectural concept in use today. This is good news, because the creation of access-policy-based separation domains will be essential in the protection of national infrastructure. Most companies today will typically use firewalls to create perimeters around their presumed enterprise, and access decisions are embedded in the associated rules sets. This use of enterprise firewalls for separation is complemented by several other common access techniques:

- *Authentication and identity management*—These methods are used to validate and manage the identities on which separation decisions are made. They are essential in every enterprise but cannot be relied upon solely for infrastructure security. Malicious insiders, for example, will be authorized under such systems. In addition, external attacks such as DDOS are unaffected by authentication and identity management.
- *Logical access controls*—The access controls inherent in operating systems and applications provide some degree of separation, but they are also weak in the presence of compromised insiders. Furthermore, underlying vulnerabilities in applications and operating systems can often be used to subvert these methods.
- *LAN controls*—Access control lists on local area network (LAN) components can provide separation based on information such as Internet Protocol (IP) or media access control (MAC) address. In this regard, they are very much like firewalls but typically do not extend their scope beyond an isolated segment.
- *Firewalls*—For large-scale infrastructure, firewalls are particularly useful, because they separate one network from another. Today, every Internet-based connection is almost certainly protected by some sort of firewall functionality. This approach worked especially well in the early years of the Internet, when the number of Internet connections to the enterprise was small. Firewalls do remain useful, however, even with the massive connectivity of most groups to the Internet. As a result, national infrastructure should continue to include the use of firewalls to protect known perimeter gateways to the Internet.

Given the massive scale and complexity associated with national infrastructure, three specific separation enhancements are required, and all are extensions of the firewall concept.

Required Separation Enhancements for National Infrastructure Protection

1. The use of network-based firewalls is absolutely required for many national infrastructure applications, especially ones vulnerable to DDOS attacks from the Internet. This use of network-based mediation can take advantage of high-capacity network backbones if the service provider is involved in running the firewalls.
2. The use of firewalls to segregate and isolate internal infrastructure components from one another is a mandatory technique for simplifying the implementation of access control policies in an organization.

When insiders have malicious intent, any exploit they might attempt should be explicitly contained by internal firewalls.

3. The use of commercial off-the-shelf firewalls, especially for SCADA usage, will require tailoring of the firewall to the unique protocol needs of the application. It is not acceptable for national infrastructure protection to retrofi t the use of a generic, commercial, off-the-shelf tool that is not optimized for its specific use (see Figure 1.6)

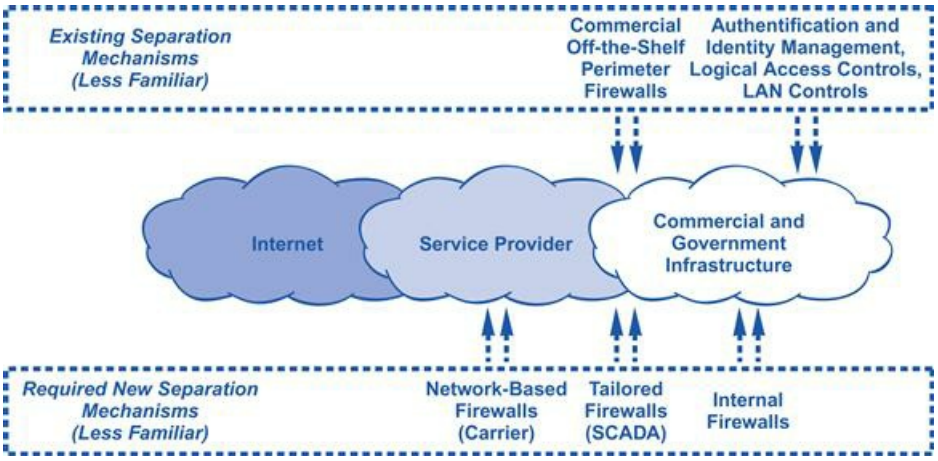


Figure 1.6 Firewall enhancements for national infrastructure.

With the advent of cloud computing, many enterprise and government agency security managers have come to acknowledge the benefits of network-based firewall processing. The approach scales well and helps to deal with the uncontrolled complexity one typically finds in national infrastructure. That said, the reality is that most national assets are still secured by placing a firewall at each of the hundreds or thousands of presumed choke points. This approach does not scale and leads to a false sense of security. It should also be recognized that the firewall is not the only device subjected to such scale problems. Intrusion detection systems, antivirus filtering, threat management, and denial of service filtering also require a network-based approach to function properly in national infrastructure.

An additional problem that exists in current national infrastructure is the relative lack of architectural separation used in an internal, trusted network. Most security engineers know that large systems are best protected by dividing them into smaller systems. Firewalls or packet filtering routers can be used to segregate an enterprise network into manageable domains. Unfortunately, the current state of the practice in infrastructure protection rarely includes a disciplined approach to separating internal assets. This is unfortunate, because it allows an intruder in one domain to have access to a more expansive view of the organizational infrastructure. The threat increases when the firewall has not been optimized for applications such as SCADA that require specialized protocol support.

Parceling a network into manageable smaller domains creates an environment that is easier to protect.

Diversity

The principle of *diversity* involves the selection and use of technology and systems that are intentionally different in substantive ways. These differences can include technology source, programming language, computing platform, physical location, and product vendor. For national infrastructure, realizing such diversity requires a coordinated program of procurement to ensure a proper mix of technologies and vendors. The purpose of introducing these differences is to deliberately create a measure of non-interoperability so that an attack cannot easily cascade from one component to another through exploitation of some common vulnerability. Certainly, it would be possible, even in a diverse environment, for an exploit to cascade, but the likelihood is reduced as the diversity profile increases.

This concept is somewhat controversial, because so much of computer science theory and information technology practice in the past couple of decades has been focused on maximizing interoperability of technologies. This might help explain the relative lack of attentiveness that diversity considerations receive in these fields. By way of analogy, however, cyber attacks on national infrastructure are mitigated by diversity technology just as disease propagation is reduced by a diverse biological ecosystem. That is, a problem that originates in one area of infrastructure with the intention of automatic propagation will only succeed in the presence of some degree of interoperability. If the technologies are sufficiently diverse, then the attack propagation will be reduced or even stopped. As such, national asset managers are obliged to consider means for introducing diversity in a cost-effective manner to realize its security benefits (see Figure 1.7).

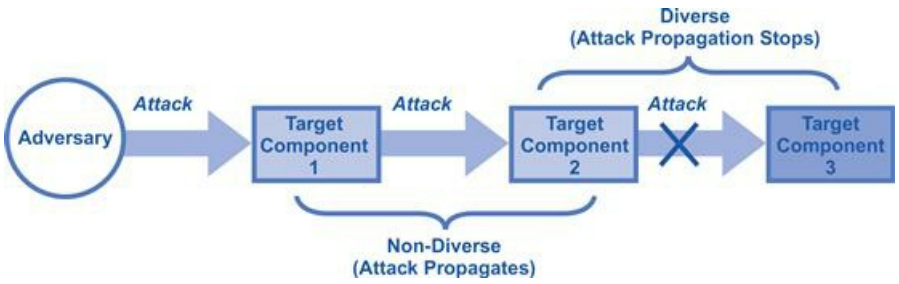


Figure 1.7 Introducing diversity to national infrastructure.

Diversity is especially tough to implement in national infrastructure for several reasons. First, it must be acknowledged that a single, major software vendor tends to currently dominate the personal computer (PC) operating system business landscape in most government and enterprise settings. This is not likely to change, so national infrastructure security initiatives must simply accept an ecosystem lacking in diversity in the PC landscape. The profile for operating system software on computer servers is slightly better from a diversity perspective, but the choices remain limited to a very small number of available sources. Mobile operating systems currently offer considerable diversity, but one cannot help but expect to see a trend toward greater consolidation.

Second, diversity conflicts with the often-found organizational goal of simplifying supplier and vendor relationships; that is, when a common technology is used throughout an organization, day-to-day maintenance, administration, and training costs are minimized. Furthermore, by purchasing in bulk, better

terms are often available from a vendor. In contrast, the use of diversity could result in a reduction in the level of service provided in an organization. For example, suppose that an Internet service provider offers particularly secure and reliable network services to an organization. Perhaps the reliability is even measured to some impressive quantitative availability metric. If the organization is committed to diversity, then one might be forced to actually introduce a second provider with lower levels of reliability.

Enforcing diversity of products and services might seem counterintuitive if you have a reliable provider.

In spite of these drawbacks, diversity carries benefits that are indisputable for large-scale infrastructure. One of the great challenges in national infrastructure protection will thus involve finding ways to diversify technology products and services without increasing costs and losing business leverage with vendors.

Consistency

The principle of *consistency* involves uniform attention to security best practices across national infrastructure components. Determining which best practices are relevant for which national asset requires a combination of local knowledge about the asset, as well as broader knowledge of security vulnerabilities in generic infrastructure protection. Thus, the most mature approach to consistency will combine compliance with relevant standards such as the Sarbanes–Oxley controls in the United States, with locally derived security policies that are tailored to the organizational mission. This implies that every organization charged with the design or operation of national infrastructure must have a local security policy. Amazingly, some large groups do not have such a policy today.

The types of best practices that are likely to be relevant for national infrastructure include well-defined software lifecycle methodologies, timely processes for patching software and systems, segregation of duty controls in system administration, threat management of all collected security information, security awareness training for all system administrators, operational configurations for infrastructure management, and use of software security tools to ensure proper integrity management. Most security experts agree on which best practices to include in a generic set of security requirements, as evidenced by the inclusion of a common core set of practices in every security standard. Attentiveness to consistency is thus one of the less controversial of our recommended principles.

The greatest challenge in implementing best practice consistency across infrastructure involves auditing. The typical audit process is performed by an independent third-party entity doing an analysis of target infrastructure to determine consistency with a desired standard. The result of the audit is usually a numeric score, which is then reported widely and used for management decisions. In the United States, agencies of the federal government are audited against a cyber security standard known as FISMA (Federal Information Security Management Act). While auditing does lead to improved best practice coverage, there are often problems. For example, many audits are done poorly, which results in confusion and improper management decisions. In addition, with all the emphasis on numeric ratings, many agencies focus more on their score than on good security practice.

A good audit score is important but should not replace good security practices.

Today, organizations charged with protecting national infrastructure are subjected to several types of security audits. Streamlining these standards would certainly be a good idea, but some additional items for consideration include improving the types of common training provided to security administrators, as well as including past practice in infrastructure protection in common audit standards. The most obvious practical consideration for national infrastructure, however, would be national-level agreement on which standard or standards would be used to determine competence to protect national assets. While this is a straightforward concept, it could be tough to obtain wide concurrence among all national participants. A related issue involves commonality in national infrastructure operational configurations; this reduces the chances that a rogue configuration installed for malicious purposes, perhaps by compromised insiders.

A national standard of competence for protecting our assets is needed.

Depth

The principle of *depth* involves the use of multiple security layers of protection for national infrastructure assets. These layers protect assets from both internal and external attacks via the familiar "defense in depth" approach; that is, multiple layers reduce the risk of attack by increasing the chances that at least one layer will be effective. This should appear to be a somewhat sketchy situation, however, from the perspective of traditional engineering. Civil engineers, for example, would never be comfortable designing a structure with multiple flawed supports in the hopes that one of them will hold the load. Unfortunately, cyber security experts have no choice but to rely on this flawed notion, perhaps highlighting the relative immaturity of security as an engineering discipline.

One hint as to why depth is such an important requirement is that national infrastructure components are currently controlled by software, and everyone knows that the current state of software engineering is abysmal. Compared to other types of engineering, software stands out as the only one that accepts the creation of knowingly flawed products as acceptable. The result is that all nontrivial software has exploitable vulnerabilities, so the idea that one should create multiple layers of security defense is unavoidable. It is worth mentioning that the degree of diversity in these layers will also have a direct impact on their effectiveness (see Figure 1.8).

Software engineering standards do not contain the same level of quality as civil and other engineering standards.

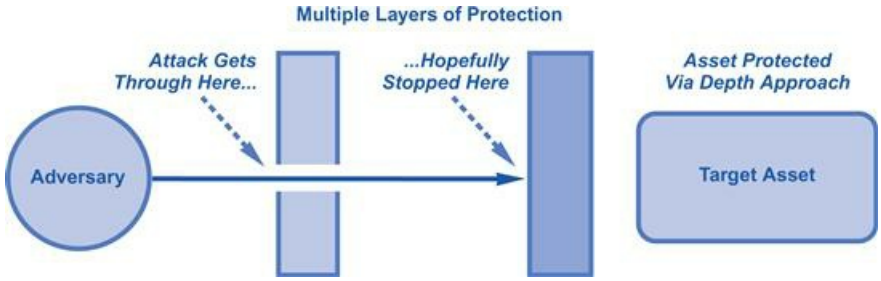


Figure 1.8 National infrastructure security through defense in depth.

To maximize the usefulness of defense layers in national infrastructure, it is recommended that a combination of functional and procedural controls be included. For example, a common first layer of defense is to install an access control mechanism for the admission of devices to the local area network. This could involve router controls in a small network or firewall access rules in an enterprise. In either case, this first line of defense is clearly functional. As such, a good choice for a second layer of defense might involve something procedural, such as the deployment of scanning to determine if inappropriate devices have gotten through the first layer. Such diversity will increase the chances that the cause of failure in one layer is unlikely to cause a similar failure in another layer.

A great complication in national infrastructure protection is that many layers of defense assume the

existence of a defined network perimeter. For example, the presence of many flaws in enterprise security found by auditors is mitigated by the recognition that intruders would have to penetrate the enterprise perimeter to exploit these weaknesses. Unfortunately, for most national assets, finding a perimeter is no longer possible. The assets of a country, for example, are almost impossible to define within some geographic or political boundary, much less a network one. Security managers must therefore be creative in identifying controls that will be meaningful for complex assets whose properties are not always evident. The risk of getting this wrong is that in providing multiple layers of defense, one might misapply the protections and leave some portion of the asset base with no layers in place.

Discretion

The principle of *discretion* involves individuals and groups making good decisions to obscure sensitive information about national infrastructure. This is done by combining formal mandatory information protection programs with informal discretionary behavior. Formal mandatory programs have been in place for many years in the U.S. federal government, where documents are associated with classifications, and policy enforcement is based on clearances granted to individuals. In the most intense environments, such as top-secret compartments in the intelligence community, violations of access policies could be interpreted as espionage, with all of the associated criminal implications. For this reason, prominent breaches of highly classified government information are not common.

Naturally, top-secret information within the intelligence community is at great risk for attack or infiltration.

In commercial settings, formal information protection programs are gaining wider acceptance because of the increased need to protect personally identifiable information (PII) such as credit card numbers. Employees of companies around the world are starting to understand the importance of obscuring certain aspects of corporate activity, and this is healthy for national infrastructure protection. In fact, programs of discretion for national infrastructure protection will require a combination of corporate and government security policy enforcement, perhaps with custom-designed information markings for national assets. The resultant discretionary policy serves as a layer of protection to prevent national infrastructure-related information from reaching individuals who have no need to know such information.

A barrier in our recommended application of discretion is the maligned notion of "security through obscurity." Security experts, especially cryptographers, have long complained that obscurity is an unacceptable protection approach. They correctly reference the problems of trying to secure a system by hiding its underlying detail. Inevitably, an adversary discovers the hidden design secrets and the security protection is lost. For this reason, conventional computer security correctly dictates an open approach to software, design, and algorithms. An advantage of this open approach is the social review that comes with widespread advertisement; for example, the likelihood is low of software ever being correct without a significant amount of intense review by experts. So, the general computer security argument against "security through obscurity" is largely valid in most cases.

"Security through obscurity" may actually leave assets more vulnerable to attack than an open approach would.

Nevertheless, any manager charged with the protection of nontrivial, large-scale infrastructure will tell you that discretion and, yes, obscurity are indispensable components in a protection program. Obscuring details around technology used, software deployed, systems purchased, and configurations managed will help to avoid or at least slow down certain types of attacks. Hackers often claim that by discovering this type of information about a company and then advertising the weaknesses they are actually doing the local security team a favor. They suggest that such advertisement is required to motivate a security team toward a solution, but this is actually nonsense. Programs around proper discretion and obscurity for infrastructure information

are indispensable and must be coordinated at the national level.

Collection

The principle of *collection* involves automated gathering of system-related information about national infrastructure to enable security analysis. Such collection is usually done in real time and involves probes or hooks in applications, system software, network elements, or hardware devices that gather information of interest. The use of audit trails in small-scale computer security is an example of a long-standing collection practice that introduces very little controversy among experts as to its utility. Security devices such as firewalls produce log files, and systems purported to have some degree of security usefulness will also generate an audit trail output. The practice is so common that a new type of product, called a *security information management system* (SIMS), has been developed to process all this data.

The primary operational challenge in setting up the right type of collection process for computers and networks has been twofold: First, decisions must be made about what types of information are to be collected. If this decision is made correctly, then the information collected should correspond to exactly the type of data required for security analysis, and nothing else. Second, decisions must be made about how much information is actually collected. This might involve the use of existing system functions, such as enabling the automatic generation of statistics on a router; or it could involve the introduction of some new type of function that deliberately gathers the desired information. Once these considerations are handled, appropriate mechanisms for collecting data from national infrastructure can be embedded into the security architecture (see Figure 1.9).

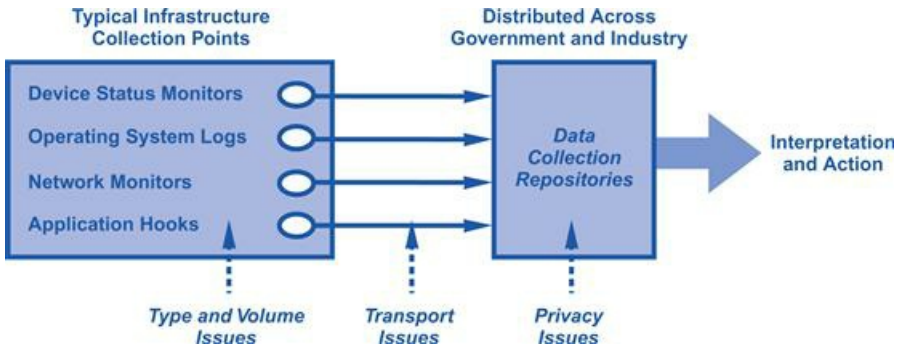


Figure 1.9 Collecting national infrastructure-related security information.

The technical and operational challenges associated with the collection of logs and audit trails are heightened in the protection of national assets. Because national infrastructure is so complex, determining what information should be collected turns out to be a difficult exercise. In particular, the potential arises with large-scale collection to intrude on the privacy of individuals and groups within a nation. As such, any initiative to protect infrastructure through the collection of data must include at least some measure of privacy policy determination. Similarly, the volumes of data collected from large infrastructure can exceed practical limits. Telecommunications collection systems designed to protect the integrity of a service provider backbone, for example, can easily generate many terabytes of data in hours of processing.

What and how much data to collect is an operational challenge.

In both cases, technical and operational expertise must be applied to ensure that the appropriate data is collected in the proper amounts. The good news is that virtually all security protection algorithms require no deep, probing information of the type that might generate privacy or volumetric issues. The challenge arises instead when collection is done without proper advance analysis which often results in the collection of more data than is needed. This can easily lead to privacy problems in some national collection repositories, so planning is particularly necessary. In any event, a national strategy of data collection is required, with the usual sorts of legal and policy guidance on who collects what and under which circumstances. As we suggested above, this exercise must be guided by the requirements for security analysis—and nothing else.

Only collect as much data as is necessary for security purposes.

Correlation

The principle of *correlation* involves a specific type of analysis that can be performed on factors related to national infrastructure protection. The goal of correlation is to identify whether security-related indicators might emerge from the analysis. For example, if some national computing asset begins operating in a sluggish manner, then other factors would be examined for a possible correlative relationship. One could imagine the local and wide area networks being analyzed for traffic that might be of an attack nature. In addition, similar computing assets might be examined to determine if they are experiencing a similar functional problem. Also, all software and services embedded in the national asset might be analyzed for known vulnerabilities. In each case, the purpose of the correlation is to combine and compare factors to help explain a given security issue. This type of comparison-oriented analysis is indispensable for national infrastructure because of its complexity.

Monitoring and analyzing networks and data collection may reveal a hidden or emerging security threat.

Interestingly, almost every major national infrastructure protection initiative attempted to date has included a fusion center for real-time correlation of data. A fusion center is a physical security operations center with means for collecting and analyzing multiple sources of ingress data. It is not uncommon for such a center to include massive display screens with colorful, visualized representations, nor is it uncommon to find such centers in the military with teams of enlisted people performing the manual chores. This is an important point, because, while such automated fusion is certainly promising, best practice in correlation for national infrastructure protection must include the requirement that human judgment be included in the analysis. Thus, regardless of whether resources are centralized into one physical location, the reality is that human beings will need to be included in the processing (see Figure 1.10).

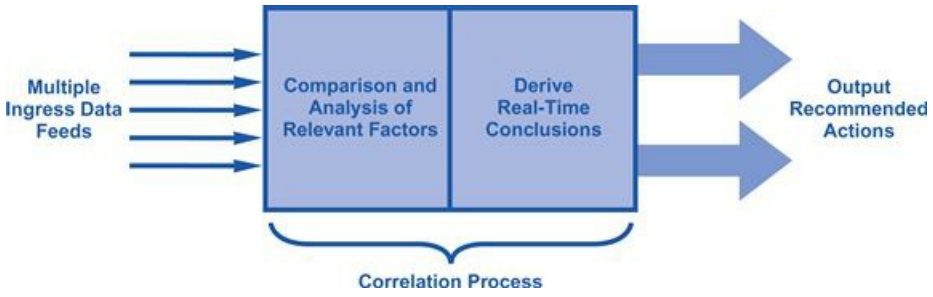


Figure 1.10 National infrastructure high-level correlation approach.

In practice, fusion centers and the associated processes and correlation algorithms have been tough to implement, even in small-scale environments. Botnets, for example, involve the use of source systems that are selected almost arbitrarily. As such, the use of correlation to determine where and why the attack is occurring has been useless. In fact, correlating geographic information with the sources of botnet activity has even led to many false conclusions about who is attacking whom. Countless hours have been spent by security teams poring through botnet information trying to determine the source, and the best one can hope for might be

information about controllers or software drops. In the end, current correlation approaches fall short.

What is needed to improve present correlation capabilities for national infrastructure protection involves multiple steps.

Three Steps to Improve Current CorrelationCapabilities

1. The actual computer science around correlation algorithms needs to be better investigated. Little attention has been placed in academic computer science and applied mathematics departments to multifactor correlation of real-time security data. This could be changed with appropriate funding and grant emphasis from the government.
2. The ability to identify reliable data feeds needs to be greatly improved. Too much attention has been placed on *ad hoc* collection of volunteered feeds, and this complicates the ability for analysis to perform meaningful correlation.
3. The design and operation of a national-level fusion center must be given serious consideration. Some means must be identified for putting aside political and funding problems in order to accomplish this important objective.

Awareness

The principle of *awareness* involves an organization understanding the differences, in real time and at all times, between observed and normal status in national infrastructure. This status can include risks, vulnerabilities, and behavior in the target infrastructure. *Behavior* refers here to the mix of user activity, system processing, network traffic, and computing volumes in the software, computers, and systems that comprise infrastructure. The implication is that the organization can somehow characterize a given situation as being either normal or abnormal. Furthermore, the organization must have the ability to detect and measure differences between these two behavioral states. Correlation analysis is usually inherent in such determinations, but the real challenge is less the algorithms and more the processes that must be in place to ensure situational awareness every hour of every day. For example, if a new vulnerability arises that has impact on the local infrastructure, then this knowledge must be obtained and factored into management decisions immediately.

Awareness builds on collection and correlation, but is not limited to those areas alone.

Managers of national infrastructure generally do not have to be convinced that situational awareness is important. The big issue instead is how to achieve this goal. In practice, real-time awareness requires attentiveness and vigilance rarely found in normal computer security. Data must first be collected and enabled to flow into a fusion center at all times so correlation can take place. The results of the correlation must be used to establish a profiled baseline of behavior so differences can be measured. This sounds easier than it is, because so many odd situations have the ability to mimic normal behavior (when it is really a problem) or a problem (when it really is nothing). Nevertheless, national infrastructure protection demands that managers of assets create a locally relevant means for being able to comment accurately on the state of security at all times. This allows for proper management decisions about security (see Figure 1.11).

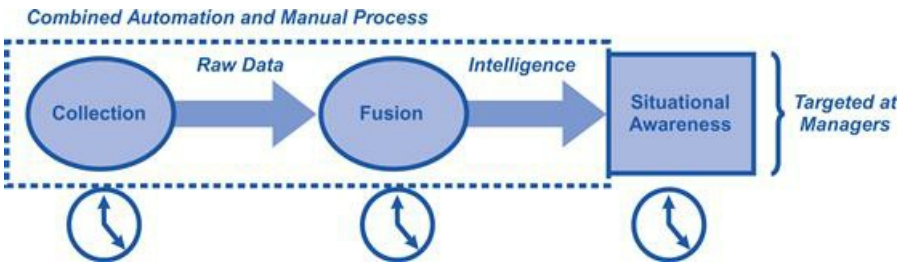


Figure 1.11 Real-time situation awareness process flow.

Interestingly, situational awareness has not been considered a major component of the computer security equation to date. The concept plays no substantive role in small-scale security, such as in a home network, because when the computing base to be protected is simple enough, characterizing real-time situational status is just not necessary. Similarly, when a security manager puts in place security controls for a small enterprise, situational awareness is not the highest priority. Generally, the closest one might expect to some degree of

real-time awareness for a small system might be an occasional review of system log files. So, the transition from small-scale to large-scale infrastructure protection does require a new attentiveness to situational awareness that is not well developed. It is also worth noting that the general notion of "user awareness" of security is also not the principle specified here. While it is helpful for end users to have knowledge of security, any professionally designed program of national infrastructure security must presume that a high percentage of end users will *always* make the wrong sorts of security decisions if allowed. The implication is that national infrastructure protection must never rely on the decision-making of end users through programs of awareness.

Large-scale infrastructure protection requires a higher level of awareness than most groups currently employ.

A further advance that is necessary for situational awareness involves enhancements in approaches to security metrics reporting. Where the non-cyber national intelligence community has done a great job developing means for delivering daily intelligence briefs to senior government officials, the cyber security community has rarely considered this approach. The reality is that, for situation awareness to become a structural component of national infrastructure protection, valid metrics must be developed to accurately portray status, and these must be codified into a suitable type of regular intelligence report that senior officials can use to determine security status. It would not be unreasonable to expect this cyber security intelligence to flow from a central point such as a fusion center, but in general this is not a requirement.

Response

The principle of *response* involves assurance that processes are in place to react to any security-related indicator that becomes available. These indicators should flow into the response process primarily from the situational awareness layer. National infrastructure response should emphasize indicators rather than incidents. In most current computer security applications, the response team waits for serious problems to occur, usually including complaints from users, applications running poorly, and networks operating in a sluggish manner. Once this occurs, the response team springs into action, even though by this time the security game has already been lost. For essential national infrastructure services, the idea of waiting for the service to degrade before responding does not make logical sense.

An additional response-related change for national infrastructure protection is that the maligned concept of "false positive" must be reconsidered. In current small-scale environments, a major goal of the computer security team is to minimize the number of response cases that are initiated only to find that nothing was wrong after all. This is an easy goal to reach by simply waiting for disasters to be confirmed beyond a shadow of a doubt before response is initiated. For national infrastructure, however, this is obviously unacceptable. Instead, response must follow indicators, and the concept of minimizing false positives must not be part of the approach. The only quantitative metric that must be minimized in national-level response is risk (see Figure 1.12).

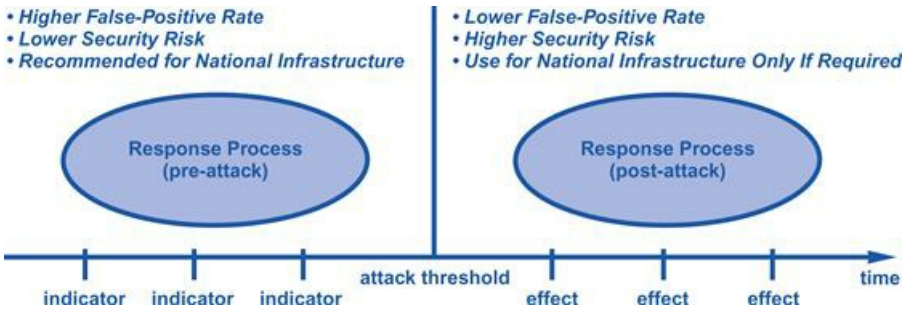


Figure 1.12 National infrastructure security response approach.

A challenge that must be considered in establishing response functions for national asset protection is that relevant indicators often arise long before any harmful effects are seen. This suggests that infrastructure protecting must have accurate situational awareness that considers much more than just visible impacts such as users having trouble, networks being down, or services being unavailable. Instead, often subtle indicators must be analyzed carefully, which is where the challenges arise with false positives. When response teams agree to consider such indicators, it becomes more likely that such indicators are benign. A great secret to proper incident response for national infrastructure is that higher false positive rates might actually be a good sign.

A higher rate of false positives must be tolerated for national infrastructure protection.

It is worth noting that the principles of collection, correlation, awareness, and response are all consistent

with the implementation of a national fusion center. Clearly, response activities are often dependent on a real-time, ubiquitous operations center to coordinate activities, contact key individuals, collect data as it becomes available, and document progress in the response activities. As such, it should not be unexpected that national-level response for cyber security should include some sort of centralized national center. The creation of such a facility should be the centerpiece of any national infrastructure protection program and should involve the active participation of all organizations with responsibility for national services.

Implementing the Principles Nationally

To effectively apply this full set of security principles in practice for national infrastructure protection, several practical implementation considerations emerge:

- *Commissions and groups*—Numerous commissions and groups have been created over the years with the purpose of national infrastructure protection. Most have had some minor positive impact on infrastructure security, but none has had sufficient impact to reduce present national risk to acceptable levels. An observation here is that many of these commissions and groups have become the *end* rather than the *means* toward a cyber security solution. When this occurs, their likelihood of success diminishes considerably. Future commissions and groups should take this into consideration.
- *Information sharing*—Too much attention is placed on information sharing between government and industry, perhaps because information sharing would seem on the surface to carry much benefit to both parties. The advice here is that a comprehensive information sharing program is not easy to implement simply because organizations prefer to maintain a low profile when fighting a vulnerability or attack. In addition, the presumption that some organization—government or commercial—might have some nugget of information that could solve a cyber attack or reduce risk is not generally consistent with practice. Thus, the motivation for a commercial entity to share vulnerability or incident-related information with the government is low; very little value generally comes from such sharing.
- *International cooperation*—National initiatives focused on creating government cyber security legislation must acknowledge that the Internet is global, as are the shared services such as the domain name system (DNS) that all national and global assets are so dependent upon. Thus, any program of national infrastructure protection must include provisions for international cooperation, and such cooperation implies agreements between participants that will be followed as long as everyone perceives benefit.
- *Technical and operational costs*—To implement the principles described above, considerable technical and operational costs will need to be covered across government and commercial environments. While it is tempting to presume that the purveyors of national infrastructure can simply absorb these costs into normal business budgets, this has not been the case in the past. Instead, the emphasis should be on rewards and incentives for organizations that make the decision to implement these principles. This point is critical because it suggests that the best possible use of government funds might be as straightforward as helping to directly fund initiatives that will help to secure national assets.

The bulk of our discussion in the ensuing chapters is technical in nature; that is, programmatic and political issues are conveniently ignored. This does not diminish their importance, but rather is driven by our decision to separate our concerns and focus in this book on the details of "what" must be done, rather than "how."

Finally, let's look at how the ever-changing policy of the United States helps prevent or minimize disruptions to the critical national infrastructure. The implementation of the policy is crucial in order to protect the public, the economy, government services, and the national security of the United States.

Protecting the Critical National Infrastructure Against Cyber Attacks

Information technology has grown to provide both government and the private sector with an efficient and timely means of delivering essential services around the world. As a result, these critical systems remain at risk from potential attacks via the Internet. It is the policy of the United States to prevent or minimize disruptions to the critical information infrastructure in order to protect the public, the economy, government services, and the national security of the United States.

The federal government is continually increasing capabilities to address cyber risk associated with critical networks and information systems. On January 8, 2008, President Bush approved the National Security Presidential Directive 54/Homeland Security Presidential Directive 23, which formalized a series of continuous efforts designed to further safeguard federal government systems and reduce potential vulnerabilities, protect against intrusion attempts, and better anticipate future threats.

While efforts to protect the federal network systems from cyber attacks remain a collaborative, government-wide effort, the Department of Homeland Security (DHS) has the lead responsibility for ensuring the security, resiliency, and reliability of the nation's information technology (IT) and communications infrastructure (see "An Agenda for Action in Preventing Cyber Attacks Methods" below).

An Agenda for Action in Preventing Cyber Attacks Methods

When completing the Preventing Cyber Attacks Methods checklist, the DHS specialist should adhere to the provisional list of actions for some of the principal cyber attack prevention methods. The order is not significant; however, these are the activities for which the research would want to provide a detailed description of procedures, review, and assessment for ease of use and admissibility. Current measures that must be adhered to in order to prevent future attacks and intrusion attempts, include (check all tasks completed):

1. Hiring additional personnel to support the U.S. Computer Emergency Readiness Team (US-CERT), DHS' 24×7 watch and warning center for the federal government's Internet infrastructure. US-CERT, a public–private partnership, operates round the clock to help government and industry analyze and respond to cyber threats and vulnerabilities.
2. Expanding the Einstein Cyber Shield to all federal departments and agencies. This will provide government officials with an early warning system to gain better situational awareness, earlier identification of malicious activity, and a more comprehensive network defense. The current version of the program is widely seen as providing meager protection against attack, but a new version being built will be more robust—largely because it is rooted in NSA technology. The program is designed to look for indicators of cyber attacks by digging into all Internet communications, including the contents of e-mails, according to a declassified summary.
3. Consolidating the number of external connections including Internet points of presence for the federal government Internet infrastructure (FGII), as part of the Office of Management and Budget's (OMB's) Trusted Internet Connections Initiative (TICI). TICI will more efficiently manage and implement security measures to help bring more comprehensive protection across the federal .gov

domains.

4. Creating a National Cyber Security Center (NCSC) to further progress in addressing cyber threats and increasing cyber security efforts. The NCSC will bring together federal cyber security organizations by virtually connecting and, in some cases, physically collocating personnel and resources to gain a clearer understanding of the overall cyber security picture of federal networks.
5. Expanding the National Cyber Investigative Joint Task Force (NCIJTF) to include representation from the U.S. Secret Service and several other federal agencies. This existing cyber investigation coordination organization overseen by the Federal Bureau of Investigation (FBI) will serve as a multiagency national focal point for coordinating, integrating, and sharing pertinent information related to cyber threat investigations.
6. Reducing the potential for adversaries to manipulate IT and communications products before they are imported into the United States. In other words, the DHS specialist must work toward a stronger supply chain defense. To address this challenge, the federal government is exploring protections into the federal acquisition process and developing a multifaceted strategy to reduce risk at the most appropriate stage of the IT and communications product lifecycle.
7. Facilitating coordination and information sharing between the federal government and private sector to reduce cyber risk, disseminate threat information, share best practices, and apply appropriate protective actions as outlined within the National Infrastructure Protection Plan (NIPP) framework. For example, DHS created the Control Systems Vulnerability Assessment Tool (CSVAT) to help all critical infrastructure sectors assess certain policies, plans, and procedures currently in place to reduce cyber vulnerabilities and leverage recognized standards.
8. Leading the nation's largest cyber security exercise, known as Cyber Storm III, in the fall of 2010, bringing together participants from federal, state, and local governments; the private sector; and the international community in order to examine and strengthen the nation's cyber security preparedness and response capabilities in response to a simulated cyber attack across several critical sectors of this nation's economy. Cyber Storm III was built upon the success of previous exercises; however, enhancements in the nation's cyber security capabilities, an ever-evolving cyber threat landscape and the increased emphasis and extent of public–private collaboration and cooperation made Cyber Storm III unique. Cyber Storm III was the primary vehicle to exercise the newly developed National Cyber Incident Response Plan (NCIRP)—a blueprint for cyber security incident response—to examine the roles, responsibilities, authorities, and other key elements of the nation's cyber incident response and management capabilities and use those findings to refine the plan. Cyber Storm III (and the upcoming Cyber Storm IV in 2012) and other exercises help ensure that public and private sectors are prepared for an effective response to attacks against this nation's critical systems and networks.
9. Partnering with academia and industry to expand cyber education for all U.S. government employees, particularly those who specialize in IT, and enhance worksite development and recruitment strategies to ensure a knowledgeable workforce capable of dealing with the evolving nature of cyber threats.
10. Increasing funding for IT security through the president's FY 2012 budget for protection efforts against cyber attacks efforts across the federal government and the private sector.

Summary

This chapter discussed how pervasive and sustained cyber attacks continue to pose a potentially devastating threat to the systems and operations of the critical national infrastructure of the United States. According to recent testimony by the Director of National Intelligence, "there has been a dramatic increase in malicious cyber activity targeting U.S. computers and networks." In addition, recent reports of cyber attacks and incidents affecting critical infrastructures illustrate the potential impact of such events on national and economic security. The nation's ever-increasing dependence on information systems to carry out essential day-to-day operations makes it vulnerable to an array of cyber-based risks. Thus, it is increasingly important that federal and nonfederal entities carry out concerted efforts to safeguard their systems and the information they contain by looking at:

- Cyber threats to cyber-reliant critical national infrastructures.
- The continuing challenges facing federal agencies in protecting the nation's cyber-reliant critical national infrastructure.

Cyber-based threats to the critical national infrastructure are evolving and growing. These threats can come from a variety of sources, including criminals and foreign nations, as well as hackers and disgruntled employees. These potential cyber attackers have a variety of techniques at their disposal that can vastly expand the reach and impact of their actions. In addition, the interconnectivity between information systems, the Internet, and other infrastructure presents increasing opportunities for such cyber attacks. Consistent with this, reports of security incidents from federal agencies are on the rise according to the Government Accounting Office (GAO), increasing over 760% over the past 6 years. In addition, reports of cyber attacks and information security incidents, affecting federal systems and systems supporting the critical national infrastructure, illustrate the serious impact such incidents can have on national and economic security, including the loss of classified information and intellectual property worth billions of dollars. The Obama administration and executive branch agencies continue to act to better protect the cyber-reliant critical national infrastructures, improve the security of federal systems, and strengthen the nation's cyber security posture, but they are still falling short of their goals. In other words, they have not yet fully implemented key actions that are intended to address threats and improve the current U.S. approach to cyber security, such as:

- Implementing near- and midterm actions recommended by the cyber security policy review directed by the president.
- Updating the national strategy for securing the information and communications infrastructure.
- Developing a comprehensive national strategy for addressing global cyber security and governance.
- Creating a prioritized national and federal research and development agenda for improving cyber security.

Federal systems continue to be afflicted by persistent information security control weaknesses. For example, as part of its audit of the fiscal year 2010 financial statements for the U.S. government, the GAO determined that serious and widespread information security control deficiencies were a government-wide material weakness. Over the past several years, GAO and agency inspectors general have made thousands of

recommendations to agencies for actions necessary to resolve prior significant control deficiencies and information security program shortfalls. The White House, the Office of Management and Budget, and selected federal agencies have undertaken additional government-wide initiatives intended to enhance information security at federal agencies. However, these initiatives face challenges, such as better defining agency roles and responsibilities, establishing measures of effectiveness, and the requirement of sustained attention, which government agencies have begun to provide. As such, the GAO continues to identify the federal government's information systems and the nation's cyber critical national infrastructure as a government-wide high-risk area.

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

Chapter Review Questions/Exercises

True/False

1. True or False? National infrastructure refers to the complex, underlying delivery and support systems for all large-scale services considered absolutely essential to a nation.
2. True or False? Vulnerabilities are more difficult to associate with any taxonomy.
3. True or False? Perhaps the most insidious type of attack that exists today is the botnet.
4. True or False? The principle of deception involves the deliberate introduction of misleading functionality or misinformation into national infrastructure for the purpose of tricking an adversary.
5. True or False? The principle of separation involves enforcement of access policy restrictions on the users and resources in a computing environment.

Multiple Choice

1. The best one can do for a comprehensive view of the vulnerabilities associated with national infrastructure is to address their relative exploitation points. This can be done with an abstract national infrastructure cyber security model that includes three types of malicious adversaries, except which two:

A. External adversary
B. Remote adversary
C. Internal adversary
D. System adversary
E. Supplier adversary

2. By using the abstract national infrastructure cyber security model, three exploitation points emerge for national infrastructure, except which two:

A. Defined methodology
B. Remote access
C. Breach of contract
D. System administration and normal usage
E. Supply chain

3. The selection and use of technology and systems that are intentionally different in substantive ways is called the principle of:

A. Consistency
B. Depth
C. Discretion
D. Collection
E. Diversity

4. The automated gathering of system-related information about national infrastructure to enable security analysis is called the principle of:

A. Correlation

B. Awareness

C. Response

D. Collection

E. Recovery

5. To effectively apply the full set of security principles in practice for national infrastructure protection, several practical implementation considerations emerge, except which one:

A. Commissions and groups

B. Information sharing

C. International cooperation

D. Technical and operational costs

E. Current correlation capabilities

Exercise

Problem

A disgruntled former hospital employee with exceptional computer skills hacks into the hospital network from their home computer and plants a very aggressive computer virus into a Computer-Aided Facility Management (CAFM) system. The computer virus activates at 1:00 a.m., shutting down the Hospital Ventilation Air Conditioning (HVAC) system, security system, building automation, and patient medical monitoring system. Please explain how the hospital's cyber security team (CST) went about resolving the problem.

Hands-On Projects

Project

Trojan Horse e-mails sent from an intruder were targeted at specific organizations and people. The Trojan Horse e-mails, when opened, compromised the system and enabled the cyber attackers to infiltrate the internal networked systems. The cyber attackers then searched the systems and network for data files and exfiltrated information through the encrypted channels. On opening the document, a real document would display, while hidden activities are executed in the background. The possibility of applications crashing is extremely high. The following is an example:

- A reverse shell leveraging port 443 (secure sockets layer [SSL]) downloaded a command and control tools from a dynamic domain. Traffic was not SSL encrypted, but was obfuscated. Obfuscated code is source or machine code that has been made difficult to understand. Programmers may deliberately obfuscate code to conceal its purpose (security through obscurity) or its logic to prevent tampering or

deter reverse engineering, or as a puzzle or recreational challenge for someone reading the source code.

- The intruder then gained access and conducted network scanning, data collection, and data exfiltration (military jargon for the removal of personnel or units from areas under enemy control by stealth, deception, surprise, or clandestine means, the opposite of infiltration).

So, how would your cyber security team go about identifying the intruder, the collection of tools used by the intruder, and recovering from the attack?

Case Projects

Problem

Let's look at a real-world scenario and how the Department of Homeland Security (DHS) plays into it. In the scenario, the United States will be hit by a large-scale, coordinated cyber attack organized by China. These attacks debilitate the functioning of government agencies, parts of the critical infrastructure, and commercial ventures. The IT infrastructure of several agencies are paralyzed, the electric grid in most of the country is shut down, telephone traffic is seriously limited and satellite communications are down (limiting the Department of Defense's [DOD's] ability to communicate with commands overseas). International commerce and financial institutions are also severely hit. Please explain how DHS should handle this situation.

Optional Team Case Project

Problem

A cadre of intruders leveraged their collective capabilities to mount a simulated coordinated cyber attack on a global scale. Although primary motives differed among the entities, a sophisticated network of relationships enabled the intruder to degrade Internet connectivity, disrupt industrial functions, and ultimately erode confidence in everyday communications. The intruder cultivated relationships with unaffiliated opportunistic intruders. Due to their critical nature and perceived vulnerabilities, the intruders specifically targeted several critical infrastructure sectors, along with state and federal agencies, the media, and foreign nations. Please identify the findings that were observed by the participants and observer/controllers through the implementation of this project.

[1] E.W. Dijkstra, *Selected Writings on Computing: A Personal Perspective*, Springer-Verlag, New York, 1982, pp. 212–213.

[2] T. Friedman, *The World Is Flat: A Brief History of the Twenty-First Century*, Farrar, Straus, and Giroux, New York, 2007. (Friedman provides a useful economic backdrop to the global aspect of the cyber attack trends suggested in this chapter.)

[3] Executive Office of the President, *Cyberspace Policy Review: Assuring a Trusted and Resilient Information and Communications Infrastructure*, U.S. White House, Washington, D.C., 2009 (http://handle.dtic.mil/100.2/ADA501541).

[4] Much of the material on botnets in this chapter is derived from work done by Brian Rexroad, David Gross, and several others from AT&T.

[5] R. Kurtz, *Securing SCADA Systems*, Wiley, New York, 2006. (Kurtz provides an excellent overview of SCADA systems and the current state of the practice in securing them.)

2

Deception

Chapter Outline

Scanning Stage
Deliberately Open Ports
Discovery Stage
Deceptive Documents
Exploitation Stage
Procurement Tricks
Exposing Stage
Interfaces Between Humans and Computers
National Deception Program
The Deception Planning Process Against Cyber Attacks
Summary
Chapter Review Questions/Exercises

> *Create a highly controlled network. Within that network, you place production systems and then monitor, capture, and analyze all activity that happens within that network Because this is not a production network, but rather our Honeynet, any traffic is suspicious by nature.*
>
> **The Honeynet Project**[1]

The use of deception in computing involves deliberately misleading an adversary by creating a system component that looks real but is in fact a trap. The system component, sometimes referred to as a *honey pot*, is usually functionality embedded in a computing or networking system, but it can also be a physical asset designed to trick an intruder. In both cases, a common interface is presented to an adversary who might access real functionality connected to real assets, but who might also unknowingly access deceptive functionality connected to bogus assets. In a well-designed deceptive system, the distinction between real and trap functionality should not be apparent to the intruder (see Figure 2.1).

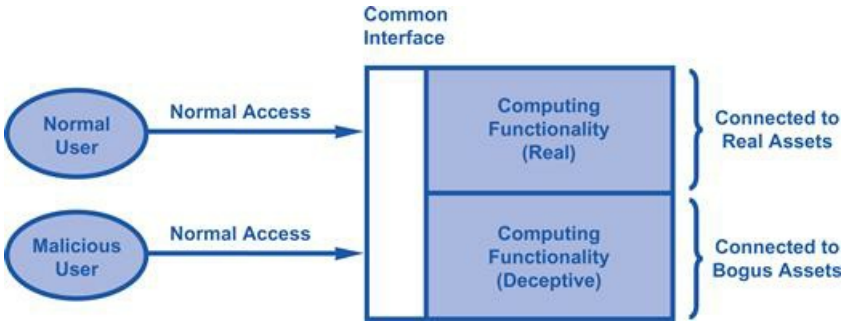


Figure 2.1 Use of deception in computing.

The purpose of deception, ultimately, is to enhance security, so in the context of national infrastructure it can be used for large-scale protection of assets. The reason why deception works is that it helps accomplish any or all of the following four security objectives:

- *Attention*—The attention of an adversary can be diverted from real assets toward bogus ones.
- *Energy*—The valuable time and energy of an adversary can be wasted on bogus targets.
- *Uncertainty*—Uncertainty can be created around the veracity of a discovered vulnerability.
- *Analysis*—A basis can be provided for real-time security analysis of adversary behavior.

The fact that deception diverts the attention of adversaries, while also wasting their time and energy, should be familiar to anyone who has ever used a honey pot on a network. As long as the trap is set properly and the honey pot is sufficiently realistic, adversaries might direct their time, attention, and energy toward something that is useless from an attack perspective. They might even plant time bombs in trap functionality that they believe will be of subsequent use in targeting real assets. Obviously, in a honey pot, this is not the case. This type of deception is a powerful deterrent, because it defuses a cyber attack in a way that could fool an adversary for an extended period of time.

The possibility that deception might create uncertainty around the veracity of a discovered vulnerability has been poorly explored to date. The idea here is that when an intruder inevitably stumbles onto an exploitable hole it would be nice if that intruder were led to believe that the hole might be a trap. Thus, under the right circumstances, the intruder might actually choose to avoid exploitation of a vulnerability for fear that it has been intentionally planted. While this might seem difficult to implement in many settings, the concept is powerful because it allows security managers to defuse existing vulnerabilities *without even knowing about them*. This is a significant enough concept that it deserves repeating: The use of deception in computing allows system security managers to reduce the risk of vulnerabilities *that they might not even know are present*.

Deception is a powerful security tool, as it protects even unknown vulnerabilities.

The fact that real-time analysis can be performed on a honey pot is reasonably well known in the computing community today. Perhaps this is because it is a widely accepted best practice that security administrators should try to observe the behavior of intruders that have been detected. Most intrusion detection systems, for example, include threat management back-end systems that are designed to support such an objective. In the best case, the forensic analysis gathered during deception is sufficiently detailed to

allow for identification of the adversary and possibly even prosecution. In the most typical case, however, accurate traceability to the original human source of a problem is rarely accomplished.

Luckily, the success of deceptive traps is assisted by the fact that intruders will almost always view designers and operators of national assets as being sloppy in their actions, deficient in their training, and incompetent in their knowledge. This extremely negative opinion of the individuals running national infrastructure is a core belief in virtually every hacking community in the world (and is arguably justified in some environments). Ironically, this low expectation is an important element that helps make stealth deception much more feasible, because honey pots do not always have to mimic a perfectly managed environment. Instead, adversaries can generally be led to find a system environment that is poorly administered, and they will not bat an eyelash. This helps the deception designer.

Honey pots should not necessarily mimic perfect environments.

The less well-understood case of openly advertised deception relies on the adversary believing that designers and operators of national assets are competent enough to plant a believable trap into a national asset. This view represents a hurdle, because the hacking community will need to see convincing evidence before they will ever believe that anyone associated with a large organization would be competent enough to manage a complex program of deceptive computing. This is too bad, because open use of deception carries great advantages, as we will explain in more detail below. In any event, the psychology of understanding and managing adversary views is not straightforward. This soft issue must become part of the national infrastructure protection equation but will obviously require a new set of skills among security experts.

Effective cyber deception involves understanding your adversary.

The most common implementation of deception involves the insertion of fake attack entry points, such as open service ports, that adversaries might expect to see in a normal system. The hope is that an adversary would discover (perhaps with a scanner) and then connect to these open service ports, which would in turn then lead to a honey pot. As suggested above, creating realism in a honey pot is not an easy task, but several design options do exist. One approach involves routing inbound open port connections to physically separate bogus systems that are isolated from real assets. This allows for a "forklift"-type copying of real functionality (perhaps with sensitive data sanitized) to an isolated, safe location where no real damage can be done.

Recall that, if the deception is advertised openly, the possibility arises that an adversary will not bother to attempt an attack. Admittedly, this scenario is a stretch, but the possibility does arise and is worth mentioning. Nevertheless, we will assume for the balance of this discussion that the adversary finds the deceptive entry point, presumes that it is real, and decides to move forward with an attack. If the subsequent deception is properly managed, then the adversary should be led down a controlled process path with four distinct attack stages: *scanning*, *discovery*, *exploitation*, and *exposing* (see Figure 2.2).

Bear in mind that a cyber honey pot might require coordination with a tangible exploitable point outside the cyber world.

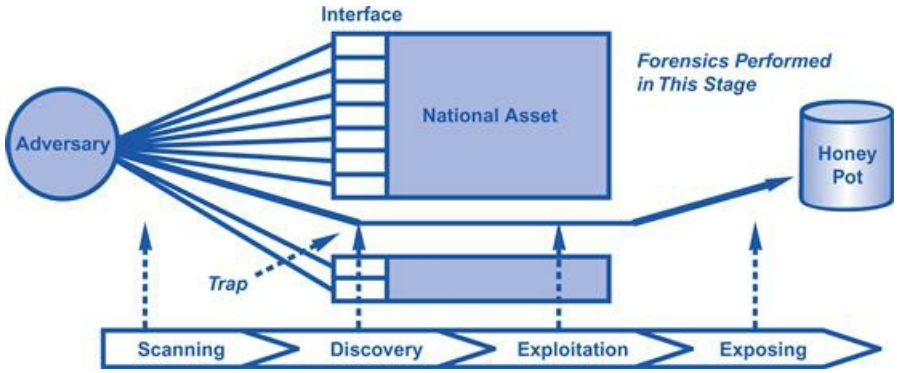


Figure 2.2 Stages of deception for national infrastructure protection.

During the initial scanning stage, an adversary is searching through whatever means is available for exploitable entry points. The presumption in this stage is that the service interface includes trap functionality, such as bogus links on proxied websites that lead to a honey pot for collecting information. It is worth noting, however, that this "searching" process does not always imply the use of a network by an adversary. Instead, the adversary might be searching for exploitable entry points in contracts, processes, locked cabinets, safes, or even relationships with national infrastructure personnel. In practice, one might even expect a combination of computing and noncomputing searches for information about exploitable entry points. The deception must be designed accordingly.

During the discovery phase, an adversary finds an exploitable entry point, which might be real or fake. If the vulnerability is real, then one hopes that good back-end security is in place to avoid an infrastructure disaster. Nevertheless, the decision on the part of the intruder to exploit a discovered vulnerability, real or fake, is an important trigger point. Good infrastructure security systems would need to connect this exploitation point to a threat management system that would either open a security trouble ticket or would alert a security administrator that an intruder has either started an attack or fallen for the deceptive bait. Obviously, such alerts should not signal an intruder that a trap is present.

During the exploitation stage, the adversary makes use of the discovered vulnerability for whatever purposes they might have. If the vulnerability is real, then the usual infrastructure break-in scenario results. If the vulnerability is a trap, however, then its effectiveness will be directly related to the realism of the honey pot. For both stealth and non-stealth deception, this is the initial stage during which data becomes available for forensic analysis. A design consideration is that the actual asset must never become compromised as a result of the trap. This requirement will likely result in deceptive functionality running on computing "islands" that are functionally separated from the real assets.

Actual assets must remain separate and protected so they are not compromised by a honey pot trap.

During the exposing stage in deception, adversary behavior becomes available for observation. Honey pots should include sufficient monitoring to expose adversary technique, intent, and identity. This is generally the stage during which management decisions are made about whether response actions are warranted. It is also a stage where real-time human actions are often required to help make the deceptive functionality look real. As we stated above, a great advantage that arises here is the low expectation the adversary will have

regarding system administrative competency on the part of the infrastructure team. This allows the security team to use the excuse of poor setup to cover functional gaps that might exist in the deception.

Monitoring honey pots takes security to the next level: potential for responsive action.

Any one of the four stages of deception can raise significant legal and social issues, so any program of national infrastructure protection must have participation from the national legal community to determine what is considered acceptable. The difference between a passive trap and an active lure, for example, is subtle and must be clarified before a live deployment is made into infrastructure. From a social perspective, one might hope that the acceptance that exists for using deception to catch online stalkers would be extended to the cyber security community for catching adversaries targeting national infrastructure.

Scanning Stage

In this first stage, the presumption is that an adversary is scanning whatever is available to find exploitation points to attack national infrastructure. This scanning can include online searches for web-based information, network scans to determine port availability, and even offline searches of documents for relevant information. Deception can be used to divert these scanning attempts by creating false entry points with planted vulnerabilities. To deal with the offline case, the deception can extend to noncomputing situations such as intentionally leaving a normally locked cabinet or safe door open with bogus documents inserted to deceive a malicious insider.

The deceptive design goal during scanning is to make available an interface with three distinct components: *authorized services*, *real vulnerabilities*, and *bogus vulnerabilities*. In a perfect world, there would be no vulnerabilities, only authorized services. Unfortunately, given the extreme complexity associated with national infrastructure services, this is an unrealistic expectation, so real vulnerabilities will always be present in some way, shape, or form. When deception is used, these real vulnerabilities are complemented by fake ones and should be indistinguishable. Thus, an adversary will see three components when presented with a national asset interface with deception (see Figure 2.3).

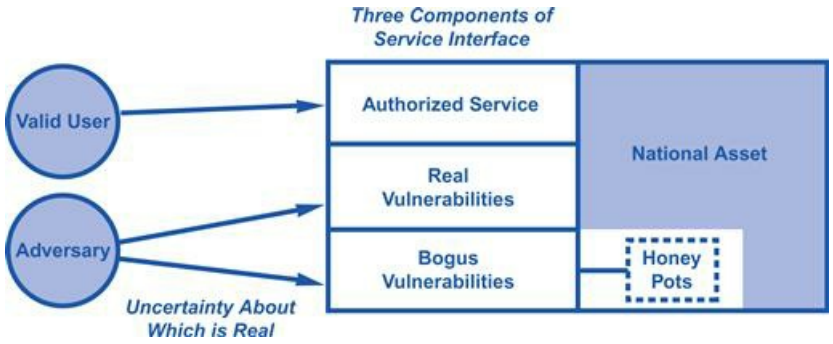


Figure 2.3 National asset service interface with deception.

Bogus vulnerabilities will generally be inserted based on the usual sorts of problems found in software. This is one of the few cases where the deficiencies of the software engineering discipline can actually be put to good use for security. One might imagine situations where new vulnerabilities are discovered and then immediately implemented as traps in systems that require protection. Nevertheless, planted holes do not always have to be based on such exploitable software bugs or system misconfigurations. In some cases, they might correspond to properly administered functionality, but that might not be considered acceptable for local use.

Honey Pots can be Built into Websites

A good example of a trap based on properly administered functionality might be a promiscuous tab on a website that openly solicits leaks of information; this is found sometimes on some of the more controversial blog sites. If legal and policy acceptance is given, then these links might be connected in a local proxied

Intranet to a honey pot collection site. Insiders to an organization might then consider leaking information directly using this link to the seemingly valid Internet site, only to be duped into providing the leak to the local security team. Again, this should only be considered for deployment if all legal and policy requirements are met, but the example does help illustrate the possibilities.

A prominent goal of deception is to observe the adversary in action. This is done via real-time collection of data about intruder activity, along with reasoned analysis about intent. For example, if the intruder seems to be guessing passwords over and over again to gain access to a honey pot system, the administrator might decide in real time to simply grant access. A great challenge is that the automation possibilities of such response are not currently well understood and are barely included in security research programs. This is too bad, because such cases could really challenge and ultimately improve the skills of a good security administrator. One could even imagine national groups sponsoring contests between live intruders and live administrators who are battling against each other in real time in a contrived honey pot.

Allowing an intruder access increases your risk level but also allows the security administrator to monitor the intruder's moves.

Deliberately Open Ports

Intruders routinely search the Internet for servers that allow connections to exploitable inbound services. These services are exploitable generally because they contain some weakness such as a buffer overflow condition that can be tripped to gain privileged access. Once privileged access is obtained, the intruder can perform administrative tasks such as changing system files, installing malware, and stealing sensitive information. All good system administrators understand the importance of *hardening* servers by disabling all exploitable and unnecessary services. The problem is that hardening is a complex process that is made more difficult in environments where the operating system is proprietary and less transparent. Amazingly, most software and server vendors still deliver their products in configurations that include most services being default enabled.

The deliberate insertion of open service ports on an Internet-facing server is the most straightforward of all deceptive computing practices. The deliberately open ports are connected to back-end honey pot functionality, which is connected to monitoring systems for the purpose of observation and analysis. The result is that servers would thus present adversaries of national infrastructure with three different views of open service ports: (1) valid open ports one might expect, such as HTTP, DNS, and SMTP; (2) open ports that are inadvertently left open and might correspond to exploitable software; and (3) open ports that are deliberately inserted and connected to bogus assets in a honey pot. As long as it is generally understood that deception could *potentially* be deployed, there could be some uncertainty on the part of the adversary about which open ports are deliberate and which are inadvertent (see Figure 2.4).

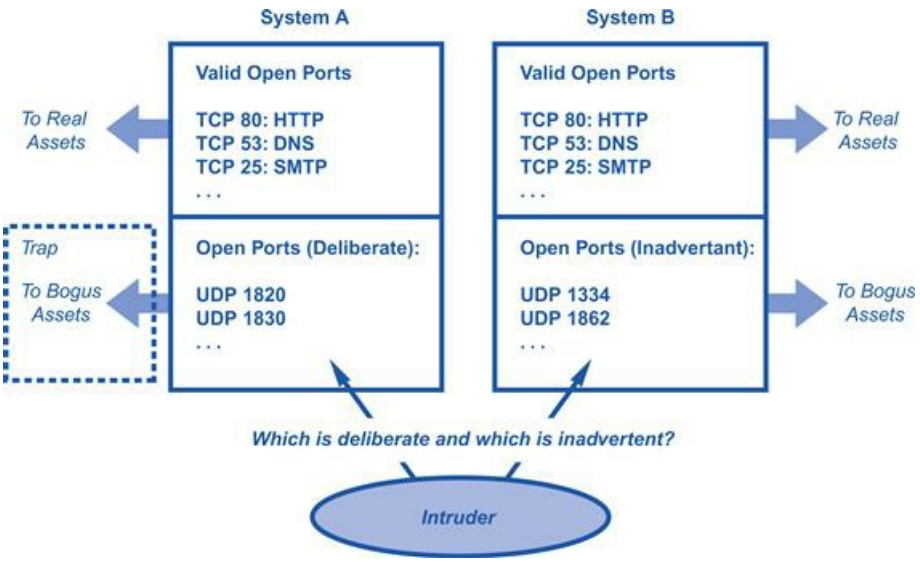


Figure 2.4 Use of deceptive open ports to bogus assets.

Security managers who use port scanners as part of a normal program of enterprise network protection often cringe at this use of deception. What happens is that their scanners will find these open ports, which will result in the generation of reports that highlight the presumed vulnerabilities to managers, users, and auditors. Certainly, the output can be manually cropped to avoid such exposure, but this might not scale well

to a large enterprise. Unfortunately, solutions are not easily identified that solve this incompatibility between the authorized use of port scanners and the deliberate use of open ports as traps. It represents yet another area for research and development in deceptive computing.

Another challenge is for security managers to knowingly keep open ports after running scanners that highlight these vulnerabilities.

An additional consideration with the deliberate use of open ports is that care must be taken on the back end to ensure that real assets cannot be exploited. Not surprisingly, practical techniques for doing this are not well known. For example, if the back-end deceptive software connected to deliberately open ports shares resources with valid assets, then the potential exists for negative side effects. The only reasonable approach today would involve deliberately open ports on bogus servers that are honey pots with no valid resources. These servers should be subtly embedded into server complexes so they look normal, but they should be hardwired to separate honey pot assets. This reduces the likelihood of negative side effects on normal servers (see Figure 2.5).

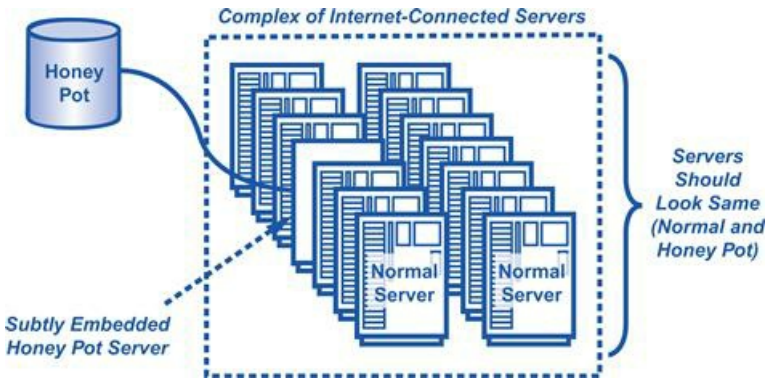


Figure 2.5 Embedding a honey pot server into a normal server complex.

In practice, the real challenge to the deceptive use of open ports is creating port-connected functionality that is sufficiently valid to fool an expert adversary but also properly separated from valid services so no adversary could make use of the honey pot to advance an attack. Because computer science does not currently offer much foundational assistance in this regard, national infrastructure protection initiatives must include immediate programs of research and development to push this technique forward.

Discovery Stage

The discovery stage corresponds to the adversary finding and accepting the security bait embedded in the trap. The two corresponding security goals during this stage are to make an intruder believe that real vulnerabilities could be bogus and that bogus vulnerabilities could be real. The first of these goals is accomplished by making the deception program well-established and openly known. Specific techniques for doing this include the following:

- *Sponsored research*—The use of deception in national infrastructure could become generally presumed through the open sponsorship and funding of unclassified research and development work in this area.
- *Published case studies*—The open publication of case studies where deception has been used effectively in national asset protection increases the likelihood that an adversary might consider a found vulnerability to be deliberate.
- *Open solicitations*—Requests for Information (RFIs) and Requests for Proposals (RFPs) should be openly issued by national asset protectors. This implies that funding must be directed toward security projects that would actually use deceptive methods.

Interestingly, the potential that an adversary will hesitate before exploiting a real vulnerability increases only when the use of deception appears to be a real possibility. It would seem a hollow goal, for example, to simply announce that deception is being used without honest efforts to really deploy such deceptions in national infrastructure. This is akin to placing a home protection sign in the landscaping without ever installing a real security system. For openly advertised deception to work, the national infrastructure team must be fully committed to actually doing the engineering, deployment, and operation.

Openly advertised use of deception may cause adversaries to question whether a discovered vulnerability is valid or bogus.

The second goal of making bogus vulnerabilities look real will be familiar to computer security experts who have considered the use of honey pots. The technique of duplication is often used in honey pot design, where a bogus system is a perfect copy of a real one but without the back-end connectivity to the real asset being protected. This is generally done by duplicating the front-end interface to a real system and placing the duplicate next to a back-end honey pot. Duplication greatly increases realism and is actually quite easy to implement in practice (see Figure 2.6).

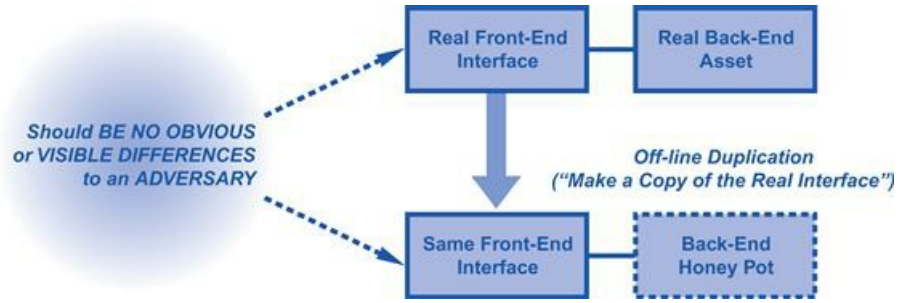


Figure 2.6 Duplication in honey pot design.

As suggested above, the advantage of duplication in honey pot design is that it maximizes authenticity. If one finds, for example, a real vulnerability in some front-end server, then an image of that vulnerable server could be used in future deceptive configurations. Programs of national infrastructure protection should thus find ways to effectively connect vulnerability discovery processes to honey pot design. Thus, when a truly interesting vulnerability is found, it can become the front end to a future deceptive trap.

Turn discovered vulnerabilities into advantages by mimicking them in honey pot traps.

Deceptive Documents

The creation and special placement of deceptive documents is an example method for tricking adversaries during discovery. This technique, which can be done electronically or manually, is especially useful for detecting the presence of a malicious insider and will only work under two conditions:

- *Content*—The bogus document must include information that is convincingly realistic. Duplication of a valid document with changes to the most sensitive components is a straightforward means for doing this.
- *Protection*—The placement of the bogus document should include sufficient protections to make the document appear truly realistic. If the protection approach is thin, then this will raise immediate suspicion. Sabotage can be detected by protecting the bogus document in an environment that cannot be accessed by anyone other than trusted insiders.

An illustrative approach for national infrastructure protection would follow these steps: First, a document is created with information that references a specially created bogus asset, such as a phone number, physical location, or server. The information should never be real, but it should be very realistic. Next, the document is stored in a highly protected location, such as a locked safe (computer or physical). The presumption is that under normal circumstances the document should sit idly in the locked safe, as it should have no real purpose to anyone. Finally, the specially created bogus asset is monitored carefully for any attempted compromise. If someone finds and grabs the document, then one can conclude that some insider is not to be trusted.

Steps to Planting a Bogus Document

To effectively plant a bogus document, consider following these steps:

1. Create a file with instructions for obtaining what would appear to be extremely sensitive information. The file could include a phone number, an Internet address for a server, and perhaps a room location in some hotel.
2. Encrypt the file and store it on a server (or print and lock it in a safe) that one would presume to be protected from inside or outside access.
3. Put monitoring of the server or safe in place, with no expectation of a time limit. In fact, the monitoring might go on indefinitely, because one would expect to see no correlative behavior on these monitored assets (see Figure 2.7).

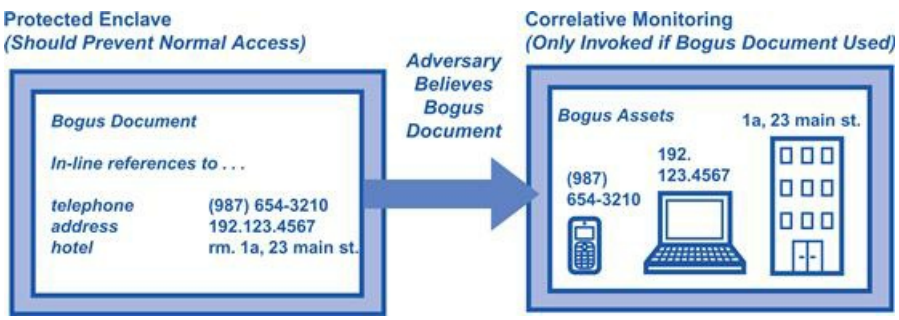


Figure 2.7 Planting a bogus document in a protected enclave.

It should be obvious that the example scheme shown in Figure 2.7 works as well for an electronic document protected by encryption and access control as for a manual paper document locked in a protected safe. In both cases, one would expect that no one would ever correlate these bogus references. If it turns out that the monitoring shows access to these bogus assets in some related way, then one would have to assume that the protected enclave has been compromised. (Monitoring a hotel might require complex logistics, such as the use of hidden cameras.) In any event, these assets would provide a platform for subsequent analysis of exploitation activity by the adversary.

Exploitation Stage

The third stage of the deception lifecycle for an adversary involves exploitation of a discovered vulnerability. This is a key step in the decision process for an adversary because it is usually the first stage in which policy rules or even laws are actually violated. That is, when an intruder begins to create a cyber attack, the initial steps are preparatory and generally do not violate any specific policy rules or laws. Sometimes security experts refer to this early activity as *low radar actions*, and when they are detected they are referred to as *indications and warnings*. Determining whether to respond to indications and warnings is a challenge, because response requires time and energy. If the track record of the security team involves many response actions to indications and warnings that are largely false positives, then the organization is often tempted to reduce the response trigger point. This is a bad idea for national infrastructure, because the chances increase that a real event will occur that is not responded to promptly.

Responding to a large number of false positives is necessary to adequately protect national infrastructure.

As such, the protection of national infrastructure should involve a mind shift away from trying to reduce false positive responses to indications and warnings. Instead, the goal should be to deal with all instances in which indication and warning actions would appear to be building up to the threshold at which exploitation begins. This is especially important, because this threshold marks the first stage during which real assets, if targeted, might actually be damaged (see Figure 2.8).

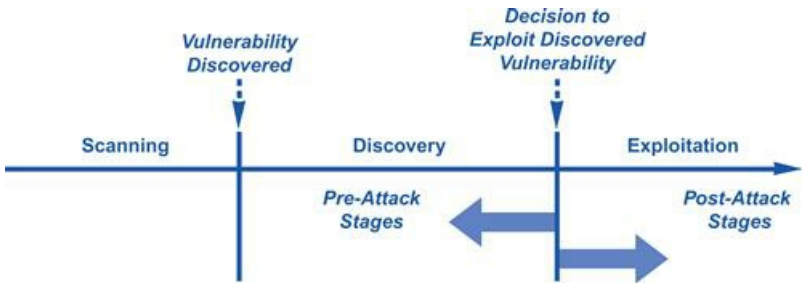


Figure 2.8 Pre- and post-attack stages at the exploitation stage.

The key requirement at this decision point is that any exploitation of a bogus asset must not cause disclosure, integrity, theft, or availability problems with any real asset. Such non-interference between bogus and real assets is easiest to accomplish when these assets are kept as separate as possible. Physical separation of assets is straightforward; a real software application with real data, for example, could be separated from a bogus application with fake data by simply hosting each on different servers, perhaps even on different networks. This is how most honey pots operate, and the risk of interference is generally low.

Achieving noninterference in an environment where resources are shared between real and fake assets is more challenging. To accomplish this goal, the deception designer must be creative. For example, if some business process is to be shared by both real and fake functionality, then care must be taken by the deception operators to ensure that real systems are not degraded in any way. Very little research has been done in this

area, especially for availability threats. Allowing a malicious adversary to execute programs on a live, valid system, for example, would provide opportunities for malicious resource exhaustion. Nevertheless, the general approach has considerable promise and deserves more attention.

When bogus and real assets reside on the same server, vulnerability risk increases dramatically.

A related issue involves the possibility that intrusion detection and incident response systems might be fooled during exploitation into believing that trap functionality is real. White hat teams in companies have dealt with this problem for years, and they must coordinate with security teams to ensure that their activity does not cause a false alarm. This can be accomplished in several ways:

- *Process coordination*—This involves the honey pot team coordinating their activities in advance with the infrastructure security teams. The problem is that it tends to highlight the deception and might destroy some hidden transparency, especially if the deception is designed to detect malicious insiders.
- *Trap isolation*—This involves making certain that security systems, such as intrusion detection probes, are not in the deceptive path. Ongoing architectural analysis is required in this case to make sure that the condition holds throughout the system lifecycle.
- *Back-end insiders*—If the security team includes some trusted insiders involved in response activities, then they might be able to ensure that a response to deceptive action does not waste time and resources. This works best if the insider is a decision-maker.
- *Process allowance*—In this case, the detection and response activities are allowed to proceed. This is the recommended case if the deception is considered sensitive and must be totally hidden from insiders.

The exploitation stage is the most dangerous part of the deceptive lifecycle because it involves real malicious activity from an adversary. Infrastructure teams must learn to understand and respect the potential for real damage if this stage is not managed carefully.

Procurement Tricks

One way to understand adversary behavior is to compare it in different environments. An example involves a deception designer creating two separate supplier solicitations for a given product or service. One solicitation would be for a benign, noncritical, nonsensitive application; the other would be for an obviously sensitive, nationally critical application. In both cases, *exactly the same product or service* would be requested, but when they are delivered to each application detailed examination would be initiated to identify differences. Any differences found in the delivered product by the intentionally planted "critical buyer" and "benign buyer" would be viewed as potential exploits (see Figure 2.9).

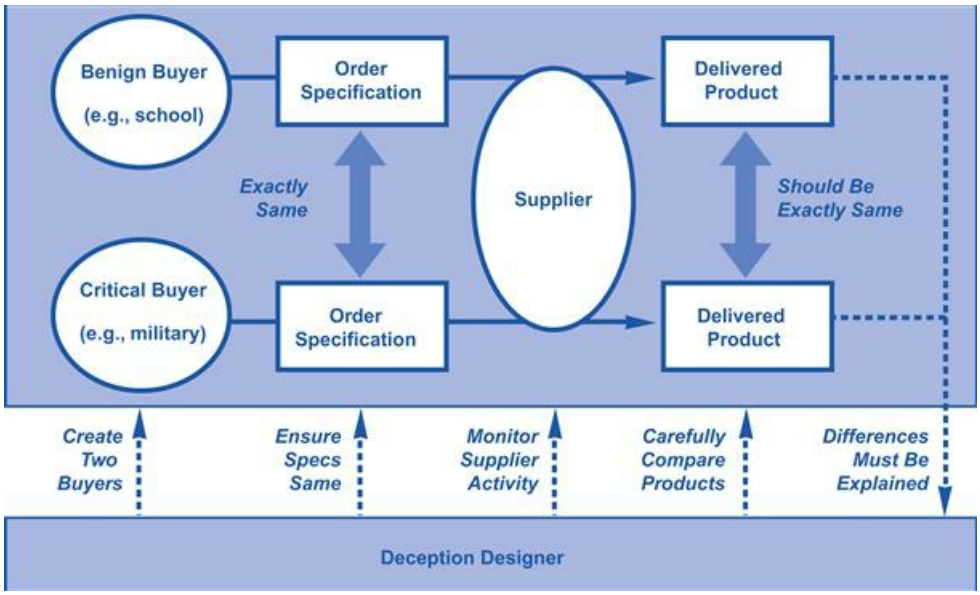


Figure 2.9 Using deception against malicious suppliers.

The deception described above only works if sufficient forensic capability exists to compare the two delivered products. For any product or service, this could include comparison of relative software size, system performance, product documentation, service functionality, or technical support. One could even imagine a second level of deception using social engineering where an impassioned plea would be made to the supplier for some undocumented type of emergency support, usually remote administration. If either of the delivered products is set up for such remote administration, then the national asset manager would know that something is wrong.

The procurement lifecycle is one of the most underestimated components in national infrastructure protection from an attack perspective. Generally, security teams focus on selecting, testing, installing, and operating functionality, with seemingly mundane procurement tasks left to the supply chain team. This is a huge mistake, and adversaries understand this point well. Thus, national infrastructure protection initiatives must extend to the procurement process, and the clever use of deception is a powerful tool in this regard.

National infrastructure protection must extend from procurement to operating functionality in order to be

truly effective.

Exposing Stage

The final stage in the deception lifecycle involves the adversary exposing behavior to the deception operator. Presumably, in this stage, the adversary is now hacking away at the trap functionality, convinced that all systems and assets are real. All sorts of possibilities arise in terms of how this hacking will proceed. It could be a flurry of intense activity in a short period of time or it could be a drawn-out process of low and slow actions, so the deception team must have patience. Also, during this stage, the adversary might expose the use of well-known hacking techniques and tools or, alternatively, could demonstrate use of techniques not previously seen by the security team (see Figure 2.10).

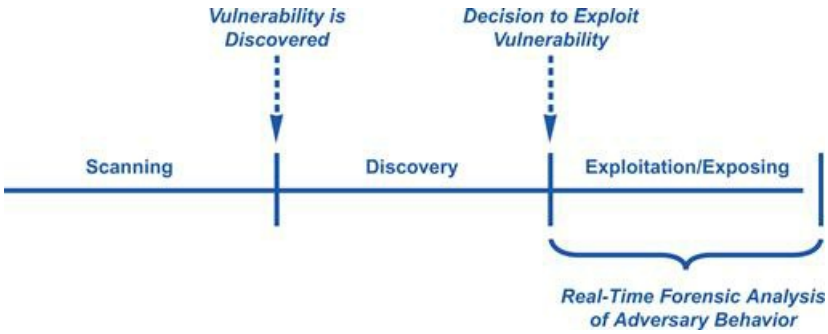


Figure 2.10 Adversary exposing stage during deception.

The challenge in this stage is that the deception must allow a window for observation of intruder activity, but must also be hidden. It must provide a convenient interface for the deception team to collect data but also must provide a way for commands to be issued or changes to be made in real time. Unfortunately, few commercial products exist that are designed to support these features. Specific functional requirements for the monitoring environment during the exposing stage of deception include the following:

Observing intruder activity can be an informative but risky process during the exposure stage.

- *Sufficient detail*—The monitoring environment must provide sufficient detail so the deception operator can determine what is going on. For example, overly cryptic audit logs in terse format with gaps in certain areas would not be the best approach. The usual administrative interface to an operating system (generally through a command interface) is often the most desirable approach. One should not expect fancy, colorful displays for the monitoring task because most security analysts prefer the convenience of a terse command line interface.
- *Hidden probes*—Monitoring in this stage of the deception works only if it is completely hidden. If an adversary figures out that someone is watching, then behavior modification would occur immediately. Simple tasks must therefore be implemented such as suppressed listing of any processes launched by the deception team (unless desired). The art of creating realistic functionality to hide probes requires support and nurturing in the security community.
- *Real-time observation*—The deception operator should have access to information about exposed

behavior as it happens. The degree of real time for such monitoring (e.g., instantaneous, within seconds, within minutes) would depend on the local circumstances. In most cases, this observation is simply done by watching system logs, but more advanced tools are required to record and store information about intruder behavior.

As we suggested above, in all cases of deception monitoring the key design goal should be to ensure a believable environment. No suspicious or unexplainable processes should be present that could tip off an intruder that logging is ongoing. Fake audit logs are also a good way to create believability; if a honey pot is developed using an operating system with normal audit logging, then this should be enabled. A good adversary will likely turn it off. The idea is that hidden monitoring would have to be put in place underneath the normal logging—and this would be functionality that the adversary could not turn off.

Interfaces Between Humans and Computers

The gathering of forensic evidence during the analysis of intruder behavior in a honey pot often relies on detailed understanding of how systems, protocols, and services interact. Specifically, this type of communication can be performed in four different ways: *human-to-human*, *human-to-computer*, *computer-to-human*, and *computer-to-computer*. If we take the first term (human or computer) to mean the intruder and we take the second term to mean the honey pot manager, then we can make some logical distinctions.

First, it should be obvious that, in an automated attack such as a botnet, the real-time behavior of the attack system will not change based on some subjective observation of honey pot functionality. Certainly, the interpretation of the results of the botnet could easily affect the thinking of the botnet operator, but the real-time functionality is not going to be affected. As such, the most powerful cases in real-time forensic analysis of honey pot behavior will be the cases where human-to-human and human-to-computer interactions are being attempted by an intruder. Let's examine each in turn.

Real-time forensic analysis is not possible for every scenario, such as a botnet attack.

The most common human-to-human interaction in national infrastructure involves help desk or customer care support functions, and the corresponding attack approach involves social engineering of such activity. The current state of the art in dealing with this vulnerability is to train operators and customer care personnel to detect attempts at social engineering and to report them to the security team. Deception, however, introduces a more interesting option. If the likelihood is high that social engineering is being attempted, then an advanced approach to protection might involve deceiving the adversary into believing that they have succeeded. This can be accomplished quite easily by simply training operators to divert social engineering attempts to specially established help desks that are phony. The operators at these phony desks would reverse social engineer such attackers to get them to expose their identity or motivation (see Figure 2.11).



Figure 2.11 Deceptively exploiting the human-to-human interface.

The most common human-to-computer interaction occurs when an intruder is trying to gain unauthorized access through a series of live, interactive commands. The idea is that intruders should be led to believe that their activity is invoking services on the target system, as in the usual type of operating system hacking. A good example might involve an intruder repeatedly trying to execute some command or operation in a trap system. If the security team notices this intent and can act quickly enough, the desired command or operation could be deliberately led to execute. This is a tricky engagement, because an expert adversary might

notice that the target configuration is changing, which obviously is not normal.

An expert adversary may become aware of the security team observing the attempted intrusion.

National Deception Program

One might hope that some sort of national deception program could be created based on a collection of traps strategically planted across national infrastructure components, tied together by some sort of deception analysis backbone. Such an approach is unlikely, because deception remains a poorly understood security approach, and infrastructure managers would be very hesitant to allow traps to be implanted in production systems. These traps, if they malfunction or do not work as advertised, could trick authorized users or impede normal operations.

Any realistic assessment of current security and information technology practice suggests that large-scale adoption of deception for national infrastructure protection would not be widely accepted today. As a result, programs of national deception would be better designed based on the following assumptions:

- *Selective infrastructure use*—One must assume that certain infrastructure components are likely to include deceptive traps but that others will not. At the time of this writing, many infrastructure teams are still grappling with basic computer security concepts; the idea that they would agree to install traps is not realistic. As such, any program of national deception must assume that not all components would utilize honey pots in the same manner.
- *Sharing of results and insights*—Programs of national deception can and should include a mechanism for the sharing of results and insights gained through operational use of traps and honey pots. Certainly, insight obtained through forensic analysis of adversary behavior can be shared in a structured manner.
- *Reuse of tools and methods*—National deception programs could serve as means for making honey pot and trap software available for deployment. In some cases, deception tools and methods that work in one infrastructure area can be reused in another.

The most common criticism of deception in large-scale national security is that automated tools such as botnets are not affected by trap functionality. While it is true that botnets attack infrastructure in a blindly automated manner regardless of whether the target is real or fake, the possibility remains that trap functionality might have some positive impact. A good example might be national coordination of numerous bogus endpoints that might be ready and willing to accept botnet software. If these endpoints are designed properly, one could imagine them being deliberately designed to mess up the botnet communication, perhaps by targeting the controllers themselves. This approach is often referred to as a *tarpit*, and one might imagine this method being quite interesting for degrading the effectiveness of a botnet.

Finally, let's briefly look at how to improve the plans to shore up the defenses against cyber attacks. Amateur cyber attacks are expected to show more deception, and cyberwarfare attacks can be expected to show more too.

The Deception Planning Process Against Cyber Attacks

Cyber attacks are increasing in technical sophistication, as easier attacks are being blocked or foiled. Deception can be a useful force multiplier for mission plans in cyberspace just as in real battle spaces. Many new deceptions are expected, since very few of the possible ploys have been explored, and it will become very easy for deceptions to succeed: Defenses have come to a screeching halt, and defenders are becoming less aware of deceptions being practiced, so the pool of potential victims for many attacks is increasing.

Nevertheless, successful deception starts with a deception plan. A successful deception process is one in which the ends dictate the means. This is reinforced by the fact that deception plans are driven by the desired effect on the target. Deception planners are guided toward a successful deception planning process that requires command involvement and approval at each stage of the process. This process could be a 34-part step-by-step planning process for deception to increase the probability of success (see "An Agenda for Action in the Planning Process for Deception" below):

An Agenda for Action in the Planning Process for Deception

When completing the Planning Process for Deception checklist, the deception planner should adhere to the provisional list of actions to prepare for contingencies in the event that the deception fails. During the course of the deception, the planner also seeks feedback to ensure that the target is responding in the expected way. The order is not significant; however, these are the activities for which the research would want to provide a detailed description of procedures, review, and assessment for ease of use and admissibility. Current measures that must be adhered to, in order to plan for deception, include (check all tasks completed):

1. Identifying the strategic goal.
2. Deciding how the target should react.
3. Determining what the target should perceive.
4. Deciding what to hide and show.
5. Analyzing the pattern for hiding.
6. Analyzing the pattern for showing.
7. Designing the desired effect with the hidden method.
8. Selling the effect to those who are executing the deception.
9. Deciding the communications channels to transmit the deception.
10. Making sure that the target buys the effect and falls for the deception.
11. Pretending to be a naive victim to entrap deceptive cyber attackers.
12. Camouflaging key targets or make them look unimportant or disguise software as different software.
13. Doing something an unexpected way.
14. Inducing the cyber attacker to download a Trojan horse.
15. Secretly monitoring attacker activities.
16. Transfering Trojan horses back to cyber attacker.
17. Trying to frighten the cyber attacker with false messages from authorities (like "we know where you are, and a drone is coming to take you out," etc. …).

18. Transfering the cyber attack to a safer machine like a honeypot.
19. Swamping the cyber attacker with messages or requests.
20. Associating false times with files.
21. Falsifying file-creation times.
22. Falsifying file-modification times.
23. Deliberately delaying processing commands.
24. Lying that you cannot do something or do something unrequested.
25. Lying that a suspicious command succeeded.
26. Lying about reasons for asking for an additional password.
27. Planting disinformation, redefining executables, and giving false system data.
28. "Emulating" hardware of a machine in software for increased safety.
29. Sending data too large or requests too hard back to the cyber attacker.
30. Systematically misunderstanding cyber attacker commands, as by losing characters.
31. Being a decoy site for the real site.
32. Asking questions that include a few cyber attacker–locating ones.
33. Giving false excuses why you cannot execute cyber attacker commands.
34. Pretending to be an inept defender, or have easy-to-subvert software.

Summary

Deception occurs in cyberspace. An analysis of how deception is used in cyber attacks can help in understanding them, with the goal of developing effective defenses for future cyber attacks against the critical national infrastructure. The deception methods described in this chapter are not difficult to use. While there have not been confirmed instances of cyberwar using deception, cyberwarfare specialists are developing cyberweapons using these methods. However, a wide variety of deception methods can be used to ensure that particular cyber-attack deceptions against a particular target are totally ineffective:

- Buffer overflows can be done by sending insincere large inputs to programs.
- To achieve surprise, cyber attacks can involve rarely used software, ports, or network sites.
- Cyber attacks can have surprising targets such as little-used software features.
- Cyber attacks can occur at surprising times.
- Cyber attacks can occur from surprising sites.
- To maximize concealment, cyber attacks can be done very slowly, as by sending one command a day to a victim computer.
- Cyber attacks can modify file or audit records in time and details to make cyber attackers appear to have been doing something different at a different time.
- Cyber attacks can claim abilities that they do not possess for purposes of extortion, such as the ability to disable a computer system.

Nonetheless, the diversity of deceptions should increase in the future as the continued development of automated tools will permit attackers to try many methods at once. But, diversity in defenses against deceptions should also increase. Deception will be increasingly common in asymmetric cyberwar, as it is in asymmetric conventional warfare, for tactics and strategies by the weaker participant.

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

Chapter Review Questions/Exercises

True/False

1. True or False? The use of deception in computing involves deliberately misleading an adversary by creating a system component that looks fake but is in fact a trap.
2. True or False? Deception can be used to divert scanning attempts by creating false entry points with planted vulnerabilities.
3. True or False? A secondary goal of deception is to observe the adversary in action.
4. True or False? The deliberate insertion of closed service ports on an Internet-facing server is the most straightforward of all deceptive computing practices.
5. True or False? The discovery stage corresponds to the adversary finding and accepting the security bait embedded in the trap.

Multiple Choice

1. The reason why deception works is that it helps accomplish any or all of the following four security objectives:

A. Attention, energy, uncertainty, and analysis
B. Attention, vulnerability, uncertainty, and analysis
C. Attention, energy, honey pot, and analysis
D. Attention, energy, uncertainty, and traceability
E. Implementation, energy, uncertainty, and analysis

2. If the deception is properly managed, then the adversary should be led down a controlled process path with four distinct attack stages, except which one:

A. Scanning
B. Functionality
C. Exploitation
D. Discovery
E. Exposing

3. Honey pots should include sufficient monitoring to expose which of the following three:

A. Adversary technique
B. Depth
C. Intent
D. Identity
E. Diversity

4. The deceptive design goal during scanning is to make available an interface with which three distinct components:

A. Authorized services
B. Real vulnerabilities
C. Unrealistic expectations
D. System misconfigurations
E. Bogus vulnerabilities

5. Servers will present adversaries of the national infrastructure with which three different views of open service ports:

A. There could be some uncertainty on the part of the adversary about which open ports are deliberate and which are inadvertent.
B. Valid open ports one might expect, such as HTTP, DNS, and SMTP.
C. If the back-end deceptive software connected to deliberately open ports shares resources with valid assets, then the potential exists for negative side effects.
D. Open ports that are inadvertently left open and might correspond to exploitable software.
E. Open ports that are deliberately inserted and connected to bogus assets in a honey pot.

Exercise

Problem

A diversified Fortune 500 corporation that provides products and services to domestic and foreign governments and commercial customers suspected that a deceptive intruder was in their network; however, they knew neither the extent of the compromise, nor what (if any) data had been breached. The persistent deceptive intruders used tools and techniques that left trace evidence on each computer system they compromised. These host-based indicators of compromise are present every time the intruders attack a network. The corporation (client) called a team of advanced persistent threat (APT) experts to validate their concerns, scope the intrusion, and provide a remediation strategy. APTs are used to identify, scope, and remediate the APT in the government and defense industrial base. The APT consists of skilled and sophisticated deceptive hackers who deploy a complex arsenal of deception malware against specific targets in the Defense Industrial Base (DIB), financial, manufacturing, and research industries. Please explain how the APT went about resolving the problem.

Hands-On Projects

Project

The Defense Information Systems Agency (DISA) within the U.S. Department of Defense (DoD) has parallel missions to evaluate new network defense technologies, policies, and tactics and to train DoD personnel to repel deception attacks upon critical cyber national infrastructures. DISA deployed a fully operational evolved cyber range to measure the resiliency (deception, performance, security, and stability) of its network and data center infrastructures, conduct advanced research and development, and train its cyber

warriors. Nevertheless, DISA needed the full functionality of a cyber range to carry out its missions. Just as traditional soldiers need a firing range to hone their skills using the latest weaponry, cyber warriors need a similar environment in the virtual world to train for deception in cyber attacks. Yet, the agency could not wait several more years for the launch of the DoD's National Cyber Range, nor could it spend the millions of dollars required for traditional custom-built cyber ranges. A new, evolved model of cyber range was required, one that could be set up in hours with minimal infrastructure, then customized for cyberwarfare scenarios within minutes. So, how would DISA's cyber security team go about creating the cyber range model?

Case Projects

Problem

Let's look at a real-world scenario of how one of the world's largest banks was challenged to harden network and data center critical infrastructure security deception measures without degrading the high performance required in the financial services industry. The bank's network security team (NST) was charged with institutionalizing a network security certification process to measure the resiliency (performance, security, and stability) of every element of the network before and after deployment. The goal for the team was to right size the critical infrastructure for each line of business without introducing risk, ensuring that they did not over- or underinvest in the network infrastructure. The team used a standardized and repeatable program to certify that devices are able to:

- Protect sensitive customer data from external deception attacks and insider threats.
- Ensure cyber secure, rapid financial transactions.
- Reduce the risk of legal liabilities associated with noncompliance.

Explain how the bank's network security team should handle this situation.

Optional Team Case Project

Problem

Yahoo! is focused on delivering fast and reliable commerce, communications, and social networking services to millions of users around the world. With one of the world's largest network and cloud infrastructures, Yahoo! faces unique challenges as it fulfills its vision to be the center of people's online lives by delivering personally relevant, meaningful Internet experiences. Yahoo!'s traffic volume and application complexity has grown rapidly over the past decade, driving the company to build out a massive network and application infrastructure to support more and more load. The company continues to invest in high-capacity servers, load balancers, routers, and switches, plus massive firewalls. Previously, however, Yahoo!'s security team had no way to stress this enormous critical infrastructure and measure its resiliency to ensure performance, stability, and cyber security. Yahoo! needed a solution to validate the performance, functionality, and capacity of its systems under a wide mix of real-world traffic, including video, instant messaging, and web applications, as well as live cyber security attacks and load from millions of users. Please identify how Yahoo!'s security team stressed their enormous critical infrastructure and measured its resiliency to ensure performance, stability, and

cyber security.

[1] The Honeynet Project, *Know Your Enemy: Revealing the Security Tools, Tactics, and Motives of the Blackhat Community*, Addison–Wesley Professional, New York, 2002. (I highly recommend this amazing and original book.) See also B. Cheswick and S. Bellovin, *Firewalls and Internet Security: Repelling the Wily Hacker*, 1st ed., Addison–Wesley Professional, New York, 1994; C. Stoll, *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage*, Pocket Books, New York, 2005.

3

Separation

Chapter Outline

What Is Separation?
Functional Separation
National Infrastructure Firewalls
DDOS Filtering
SCADA Separation Architecture
Physical Separation
Insider Separation
Asset Separation
Multilevel Security (MLS)
Protecting the Critical National Infrastructure Through Use of Separation
Summary
Chapter Review Questions/Exercises

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

Dorothy Denning[1]

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

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

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

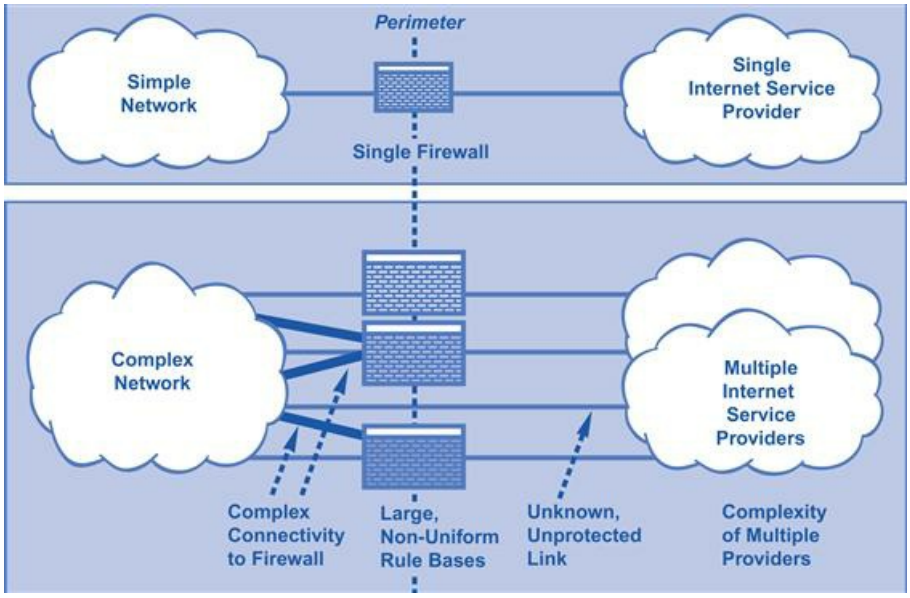


Figure 3.1 Firewalls in simple and complex networks.

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

- *Network-based separation*—Because the perimeter of any complex national infrastructure component will be difficult to define accurately, the use of separation methods such as network-based firewalls is imperative. Such cloud-based functionality allows a broader, more accurate view of the egress and ingress activity for an organization. It also provides a richer environment for filtering high-capacity attacks. The filtering of denial of service attacks aimed at infrastructure, for example, can only be stopped with special types of cloud-based filtering firewalls strategically placed in the network.
- *Internal separation*—National infrastructure protection will require a program of internal asset separation using firewalls strategically placed in infrastructure. This type of separation of internal assets using firewalls or other separation mechanisms (such as operating system access controls) is not generally present in most infrastructure environments. Instead, the idea persists that insiders should have unrestricted access to internal resources and that perimeter firewalls should protect resources from untrusted, external access. This model breaks down in complex infrastructure environments because it is so easy to plant insiders or penetrate complex network perimeters.
- *Tailored separation*—With the use of specialized protocols in national infrastructure management, especially supervisory control and data acquisition (SCADA), tailoring firewalls to handle unique

protocols and services is a requirement. This is a challenge because commercial firewalls are generally designed for generic use in a wide market and tailoring will require a more focused effort. The result will be more accurate firewall operation without the need to open large numbers of service ports to enable SCADA applications.

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

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

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

What Is Separation?

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

- *Adversary separation*—The first separation goal involves separating an asset from an adversary to reduce the risk of direct attack. Whatever implementation is chosen should result in the intruder having no direct means for accessing national assets.
- *Component distribution*—The second separation goal involves architecturally separating components in an infrastructure to distribute the risk of compromise. The idea here is that a compromise in one area of infrastructure should not be allowed to propagate directly.

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

A Working Taxonomy of Separation Techniques

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

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

Figure 3.2 Taxonomy of separation techniques.

The first column in the taxonomy shows that separation controls are focused on keeping either insiders or outsiders away from some asset. The key difference here is that insiders would typically be more trusted and would have more opportunity to gain special types of access. The second column indicates that the separation

controls are focused on either keeping an adversary away from some asset or inherently separating components of the actual asset, perhaps through distribution. The third column identifies whether the separation approach uses computing functionality or would rely instead on some tangible, physical control.

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

Functional Separation

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

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

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

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

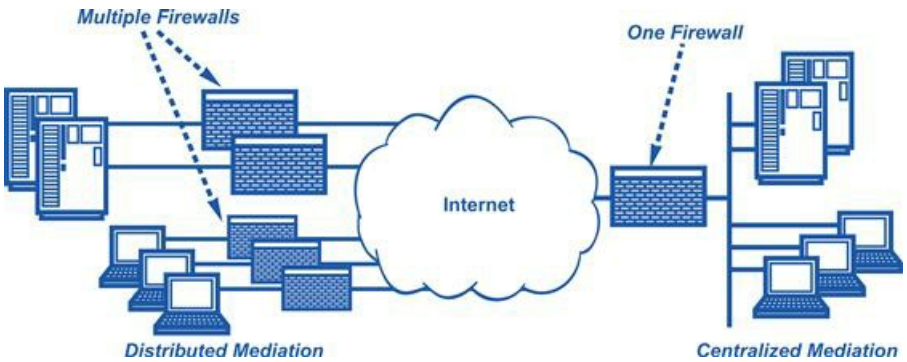


Figure 3.3 Distributed versus centralized mediation.

The distributed approach has had considerable appeal for the global Internet community to date. It avoids the problem of having to trust a large entity with mediation decisions, it allows for commercial entities to market their security tools on a large scale to end users, and it places control of access policy close to the

asset, which presumably should increase the likelihood that the policy is appropriate. The massive global distribution of computer security responsibility to every owner of a home personal computer is an example of this approach. End users must decide how to protect their assets, rather than relying on some centralized authority.

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

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

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

National Infrastructure Firewalls

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

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

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

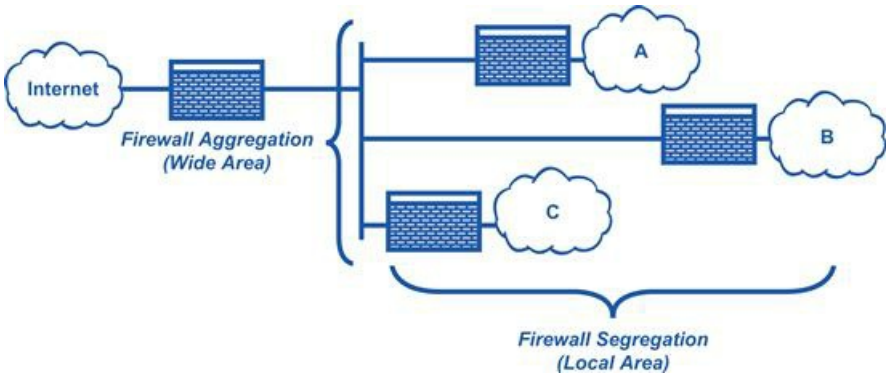


Figure 3.4 Wide area firewall aggregation and local area firewall segregation.

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

A major challenge to national infrastructure comes with the massive increase in wireless connectivity that must be presumed for all national assets in the coming years. Most enterprise workers now carry around some sort of smart device that is ubiquitously connected to the Internet. Such smart devices have begun to resemble

computers in that they can support browsing, e-mail access, and even virtual private network (VPN) access to applications that might reside behind a firewall. As such, the ease with which components of infrastructure can easily bypass defined firewall gateways will increase substantially. The result of this increased wireless connectivity, perhaps via 4 G deployment, will be that all components of infrastructure will require some sort of common means for ensuring security.

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

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

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

- *Vantage point*—The network service provider has a wide vantage point that includes all customers, peering points, and gateways. Thus, if some incident is occurring on the Internet, the service provider will observe its effects.
- *Operations*—Network service providers possess the operational capability to ensure up-to-date coverage of signatures, updates, and new security methods, in contrast to the inability of most end users to keep their security software current.
- *Investment*—Where most end users, including enterprise groups, are unlikely to have funds sufficient to install multiple types of diverse or even redundant security tools, service providers can often support a business case for such investment.

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

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

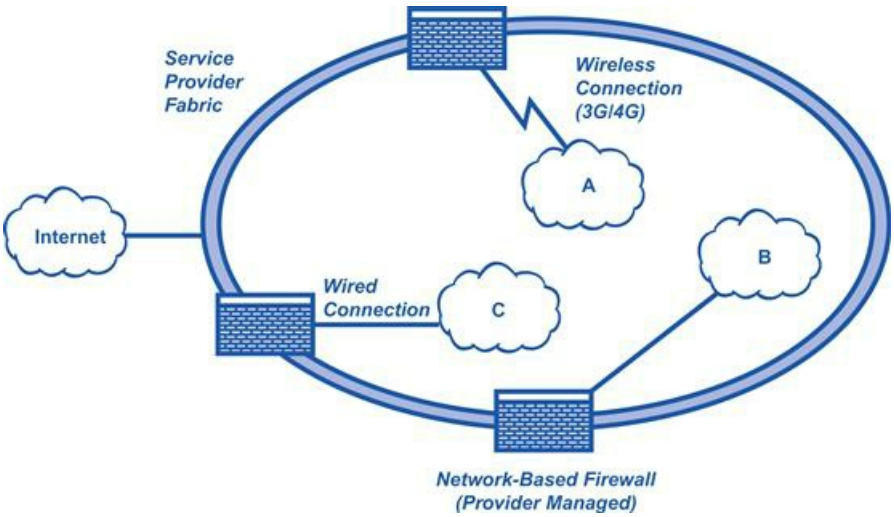


Figure 3.5 Carrier-centric network-based firewall.

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

DDOS Filtering

A major application of the network-based firewall concept includes a special type of mediation device embedded in the wide area network for the purpose of throttling distributed denial of service (DDOS) attacks. This device, which can be crudely referred to as a *DDOS filter*, is essential in modern networking, given the magnified risk of DDOS attacks from botnets. Trying to filter DDOS attacks at the enterprise edge does not make sense given the physics of network ingress capacity. If, for example, an enterprise has a 1-Gbps ingress connection from the Internet, then a botnet directing an inbound volume of anything greater than 1 Gbps will overwhelm the connection.

The risk of DDOS attacks must be effectively addressed.

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

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

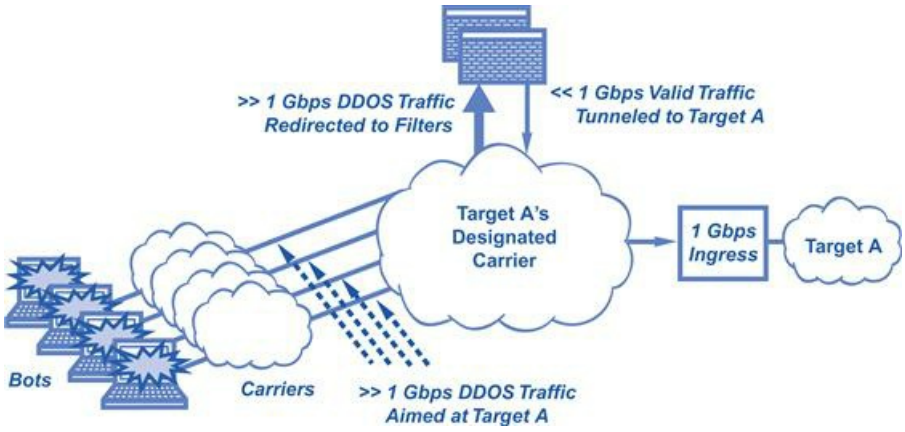


Figure 3.6 DDOS filtering of inbound attacks on target assets.

Quantitative analysis associated with DDOS protection of national infrastructure is troubling. If, for example, we assume that bots can easily steal 500 Kbps of broadband egress from the unknowing infected computer owner, then it would only require three bots to overwhelm a T1 (1.5-Mbps) connection. If one

carries out this argument, then botnets with 16,000 bots are sufficient to overwhelm a 10-Gbps connection. Given the existence of prominent botnets such as Storm and Conficker, which some experts suggest could have as many as 2 or 3 million bots, the urgency associated with putting DDOS filtering in place cannot be understated. An implication is that national infrastructure protection initiatives must include some measure of DDOS filtering to reduce the risk of DDOS attacks on national assets.

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

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

The Great Challenge of Filtering Out DDOS Attacks

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

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

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

SCADA Separation Architecture

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

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

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

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

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

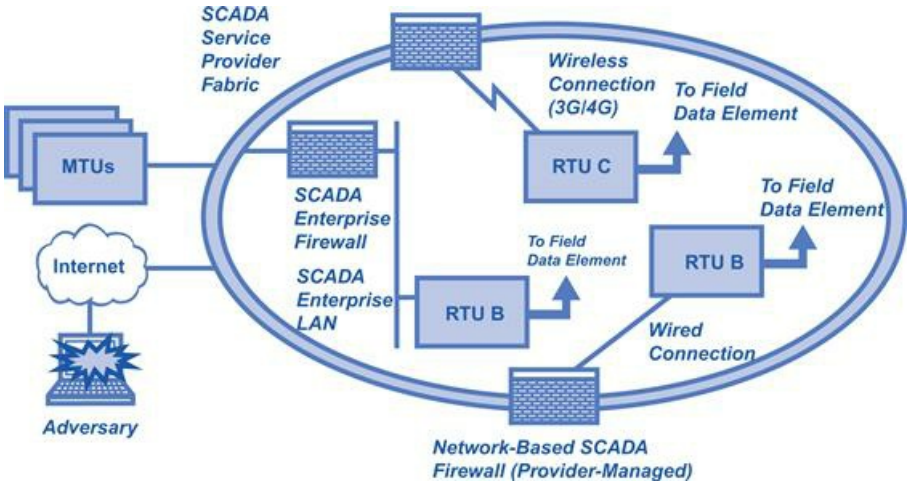


Figure 3.7 Recommended SCADA system firewall architecture.

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

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

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

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

Physical Separation

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

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

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

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

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

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

Perhaps the biggest threat to physical network isolation involves dual-homing a system to both an enterprise network and some external network such as the Internet. Such dual-homing can easily arise where

an end user utilizes the same system to access both the isolated network and the Internet. As laptops have begun to include native 3 G wireless access, this likelihood of dual-homing increases. Regardless of the method, if any sort of connectivity is enabled simultaneously to both systems, then the end user creates an inadvertent bridge (see Figure 3.8).

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

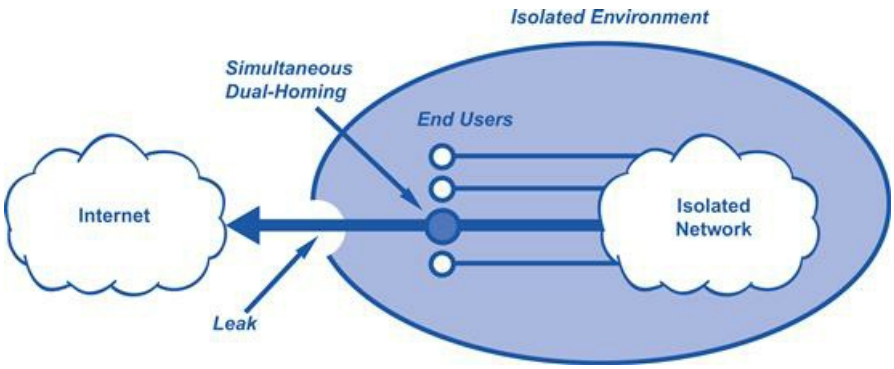


Figure 3.8 Bridging an isolated network via a dual-homing user.

It is worth mentioning that the bridge referenced above does not necessarily have to be established simultaneously. If a system connects to one network and is infected with some sort of malware, then this can be spread to another network upon subsequent connectivity. For this reason, laptops and other mobile computing devices need to include some sort of native protection to minimize this problem. Unfortunately, the current state of the art for preventing malware downloads is poor.

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

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

Insider Separation

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

An adversarial threat may come from a trusted partner.

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

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

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

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

identifiable information (PII).

- *Data leakage protection (DLP) systems*—Techniques for sniffing gateway traffic for sensitive or inappropriate materials are becoming common. Tools called DLP systems are routinely deployed in companies and agencies. At best, they provide weak protection against insider threats, but they do help identify erroneous leaks. Once deployed, they provide statistics on where and how insiders might be using corporate systems to spill information. In practice, however, no knowledgeable insider would ever be caught by a data leakage tool. Instead, the leak would be done using non-company-provided computers and networks.

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

Segregation of duties offers another layer of protection.

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

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

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

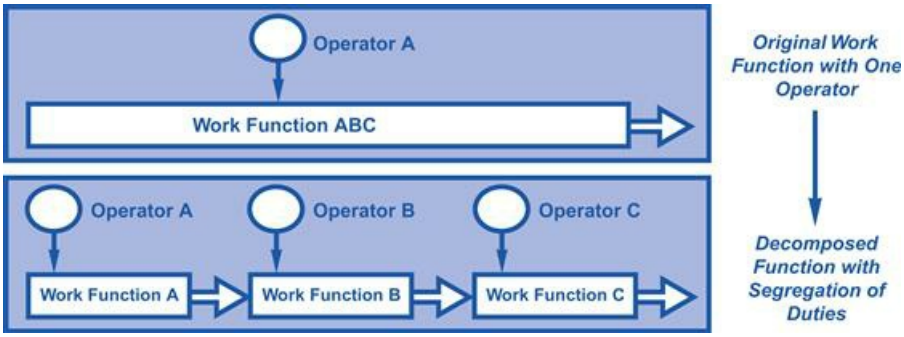


Figure 3.9 Decomposing work functions for segregation of duty.

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

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

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

Asset Separation

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

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

Segregation is one method of separation.

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

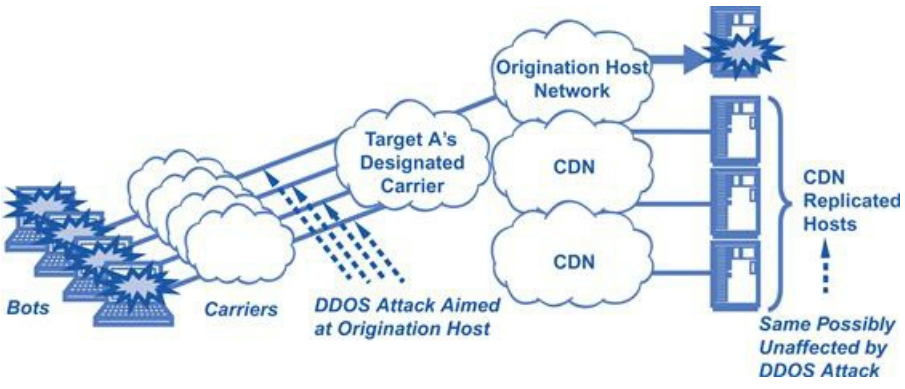


Figure 3.10 Reducing DDOS risk through CDN-hosted content.

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

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

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

Multilevel Security (MLS)

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

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

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

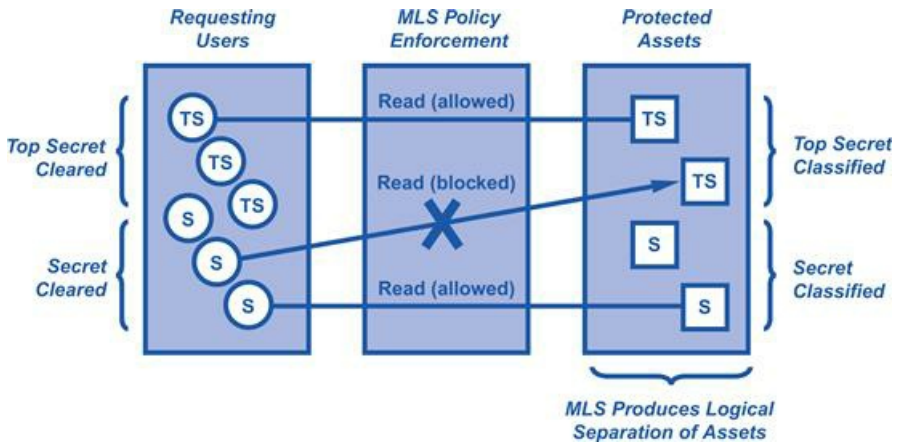


Figure 3.11 Using MLS logical separation to protect assets.

Several models of computer system behavior with such MLS functionality were developed in the early years of computer security. The Bell-La Padula disclosure and Biba integrity models are prominent examples. Each of these models stipulated policy rules that, if followed, would help to ensure certain desirable security properties. Certainly, there were problems, especially as networking was added to isolated secure systems, but, unfortunately, most research and development in MLS dissolved mysteriously in the mid-1990s, perhaps as a result of the economic pull of the World Wide Web. This is unfortunate, because the functionality inherent in such MLS separation models would be valuable in today's national infrastructure landscape. A renewed

interest in MLS systems is thus strongly encouraged to improve protection of any nation's assets.

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

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

Implementing a National Separation Program

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

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

Finally, let's briefly look at some practical ways to protect the critical national infrastructure through use

of separation techniques. Current threats and vulnerabilities are also covered.

Protecting the Critical National Infrastructure Through Use of Separation

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

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

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

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

1. Implement real-time threat protection.
2. Separate and protect critical assets from interconnected networks by taking the following actions:

a. Control port access based on a positive security model (i.e., they deny all access except that which is explicitly allowed).

b. Operate at gigabit speed and, therefore, do not interfere with control system availability and integrity standards.

c. Include specific capabilities designed for control systems.

d. Deliver a truly hardened operating system (not just a modified commercial one) that can defend itself from attacks, prevent or eliminate root access, and restrict access escalation or arbitrary code execution by any outside party.

e. Eliminate all unconstrained privileges and extraneous services, including network stack separation and control of super-user privileges, while providing triggers for intrusion detection.

f. Provide easy-to-deploy and manage architecture with central policies, reporting, and strong forensics.

g. Automatically filter out connections from locations that are suspicious or unnecessary to normal operations.

h. Scan encrypted traffic (HTTPS, SSL, SSH, SFPT, SCP, etc.) to uncover and block hidden attacks

i. Provide strong industry and government certifications and references (Common Criteria certification of EAL4+ is the minimum level suggested).

j. Deliver a security architecture that has a long and proven history of never being breached or hacked.

3. Provide a suitable intrusion prevention security (IPS) solution by taking the following actions:

a. Provide real-time protection from known, zero-day, DoS, DDoS, SYN flood, and encrypted attacks, as well as threats such as spyware, VoIP vulnerabilities, botnets, malware, worms, Trojans, phishing, and peer-to-peer tunneling.

b. Maximize accuracy by using multiple advanced detection methods, including signature, application, and protocol anomaly; shell-code detection algorithms; and next-generation DoS and DDoS prevention.

c. Parse over 200 protocols and review over 6,000 high-quality, multitoken, multitrigger signatures with stateful traffic inspection.

d. Offer proactive, out-of-the-box blocking for hundreds of attacks by featuring preconfigured "recommended for blocking" policies.

e. Receive continuous threat updates 24/7 from global research teams.

4. Control user access and network activities.

5. Protect information about critical assets from data leakage by taking the following actions:

a. Drop

b. Blind copy

c. Replace

d. Drop a portion or even the entire message

e. Forward in line or as an attachment

f. Quarantine

g. Reroute

h. Prepend

i. Log

j. Encrypt for secure delivery

k. Rewrite the subject line

l. Notify employees, managers, compliance officers, etc.

m. Archive

n. Educate users on rules

6. Implement strong security without jeopardizing availability, integrity, and reliability requirements by taking the following actions:

a. Perform automatic updates that don't require critical assets be taken off line

b. Support the long asset lifetimes of critical assets

c. Minimize the need for extensive testing and downtime before patches can be applied

d. Protect against threats that have yet to be identified

e. Prevent privilege escalation vulnerabilities

f. Support the custom and relevant signatures specific to critical networks

g. Perform security at speeds that won't impact network performance

h. Deploy a trusted security model based on reputation and an in-depth understanding of applications

Summary

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

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

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

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

Chapter Review Questions/Exercises

True/False

1. True or False? The separation of network assets from malicious intruders using a firewall is perhaps the least familiar protection approach in all of computer security.
2. True or False? Functional separation of an adversary from any computing asset is most commonly achieved using an access control mechanism with the requisite authentication and identity management.
3. True or False? The most common application of a firewall involves its placement between a system or enterprise to be protected and some trusted network such as the Internet.
4. True or False? A major application of the network-based firewall concept includes a special type of mediation device embedded in the wide area network for the purpose of throttling distributed denial of service (DDOS) attacks.
5. True or False? Many critical national infrastructure systems do not include supervisory control and data acquisition (SCADA) functionality.

Multiple Choice

1. The general structure of SCADA systems includes the following components, except which one?

A. Clear policy
B. Human-machine interface (HMI)
C. Master terminal unit (MTU)
D. Remote terminal unit (RTU)
E. Field control systems

2. To create a truly air-gapped network, which one of the following basic principles is needed?

A. Scanning
B. Boundary scanning
C. Exploitation
D. Discovery
E. Exposing

3. Techniques for reducing the risk of unauthorized insider access do exist that can be embedded in the design and operation of national infrastructure operation. These techniques include the following, except which one?

A. Internal firewalls
B. Deceptive honey pots
C. Segregation of duties
D. Enforcement of data markings
E. Data leakage protection (DLP) systems

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

A. Distribution
B. Replication
C. Decomposition
D. Misconfiguration
E. Segregation

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

A. Investment separation
B. Operations separation
C. Accuracy separation
D. Coverage separation
E. Internet separation

Exercise

Problem

Recently, the Pentagon has concluded that computer sabotage coming from another country can constitute an act of war, a finding that for the first time opens the door for the United States to respond using traditional military force, which would also include the use of nuclear tactical weapons, as well as armed unmanned drones. The Pentagon's first formal cyber strategy, unclassified portions of which became public in June of 2011, represents an early attempt to grapple with a changing world in which a hacker could pose as significant a threat to U.S. nuclear reactors, subways, or pipelines as a hostile country's military. In part, the Pentagon intends its plan as a warning to potential adversaries of the consequences of attacking the United States in this way. In other words, if a foreign state (China, North Korea, Iran, etc.) shuts down the United States' power grid, maybe the U.S. military will put an electromagnetic pulse (EMP) missile down on the foreign country's power grid and take it out. Recent cyber attacks on the Pentagon's own systems by another foreign country (which resulted in the penetration and extraction of very sensitive military information), have given new urgency to U.S. efforts to develop a more formalized approach to cyber attacks. Please explain what type of military response is appropriate in resolving this cyber attack problem.

Hands-On Projects

Project

The Department of Homeland Security's Cyber Storm IT security exercise found problems and some strengths in the United States' ability to respond to simulated attacks on the electronic critical national infrastructure, highlighting areas where the government and private organizations must improve their responsiveness to emerging IT-related threats. It was the largest and most complex multinational, government-led cyber exercise to examine response, coordination, and recovery mechanisms to a simulated cyber event.

The Cyber Storm test was launched to help gauge the information-sharing capabilities and IT attack readiness of government branches on the federal, state, and local level. Also, part of the study was those groups' abilities to cooperate with foreign nations and private sector organizations in the event of a major attack or natural disaster. In other words, Cyber Storm was meant to provide participants with a controlled environment in which they could simulate the coordination that would be necessary during a cyber-related incident of national significance, such as an attack on the critical national infrastructure supporting the nation's Internet operations or a natural disaster like Hurricane Katrina.

Funded by the federal government and mandated by Congress, the test included over 200 public and private organizations at over 70 locations in six countries that collaborated as they would in the case of such a crisis. The exercise was meant to recreate the conditions a cyber attack or disaster could have on operations related to the nation's energy, IT, transportation, and telecommunications sectors.

In many ways, this exercise was designed to push the system to the maximum edge. That allows the participants to identify the greatest points of vulnerability.

Parties involved in the test staged primary cyber attacks targeting the energy, transportation, and IT/telecommunications sectors that were intended to disrupt certain elements of critical national infrastructure. The cyber attacks were meant to touch off potentially cascading effects within other elements of the United States and participating countries' economic, social, and governmental structures.

Some of the cyber attacks in the exercise were aimed specifically at disrupting government operations that would be used to respond to a cyber threat in the name of undermining public confidence in those entities. So, how would you go about creating such an exercise project with the use of separation techniques to protect the critical national infrastructure from cyber attacks?

Case Projects

Problem

Let's look at a real-world scenario of how both government and the private sector are struggling to provide a secure, efficient, timely, and separate means of delivering essential services around the world. As a result, these critical national infrastructure systems remain at risk from potential attacks via the Internet. It is the policy of the United States to prevent or minimize disruptions to the critical national information infrastructure in order to protect the public, the economy, government services, and the national security of the United States.

The Federal Government is continually increasing capabilities to address cyber risk associated with critical networks and information systems. Please explain how you would reduce potential vulnerabilities, protect against intrusion attempts, and better anticipate future threats.

Optional Team Case Project

Problem

Countries need to step up international cooperation to protect the critical national infrastructure against increasingly sophisticated cyber threats. Please identify how you would go about using the separation of critical assets to protect the critical infrastructure and measure its resiliency to ensure performance, stability, and cyber security.

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

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

4

Diversity

Chapter Outline

Diversity and Worm Propagation
Desktop Computer System Diversity
Diversity Paradox of Cloud Computing
Network Technology Diversity
Physical Diversity
National Diversity Program
Critical Infrastructure Resilience and Diversity Initiative
Summary
Chapter Review Questions/Exercises

We are looking at computers the way a physician would look at genetically related patients, each susceptible to the same disorder.

Mike Reiter, professor of electrical and computer engineering and computer science at Carnegie-Mellon University[1]

Making national infrastructure more diverse in order to create greater resilience against cyber attack seems to be a pretty sensible approach. For example, natural scientists have known for years that a diverse ecosystem is always more resilient to disease than a monoculture. When a forest includes only one tree, the possibility arises that a single disease could wipe out the entire ecosystem. This type of situation arises even in business. Certain airlines, for example, have decided to use only one model of aircraft. This reduces the cost of maintenance and training but does create a serious risk if that particular aircraft were grounded for some reason. The airline would be out of business—a risk that is avoided by a diversity approach.

So it would stand to reason that the process of securing any set of national assets should always include some sort of diversity strategy. This diversity should extend to all applications, software, computers, networks, and systems. Unfortunately, with the exception of familiar geographic requirements on network routes and data centers, diversity is not generally included in infrastructure protection. In fact, the topic of deliberately introducing diversity into national infrastructure to increase its security has not been well explored by computer scientists. Only recently have some researchers begun to investigate the benefits of diversity in software deployment.

Introducing diversity at all levels of functionality has not been properly explored as a protection strategy.

Diversity in national infrastructure involves the introduction of intentional *differences* into systems. Relevant differences include the vendor source, deployment approach, network connectivity, targeted standards, programming language, operating system, application base, software version, and so on. Two systems are considered diverse if their key attributes differ, and nondiverse otherwise (see Figure 4.1).

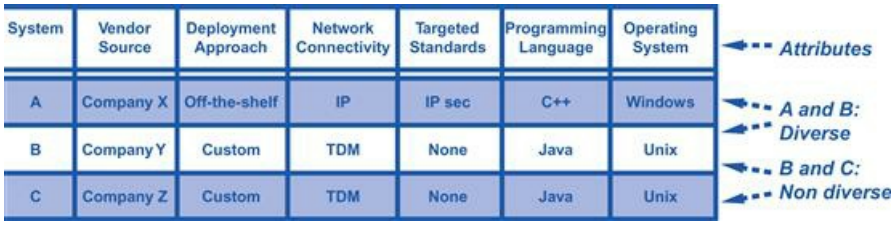


Figure 4.1 Diverse and nondiverse components through attribute differences.

The general idea is that an adversary will make assumptions about each of the relevant attributes in a target system. In the absence of diversity, a worst-case scenario results if the adversary makes the right assumptions about each attribute. If, for example, the adversary creates an attack on a set of computers that assumes an underlying Microsoft® operating system environment, and the national asset at risk employs only these types of systems, then the effect could be significant. In the presence of diversity, however, it becomes much more difficult for an adversary to create an attack with maximal reach. This is especially relevant for attacks that are designed to automatically propagate. Eventually, the attack will reach a point where it can no longer copy itself or remotely execute, and the process will cease.

Diversity increases the number of assumptions an adversary has to make about the system and creates more potential for an adversary's plan to fail.

Standardized operations are important for compliance but are at odds with diversity.

Why, then, is diversity so underrepresented in national infrastructure protection? To understand this, one must first recognize the near-obsessive goal of enforcing sets of common standards that the information technology and security communities have attempted to achieve. In nearly every facet of computing, sets of standard, auditable practices have been defined and backed by powerful organizations. In the United States, the Sarbanes- Oxley standard has had a profound influence on the operation of every major corporation in the country, leading to more common approaches to financial systems operation. Commonality, as we discuss in the next chapter, is somewhat at odds with diversity.

This focus on maintaining common, standard operating environments should not come as a surprise. The rise of the Internet, for example, was driven largely by the common acceptance of a single protocol suite. Even the provision of Internet-based services such as websites and mail servers requires agreement among system administrators to follow common port assignments. Chaos would ensue if every administrator decided to assign random ports to their Internet services; end users would not be able to easily locate what they need, and the Internet would be a mess (although this would certainly complicate broad types of attacks). So, the

result is general agreement on common computing configurations.

Another key motivation to avoid diversity for most system managers is the costs involved. Typical computing and networking management teams have created programs focused on removing differences in enterprise systems in order to reduce operating expenses. Clearly, nondiverse information technology systems simplify platform deployment, end-user training, system administrative practices, and system documentation. For these cost-related reasons, diversity is generally not a prominent goal in most current national infrastructure settings. The result is less secure infrastructure.

Diversity currently competes with commonality and cost savings.

Diversity and Worm Propagation

The self-propagation of a computer worm is a good example of an attack that relies on a nondiverse target environment to function properly. The box shows how relatively simple an attack can be.

Worm Functionality in Three Easy Steps

The functionality of a typical, generic computer worm is quite straightforward (only three steps) and can be described in simple pseudo-code terms as follows:

Program: *Worm*
Start
Step 1. Find a target system on the network for propagation of Program Worm.
Step 2. Copy Program *Worm* to that target system.
Step 3. Remotely execute Program *Worm* on that target system.

Repeat Steps 1 through 3.

As you can see, a worm program relies on the ability to find common, reachable, interoperable systems on the network that will accept and execute a copy of the worm program. In the early days of the Internet, this would be accomplished by checking a local file that would include a list of systems that were reachable. Today, it's done by creating batches of Internet Protocol addresses. Also, in those early days, it was quite easy to copy and execute programs from one system to another, because no one had yet invented the firewall.

A worm propagates by finding interoperable systems to target.

One would have hoped that the global deployment of firewalls would have stopped the ability of adversaries to create worms, but sadly it has not. Instead, vulnerabilities or services open through the firewalls are used as the basis for worms. Nondiversity in such setups is also the norm. This is unfortunate, because if a worm operates in a diverse environment, and thus cannot find systems that consistently meet one or more of these criteria, then its propagation will cease more rapidly. This can be depicted in a simple reachability diagram showing the point of initiation for the worm through its propagation to the final point at which the activity ceases as a result of diversity. As the worm tries to propagate, diversity attributes that reduce its ability to locate reachable systems, make copies, and remotely execute are the most effective (see Figure 4.2).

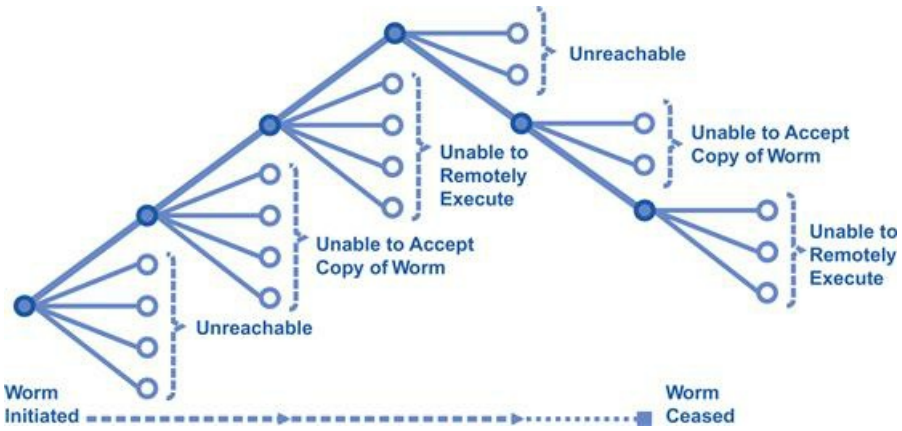


Figure 4.2 Mitigating worm activity through diversity.

Obviously, all worms will eventually cease to propagate, regardless of the degree of diversity in a given network. The security advantage one gains with diversity is that the worm is likely to cease more quickly and perhaps without human intervention. Empirical experience in the global security community dealing with worms such as the SQL/Slammer and Blaster worms of 2003 and the Sasser worm of 2004 suggest that significant human intervention is required to halt malicious operation. During the early hours of the SQL/Slammer worm, most of the security incident response calls involved people trying to figure out what to do. Eventually, the most effective solution involved putting local area network blocks in place to shut down the offending traffic. By the time the event died down, many millions of hours of global labor had been expended working on the problem. By increasing diversity, one should expect to reduce response costs around the world associated with fighting worms.

Although introducing security can seem expensive, one should expect to save money on response costs with an effective diverse environment.

The real challenge here is that both the Internet and the networks and systems being run by companies and agencies charged with national infrastructure are simply not diverse—and there is little discussion in place to alter this situation. As we suggested earlier, this is driven largely by the goal to maximize interoperability. There are some exceptions in the broader computing community, such as digital rights management (DRM)-based systems that have tended to limit the execution of certain content applications to very specific devices such as the iPod® and iPhone®. The general trend, however, is toward more open, interoperable computing. What this means is that, for national infrastructure components that must be resilient against automated attacks such as worms, the threat will remain as long as the networking environment is a monoculture.

Desktop Computer System Diversity

Typical individual computer users in the home or office, regardless of their location in the world, are most likely to be using a commercial operating system running on a standard processor platform and utilizing one of a couple of popular browsers to perform searches on a popular search engine. This might seem an obvious statement, but in the early days of computing there were many users on home-grown or proprietary systems using all sorts of software that might only be known locally.

The average home PC user is working in a highly predictable computing environment.

Today, however, the most likely configuration would be a Windows®-based operating system on an Intel® platform with Internet Explorer® being used for Google® searches. We can say this confidently, because almost all current estimates of market share list these products as dominant in their respective fields. Certainly, competing platforms and services from Apple® and others have made inroads, but for the most part, especially in business and government environments, the desktop configuration is highly predictable (see Figure 4.3).

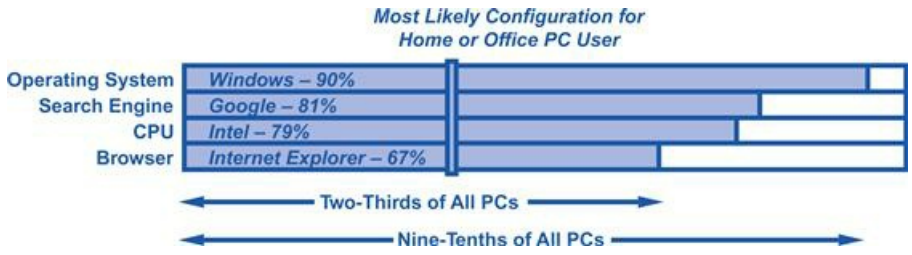


Figure 4.3 Typical PC configuration showing nondiversity.

This dominant position for these few companies has admittedly led to a number of positive results. It has, for instance, pushed a deeper common understanding of computing among individuals around the world. Different people from different cultures around the world can share their experiences, recommendations, and suggestions about operating systems, search engines, CPUs, and browsers, and the likelihood of applicability is high. The dominant position of these respective products has also helped the software development industry by creating rich and attractive common target markets. Developers generally love to see a dominant platform configuration, because it increases their potential profits through maximal usage. So, computing certainly has moved forward as a result of commonality; not much disagreement exists on this point.

The drawback from a national infrastructure perspective, however, is that adversaries will have an easier time creating attacks with significant reach and implication. Just as a game of dominoes works best when each domino is uniformly designed and positioned, so does common infrastructure become easier to topple with a single, uniform push. In some cases, the effect is significant; the operating system market on desktop PCs, for example, is dominated by Microsoft® to the point where a well-designed Windows®-based attack could be applicable to 90% of its desktop targets.

Targeting the most popular operating system software with a worm attack could bring the majority of PCs to a standstill.

More likely, however, is the situation where the creation of a botnet becomes much easier given the nondiversity of PC configurations. When a botnet operator conceptualizes the design of a new botnet, the most important design consideration involves reach. That is, the botnet operator will seek to create malware that has the maximal likelihood of successfully infecting the largest number of target PCs. As such, the nondiversity of end-user configurations plays right into the hands of the botnet operator. Combine this with the typically poor system administrative practices on most PCs, and the result is lethal. Worse, many security managers in business and government do not understand this risk. When trying to characterize the risk of attack, they rarely understand that the problem stems from a global set of nondiverse end-user PCs being mismanaged by home and office workers.

Security managers are unlikely to consider the home PC user when assessing risk.

In response to this threat, national infrastructure protection requires a deliberate and coordinated introduction of diversity into the global desktop computing environment. Enterprise attention is obviously different than that of individuals in homes, but the same principle applies. If the desktop computing assets that can reach a national asset must be maximally resilient, then desktop diversity is worth considering. The most obvious challenge here is related to the consumer marketplace for PCs; that is, the reason why consumers use the same platform is because they prefer it and have chosen to purchase it. If Microsoft® and Intel®, for example, were not providing value in their products, then people would buy something else. The biggest hurdle, therefore, involves enabling nondiversity without altering the ability of companies to provide products that people like to use. Perhaps this goal could be accomplished via diversity elements coming from within the existing vendor base.

Desktop Diversity Considerations

Additional issues that arise immediately with respect to desktop diversity programs include the following:

- *Platform costs*—By introducing multiple, diverse platforms into a computing environment, the associated hardware and software costs might increase. This is a common justification by information technology (IT) managers for avoiding diversity initiatives. Certainly, the procurement of larger volumes of a given product will reduce the unit cost, but by introducing competition into the PC procurement arena increased costs might be somewhat mitigated.
- *Application interoperability*—Multiple, diverse platforms will complicate organizational goals to ensure common interoperability of key applications across all platforms. This can be managed by trying to match the desktop platform to local needs, but the process is not trivial. The good news is that most web-based applications behave similarly on diverse platforms.
- *Support and training*—Multiple, diverse platforms will complicate support and training processes by adding a new set of vendor concerns. In practical terms, this often means introducing a platform such as Mac OS® to a more traditional Windows®-based environment. Because many consumers are

comfortable with both platforms, especially youngsters who tend to be more diverse in their selections, the problem is not as intense as it might be.

For national infrastructure protection, desktop diversity initiatives that are focused on ensuring enterprise differences in companies and agencies have a good chance of success. Rewards and incentives can be put in place to mix up the desktop platforms in a given enterprise. The problem is that this will have only limited usefulness from the perspective of botnet design and recruitment. The real advantage would come from diversity in broadband-connected PCs run by consumers around the world. Unfortunately, this is not something that can be easily controlled via an initiative in any country, including the United States.

Global diversity in broadband-connected home PCs would stymie many botnet attacks.

Interestingly, a related problem that emerges is the seemingly widespread software piracy one finds in certain areas of the globe. Software piracy on the desktop introduces the problem of security updates; that is, depending on the specifics of the theft, it is often difficult for pirated PCs to be properly protected with required patches. When many millions of PCs are in this state, the problem of nondiversity becomes all the more severe.

Diversity Paradox of Cloud Computing

To better understand how diversity goals can be accomplished, it helps to introduce a simple model of desktop computing systems. The model is represented as a linear spectrum of options related to the degree to which systems are either diverse or nondiverse. As such, the two ends of the model spectrum are easy to identify for a given environment. On one side of the spectrum would be the option of complete nondiversity, where every desktop system in the organization, enterprise, or group is exactly the same. On the other side of the spectrum would be the option of complete diversity across the organization, where no two desktop systems are the same. In the middle of the spectrum would be the usual types of settings, where some minor degree of diversity exists, but with a clearly dominant platform.

The model spectrum is useful because it allows illustration of our basic infrastructure security proposition around PCs—namely, as diversity increases, desktop attacks, including the use of worms to create a local denial of service condition, are more difficult to accomplish. One might also suggest that the creation and use of botnets would also be more difficult, but this benefit might be more modest (see Figure 4.4).



Figure 4.4 Spectrum of desktop diversity options.

In fact, diverse desktops are tougher to uniformly compromise, because they are less conducive as a group to a scalable, self-propagating attack. For example, if a company has half of its PCs running Windows®-based operating systems and half running Mac OS®-based operating systems, then this will clearly be more challenging for an automatically propagating attack. Hence, the level of diversity and the associated difficulty of attack appear to correlate. A challenge with this view, however, is that it does not properly characterize the optimal choice in reducing desktop attack risk—namely, the *removal* of desktops from the target environment. After all, one cannot attack systems that are not even there. This suggests a new (and admittedly theoretical) diversity and attack difficulty spectrum (see Figure 4.5).

As the level of diversity increases, the level of difficulty for an attack likewise increases.

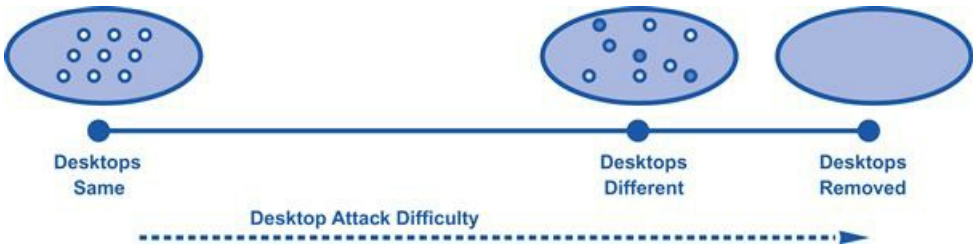


Figure 4.5 Diversity and attack difficulty with option of removal.

This suggests that the ultimate (albeit impossible) option for making desktops more secure involves their removal. Obviously, this is not a practical goal, but computer security objectives are often made more tractable via clear statements of the ideal condition. So, while current enterprise or home computing architectures do not include the option of having no desktop computers, older readers will remember the days when desktops did not exist. Rather, people used computer terminals to access information on mainframes, and security benefits were certainly present in such a setup. This included no need for end-user software patching, as well as no end-user platform for targeted malware. One great irony in the present deployment of desktops to every man, woman, and child on the planet is that most people really do not need such computing power. It is likely that they would be just fine with a keyboard, screen, and mouse connected to network-resident applications that are ubiquitously available via the Internet.

The global proliferation of home PCs has increased the risk of malware attacks.

In modern computing, the closest thing we have to this arrangement is virtualized, cloud-based computing. In such a setup, computing power and application intelligence move to a centralized complex of servers, accessible via light clients. In fact, handheld mobile devices provide the equivalent of a desktop computer in such a cloud environment. One should therefore presume, from the diagram in Figure 4.5, that cloud computing would provide considerable security benefits by removing nondiverse desktops from the environment. This is most likely true, as long as the infrastructure supporting the cloud applications is properly secured, as per the various principles described in this book. If this is not the case, then one is simply moving nondiversity vulnerabilities from the desktops to the servers.

Cloud computing may offer home PC users the diverse, protected environment they cannot otherwise access.

Network Technology Diversity

Modern telecommunications network systems can be viewed as consisting of the following two basic types of technologies:

- *Circuit-switched*—This includes legacy, circuit-switched systems that support traditional plain old telephone services (POTS) and related voice and data services. The public switched telephone network (PSTN) is the most significant example of deployed circuit-switched technology.
- *Packet-switched*—This includes more modern, packet-switched systems that support Internet Protocol (IP) and related voice, data, and multimedia services. In addition to the Internet as the most obvious example of packet switching, the signaling network controlling the PSTN is itself a packet-switched system.

For the most part, both logical and physical diversity naturally exist between these two types of services, largely due to technology interoperability. That is, the vast majority of equipment, software, processes, and related infrastructure for these services are fundamentally different. Packets cannot accidentally or intentionally spill into circuits, and *vice versa*.

Circuit-switched and packet-switched systems automatically provide diversity when compared to one another.

From a networking perspective, what this means is that a security event that occurs in one of these technologies will generally not have any effect on the other. For example, if a network worm is unleashed across the Internet, as the global community experienced so severely in the 2003–2004 time frame, then the likelihood that this would affect traditional time-division multiplexed (TDM) voice and data services is negligible. Such diversity is of significant use in protecting national infrastructure, because it becomes so much more difficult for a given attack such as a worm to scale across logically separate technologies (see Figure 4.6).

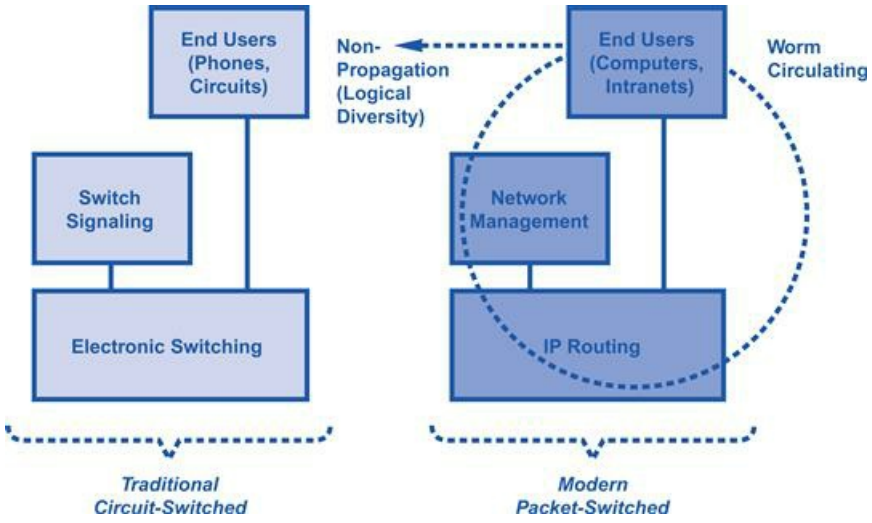


Figure 4.6 Worm nonpropagation benefit from diverse telecommunications.

Even with the logical diversity inherent in these different technologies, one must be careful in drawing conclusions. A more accurate view of diverse telecommunications, for example, might expose the fact that, at lower levels, shared transport infrastructure might be present. For example, many telecommunications companies use the same fiber for their circuit-switched delivery as they do for IP-based services. Furthermore, different carriers often use the same right-of-way for their respective fiber delivery. What this means is that in many locations such as bridges, tunnels, and major highways, a physical disaster or targeted terrorist attack could affect networks that were designed to be carrier diverse.

Unfortunately, vulnerabilities will always be present in IP-based and circuit-switched systems.

While sharing of fiber and right-of-way routes makes sense from an operational implementation and cost perspective, one must be cognizant of the shared infrastructure, because it does change the diversity profile. As suggested, it complicates any reliance on a multivendor strategy for diversity, but it also makes it theoretically possible for an IP-based attack, such as one producing a distributed denial of service (DDOS) effect, that would have negative implications on non-IP-based transport due to volume. This has not happened in practical settings to date, but because so much fiber is shared it is certainly a possibility that must be considered (see Figure 4.7).

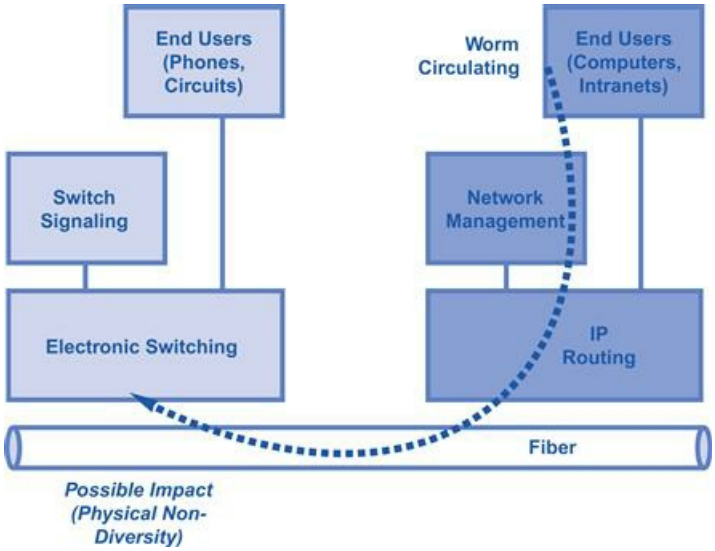


Figure 4.7 Potential for impact propagation over shared fiber.

A more likely scenario is that a given national service technology, such as modern 2G and 3G wireless services for citizens and business, could see security problems stemming from either circuit- or packet-switched-based attacks. Because a typical carrier wireless infrastructure, for example, will include both a circuit- and packet-switched core, attacks in either area could cause problems. Internet browsing and multimedia messaging could be hit by attacks at the serving and gateway systems for these types of services; similarly, voice services could be hit by attacks on the mobile switching centers supporting this functionality. So, while it might be a goal to ensure some degree of diversity in these technology dependencies, in practice this may not be possible.

Diversity may not always be a feasible goal.

What this means from a national infrastructure protection perspective is that maximizing diversity will help to throttle large-scale attacks, but one must be certain to look closely at the entire architecture. In many cases, deeper inspection will reveal that infrastructure advertised as diverse might actually have components that are not. This does not imply that sufficient mitigations are always missing in nondiverse infrastructure, but rather that designers must take the time to check. When done properly, however, network technology diversity remains an excellent means for reducing risk. Many a security officer will report, for example, the comfort of knowing that circuit-switched voice services will generally survive worms, botnets, and viruses on the Internet.

Physical Diversity

The requirement for physical diversity in the design of computing infrastructure is perhaps the most familiar of all diversity-related issues. The idea is that any computing or networking asset that serves as an essential component of some critical function must include physical distribution to increase its survivability. The approach originated in the disaster recovery community with primary emphasis on natural disasters such as hurricanes and fires, but, as the security threat has matured, infrastructure managers have come to recognize the value of providing some degree of physical diversity. This reduces, for example, reliance on a single local power grid, which is a valued cyber attack target for adversaries. It also greatly reduces the chances of a physical or premise-based attack, simply because multiple facilities would be involved.

Physical diversity adds another important layer of protection against cascading effects.

These issues are not controversial. In fact, for many years, procurement projects for national asset systems, in both government and industry, have routinely included the demand that the following physical diversity issues be considered:

- *Backup center diversity*—If any major center for system, network, or application management is included in a given infrastructure component, then it is routinely required that a backup center be identified in a physically diverse location. Few would argue with this approach; if properly applied, it would ensure that the two centers are in different weather patterns and power grid segments.

Physical diversity has been incorporated into the national asset system for many years.

- *Supplier/vendor diversity*—Many organizations dictate that for critical infrastructure components, some degree of diversity must be present in the supplier and vendor mix. This reduces the likelihood that any given firm would have too much influence on the integrity of the infrastructure. It also reduces the likelihood of a cascading problem that might link back to some common element, such as a software routine or library, embedded in one vendor's product portfolio.
- *Network route diversity*—When network infrastructure is put in place to support national infrastructure, it is not uncommon to demand a degree of network route diversity from the provider or providers. This helps reduce the likelihood of malicious (or nonmalicious) problems affecting connectivity. As mentioned above, this is complicated by common use of bridges, tunnels, or highways for physical network media deployments from several different vendors.

Achieving Physical Diversity via Satellite Data Services

A good example application that demonstrates physical diversity principles is the provision of certain types of SCADA systems using IP over satellite (IPoS). Satellite data services have traditionally had the great advantage of being able to operate robustly via the airwaves in regions around the globe where terrestrial network construction would be difficult. Generally, in such regions commercial wireless coverage is less

ubiquitous or even completely unavailable. Some SCADA applications have thus taken advantage of this robust communication feature in satellite systems to connect remote end-user terminals to the SCADA host system, but the requirement remains that some degree of diversity be utilized. As suggested above, most of this diversity emphasis has been driven largely by concerns over natural and physical disasters, but a clear cyber security benefit exists as well.

Generally, the setup for satellite-connected SCADA involves end users connecting to a collection of physically diverse hubs via IPoS. These diverse hubs are then connected in a distributed manner to the SCADA hosts. An adversary seeking to attack these hubs would have to use either logical or electronic means, and a great degree of logistic effort would be required, especially if the hubs are located in different parts of the world. The Hughes Corporation, as an example, has been aggressive in marketing these types of configurations for SCADA customers. Their recommended remote access configuration for diverse SCADA system control is shown in Figure 4.8.

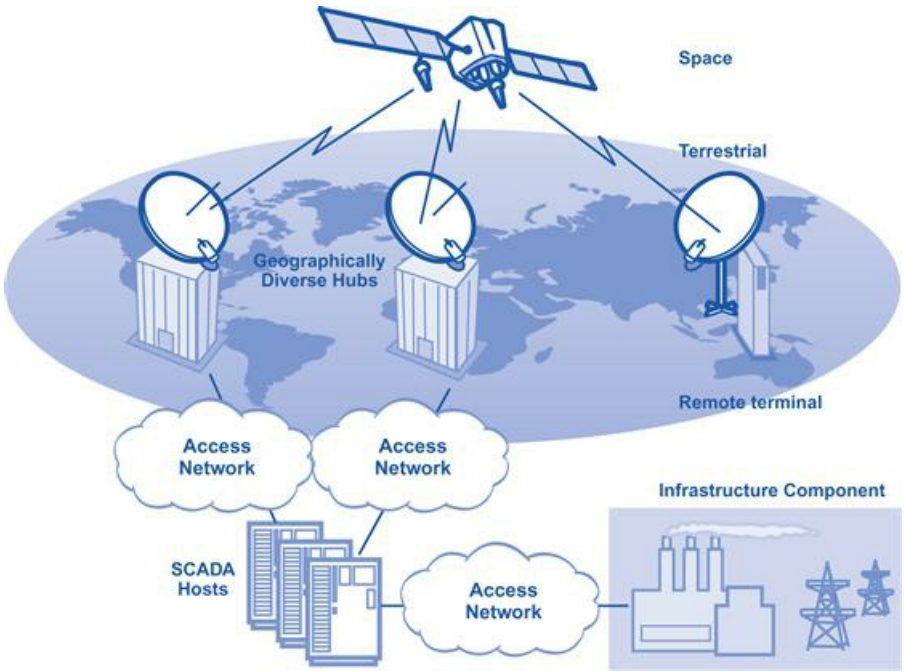


Figure 4.8 Diverse hubs in satellite SCADA configurations.

The advantage of diverse hubs is obvious; if any should be directly compromised, flooded, or attacked (physically or logically), then the SCADA hosts are still accessible to end users. In addition, attacks on local infrastructure components on which the SCADA operation depends, such as power, will not have a cascading effect. Such an approach only works, however, if all diverse components operate at a common service level. For example, if one service provider offers highly reliable, secure services with historical compliance to advertised service level agreements (SLAs), then introducing a diverse provider with poor SLA compliance might not be such a good idea. This is a key notion, because it is not considered reasonable to take a highly functioning system and make it diverse by introducing an inferior counterpart. In any event, this general concept of diverse relay between users and critical hosts should be embedded into all national infrastructure

systems.

National Diversity Program

The development of a national diversity program would require coordination between companies and government agencies in the following areas:

- *Critical path analysis*—An analysis of national infrastructure components must be made to determine certain critical paths that are required for essential services. For example, if a military group relies on a specific critical path to complete some logistic mission, then assurance should exist that this critical path is supported by diverse vendors, suppliers, support teams, and technology.
- *Cascade modeling*—A similar analysis is required to identify any conditions in a national infrastructure component where a cascading effect is possible due to nondiversity. If, for example, 100% of the PCs in an organization are running in exactly the same configuration, then this poses a risk. Admittedly, the organization might choose to accept the risk, but this should be done explicitly after a security investigation, rather than by default.
- *Procurement discipline*—The selection and procurement of technology by organizations charged with critical infrastructure should include a degree of diversity requirements. This generally occurs naturally in most large organizations, so the urgency here might not be as intense but the security benefits are obvious.

The decision of whether to provide rewards and incentives for diversity versus a stricter approach of requiring evidence of some targeted percentage of diversity must be driven by the local environment and culture. The threat environment in a military setting is considerably different than one might find in telecommunications or transportation, so it would seem prudent to make such implementation decisions locally.

Finally, let's briefly look at some practical ways to make the critical national infrastructure more diverse in order to create greater resilience against cyber attacks. The critical national infrastructure resiliency and diversity agenda for action focuses on telecommunications, natural gas, electric energy, and transportation.

Critical Infrastructure Resilience and Diversity Initiative

The economy, national security, welfare, opportunities, and freedoms afforded to U.S. citizens are all highly dependent upon a vast network of highly complex, automated, largely privately owned, and inextricably and interdependently operated national and global critical infrastructure systems and services. These critical cyber and physical infrastructures produce and distribute energy; enable communications; control transportation; ensure the availability of food, water, and emergency care; and moreover, provide every service and support, and every activity that defines and empowers the United States. The United States must provide an objective foundation for investment in and management of the effectiveness and efficiency of critical national infrastructure resiliency and diversity efforts (see "An Agenda for Action in Using the Critical National Infrastructure Resiliency and Diversity Initiative").

An Agenda for Action in Using the Critical National Infrastructure Resiliency and Diversity Initiative

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

1. Identifying regional critical national infrastructure concentrations and chokepoints whose operational disruption (regardless of cause) could adversely affect regional and national businesses and security.
2. Establishing desired infrastructure resiliency standards.
3. Establishing a regional rating as a result of the region's identified standards, quantified threats, vulnerabilities, and consequences.
4. Raising awareness of critical national infrastructure operational and interdependency issues and their potentially catastrophic consequences.
5. Identifying public information challenges and ways to improve dissemination of timely, actionable threat and warning information, and lessons learned/best practices to support infrastructure resiliency efforts and foster trust and coordination between all levels of government, business partners, and critical national infrastructure owners and operators.
6. Identifying critical national infrastructure technology development needs to provide for 21st-century infrastructure management.
7. Conducting economic probabilistic modeling of regional incidents; critical national infrastructure failures; and resulting societal, business, and economic consequences on the critical infrastructure, business, and government operations and regions being studied.
8. Promulgating critical national infrastructure resilience (CNIR) as the top-level strategic objective (the desired outcome) to drive national policy and planning.
9. Aligning policy and implementing directives for risk-based decision making with the CIR objective within the broad context of the homeland security mission.

10. Creating a framework of cascading national goals flowing from the top-level CIR objective.
11. Establishing and institutionalizing proactive mechanisms to ensure that critical national infrastructure policy and planning guidance continually evolves.
12. Establishing a governance structure that supports the diversity of stakeholders within and among sectors, as well as the realities of infrastructure placement and operation within communities.
13. Establishing an information-sharing regime explicitly linked to critical national infrastructure resiliency goals and governance—but integrated within an enterprise-wide information architecture.
14. Identifying and building a general equation for the economic value/loss of a critical national infrastructure failure including the cost and the loss of benefits, specifically by taking the following actions:

a. Identifying the key factors that determine the business economic valuation (BEV) of the critical national infrastructure to the business, region, and potentially the nation.
b. Identify the cost and loss of the services not delivered when there is a critical national infrastructure failure, and as a result government and/or businesses failure.
c. Identify all of the external relationships affected by this failure and quantify the impact on important constituencies such as customers, suppliers, banks, business partners, and others.

Summary

This chapter focused on practical ways to make the critical national infrastructure more diverse in order to create a greater resilience against cyber attacks. The objective of this chapter is to advance the national policies and strategies that will foster the development of more resilient critical national infrastructures. The recommendations contained herein leverage the foundation built by prior and ongoing critical national infrastructure protection programs, but assert that a future focus on resilience would establish a more appropriate basis for risk-based decision making.

This nation's critical national infrastructures (cyber and physical) empower and enable every aspect of society and economy. From a homeland security perspective, fully functioning critical national infrastructures are fundamental to all preparedness efforts. Consequently, the critical national infrastructures represent attractive targets to adversaries. At the same time, critical national infrastructures are inherently vulnerable to natural disasters, accidents, and other hazards that are a part of daily life. Given this diverse spectrum of potential threats, coupled with the reality that resources are limited, this chapter concluded that policies and strategies focusing on achieving resilience would be more robust than current guidance, which focuses primarily on protection. Specifically, this chapter observed that while protection is a necessary component of building resilience, resilience is not an inevitable outcome of strategies that focus on protection.

Furthermore, there are technologies that could mitigate some concerns through the replacement of obsolete equipment. Other modern technologies could be used to enable operators to rapidly detect or anticipate impending failures. Current and emerging modeling and simulation tools are available to help analyze interdependencies and their consequences, as well as to investigate risk mitigation options. But all of these technologies, tools, and techniques are of value only if applied within the context of a clear objective—a desired outcome that is measurable.

The time for major investment in this nation's critical infrastructures is long overdue. But such investment is necessarily a shared responsibility and therefore requires the full support of the private and public sector stakeholders. Such support will not be obtained without a shared objective that is aligned with the interests of all stakeholder communities.

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

Chapter Review Questions/Exercises

True/False

1. True or False? Making the critical national infrastructure more diverse in order to create a greater resilience against cyber attack seems to be a pretty sensible approach.
2. True or False? The self-propagation of a computer worm is a good example of an attack that relies on a diverse target environment to function properly.
3. True or False? Typical individual computer users in the home or office, regardless of their location in the world, are less likely to be using a commercial operating system running on a standard processor platform and utilizing one of a couple of popular browsers to perform searches on a popular search engine.
4. True or False? To better understand how diversity goals can be accomplished, it helps to introduce a simple model of desktop computing systems.
5. True or False? Circuit-switched and packet-switched systems automatically provide diversity when compared with one another.

Multiple Choice

1. Additional issues that arise immediately with respect to desktop diversity programs include the following, except which ones:

A. Platform costs
B. Human-machine interface (HMI)
C. Application interoperability
D. Support and training
E. Field control systems

2. Modern telecommunications network systems can be viewed as consisting of which of the following two basic types of technologies:

A. Circuit-switched
B. Boundary scanning
C. Packet-switched
D. Discovery
E. Exposed-switched

3. For many years, procurement projects for national asset systems, in both government and industry, have routinely included the demand that the following physical diversity issues be considered, except which ones:

A. Backup center diversity
B. Supplier/vendor diversity
C. Network route diversity

D. Enforcement of data diversity
E. Data leakage diversity

4. The development of a national diversity program would require coordination between companies and government agencies in the following areas, except which ones?
A. Critical path analysis
B. Cascade modeling
C. Decomposition analysis
D. Misconfiguration modeling
E. Procurement discipline

5. Introducing diversity at all levels of functionality has not been properly explored as:
A. Investment strategy
B. Operations strategy
C. Accuracy strategy
D. Coverage strategy
E. Protection strategy

Exercise

Problem

The critical national infrastructure the U.S. military depends on extends to private companies that build DOD's equipment and technology. It is a significant concern that over the past decade terabytes of data have been extracted by foreign intruders from the corporate networks of defense companies. In a recent single intrusion, 35,000 files were taken. The stolen data ranged from specifications for small parts of tanks, airplanes, and submarines to aircraft avionics, surveillance technologies, satellite communications systems, and network security protocols. Current countermeasures have not stopped this outflow of sensitive information. Identify what type of new countermeasures should be implemented to resolve this cyber attack problem.

Hands-On Projects

Project

Network and data center infrastructures are evolving rapidly to support a dynamic mix of high-volume application traffic and to defend against cyber security attacks. Without current and measurable insight into network resiliency, it is simply not possible to accurately assess the performance, security, and stability of cyber infrastructure elements and systems. Given the scope and persistence of the cyber challenges facing government agencies, enterprises, and service providers, it is imperative that network and data center equipment be examined with resiliency in mind before it is installed. Only this approach can uncover the weaknesses lurking within critical national infrastructures—before it's too late. So, with the preceding in

mind, how would you go about creating an exercise project with the use of diversity and resilience techniques to protect the critical national infrastructure from cyber attacks?

Case Projects

Problem

Let's look at a real-world scenario of a large company's journey in addressing a web application infrastructure security incident that led to a deep analysis and a change in how a development organization builds cyber security into their software development process to prevent cyber attacks. The company is a global leader in producing energy from diversified fuel sources for the U.S. and U.K. consumer markets with approximately 8.9 million electricity and gas customers worldwide. Recently, the company's website was under attack from a botnet titled fringe47. Botnets are networks of compromised computers controlled by hackers known as bot-herders and have become a serious problem in cyberspace.

The company has a long tradition of customer service, so this was a very important issue to them. They surveyed industry best practices and chose a resilient and diverse process for developing secure software and changing their engineering practices.

Explain how you would reduce potential vulnerabilities, protect against intrusion attempts, and better anticipate future threats.

Optional Team Case Project

Problem

Recently, Estonia was bombarded by cyber attacks from all over the world. Some were hosted by Russian state servers. Estonia's foreign ministry published a list of IP addresses from where the cyber attacks came from. The cyber attacks came mainly in the form of DDOS attacks, primarily targeting Estonian government and police sites. Private sector banking and online media were also heavily targeted. The cyber attacks also affected the functioning of the rest of the network infrastructure in Estonia. The cyber attacks against government websites came in waves: They started and ended, and then started again after a few days' break.

Estonia's second-biggest bank, Swedish-owned SEB Eesti Uhispank, was forced to block access from abroad to its online banking service after it came under a massive cyber attack. This was after Hansapank, the biggest bank in Estonia, also came under attack. The first wave of cyber attacks against official websites fizzled out after Estonian Foreign Minister Urmas Paet publicly declared that many of the attacks had originated from Russian government computers. The new wave of attacks came from around the world.

Computers as far away as Vietnam had been involved in cyber attacks against Estonia. The attackers tried to restrict access to Estonian websites and, in some cases, tried to change the information on the website they had attacked. Some sites were defaced to redirect users to images of Soviet soldiers and quotations from Martin Luther King about resisting evil. And hackers who hit the ruling Reform Party's website at the height of the tension left a spurious message that the Estonian prime minister and his government were asking for forgiveness. The cyber attacks might have originated in computers around the world, but they still had Russian roots. The Internet has been full of Russian language instructions on how to inflict damage on

Estonian cyberspace.

NATO and EU Internet experts are all helping to track down the culprits, but Estonian officials had no cooperation from Russia. While the initial wave of cyber attacks came from official structures in Russia, it might be very difficult to track the perpetrators down. Botnets (the term given to the groups of computers that mount denial-of-service attacks) can be located across several countries, or even continents. Russia (which has a large community of hackers and computer virus writers) has been accused of mounting such cyber attacks before, in the United States and the Ukraine.

So, in keeping the preceding in mind, identify how the computer emergency response team (CERT) would go about tackling security incidents in Estonia's Internet domain.

[1] Quoted in "Taking Cues from Mother Nature to Foil Cyber Attacks" (press release), Office of Legislative and Public Affairs, National Science Foundation, Washington, D.C., 2003 (http://www.nsf.gov/od/lpa/news/03/pr03130.htm).

5

Commonality

Chapter Outline

Meaningful Best Practices for Infrastructure Protection
Locally Relevant and Appropriate Security Policy
Culture of Security Protection
Infrastructure Simplification
Certification and Education
Career Path and Reward Structure
Responsible Past Security Practice
National Commonality Program
How Critical National Infrastructure Systems Demonstrate Commonality
Summary
Chapter Review Questions/Exercises

The only truly secure system is one that is powered off, cast in a block of concrete, and sealed in a lead-lined room with armed guards—and even then I have my doubts.

Eugene Spafford, Executive Director of the Purdue University Center for Education and Research in Information Assurance and Security (CERIAS)[1]

Now that we have outlined our proposal in the previous chapter for national infrastructure systems to include diversity, we can discuss the seemingly paradoxical requirement that infrastructure systems must also demonstrate a degree of *commonality*. In particular, certain desirable security attributes must be present in all aspects and areas of national infrastructure to ensure maximal resilience against cyber attack. Anyone who has worked in the security field understands this statement and is likely to agree with its basis. The collection of desirable security attributes is usually referred to collectively as *security best practices*. Example best practices include routine scanning of systems, regular penetration testing of networks, programs for security awareness, and integrity management checking on servers.

When security best practices are easily identified and measurable, they can become the basis for what is known as a *security standard*. A security standard then becomes the basis for a process known as a *security audit*, in which an unbiased third-party observer determines based on evidence whether the requirements in the standard are met. The key issue for national infrastructure protection is that best practices, standards, and audits establish a low-water mark for all relevant organizations (see Figure 5.1).



Figure 5.1 Illustrative security audits for two organizations.

Organizations that are below a minimally acceptable security best practices level will find that security standards audits introduce new practices, in addition to revisiting existing practices. The desired effect is that the pre-audit state will transition to an improved post-audit state for all practices. This does not always happen, especially for organizations that have a poor environment for introducing new security practices, but it is the goal. For organizations that are already above the minimally acceptable level, perhaps even with world-class features, the audit will rarely introduce new practices but will instead revisit existing ones. The desired effect here is that these practices would be strengthened, but, again, this does not always work perfectly, especially if the auditors are less familiar with the world-class security features already in place. Some common security-related best practices standards that one will find in national infrastructure settings are listed in the box.

The purpose of a security audit is to raise the level of security features currently in place.

Common Security-Related Best Practices Standards

- *Federal Information Security Management Act (FISMA)*—FISMA sets minimal standards for security best practices in federal environments. It is enforced by congressional legislation and involves an annual letter grade being assigned to individual agencies. The following departmental agencies received an "F" for their FISMA rating in 2007: Defense, Commerce, Labor, Transportation, Interior, Treasury, Veterans Affairs, and Agriculture (so did the Nuclear Regulatory Commission).
- *Health Insurance Portability and Accountability Act (HIPAA)*—Title II of HIPAA includes recommended standards for security and privacy controls in the handling of health-related information for American citizens. It is also enforced by congressional legislation.
- *Payment Card Industry Data Security Standard (PCI DSS)*—This security standard was developed by the PCI Security Council, which includes major credit card companies such as Visa® Card, Discover® Card, American Express®, and MasterCard®. It includes requirements for encrypting sensitive customer data.
- *ISO/IEC 27000 Standard (ISO27K)*—The International Organization for Standardization (ISO) and

International Electrotechnical Commission (IEC) evolved a British Security Standard known as BS-7799 into an internationally recognized set of auditable security best practices. Some security experts believe that the ISO27K family of security standards is the most global and generally agreed upon set of best practices.

All of these standards, and the many additional ones that are not mentioned above, include a large subset of security and functional requirements that are virtually the same. For example, each standard requires carefully documented policies and procedures, authentication and authorization controls, data collection systems, and embedded encryption. Each standard also requires management oversight, ongoing security monitoring, compliance scores issued by designated auditors, and some form of fines or punishment if the standard best practices are not met.

With such redundancy in security standards and compliance, one would guess that the principle of commonality would be largely met in national infrastructure protection. For example, some organizations might be required to demonstrate compliance to dozens of different security standards. One would expect that such intense and focused attention on security would lead to a largely common approach to security around the globe. Sadly, the belief here is that in spite of the considerable audit and compliance activity around the world, most of it does not address the type of security commonality that will make a positive difference in national infrastructure protection. The activity instead tends to focus on requirements that have some value but do not address the most critical issues. In fact, most of these practices exist in the category of state-of-the-art security, far beyond the minimally acceptable levels addressed in most audits.

The audit problem stems from the inherent differences between *meaningful* and *measurable* security best practices. There's an old dumb joke about a man looking for his lost money on 42nd and Eighth. When a passerby asks whether the money was actually lost at that spot, the man looks up and says that the money was actually lost over on 41st and Tenth but the light is much better here. Security audit of best practices is often like this; the only practices that can be audited are ones where the light is good and measurable metrics can be established. This does not, however, imply that such metrics are always meaningful (see Figure 5.2).



Figure 5.2 Relationship between meaningful and measurable requirements.

The example requirements shown in Figure 5.2 provide a hint as to the types of requirements that are likely to be included in each category. One can easily levy a measurable requirement on password length, for example, even though this is generally a less useful constraint. This could be viewed as an example that is measurable but not meaningful. Conversely, one can levy the important requirement that a strong culture of security be present in an environment. This is a meaningful condition but almost impossible to measure. The example requirement that a security policy be present is both meaningful and measurable. It demonstrates that there are certainly some requirements that reside in both categories.

Ideally, security practices are both meaningful *and* measurable.

Meaningful Best Practices for Infrastructure Protection

A provocative implication here is that the ability to audit a given best practice does not determine or influence whether it is useful for infrastructure protection. In fact, the primary motivation for proper infrastructure protection should not be one's audit score; rather, the motivation should be success based and economic. The fact is that companies, agencies, and groups with responsibility for infrastructure protection will eventually fail if they do not follow the best available recommendations for security best practices. Unfortunately, the best recommendations come not from the security standards and audit community but from practical experience.

A great audit score does not necessarily guarantee successful infrastructure protection.

If you do not agree, then please consider that security standards backed by powerful and authoritative groups have existed for many decades. In addition, security auditors have been in business for decades, performing diligent analysis and issuing embarrassing failure grades to security teams around the world. Our earlier reference to FISMA, for example, included failing grades for many prominent government agencies in the United States. In spite of all this activity and reporting, however, nothing truly material has changed during these past decades in the way computer and network systems are secured. In fact, one could easily make the claim that national infrastructure is more vulnerable to attack today than it was 20 years ago. What makes one think that more stringent security standards and audit processes are going to change this now?

Based on this author's experiences managing the security of major critical infrastructure components for many years, the answer lies in a two-step methodology:

- *Step 1. Standard audit*—The first step is conventional, in that it recommends that every organization submit to a standard audit to ensure that no group is operating below the minimally acceptable threshold. While most organizations would claim to already have this step ongoing, the goal here is to be given a desirable rating or score, rather than a failing one. So, even if a company or agency has ongoing audits, the goal here is to *pass* these audits. Any one of the major audit standards mentioned above is probably acceptable; they all roughly direct the same sort of minimal practices.

A successful protection strategy should start with at least a passing score on a standard security audit.

- *Step 2. World-class focus*—The second step involves a more intense focus on a set of truly meaningful national infrastructure protection practices. These practices are derived largely from experience. They are consistent with the material presented in this book, and they will only be present in pieces in most existing security audit standards. The greatest success will typically come from organizations self-administering this new focus, especially because these practices are not easy to measure and audit (see Figure 5.3).



Figure 5.3 Methodology to achieve world-class infrastructure protection practices.

For the first step, an important issue involves ensuring that the audit does not cause more harm than good. For example, suppose that a competent and trustworthy system administrator has been charged with a bevy of responsibilities for an infrastructure component and that she has demonstrated excellent results over a long period of time, with no security problems. This is a common situation, especially in companies and agencies that take system administration seriously. Unfortunately, a security auditor would look at such a setup with horror and would deem it a clear violation of least privilege, separation of duties, and so on.

Sometimes security audit standards and best practices proven through experience are in conflict.

In the United States, if the component being administered was a financial one in a public company, then this would be a violation of the Sarbanes-Oxley segregation of duties requirements. The auditor would typically require that the single competent administrator be replaced by a bureaucratic process involving a team of potentially inferior personnel who would each only see a portion of the total task. It is not difficult to imagine the component being more poorly managed and, hence, less secure. This is the worst case in any audit and must be explicitly avoided for national infrastructure protection.

For the second step, the box lists specific meaningful security best practices, six in total, for national infrastructure protection. These six best practices do not contradict current auditing processes and standards, but they are certainly not designed for easy audit application; for example, it is difficult to validate whether something is "appropriate" or "simplified." Nevertheless, our strong advice is that attentiveness to ensuring commonality across national infrastructure with these six practices will yield significant benefits.

Six Best Practices for National Infrastructure Protection

- *Practice 1. Locally relevant and appropriate security policy*—Every organization charged with the design or operation of national infrastructure must have a security policy that is locally relevant to the environment and appropriate to the organizational mission. This implies that different organizations should expect to have different security policies. The good news is that this policy requirement is largely consistent with most standards and should be one of the more straightforward practices to understand.
- *Practice 2. Organizational culture of security protection*—Organizations charged with national infrastructure must develop and nurture a culture of security protection. The culture must pervade the organization and must include great incentives for positive behavior, as well as unfortunate consequences for negative. No security standard currently demands cultural attentiveness to security, simply because it cannot be measured.

- *Practice 3. Commitment to infrastructure simplification*—Because complexity is arguably the primary cause of security problems in most large-scale environments, a commitment to simplifying infrastructure is critical to ensuring proper security. Determining what "simplification" means is a subjective, local concept that is dependent on the specifics of the target environment. No current security standards demand infrastructure simplification.
- *Practice 4. Certification and education program for decision-makers*—A program of professional certification and security education must be present for those who are making decisions about national infrastructure or who are directly charged with their implementation. Ideally, this should not have to include end users, because this greatly reduces the chances of proper coverage.
- *Practice 5. Career path and reward structure for security teams*—Those performing security in national infrastructure environments must have clearly defined career paths and desirable rewards as part of their professional journey. In the absence of these enticements, important security work is often handled by people who are untrained and poorly motivated. This requirement is generally more meaningful in larger organizations.
- *Practice 6. Evidence of responsible past security practice*—Just as most craftsmen go through a period of apprenticeship to learn and to demonstrate proper skills, so should an organization have to demonstrate a period of learning and attainment of proper skills before being charged with national infrastructure protection. It is amazing that existing security audits generally do not include a careful inspection of past security practices in dealing with live cyber attacks.

Readers familiar with standards and audits will recognize immediately the challenges with the subjective notions introduced in the box. For this reason, the only way they can be applied appropriately is for security managers to understand the purpose and intent of the requirements, and to then honestly self-administer a supporting program. This is not optimal for third-party assurance, but it is the only reasonable way to reach the level of world-class security best practices.

Locally Relevant and Appropriate Security Policy

Any commercial or government organization that is currently developing or managing national infrastructure already has some sort of security policy. So the question of whether to develop a policy is not relevant; every organization has *something*. The real question instead for most organizations in national infrastructure roles is how to make the policy more relevant and appropriate to the local environment. Specifically, four basic security policy considerations are highly recommended for national infrastructure protection:

The question is not *whether* to develop a security policy, but rather what that policy will entail.

- *Enforceable*—Most security policies are easy to write down but are not easy to enforce. Organizations must therefore spend a great deal of time on the issue of security policy enforcement. The local threat environment must be a consideration here, because the employees of some companies and agencies are more apt to follow security policy rules than others. Nevertheless, a policy is only as good as its degree of enforceability, so every organization should be able to explicitly describe their enforcement strategy.
- *Small*—Most security policies are too large and complex. If there is one exercise that would be the healthiest for national infrastructure teams, it would be to go through existing policy language to prune out old references, obsolete statements, and aged examples. Large, complex security policies with too much detail are to be avoided. A key issue is the direction in which one's policy is headed; it is either staying the same (stagnant), getting more complex (unhealthy), or becoming smaller and more compact (healthy).
- *Online*—Policy language must be online and searchable for it to be truly useful in national infrastructure settings. Teams must be able to find relevant requirements easily and should have the ability to cut and paste the relevant statements into their project or process documentation. The old days of printing and distributing a security policy with a fancy cover should be long gone.
- *Inclusive*—Policy must be inclusive of the proper computing and networking elements in the local national infrastructure environment. This can only be determined by an analysis. Unfortunately, this analysis can be somewhat time consuming and tedious, and without proper attention it could result in an overly complex policy. Considerable skill is required to write policy that is inclusive but not too complicated.

These four requirements for security policies in groups charged with national infrastructure can be subjected to a simple decision analysis that would help determine if the local policy is relevant and appropriate to the mission of the organization; this decision process is shown in Figure 5.4.



Figure 5.4 Decision process for security policy analysis.

It's worth mentioning that, as will be seen in the next section, the culture of the local environment can really have an impact on the development of security policy. In an environment where technology change is not dramatic and operational skills are mature (e.g., traditional circuit-switched telephony), policy language can be less detailed and used to identify unexpected procedures that might be required for security. In an environment where technology change is dramatic and operational skills might be constantly changing (e.g., wireless telephony), then policy language might have to be much more specific. In either case, the issue is not whether the policy has certain required elements, but rather whether the policy is locally relevant and appropriate.

Culture of Security Protection

Our second recommended common practice involves creation of an organizational culture of security protection. When an organization has such a culture of security protection, the potential for malicious exploitation of some vulnerability is greatly reduced for two reasons: First, the likelihood for the vulnerability itself to be present is reduced, as local diligence will weigh in favor of more secure decision-making. Second, real-time human vigilance in such a culture often helps avoid exploitation. Time after time, the alertness of human beings in a culture of security is effective in helping to avoid malicious attacks. (Readers will remember that the only effective security measures that took place on September 11, 2001, were the ones initiated by human beings.)

Here's a simple test to determine if a given organization has a culture of security protection. Go to that organization's local facility and observe how carefully the physical premises are policed for unauthorized entry. If an electronic door is used to authenticate entry, followed by a guard eyeballing every visitor, then chances are pretty good that the culture is one of protection. If, however, the person in front of you holds the door open for you to enter without bothering to check for your credentials or, worse, the door itself is propped open, then the culture is probably more open. A culture of security certainly does not imply that things will be perfectly secure, but such a culture is essential in the protection of national assets.

An organization with a culture of security is one in which standard operating procedures work to provide a secure environment.

Unfortunately, most of us tend to equate an organizational culture of security with a rigid, paranoid, authoritative, perhaps even military environment. Furthermore, a culture of security is generally associated with managers who avoid risks, stay away from the media, dislike remote access or telecommuting, and demonstrate little comfort with new technologies such as social networking. Similarly, one would equate a nonculture of security with a young, dynamic, creative, open, and egalitarian environment. In such a culture, managers are generally viewed to be comfortable with risk, open in speaking to outsiders about their work, in love with every new technology that comes along, and supportive of remote access and telecommuting.

The reality is that neither stereotype is accurate. Instead, the challenge in promoting a culture of security is to combine the best elements of each management approach, without the corresponding weaknesses. The idea is to nurture any positive environmental attributes, but in a way that also allows for sensible protection of national assets; that is, each local environment must have a way to adapt the various adjectives just cited to their own mission. For example, no group generally wants to be referred to as closed and paranoid, but a military intelligence group might have no choice. Similarly, no group wants to be referred to as being loose with security, but certain creative organizations, such as some types of colleges and universities, make this decision explicitly.

An ideal security environment can marry creativity and interest in new technologies with caution and healthy risk aversion.

As such, organizations must consider the spectrum of options in developing a suitable local culture. This spectrum acknowledges how straightforward it can be to assume an inverse relationship between organizational rigidity and security. It's easy to just make everything rigid and authoritative and hope that a culture of increased security will develop. The challenge, however, lies in trying to break up this relationship by allowing open, creative activity in a way that does not compromise security. This might result in some aspects of the environment being more secure and others being less so. Such a combined cultural goal should be viewed as a common requirement for all groups involved with national assets (see Figure 5.5).

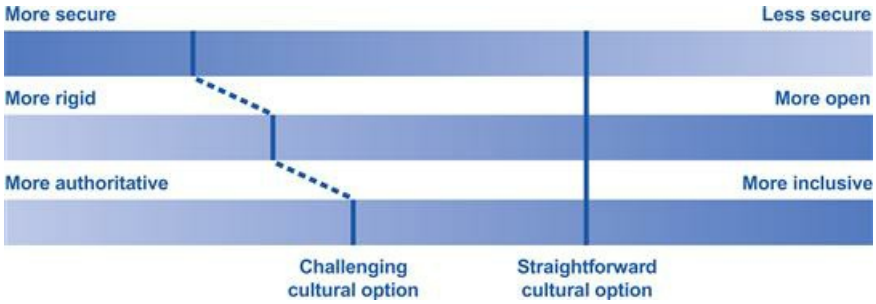


Figure 5.5 Spectrum of organizational culture of security options.

So an obvious question one might ask from the perspective of national infrastructure protection is why the highest level of security culture should not be required in all cases, regardless of any cultural goals of being open, creative, and willing to interact publicly. The U.S. military, for example, might exemplify such a level of rigid cultural commitment to security. One answer, as we've discussed above, is that it is difficult to *require* that a culture be in place in an organization. Specific aspects of a culture might be required such as strong policy, tough enforcement, and so on, but to require the presence of a culture is easy to confirm. Nevertheless, the premise is correct; that is, for national infrastructure, certain security standards are required that can only be met in an environment where a culture of security protection is met. This demands the uncomfortable situation in which local managers must honestly work to create the appropriate culture, which in some cases might require decades of attention.

Implementation of a true culture of security cannot happen overnight; it may take years to develop.

An important element of security culture is the symbolism that management can create by its own behavior. This means that when senior executives are given passes that allow policy violations, this is a serious error as it detracts from the cultural objectives. Unfortunately, the most senior executives almost always outrank security staff, and this practice of senior exemption is all too common. Perhaps major national infrastructure solicitations should include questions about this type of senior executive practice before contracts can be granted to an organization. This might give the security team more concrete ammunition to stop such exemptions.

A true culture of security must be implemented at all levels of an organization—including the most senior executives.

Infrastructure Simplification

Our third recommended common practice involves an explicit organizational commitment to infrastructure simplification. Defining what we mean by simplification in the context of infrastructure requires that we use subjective language. Simpler infrastructure is easier to understand, less cumbersome, and more streamlined. As such, simplification initiatives will be subjective and much more difficult to measure using some quantitative metric. To illustrate this process of simplification, let's look at a typical sort of cluttered engineering schematic that one might use to describe network infrastructure. The chart shown in Figure 5.6 is derived from the design documentation embedded in an infrastructure project with which this author was recently involved. This diagram suffers from the typical sorts of issues that one finds in the design and operation of national infrastructure:

- *Lack of generalization*—Systems in the diagram are not viewed in a generalized manner. The same thing is shown multiple times in different places in the diagram (e.g., servers), rather than just generalizing one component to depict both.
- *Clouding the obvious*—Interfaces in the diagram are not depicted obviously. Lines are cluttered across the drawing, and simple interfaces are clouded to avoid what is actually quite obvious connectivity.
- *Stream-of-consciousness design*—The diagram seems to be the product of first-draft, stream-of-consciousness thinking rather than a carefully planned layout. Too often, infrastructure is put in place in a first draft without anyone taking the time to review and revise.
- *Nonuniformity*—Objects are not referred to uniformly; the references to an IP-based network are slightly different, and in fact should just reference the Internet anyway.

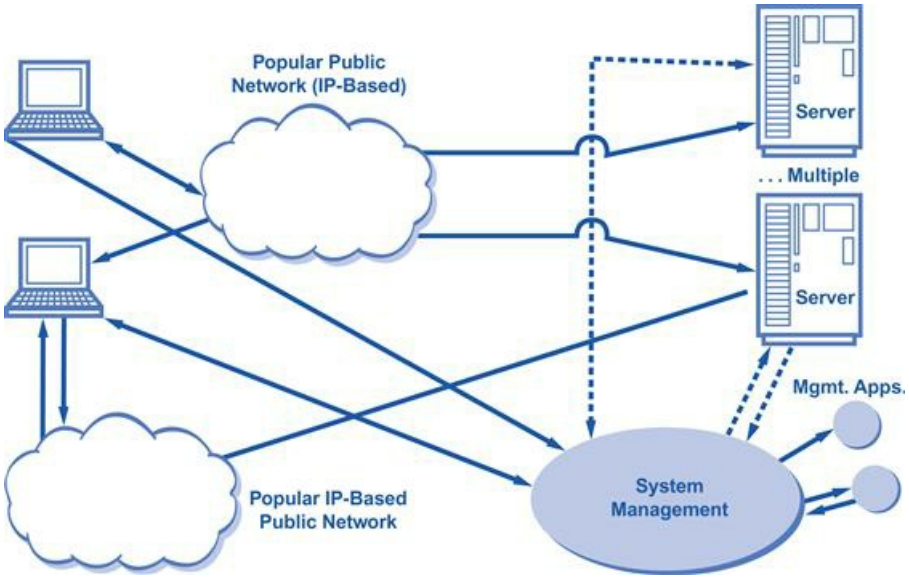


Figure 5.6 Sample cluttered engineering chart.

If one applies some rational simplification to the design in the cluttered chart shown above, with

attention to each of the elements just mentioned, then the resultant functionally equivalent product is much easier to understand. The more improved diagram requires that you go back and confirm that it really does describe the same function, but in fact it does (see Figure 5.7).

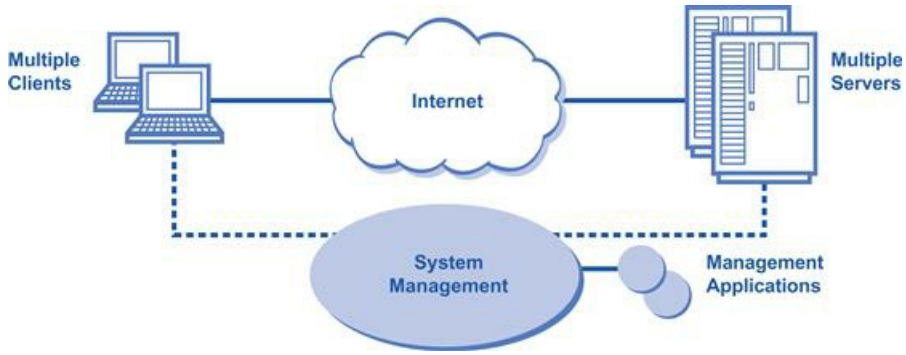


Figure 5.7 Simplified engineering chart.

Analysis of how we simplified the cluttered diagram into something more easily understood highlights some of the techniques that can be useful in simplifying a national infrastructure component environment (see box).

How to Simplify a National Infrastructure (or Otherwise Complex) Environment

- *Reduction in size*—The second diagram is smaller than the first one. Relevance of such action to national infrastructure should be obvious. Simplification should include reduction wherever possible. Less code, fewer interfaces, and reduced functionality are all healthy simplification objectives that will almost certainly improve security. In fact, a requirement for national infrastructure should be demonstrable evidence of software removal or reduction initiatives. The only truly secure piece of code is the one that you have removed.
- *Generalization of concepts*—The second diagram generalizes concepts more effectively than the first. This should be true in national infrastructure as well. Rather than managing dozens or hundreds or thousands of special cases, it is more effective to have a planned generalization strategy that allows for simpler management. Obviously, this requires some balancing with local diversity requirements.
- *Cleaner interfaces*—Perhaps the most obvious difference between the two diagrams is the much cleaner view of interfaces that the second one provides. Because national infrastructure will include complex interfaces between systems, initiatives to simplify these interfaces must be present to optimize security for national assets.
- *Highlighting of patterns*—The second diagram demonstrates functional and data flow patterns in an obvious manner. This simplifies any changes that might have to be made to the architecture. Infrastructure should also be designed in a manner to highlight important patterns in data or processing.
- *Reduction of clutter*—The first diagram is as cluttered as one can imagine, and this generally indicates a stream-of-consciousness design process. Too often, national infrastructure emerges in the same manner, with one system being put in place and then another, they are then connected to something

else, and on and on. The result is usually not optimal from a security perspective.

The process of auditing these subjective goals will be challenging, if not intractable, but this does not reduce the importance of trying to attain each goal in national infrastructure. Infrastructure simplification could, in fact, be argued to be the most important single goal in the protection of national assets. One bright spot here is that security managers will find kindred spirits with most information technology managers, although it is the rare CIO who truly knows how to manage the simplification and reduction of infrastructure. A good sign that the local organization is trying would be some sort of initiative focused on the reduction or removal of software applications.

Simplification may be the first and most tractable step toward creating a new, more secure infrastructure environment.

Certification and Education

Our fourth recommended common practice involves certification and education programs for key decision-makers. Most current computer security education initiatives tend to focus on teaching awareness to end users about proper selection of passwords, storage of data, handing of devices, and so on. These awareness initiatives stem from the common belief that computer and network systems would be perfectly secure if end users would just take the time to learn and follow the security policy rules. The situation is reminiscent of doctors blaming their patients for their diseases.

Security auditors generally agree with this view of end-user responsibility, and they will often perform spot checks in target environments. This usually involves quizzing random individuals about their knowledge and interpretation of the local security policy. When the inevitable bad grade occurs because high percentages of individuals do not know some of the policy rules, security teams are forced to increase the intensity of the awareness program with posters, videos, mandatory tests, and even punishments for end-user ignorance.

Based on decades of experience in performing these types of audits, supporting them, and also being subjected to them, the conclusion reached here is that the goal of reaching 100% end-user awareness of security is impractical. Certainly, security education for end users does not hurt, because everyone should be aware of the risks of any actions they might take that could damage security in the local environment. If end users are entrusted with proprietary information, for example, they need to understand the implications of allowing such information to be provided to unauthorized sources.

One hundred percent end-user awareness of security policies may remain an illusive goal.

For national infrastructure protection, however, a much more practical goal is to focus primarily on improving the security competence of decision-makers rather than on end users. The distinction here is subtle, but fundamental. Key decision-makers in national infrastructure settings include the following:

- *Senior managers*—These are the people who set financial and operational priorities affecting national infrastructure. They include the most senior managers in an organization or the highest ranking in the military.
- *Designers and developers*—These are the network, system, and application designers and developers who determine what security features and functionality are in the systems that people use. They often work in information technology groups.
- *Administrators*—These are the system and network administrators who perform the day-to-day tasks of maintaining and running the systems that people use. Too often, these folks are underpaid and poorly trained.
- *Security team members*—These are the security staff charged with the organizational systems for protecting assets. An increasing number of organizations outsource aspects of this work. There is nothing wrong with this trend, as long as the arrangement is well managed and coordinated.

These four types of key decision-makers are the people who can make the most substantive difference in the security of an organization and for whom 100% coverage should be a tractable goal. It doesn't hurt that

the size of the key decision-maker population in a company or agency will be much smaller than the total population. It also doesn't hurt that they tend to be the ones best trained to understand the importance of security. From an investment perspective, the returns on education investment look quite different for end users and decision-makers (see Figure 5.8).

Target the key decision-makers in your quest for organizational security policy awareness and competence.

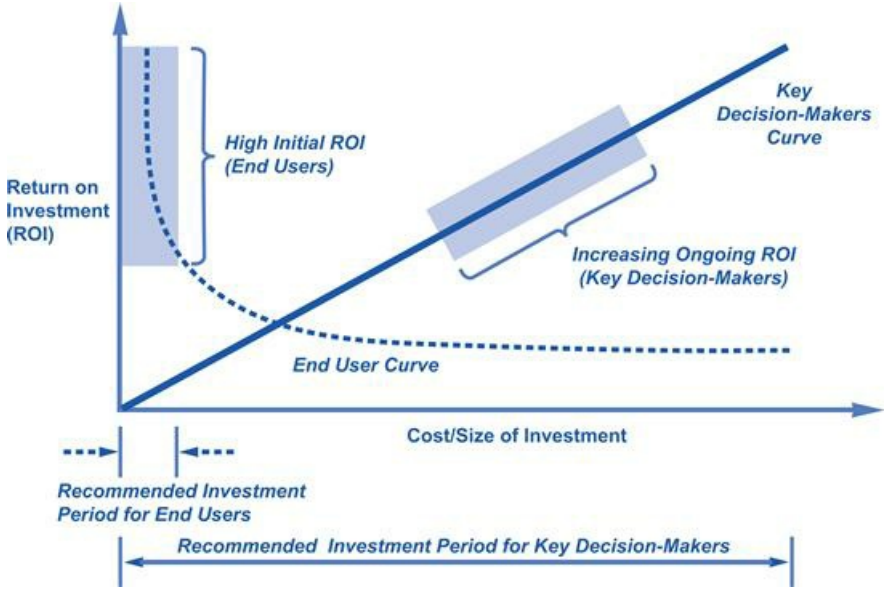


Figure 5.8 Return on investment (ROI) trends for security education.

The message embedded in the ROI curves in Figure 5.8 is that a small initial investment in security certification and education for end users produces a reasonable initial return. This return rapidly diminishes, however, because in a typical environment there is only so much an end user can do. In fact, in the best designed environments, the obligation for end users to make security decisions on their own is always minimized. For key decision-makers, the ROI is ongoing and steadily increasing throughout the investment lifecycle. Unlike end users, key decision-makers can consistently apply their increased security knowledge to infrastructure in a meaningful and scalable manner.

To summarize, our recommendation here is a twofold approach for security certification and education in a national infrastructure environment:

- *Key decision-makers*—Focus on providing ongoing, lifecycle programs for decision-makers in security certification and education. By focusing on key decision-makers, the returns will be consistent, increasing, and scalable.
- *End users*—Create low-cost, high-initial-return activities for certifying and educating end users. As a complement, systems must be designed that minimize the decisions end users make about security.

The specific certification and education programs for a given environment should be locally determined and appropriately applied. They are not difficult to find or create but can be misapplied without some careful

planning. Well-known security certifications, such as Certified Information Systems Security Professional (CISSP), are excellent for system or network administrators but totally unnecessary for end users. Similarly, awareness programs on selecting good passwords are fine for end users but will just annoy your system administrators.

Career Path and Reward Structure

Our fifth recommended common practice involves the creation and establishment of career paths and reward structures for security professionals. It should come as no surprise that organizations charged with national infrastructure should demonstrate some common form of career path and reward structure for security staff. This is particularly important, because to perform security tasks properly, some degree of longevity is desirable. Too often, important cyber security tasks are attempted by staff who are new to the security discipline and who are poorly compensated for their work.

Creating career paths and incentives is important in any field, no less so in security management.

Fixing this might seem obvious, but virtually no security standards used for the purposes of audit include this in a meaningful way. Elements that should be commonly present in national infrastructure environments include the following:

- *Attractive salaries*—Infrastructure organizations should demonstrate salary structure that takes into account the specialized skills associated with cyber security. Salaries should be above industry averages, a metric that can be quantitatively audited. (Amazingly, I've never seen security staff salaries audited as part of any due diligence activity by an auditor.)
- *Career paths*—Opportunities for career advancement, promotion, and salary increase should be present in infrastructure organizations. Perhaps more than any other information technology or network-related discipline, security engineering of national infrastructure requires years of experience in order to develop proper judgment. If these years do not include attention to career issues, then the organization is unlikely to maintain the best staff.
- *Senior managers*—It is desirable for senior managers in infrastructure organizations to have some degree of heritage in the security community. This certainly will help with decision-making at the senior level, but more importantly it serves as a symbol for the security staff that senior level management is attainable from the security ranks.

These career-related organizational attributes are rarely discussed in the context of determining whether proper security is in place in an organization. Auditors never discuss these issues. This is unfortunate, as good salaries and career paths for security staff are more relevant to the overall security posture of an organization than checking for trivia such as password length, time-outs after bad login attempts, and other elements commonly found in security standards.

A strong indicator of a healthy security environment might be something that is often overlooked, such as heritage of the senior security managers in a company.

It is also worth noting that companies and agencies should not actively recruit and hire individuals who have a history of breaking laws on computers and networks. Hacking, in its original incarnation, was all about the desire to learn and share; when hackers demonstrate this type of perspective, they can easily blend into a company or agency and be productive. The associated career and reward track for such individuals is rarely

promotion or money but rather ongoing or increased access to the latest and greatest types of technologies.

Responsible Past Security Practice

Our sixth recommended common practice involves two specific actions: The first is that any company or agency being considered for national infrastructure work should be required to demonstrate past practice in live security incidents. The second is that companies and agencies must do a better job of managing their inventory of live incidents, including databases of key factors, root causes, and security learning from events. These two seemingly obvious actions are almost never performed explicitly, and most companies and agencies do not even maintain formal documentation on past security incidents.

Companies and agencies should maintain a historical record showing clear incident response documentation.

The good news is that most solicitations for national infrastructure project work do include some requirement for demonstrating past engineering practices, so there is certainly a base on which to improve matters for security. When federal agencies contract for engineering or technical work, for example, boilerplate language is usually embedded into the contract for information on previous projects, similar work activities, and lists of reference clients. This practice is appropriate and valuable, although it is usually treated too much as a generic type of information-gathering task.

For security, in particular, this practice currently involves requests for information on security policies, security architectural elements, and even specific techniques such as encryption. Such requests are important and should be highlighted for national infrastructure protection projects. The problem is that such inquiries simply do not go far enough. In particular, any organization being considered in a solicitation that involves national infrastructure should provide evidence of at least the following past practices:

- *Past damage*—The organization should be able to provide evidence of past security incidents that it dealt with that produced real malicious damage to some valued asset. Although this might seem paradoxical, the reality is that no organization can claim true skill in securing large infrastructure if it has not dealt with a real incident in the past. Groups who are forthcoming in explaining these past incidents are also generally more mature in their current security processes.

A mature security organization will admit to successful attacks against them.

- *Past prevention*—Similarly, the organization should be able to provide evidence of incidents prevented. This is tougher than one might think, because in many cases security protections have a preventive effect that is not easily determined or measured. So only the truly skilled security organizations can provide this evidence of deliberate action that prevented an attack from succeeding. A good example might be the establishment of real-time network filtering well in advance of any DDOS attack; if this filtering was actually used to stop an attack attempt, it demonstrates excellent judgment regarding the organizational priorities around security.

Providing evidence of successful preventive measures is a challenge for most organizations.

- *Past response*—This is the most commonly cited security experience component. Groups can generally

point to their response functions as being invoked during worms, viruses, and other attacks.

In any formal project solicitation, these requirements should be highlighted and assigned high priority. Few requirements can properly highlight an organization's ability to handle security situations in the future as their experiences dealing with similar matters in the past.

National Commonality Program

The challenge in creating a new national program of ensuring commonality with state-of-the-art security practices in infrastructure protection involves balancing several different concerns:

- *Plethora of existing standards*—Most organizations are already frustrated with the number of standards and audits that must be covered. The implication is that the creation of a new national security standard commensurate with the six practices described in this chapter would not be well received.
- *Low-water mark versus world class*—As we've discussed, the existing security standards and audits in place today are more focused on creating a common low-water mark, rather than pushing groups to reach for world-class status in security.
- *Existing commissions and boards*—The field is already crowded

with national commissions, working groups, and boards comprised of business and government leaders who are working to create sets of recommendations for infrastructure security. They are unlikely to go away and must be factored into any implementation plan.

Do not try to work around the existing security commissions and boards; instead, factor them into your overall security plans and policies.

While these may not be formal standards with associated audit processes, affected organizations feel the pressure to review these works and to demonstrate some degree of acceptance, if not compliance. The solution to balancing these concerns lies in several implementation approaches and hints that are based on previous experiences with multiple standards and requirements, such as the Orange Book, Red Book, and associated "security rainbow series" in the 1980 s. The first is that government really should adopt a single standard for all commercial and government security audits. It really doesn't even matter which audit standard is selected as long as it is *only one*. All subsequent government solicitations and contracts should demand compliance with this standard. Commercial entities might gradually merge toward this standard.

Second, the world-class practices described here should be embedded into all government solicitations and contracts as functional requirements on companies and agencies. This would avoid the problems of audit compliance and would push the security components into the functional category along with performance, processing, storage, and networking. Government agencies could perhaps complement this by rewarding or providing incentives for the inclusion of these requirements in private deals between companies.

Finally, let's briefly look at some practical ways of how critical national infrastructure systems demonstrate commonality. As previously stated, certain desirable security attributes must be present in all aspects and areas of the critical national infrastructure to ensure maximal resilience against cyber attacks.

How Critical National Infrastructure Systems Demonstrate Commonality

The threats to systems supporting critical national infrastructures are evolving and growing. Varying types of threats from numerous sources can adversely affect computers, software, networks, organizations, entire industries, or the Internet itself. These include both unintentional and intentional threats, and may come in the form of targeted or untargeted attacks from criminal groups, hackers, disgruntled employees, hostile nations, or terrorists. The interconnectivity and commonality between information systems, the Internet, and other infrastructures can amplify the impact of these threats, potentially affecting the operations of the critical national infrastructure, the security of sensitive information, and the flow of commerce. Recent cyber attack incidents include hackers accessing the personal information of hundreds of thousands of customers of major U.S. banks and sophisticated cyber attacks targeting control systems that are used to operate industrial processes in the energy, nuclear, and other critical sectors.

Over the past 4 years, the federal government, in partnership with the private sector, has taken a number of steps to address threats to the cyber critical national infrastructure. Recently, the White House conducted a review of the nation's cyberspace policy that addressed the missions and activities associated with the nation's information and communications infrastructure. The results of the review led, among other things, to the appointment of national Cyber Security Coordinator with responsibility for coordinating the nation's cyber security policies and activities. Also, the DHS updated its National Infrastructure Protection Plan, which provides a commonality framework for addressing threats to critical national infrastructures and relies on a public-private partnership model for carrying out these efforts. DHS has also established a communications center to coordinate national response efforts to cyber attacks and work directly with other levels of government and the private sector, and has conducted several cyber attack simulation exercises. Despite recent actions taken, a number of significant challenges remain to enhancing the security of cyber-reliant critical national infrastructures (see "An Agenda for Action in Enhancing the Security of Cyber-Reliant Critical National Infrastructures").

An Agenda for Action in Enhancing the Security of Cyber-Reliant Critical National Infrastructures

When completing the Enhancing the Security of Cyber-Reliant Critical National Infrastructures checklist, the Department of Homeland Security (DHS) should adhere to the provisional list of actions for enhancing the security of cyber-reliant critical national infrastructures. The order is not significant; however, these are the activities for which the research would want to provide a detailed description of procedures, review, and assessment for ease of use and admissibility. The significant challenges that remain to enhancing the security of cyber-reliant critical national infrastructures include (check all tasks completed):

1. Implementing actions recommended by the president's cyber security policy review.
2. Updating the national strategy for securing the information and communications infrastructure.
3. Reassessing DHS's planning approach to critical infrastructure protection.
4. Strengthening public–private partnerships, particularly for information sharing.
5. Enhancing the national capability for cyber warning and analysis.
6. Addressing global aspects of cybersecurity and governance.

7. Securing the modernized electricity grid, referred to as the smart grid.
8. Ensuring the safety and security of food, animal feed, and food-producing animals; coordinating animal and plant disease and pest response; and providing nutritional assistance.
9. Providing the financial infrastructure of the nation.
10. Transforming natural raw materials into commonly used products benefiting society's health, safety, and productivity.
11. Including prominent commercial centers, office buildings, sports stadiums, theme parks, and other sites where large numbers of people congregate to pursue business activities, conduct personal commercial transactions, or enjoy recreational pastimes.
12. Providing wired, wireless, and satellite communications to meet the needs of businesses and governments.
13. Transforming materials into finished goods.
14. Managing water retention structures, such as levees, dams, navigation locks, canals (excluding channels), and similar structures, including larger and nationally symbolic dams that are major components of other critical infrastructures that provide electricity and water.
15. Supplying the military with the means to protect the nation by producing weapons, aircraft, and ships and providing essential services, including information technology and supply and maintenance.
16. Saving lives and property from accidents and disaster.
17. Providing the electric power used by all sectors and the refining, storage, and distribution of oil and gas.
18. Ensuring the continuity of functions for facilities owned and leased by the government, including all federal, state, territorial, local, and tribal government facilities located in the United States and abroad.
19. Mitigating the risk of disasters and attacks and also providing recovery assistance if an attack occurs.
20. Producing information technology and including hardware manufacturers, software developers, and service providers, as well as the Internet as a key resource.
21. Maintaining monuments, physical structures, objects, or geographical sites that are widely recognized to represent the nation's heritage, traditions, or values, or widely recognized to represent important national cultural, religious, historical, or political significance.
22. Providing nuclear power.
23. Delivering private and commercial letters, packages, and bulk assets.
24. Enabling movement of people and assets that are vital to the economy, mobility, and security with the use of aviation, ships, rail, pipelines, highways, trucks, buses, and mass transit.
25. Providing sources of safe drinking water from community water systems and properly treated wastewater from publicly owned treatment works.

Summary

This chapter focused on how the commonality of the critical national infrastructure systems demonstrates certain desirable security attributes. These attributes must be present in all aspects and areas of the critical national infrastructure to ensure maximal resilience against cyber attacks.

Systems supporting the nation's critical national infrastructure are not sufficiently protected to consistently thwart the threats. While actions have been taken, the administration and executive branch agencies need to address the challenges in this area to improve the nation's cyber security posture, including enhancing cyber analysis, warning capabilities, and strengthening the public-private partnerships for securing cyber-critical infrastructure. Until these actions are taken, the nation's cyber-critical infrastructure will remain vulnerable.

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

Chapter Review Questions/Exercises

True/False

1. True or False? When security best practices are easily identified and measurable, they can become the basis for what is known as a security audit.
2. True or False? The ability to audit a given best practice does determine or influence whether it is useful for infrastructure protection.
3. True or False? Any commercial or government organization that is currently developing or managing a national infrastructure does not have a security policy.
4. True or False? An organization with a culture of security is one in which standard operating procedures work to provide a secure environment.
5. True or False? Oversimplification may be the first and most tractable step toward creating a new, more secure infrastructure environment.

Multiple Choice

1. Common security-related best practices standards include the following, except which one:
A. International Organization for Standardization (ISO)
B. Federal Information Security Management Act (FISMA)
C. Health Insurance Portability and Accountability Act (HIPAA)
D. Payment Card Industry Data Security Standard (PCI DSS)
E. ISO/IEC 27000 Standard (ISO27K)

2. Which of the following is one of the six best practices for national infrastructure protection?
A. Unorganized culture of security protection
B. Boundary scanning
C. No commitment to infrastructure simplification
D. Locally relevant and appropriate security policy
E. Certification and education program

3. Specifically, four basic security policy considerations are highly recommended for national infrastructure protection, except which one:
A. Enforceable
B. Diversity
C. Small
D. Online
E. Inclusive

4. A typical sort of cluttered engineering schematic that one might use to describe network infrastructure suffers from the following issues that one finds in the design and operation of the national

infrastructure, except which one:

A. Lack of generalization

B. Clouding the obvious

C. Stream-of-consciousness design

D. Nonuniformity

E. Procurement discipline

5. Which of the following is one of the ways to simplify a National Infrastructure Environment?

A. Investment size

B. Operations concept

C. Accuracy interface

D. Coverage patterns

E. Reduction in size

Exercise

Problem

Imagine a real-life cyber attack where a downloadable application turns smartphones into network-clogging bots, causing the U.S. critical mobile phone network infrastructures to fail, and eventually spreads to the wireless Internet. Then, it starts to spread to the energy grid on the eastern seaboard, where it begins to fail. In other words, the cyber attack started small, with a downloadable application infecting smartphones. But the number of infected phones grew, and the malware started attacking the wireless Internet as smartphone users synched their phones with their computers. The malware began sending huge video files across the Internet, crippling both mobile networks and the wireless Internet. Identify what type of new countermeasures should have been implemented to prevent this cyber attack from occurring.

Hands-On Projects

Project

This scenario details a cyber attack that resulted in a self-propagating virus spreading across a bank's networks, leading to a steadily increasing number of files to become encrypted and, thereby, inaccessible to the bank. These issues were further complicated by an extortion demand and the execution of a successful denial-of-service attack. So, with the preceding in mind, how would you go about creating an exercise project with the paradoxical requirement that infrastructure systems must also demonstrate a degree of commonality?

Case Projects

Problem

This project concerns cyber attacks affecting the availability of the Internet in several European countries. The

basic idea is that Internet interconnectivity between countries becomes gradually unavailable. As a result citizens, businesses, and public institutions will have difficulties in accessing critical infrastructure online services, unless the traffic from affected interconnections is rerouted. As the cyber attacks continue, one country after the other will increasingly throughout the day suffer from this problem, over phone and mails. Explain how you would reduce potential vulnerabilities, protect against intrusion attempts, and better anticipate future threats.

Optional Team Case Project

Problem

This scenario is an intentional cyber-security attack on the SCADA system of water or wastewater utilities. It occurs during the summer in Fringe City. A disgruntled utility worker, laid off due to recent budgetary cutbacks, decides to infiltrate the SCADA system from a dial-in connection from his home computer. He infects the SCADA system with a virus that hinders its operation. The system begins to issue alarms that inform the utility operators that various systems in the treatment process are malfunctioning, and that the water or wastewater leaving the plant is not meeting water quality standards. So, in keeping the preceding in mind, identify how the utility would go about tackling this type of cyber attack.

[1] Quoted in A. K. Dewdney, "Computer recreations: of worms, viruses and Core War," *Sci. Am.*, 260(3), 90–93, 1989.

6

Depth

Chapter Outline

Effectiveness of Depth
Layered Authentication
Layered E-Mail Virus and Spam Protection
Layered Access Controls
Layered Encryption
Layered Intrusion Detection
National Program of Depth
Practical Ways for Achieving Information Assurance in Infrastructure Networked Environments
Summary
Chapter Review Questions/Exercises

Sun myth: If a person is wearing a foundation makeup with SPFs of #4 or #8, then she won't need additional sunscreen or sunblock.

http://www.ultimate-cosmetics.com

The general security strategy of *defense in depth* is based on the observation that any given layer of protection can fail at any time. As such, defense in depth involves the deliberate introduction of multiple layers of defense in order to increase the likelihood that a given attack will be stopped or at least slowed down. This likelihood is dependent upon the quality and relative attributes of the various defensive layers. Cost and end-user experience issues usually create constraints on just how strong the various layers can actually be in practice. Most security experts understand this strategy of defense in depth, but evidence of its use in national infrastructure settings is often lacking. This is too bad, because the protection of national infrastructure lends itself naturally to multiple layers of defense.

The general schema associated with layered defense is that a series of protective elements is located between an asset and the adversary. Obviously, it would be best if the series is actually that—a serial collection of protective elements that must each be traversed successfully to gain access to a protected resource. Most of the time, however, the layering is not so efficient and may include different combinations of elements between an asset and an adversary. The strategic goals in such cases are to detect and remove any single-layer access paths and, obviously, to avoid situations where the layers might be conflicting. For national infrastructure, the goal is to place multiple security layers in front of all essential services (see Figure 6.1).

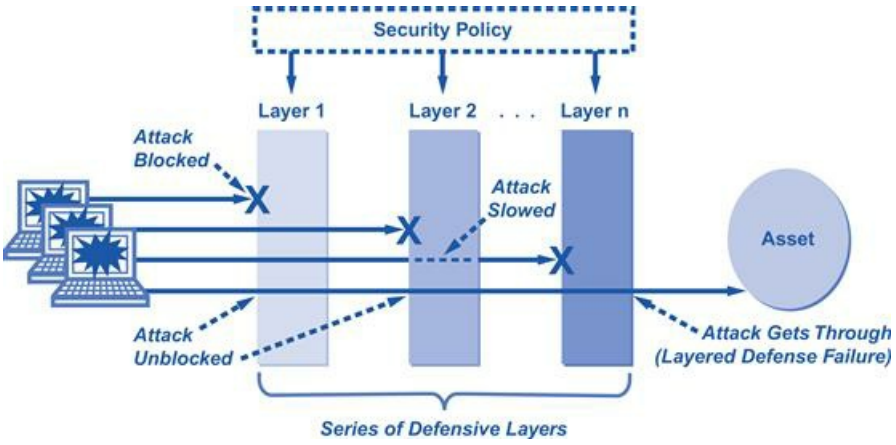


Figure 6.1 General defense in depth schema.

The security intent for any series of layers is to enforce policy across all possible access paths to the target asset. Thus, if an asset is accessible through a single entry point, then the layers only need to enforce policy at that point. If an asset is broadly accessible from a collection of different entry points, then the layered defense needs to fan out across these points to enforce policy. Defense in depth methods are said to fail if all of the layers do not either block or sufficiently degrade attack attempts at the protected asset, resulting in security policy violations by an adversary. It is relatively easy to determine that a failure has occurred when an attack is detected; however, when an attack goes unnoticed or when the forensic analysis after an attack cannot determine the point of exploitation, then holes in layered defenses might remain indefinitely.

If layered defenses are penetrated, it is crucial to identify the entry point used by the attacker.

Defense in depth implementations are sometimes inappropriately or even maliciously bypassed by presumably trusted users, generally insiders to a company or agency, including its employees, contractors, and partners. For example, an infrastructure organization might create diverse layers of security functionality to ensure that intruders cannot compromise assets from an external environment such as the Internet. Problems arise, however, if malicious insiders can directly access and compromise assets. This implies that great rigor and discipline are required to ensure that defense in depth truly surrounds an asset, both internally to an organization, as well as externally on the Internet. This generally requires additional functional controls on the local enterprise network to protect assets from insiders.

Do not overlook the need for protection against both internal and external adversaries.

Depth strategies sometimes involve the familiar military notion of one protection layer slowing down an intruder. It turns out that throttling does not always extrapolate well to cyber security. In practice, cyber security methods tend to be binary in their functionality; that is, a protection will either work or it will not. Debates thus arise around how long an approach will hold off attackers, as in the selection of cryptographic key length. Similarly, network attacks are often dealt with by throttling or rate-limiting the traffic allowable into a target asset environment. These approaches might work to a degree, but they are the exceptions, and it

is recommended that cyber security architectures for national infrastructure not rely on any element having only a partial effect on a given attack.

Ideal defensive strategies will stop—not slow down—an adversary.

Effectiveness of Depth

Academics formally model the effectiveness of a collection of defensive layers using mathematical probability. Such an approach requires that one quantitatively measure the relative dependencies between the layers, as well as the probability of effectiveness for any given layer. Unfortunately, in any nontrivial environment, both of these estimates are unlikely to be more than just an educated guess. We know, for example, that the success of access controls for enterprise applications is dependent on the success of strong authentication for remote access. Trying to accurately quantify this dependency for probabilistic analysis is a waste of time and will not result in any estimate better than an expert guess.

How can effectiveness of a security layer be measured or quantified?

Thus, from a practical perspective, and in the context of real national infrastructure protection, determining the effectiveness of a defense in depth scheme must be done via educated guesses. We can make this sound better by referring to it as *informal subjective reasoning based on relevant security factors*, but it is still just a guess. The relevant factors for estimating effectiveness of a layer include the following:

- *Practical experience*—One can certainly analyze practical experience and past results for a given security method. This is dangerous if taken too literally, because many attacks are missed, and seemingly correct, but actually vulnerable, defenses might be dormant for a period of time before an attack.
- *Engineering analysis*—Experienced security engineers will use their knowledge and expertise to provide excellent judgment on whether a given layer will be effective. Vendors and salespeople are to be avoided in this process, because they will invariably distort their product and service capability.
- *Use-case studies*—Providing some rigor to the engineering analysis is a good idea, and the familiar use-case methodology is especially appropriate for security layers. It is really a form of testing.
- *Testing and simulation*—Actual testing of a layer in a controlled setting will provide good information on its effectiveness. Simulation is also a good idea in cases where a defensive layer protects against something not easily tested, such as a massive denial of service attack.

To illustrate this approach, let's start with a simple setup, as shown in Figure 6.2. Specifically, a single layer of protection depth is depicted and is estimated to have "moderate" effectiveness. We can assume that some subset of the factors described above was used to make this determination. Maybe some team of experts analyzed the protection, looked at its effectiveness in similar settings, and performed a series of tests and simulations. In any event, let's assume that they decided that a given protection would be moderately effective against the types of attacks to be expected in the local threat environment.

A moderately effective defense strategy will stop most, but not all, attacks.

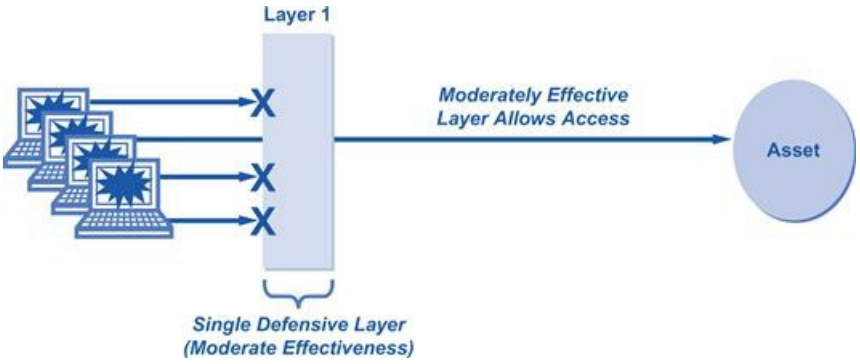


Figure 6.2 Moderately effective single layer of protection.

The determination that this single layer is "moderately" effective is nothing more than a subjective guess in most cases. It is, however, an important piece of information for national infrastructure protection because it implies that the protection will not work in all cases; that is, the experts have determined that some types of attacks will bypass or break the protection and will thus expose the asset to malicious intruders. As a result, when a given protection layer does not address all known attacks, then we can conclude the following:

- *Flaws*—The protection might be flawed. This could be some minor issue such as an obscure bug that would allow certain types of attacks or it could be potentially front-page news with major implications. In either case, flaws in protections require either that they be fixed or that they be mitigated by a complementary layer of protection.
- *Suitability*—The protection might be unsuited to the target environment; for example, it might be intended to prevent events A and B in an environment where the real threat is event C. Such scenarios are commonly found during incident response, when some event has occurred and the presumed protections are discovered to have had little effect, simply because of a mismatch. This is fixed by either changing the layer or complementing it with another.

Whether the layer is flawed or mismatched, the situation is made worse if the adversary has knowledge of the situation. Regardless of the common argument by hackers that exposing problems in a protection method should always be reported, the reality is that such information generally does more harm than good. Certainly, if an organization is lax in fixing a problem with broad implications, this is unacceptable, but the technique of extorting that group into taking immediate action is not always in everyone's best interests. The hacker who exposes vulnerabilities in a moderately effective mobile telephony control, for example, without first alerting the service provider, might be guilty of degrading essential communication services that might affect human lives.

Multiple layers of protection will mitigate the effects of flaws or protections that are unsuited to the target environment.

Assuming an organization is diligent and chooses to improve or fix a moderately effective protection, the result will be that the new estimate or guess might be "highly" effective. For example, suppose that some home-grown intrusion detection system is becoming difficult to maintain. The local team might thus

determine that it is only moderately effective and might replace it with a vendor-supported product. In most cases, the new system would now be viewed as highly effective (with the caveat that no intrusion detection systems ever seem to work as well as they should). The end result is that the layer has now been improved from moderately to highly effective. It should be obvious that even in a highly effective protection environment, there will always be exceptional conditions where the protection may fail (see Figure 6.3).

A protection layer can be improved to become "highly" effective, but no layer is 100% effective all of the time.

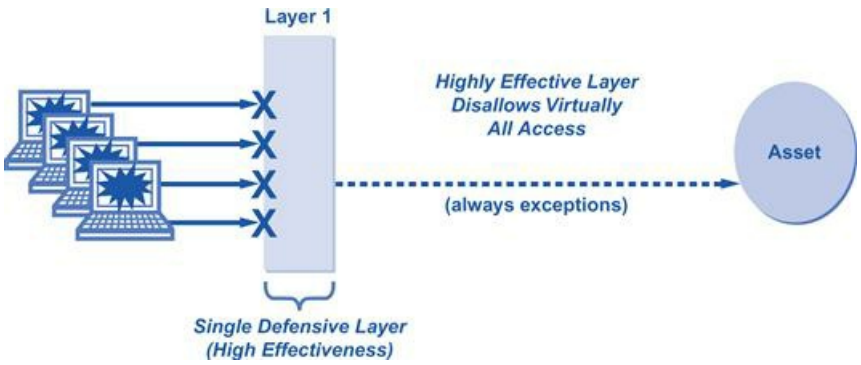


Figure 6.3 Highly effective single layer of protection.

Improving one layer is not, however, the only option available. An alternative would be for the moderately effective control to be left in place and complemented with another layer of protection. This has certain advantages, including reducing the cost and risk of forklifting out a security protection layer and replacing it with a new one. The result of complementing one moderately effective protection layer with another is that the end result should mitigate a larger set of attacks. This does introduce an odd sort of calculus to the security manager, where decisions are required around whether some number of moderately effective protections is better or worse than a smaller number of stronger protections (see Figure 6.4).

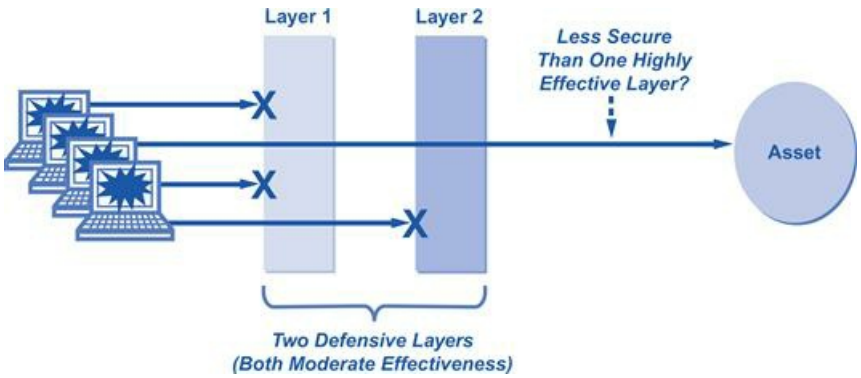


Figure 6.4 Multiple moderately effective layers of protection.

The answer to whether multiple moderately effective layers outperform fewer highly effective ones will depend on aggregation considerations. That is, if two moderate protections complement each other by

balancing each respective weakness, then the composite protection will be quite good. If, on the other hand, multiple moderate protections suffer from similar weaknesses, then the weakness will remain in the aggregate protection. In practice, security managers generally should look for a diverse set of protections that are as strong as possible and that balance weaknesses in some demonstrable manner. For national infrastructure protection, this will typically involve layers of protection in authentication, malware protection, access controls, encryption, and intrusion detection.

Diversity of protection layers—including diversity of weaknesses—is critical in maintaining successful protection against attacks.

Layered Authentication

Most information technology (IT) and security teams in government and industry are committed to reducing the number of passwords, passphrases, handheld tokens, certificates, biometrics, and other validation tokens that exist in their environment. These initiatives are generally met with great enthusiasm among end users, because they result in simpler, cleaner infrastructure and much less for end users to have to remember, protect, or write down. One cannot deny that such simplification has a beneficial impact on overall security. For these reasons, various proposals have been made for national authentication systems run by government and that would include every citizen.

Single sign-on (SSO) initiatives are generally used to accomplish this authentication simplification objective. SSO is accomplished by the use of a single, common identification and authentication system for all relevant applications. This common system is then embedded into one identity management process so reported identities can be administered and protected uniformly. The simplification inherent in SSO is desirable from a security perspective, because it reduces the likelihood of errors that result when multiple complex login systems are present. Common identity management is thus generally desirable from a security perspective, especially in enterprise settings.

End users will embrace authentication simplification initiatives, and these are certainly easier to monitor from a security management standpoint.

Problems can arise, however, in national infrastructure protection environments if the process of streamlining authentication goes too far. Even the staunchest advocate of SSO must agree that, for certain applications, a properly managed, properly designed, and diverse series of authentication challenges that are reliant on separate proof factors will be more secure than a comparable SSO system. The diverse series of authentication steps will certainly be less convenient for end users but, if run correctly, will be more secure. This is because such a scheme avoids the nightmarish scenario where a single login provides an adversary with common access across multiple national infrastructure systems. This attack scenario is so unacceptable at the national level that it dictates special consideration.

Single sign-on initiatives may be embraced by end users but may not provide the ideal level of security protection.

Specifically, for national infrastructure management, organizations can acceptably maintain the goal of balancing the risks and rewards of SSO for all enterprise-grade applications such as business e-mail, routine applications, and remote access. As long as no national assets can be directly compromised with SSO access, this is fine. Companies and agencies charged with national infrastructure can and should move to an SSO scheme with corresponding identity management. For critical national services and applications, however, a more complex, defense in depth scheme is highly recommended for end-user authentication (see box).

Factors of a Successful National Infrastructure SSO Access System

Critical national infrastructure services need a defense in depth scheme that is developed with the following considerations:

- *Diversity with single sign-on*—Authentication systems for national asset protection must be different from the SSO scheme used for enterprise access. This implies that a separate technology, vendor, and management process should be considered between enterprise SSO and national infrastructure authentication. The goal is to ensure that flaws in one authentication system are not present in the other.
- *Diversity of proof factors*—Similarly, the familiar proof factors:
- "Something you know"
- "Something you have"
- "Something you embody (biometrics)"
- "Somewhere you are"

should be diverse for national assets from any SSO proof factors. This implies that employees should not be handed a single handheld authenticator that can be used to gain access to e-mail and also to some critical infrastructure operational component.

- *Emphasis on security*—While it is acceptable to emphasize usability in enterprise SSO initiatives, the emphasis of national infrastructure protection should shift squarely toward security. The only relevant end-user issues are ones that simplify usage to reduce errors. Convenience should not necessarily be a major goal, as long as the authentication scheme does not drive bad behavior such as sharing tokens or writing down passwords.

A resultant typical defense in depth scheme for national infrastructure organizations would include SSO for enterprise-grade applications and access and a subsequent, diverse authentication process for all national assets. The result is that end users would need to be authenticated twice before gaining access to a critical asset. Correspondingly, intruders would have to break through two authentication systems to gain malicious access to the target asset. End users probably would not like this and the costs are higher, but the increased security is worth the trouble (see Figure 6.5).

Single sign-in access can be part of a multilayered defense in depth strategy.

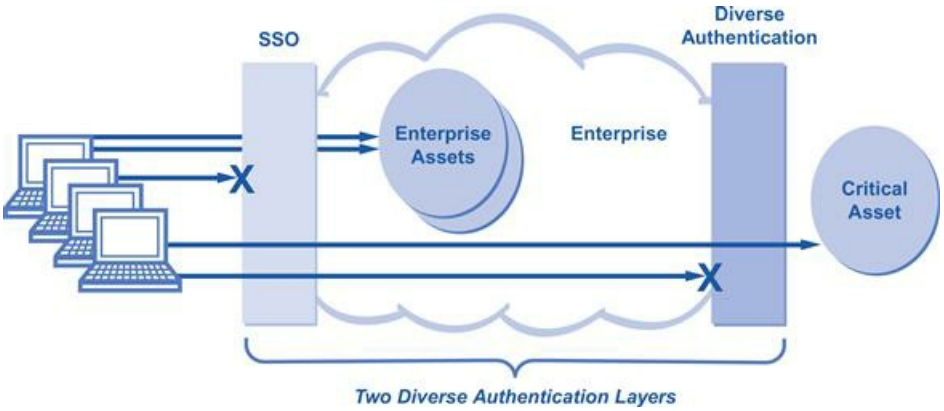


Figure 6.5 Schema showing two layers of end-user authentication.

For multiple critical national assets in an infrastructure environment, the depth strategy should include maximal diversity for each asset. That is, the general computing characteristics and source of the authentication functionality should be diverse. Furthermore, the factors used in establishing proof of identity for critical assets should be stronger than simple passwords; handheld authentication or biometrics would be recommended. An implication here is that the underlying infrastructure be operated with the greatest precision and correctness. Administrative procedures for obtaining an authentication token, restoring access when a token or password is lost, and providing assistance to confused end users must be carefully designed to avoid social engineering attacks. At the national level, this would require frequent testing.

A key modern consideration for enterprise authentication is the degree to which mobile access to infrastructure potentially changes security posture. As an example, consider that most organizations go to great lengths to ensure that several layers of authentication reside between remote workers and sensitive applications such as enterprise e-mail. In fact, see the box to follow the experience most people have when trying to get their enterprise e-mail from a remote location using a laptop.

The example in the box also highlights the importance of recognizing trends in technology as national infrastructure protection initiatives are considered. For the enterprise, the old notion of protected perimeter thus disappears with the advent of mobile access across wireless carrier infrastructure. One still finds architectures where users must "hairpin" their mobile access to the enterprise and then through a firewall to the target application, but this practice is likely to wane (see Figure 6.6).

Unfortunately, mobile devices eliminate the multi-layered protection most companies build into their remote network access.

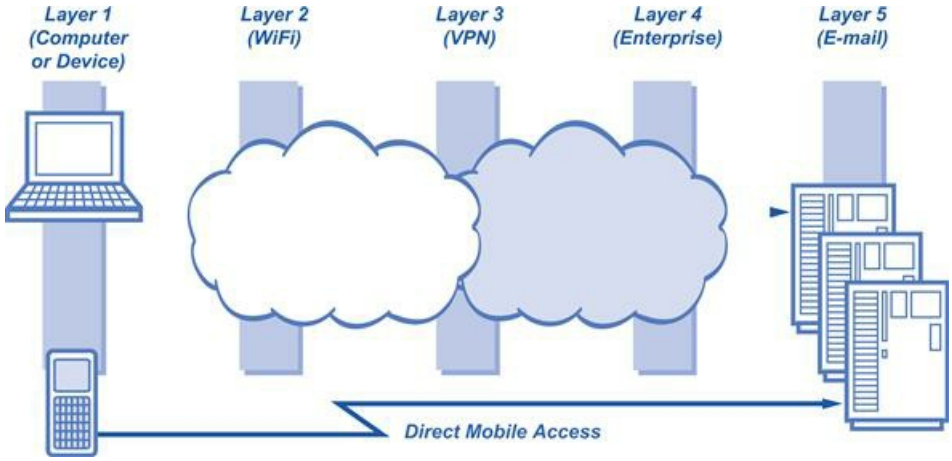


Figure 6.6 Authentication options including direct mobile access.

For applications such as enterprise e-mail, this type of convenient bypass might be perfectly fine. In fact, for enterprise e-mail specifically, it would be unreasonable to expect that workers in national infrastructure settings should not be allowed mobile access. For more sensitive national infrastructure applications, however, such as those that provision or control critical systems, a threat analysis would be required before any alternative paths with mobile devices are allowed. Classified information would be another example asset that requires multiple layers without mobile access bypass. These types of requirements should find their way into any type of national infrastructure support contracts.

Multi-Layered Protection: Five Steps to Remote E-Mail Access

A typical remote worker will need to follow these steps to access their enterprise e-mail account:

- *Authentication layer 1.* The user must first login to the computer. Presumably, this is done using a password that is set by the enterprise information technology or security group.
- *Authentication layer 2.* The user must then login to the local WiFi or broadband access network. Sometimes this is free; other times it requires a credit card, which can be viewed as an added identification step.
- *Authentication layer 3.* The user must then login to the remote access server, probably over a virtual private network (VPN). Most of the time, companies and agencies require a personal identification number (PIN), password, or handheld token to authenticate VPN access.
- *Authentication layer 4.* The user must then login to the enterprise network, probably with some sort of domain password. This is also controlled by the local information technology or security group.
- *Authentication layer 5.* The user must finally login to the specific e-mail application being used by the enterprise. Sometimes this requires another password, but often it just requires access.

On the surface, this would seem like the ultimate in layered authentication with no less than five layers! The problem is that many organizations provide their employees with means to remotely access applications such as e-mail with a handheld device. Consider, in this case, the experience most people have when trying to retrieve their enterprise e-mail using a mobile device:

- *Authentication layer 1.* The user must simply login to the mobile device, click on the e-mail icon, and

then read or create mail.

This is obviously only one layer of authentication for mobile devices, and it demonstrates the importance of recognizing that users might find more convenient paths around presumed layers of authentication.

Exposing critical national assets to mobile access (even by trusted personnel) opens a gateway for an adversarial attack.

Layered E-Mail Virus and Spam Protection

Commercial environments are increasingly turning to virtualized, in-the-cloud solutions for their gateway filtering of e-mail viruses and spam. This decision allows the organization to remove the gateway filters or to simply offload the work those filters must perform. This is a healthy decision, because a general security principle is that attacks should be stopped as close as possible to their source. The network is certainly closer than the attack target's ingress point, so virtual filtering is desirable. It is also helpful to the carrier, because it reduces the junk floating around network infrastructure, which helps carriers perform their tasks more efficiently in support of national services.

Managers of commercial environments have also come to recognize that their computing end points cannot rely solely on gateway or in-the-cloud processing. As such, the state of the practice in e-mail virus and spam protection involves a defense in depth deployment of filters to each laptop, netbook, personal computer, and server in the enterprise. The approach is even beginning to find its way to the mobile handheld device, where the threat of viruses and spam is increasing. As such, a given virus or spam e-mail sent from a malicious source will have to find its way through at least two layers of filtering in order to reach its intended source (see Figure 6.7).

Mobile devices are susceptible to viruses and spam, yet spam is more of a nuisance than an actual threat to national infrastructure.

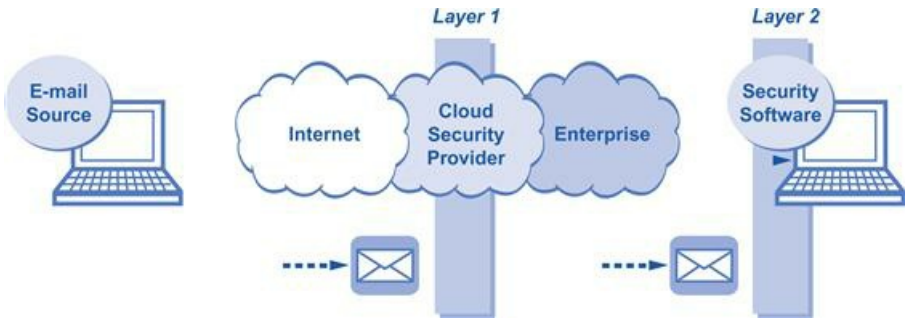


Figure 6.7 Typical architecture with layered e-mail filtering.

This cloud filtering arrangement found in most companies is acceptable for organizations charged with national infrastructure. For the most critical applications, it is recommended that a depth approach involving both in-the-cloud and perimeter processing be employed. In addition, for key executives in these companies and agencies who might be directly targeted by adversaries, additional desktop and application filtering might be prudent. Practical experience suggests that spam is more a nuisance than significant threat to national asset management, so the likelihood of attackers using spam to interrupt national services is only moderate. In addition, antivirus software has become less relevant in recent years, simply because so many software threats such as well-coded bots are not easily detected by antivirus software. Research into better techniques for detecting the presence of malware should become an immediate national priority.

Antivirus software, while still necessary, is not likely to detect such threats as a botnet attack.

Layered Access Controls

Access controls determine who can access what resources under which conditions. They are one of the most common and most mature security protection methods, dating back to the earliest electronic computers. If some asset is protected by a single set of access controls, then this is similar to using a single combination lock to protect a physical asset. That is, if an individual has the correct combination, then access is allowed. Common access controls include access control lists (ACLs) on Windows®-based operating systems and permissions vectors in UNIX®-based operating systems. These are implemented as software data structures that determine access based on some defined policy.

One approach to using defense in depth to protect a software application involves embedding one type of access control into the application environment and then hosting the application on an operating system that utilizes a different type of access control. In such a setup, access to the application can only be obtained by successfully negotiating the following layers:

- *Access control layer 1*. The user must be permitted entry to the operating system via the operating system access controls. This might be UNIX® permissions, Windows® ACLs, or something similar.

Some form of access control is present in any network connection (e.g., your personal password to access your e-mail account).

- *Access control layer 2*. The user must be permitted entry to the application via the application access controls. This is likely to be a password embedded in the application environment and controlled by the application owner.

In cases where an operating system and application cannot be remotely reached, these two layers can be augmented with additional diverse controls such as guarded access to the physical premise or to a locked data center. This implies that access to an application would require first obtaining physical access to a console before access to the operating system and application can even be attempted. These two layers of authentication are important and should be tested in every national infrastructure environment, especially ones employing supervisory control and data acquisition (SCADA), where computer security techniques have a more short-lived legacy. A caution, however, is that insiders are likely to possess both types of access, so the layers will not be helpful in stopping most forms of sabotage.

Restricting physical access to assets always adds another layer of protection from outsiders, but not from internal saboteurs.

In cases where remote access is allowed, then the use of a firewall is the most common method to ensure policy compliance for those permitted access. Such policy is almost always based on the source Internet protocol (IP) address of the requesting party. This is not the strongest of access control methods, simply because IP addresses are so easily spoofed. Also, to maintain such a scheme, a complex and potentially error-prone or socially engineered bureaucracy must be put in place that accepts and maintains access requests.

When used in conjunction with additional access control layers such as operating system and application controls, the result might be acceptable in some environments (see Figure 6.8).

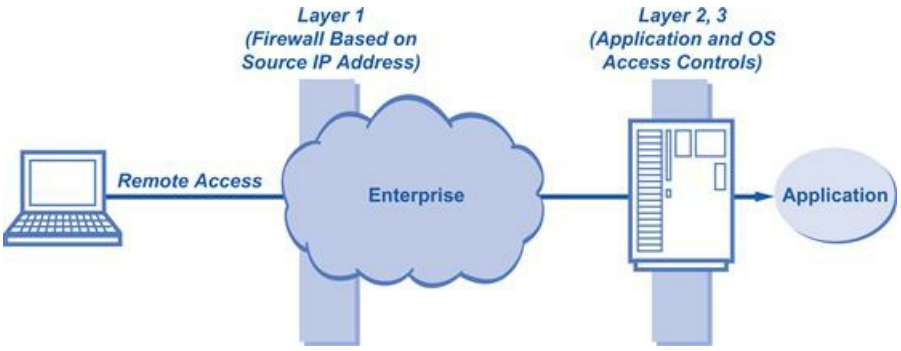


Figure 6.8 Three layers of protection using firewall and access controls.

For national infrastructure protection, critical assets should be covered by as many layers of access control as deemed feasible. As with authentication, the issue of end-user convenience must be viewed as lower priority if critical national services are at stake. Some general heuristics for protecting national infrastructure with layered access controls include the following:

The implementation of layered access controls places greater emphasis on protection than on end-user convenience.

- *Network-based firewalls*—Using cloud firewalls offers an additional blanket layer of control. This technique is useful as a complement to existing enterprise controls, especially because carrier-based systems will generally differ from whatever firewalls and related systems might be deployed in the enterprise.
- *Internal firewalls*—This provides yet another layer of protection within the enterprise to ensure that individuals with access to resource *X* only gain access to that resource and no other. Routers can often provide a simple packet-filtering capability as part of their native processing suite, which simplifies architecture and minimizes cost.
- *Physical security*—Excellent facility and premise-access security provides an additional tangible layer of protection and is essential for any national infrastructure protection initiatives. This must be complemented by selecting suitable applications and systems that can never be accessed remotely or even across a local area network.

When multiple access control systems are in place, the benefit of layering is reduced when the underlying administration function is performed by one team using a common set of tools. When this involves a protected and carefully managed security operations center the situation is acceptable, but when the management is *ad hoc* and poorly controlled the layering might be undermined by an attacker who successfully infiltrates the administration systems.

Multiple access control systems must be well managed so as not to allow an internal attacker successful

infiltration to the systems.

Layered Encryption

Encryption is an effective and well-known security control for protecting information. While mathematicians and computer scientists have created hundreds of different taxonomies for categorizing symmetric and public key systems, the box shows specific methods that are useful for the protection of national infrastructure.

Five Encryption Methods for National Infrastructure Protection

1. *Mobile device storage*—Mobile smart phones and laptops should have native encryption to protect against loss or theft and the resulting information compromise. The encryption will never be perfect but should provide useful protection in the field. Several vendors offer this type of encryption as an add-on service, but this should eventually become a native function in all mobile devices and laptops.
2. *Network transmission*—Any sensitive data being transmitted within an enterprise or between knowing partners should be encrypted. The traditional means for such encryption has been symmetric and embedded in hardware devices. More recently, the associated cryptography is often software based and involves public keys supported by public key infrastructure (PKI) tools. When network transmission occurs in an *ad hoc* manner, the practical consideration is that shared cryptography simply does not exist between organizations due to complexity. This makes it difficult to encrypt network traffic without coordinating things in advance.
3. *Secure commerce*—If an organization offers electronic commerce services over the Internet, the use of common encryption techniques such as Secure Sockets Layer (SSL) is presumed. The associated cryptography here will be public key based.
4. *Application strengthening*—E-mail is the most obvious application that can introduce secrecy and authentication properties via the use of encryption. As noted above, federating this cryptography, almost always public key based, between organizations has not been done on a wide scale to date.
5. *Server and mainframe data storage*—Encryption on servers and mainframes has received considerable attention in recent years but should be viewed with suspicion. Data at rest is poorly protected by cryptography because the associated key management systems, which require a long life, can have obvious holes. In the worst case, sloppy key management can make data less secure. Note that smart phones and laptops are different from servers because they are *moving*.

The good news is that, for the most part, these five encryption methods will not collide in practice. They can be used in combination and in cooperation, with no great functional or administrative problems expected. It is also perfectly fine to encrypt information multiple times, as long as the supporting administrative tools are working properly. As such, one can easily imagine scenarios where all five systems are in place and provide five different layers of information protection. Not all will typically reside in a perfect series, but all can be in place in one infrastructure setting providing layered security (see Figure 6.9).

Information can be encrypted multiple times to achieve layered protection.

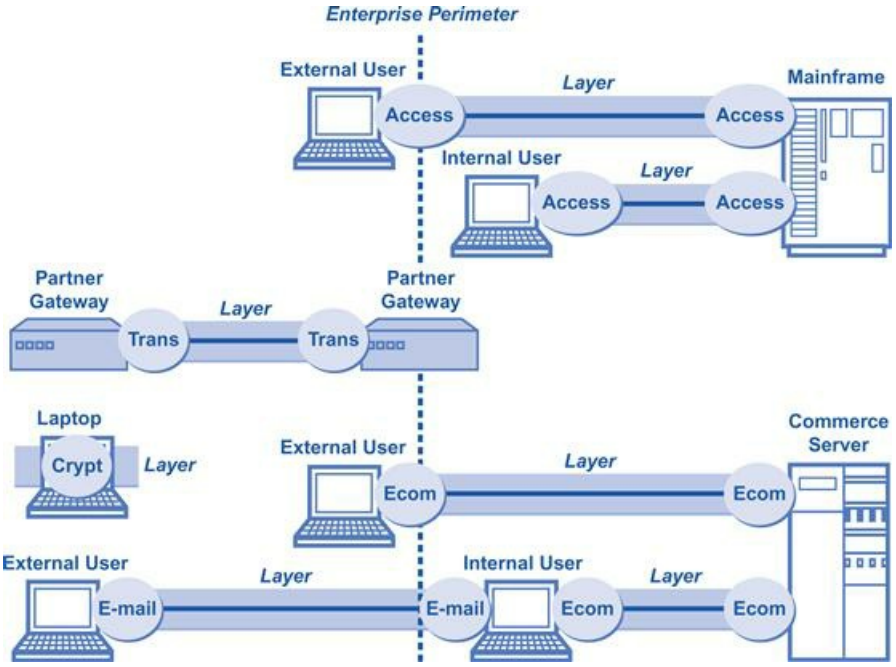


Figure 6.9 Multiple layers of encryption.

The bad news, however, is that each will typically require its own user administration and key management systems. The result is a disparate view of cryptography across the enterprise that can be seen in the somewhat scattered arrangement in Figure 6.9. This is unfortunate, because it increases complexity, which increases the chances of error or compromise, especially to underlying infrastructure. Regardless, the use of cryptography in national infrastructure protection should be encouraged, even if the layers are not optimally coordinated.

Layered Intrusion Detection

Intrusion detection was once viewed as the most promising of large-scale security techniques. Even the provocative and hopeful name "intrusion detection" suggests a powerful technology that can be inserted into an environment to alert security teams when an intrusion is imminent. While this goal has not been fully met in practice, intrusion detection does provide a useful means for detecting indicators of potentially harmful behavior. These indicators are sometimes used for early warning, but more often are used to correlate with other types of available information during an incident.

Because intrusion detection is typically performed offline, it lends itself to multiple layers of monitoring. Obviously, if the intrusion detection includes an active response—which is referred to collectively as *intrusion prevention*—the layered arrangement could be more complex, but for now let's analyze strategies for passive, offline monitoring of attack. Most organizations accomplish this task using commercial systems that include three components: monitors that are placed in strategic locations to collect data, transmission systems that move alarm information to a central location, and a master monitoring function that processes incoming data and provides some sort of correlated summary, usually in the form of an alarm to a console. When this type of intrusion detection system is in place in an enterprise, it can be viewed as an explicit layer of protection. In fact, many auditors will accept intrusion detection as a complementary control when some other protection displays weaknesses.

Intrusion detection with data security is similar to physical security intrusion detection: monitoring, an alarm system, and a central console.

One can conceptualize an alternate layer of intrusion detection being put in place at a broader level, perhaps coordinated by some government or industry group. The components of the system would be the same, but differences from the enterprise would include diverse monitor placement, different signatures of attack, and a broader base on which to perform correlation of data. An issue with this alternative layer is that the protection would likely involve network paths that are largely separate from those in specific enterprise settings. For example, an intrusion aimed at some government agency would not be detected by the intrusion detection system located within a separate enterprise. There are, however, three specific opportunities for different intrusion detection systems to provide layered protection:

- *In-band detection*—If two intrusion detection systems both have monitoring access to the same attack stream, or a related one, then they might both have the opportunity to detect the condition. Thus, if one system fails, it is possible that another might not. This is the essence of defense in depth, but it only works if the response processes for each detection system are coordinated.
- *Out-of-band correlation*—During an incident, the operators of an intrusion detection system might benefit from information that might become available from other operators. This can be intelligence about sources, methods, or techniques being used by attackers. It is usually best used if made available in real time.
- *Signature sharing*—A special case of the above correlation involves sharing of specific attack signatures by one operator that can be keyed into the systems being run by other operators. Military

organizations, for example, sometimes develop signatures that could be shared with industrial groups to improve their security.

In each of these cases, diverse intrusion detection systems can be viewed as providing a defense in depth for target assets. The result is a potentially coordinated series of intrusion detection layers that will help protect national infrastructure. This coordination usually requires sharing between different monitoring and analysis centers; that is, if one intrusion detection system notices an attack such as a botnet, then it might share this information with another system that might not have detected the condition (see Figure 6.10).

A certain amount of information sharing between government agencies may serve to increase intrusion detection effectiveness.

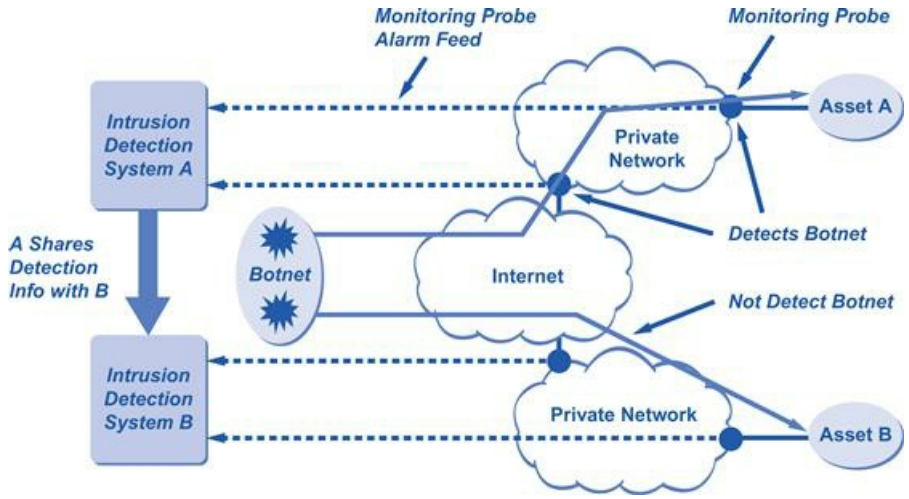


Figure 6.10 Sharing intrusion detection information between systems.

This idea of coordinated intrusion detection systems is certainly not new; for example, government cyber security commissions and groups have advocated the notion of signature sharing between government and industry for years. For whatever reason, however, such coordination has rarely occurred, but for national infrastructure protection to reach its full potential such cooperation must be encouraged and rewarded.

National Program of Depth

Creating a coordinated program of defense in depth using multiple layers of security for national infrastructure can only be ensured via careful architectural analysis of all assets and protection systems. The architectural analysis should result in a mapping, perhaps represented as a matrix, where each critical national asset is shown to be protected by certain multiple layers of security. For each layer, subjective determination of its effectiveness is also required. Once this is done, simple calculations can be performed to determine the difficulty of penetration through the various layers. This task is easier than it sounds; some of the more practical considerations that arise in such an exercise include the following:

Reviewing systems and strategies to identify existing layers of protection will create a "map" of the current depth of defensive protection.

- *Identifying assets*—This is a required step for several of our recommended national infrastructure protection principles, including, for example, deception. It is particularly important for defense in depth, because the analysis of depth effectiveness can only be measured from the specifically identified assets.
- *Subjective estimations*—The challenges inherent in this step were explained in detail above; certainly, in practice, certain conventions could arise that would help security experts arrive at common estimations of effectiveness. In the 1980 s, the U.S. Department of Defense created a set of criteria (informally called the Orange Book) for measuring the effectiveness of security in systems. Perhaps some elements of this criteria approach could be introduced to provide assistance in subjective estimations of the effectiveness of a layer.
- *Obtaining proprietary information*—If a company or agency has some defense in place (or, more importantly, perhaps some defense that may be missing) for some essential national service, then obtaining this information for broad analysis may be difficult. The goal would be to demonstrate value for organizations sharing detailed information, even if it is bad news.
- *Identifying all possible access paths*—Perhaps the toughest part of any cyber security exercise involves trying to determine means for accessing some target. If this is not done properly, then the defense in depth strategy will fall apart, so this important step requires special consideration.

These considerations can introduce significant challenges in practice. It does not help that most existing security teams, even in large-scale settings, rarely go through a local exercise of identifying defense in depth conditions. As a result, most national infrastructure protection teams would be working this exercise for the first time in the context of a national program.

Finally, let's briefly look at some practical ways for achieving information assurance in critical national infrastructure environments through the general security strategy of defense in depth. As previously stated, defense in depth involves the deliberate introduction of multiple layers of defense in order to increase the likelihood that a given attack will be stopped or at least slowed down.

Practical Ways for Achieving Information Assurance in Infrastructure Networked Environments

To effectively resist cyber attacks against its information and information systems, an organization needs to characterize its adversaries, their potential motivations, and their classes of attack. Potential adversaries might include nation states, terrorists, criminal elements, hackers, or corporate competitors. Their motivations may include intelligence gathering, theft of intellectual property, denial of service, embarrassment, or just pride in exploiting a notable target. Their classes of cyber attack may include passive monitoring of communications, active network attacks, close-in attacks, exploitation of insiders, and attacks through the industry providers of one's information technology resources. It's also important to resist the detrimental effects from nonmalicious events such as fire, flood, power outages, and user error through information assurance.

Information assurance in this case is achieved when information and information systems are protected against such cyber attacks through the application of security services such as availability, integrity, authentication, confidentiality, and nonrepudiation. The application of these services should be based on the protect, detect, and react paradigm. This means that in addition to incorporating protection mechanisms, organizations need to expect cyber attacks and include attack detection tools and procedures that allow them to react to and recover from these attacks. So, with the preceding in mind, an important principle of the defense in depth strategy is that achieving information assurance requires a balanced focus on three primary elements: people, technology, and operations.

The defense in depth strategy recommends several information assurance principles (see "An Agenda for Action for the Defense in Depth Strategy"). These include defense in multiple places and layered defenses.

An Agenda for Action for the Defense in Depth Strategy

Given that adversaries can attack a target from multiple points using either insiders or outsiders, an organization needs to deploy protection mechanisms at multiple locations to resist all classes of cyber attacks. When completing the defense in depth strategy, an organization should adhere to the provisional list of actions to increase risk (of detection) for the adversary, while reducing his or her chances of success or making successful penetrations unaffordable. The order is not significant; however, these are the activities for which the research would want to provide a detailed description of procedures, review, and assessment for ease of use and admissibility. As a minimum, these defensive "focus areas" should include (check all tasks completed):

1. Defending the networks and infrastructure.
2. Protecting the local and wide area communications networks (e.g., from denial of service attacks).
3. Providing confidentiality and integrity protection for data transmitted over these networks (using encryption and traffic flow security measures to resist passive monitoring).
4. Defending the enclave boundaries (deploying firewalls and intrusion detection to resist active network attacks).
5. Defending the computing environment (providing access controls on hosts and servers to resist insider, close-in, and distribution attacks).
6. Specifying the security robustness (strength and assurance) of each information assurance component

as a function of the value of what it is protecting and the threat at the point of application.

7. Deploying robust key management and public key infrastructures that support all of the incorporated information assurance technologies and that are highly resistant to cyber attack.
8. Deploying critical national infrastructures to detect intrusions and analyzing and correlating the results and reacting accordingly.
9. Maintaining visible and up-to-date system security policy.
10. Certifying and accrediting changes to the information technology baseline.
11. Managing the security posture of the information assurance technology (installing security patches and virus updates, maintaining access control lists).
12. Providing key management services to protect this lucrative critical national infrastructure.
13. Performing system security assessments (vulnerability scanners, RED teams) to assess the continued "security readiness."
14. Monitoring and reacting to current cyber attack threats.
15. Implementing sensing, warning, and response.
16. Recovering and reconstituting.

Summary

This chapter focused on the general security strategy of defense in depth. Defense in depth involves the deliberate introduction of multiple layers of defense in order to increase the likelihood that a given attack will be stopped or at least slowed down. This likelihood is dependent upon the quality and relative attributes of the various defensive layers.

The chapter also covered how to achieve information assurance, which begins with a senior-level management commitment (typically at the chief information officer level) based on a clear understanding of the perceived cyber threat. This must be followed through with effective information assurance policies and procedures, assignment of roles and responsibilities, commitment of resources, training of critical personnel (users and system administrators), and personal accountability. This includes the establishment of physical security and personnel security measures to control and monitor access to facilities and critical elements of the information technology environment.

Furthermore, a wide range of technologies are available for providing information assurance services and detecting intrusions. To ensure that the right technologies are procured and deployed, an organization should establish effective policy and processes for technology acquisition. These should include security policy, information assurance principles, system-level information assurance architectures and standards, criteria for needed information assurance products, acquisition of products that have been validated by a reputable third party, configuration guidance, and processes for assessing the risk of the integrated systems.

Even the best available information assurance products have inherent weaknesses. So, it is only a matter of time before an adversary will find an exploitable vulnerability. An effective countermeasure is to deploy multiple defense mechanisms between the adversary and his or her target. Each of these mechanisms must present unique obstacles to the adversary. Furthermore, each should include both "protection" and "detection" measures. These help to increase risk (of detection) for the adversary while reducing his chances of success or making successful penetrations unaffordable. Deploying nested firewalls (each coupled with intrusion detection) at outer and inner network boundaries is an example of a layered defense. The inner firewalls may support more granular access control and data filtering.

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

Chapter Review Questions/Exercises

True/False

1. True or False? The general security strategy of *defense in depth* is based on the observation that any given layer of protection cannot fail at any time.
2. True or False? Academics formally model the effectiveness of a collection of defensive layers using mathematical probability.
3. True or False? Most information technology (IT) and security teams in government and industry are committed to reducing the number of passwords, passphrases, handheld tokens, certificates, biometrics, and other validation tokens that exist in their environment.
4. True or False? Commercial environments are increasingly turning to virtualized, in-the-cloud solutions for their gateway filtering of e-mail viruses and spam.
5. True or False? Access controls determine who can access what resources under which conditions.

Multiple Choice

1. The relevant factors for estimating effectiveness of a layer include the following, except which one:

A. Practical experience
B. Informal subjective reasoning
C. Engineering analysis
D. Use-case studies
E. Testing and simulation

2. When a given protection layer does not address all known attacks, then we can conclude which of the following two?

A. Flaws
B. Boundary
C. Simplification
D. Policy
E. Suitability

3. Critical national infrastructure services need a defense in depth scheme that is developed with the following considerations, except which ones?

A. Diversity with single sign-on
B. Diversity of proof factors
C. Small factors
D. Online factors
E. Emphasis on security

4. Access to an application can only be obtained by successfully negotiating which of the following two

layers?

A. Access control layer 1
B. Access control layer 2
C. Access control layer 3
D. Access control layer 4
E. Access control layer 5

5. Some general heuristics for protecting national infrastructure with layered access controls include the following, except which two:

A. Network-based firewalls
B. Operations concept
C. Accuracy interface
D. Internal firewalls
E. Physical security

Exercise

Problem

This defense in depth exercise scenario is an intentional cybersecurity attack on the water utility's SCADA system. It occurs during the fall after a dry summer in Fringe City. The water utility's Information Technology (IT) person did not receive an expected pay raise and decides to reprogram the SCADA system to shut off the high-lift pumps. The operator's familiarity with the SCADA system allows him to reprogram the alarms that typically notify operators of a high-lift pump failure. In addition, he prevents access to the SCADA system by others. A wildfire breaks out on the outskirts of the city. Please identify what type of new countermeasures should have been implemented to prevent this cyber attack from occurring.

Hands-On Projects

Project

Trojaned e-mails were sent from an intruder and targeted at specific organizations and people. The Trojaned e-mails, when opened, compromised a system and enabled the cyber attackers to infiltrate internal networked systems. The cyber attackers then searched systems and network for data files and exfiltrated information through encrypted channels. So, how would you go about preventing the cyber attack in the first place?

Case Projects

Problem

A virus-infected laptop was introduced to the internal network, thus propagating the worm throughout the

organization. Individuals did not realize they were infected. No antivirus scanning was done prior to allowing the laptop to connect to the network. An out-of-date antivirus software was used, thus allowing for the massive infection of the network. Containment and recovery operations were a major challenge. Explain how you would reduce potential vulnerabilities, protect against intrusion attempts, and better anticipate future threats.

Optional Team Case Project

Problem

In the following web attack scenario, the cyber attackers planned ahead and identified their targets. The cyber attackers compromised the website(s) by dropping malicious code or IFRAME, e-mailing links to users in certain instances, compromising systems by using the Rifle or Shotgun approach, elevating privileges in certain instances, using password-capturing binaries, and spreading laterally to other systems from points of entry. So, in keeping the preceding in mind, please identify how the organization would go about tackling this type of cyber attack.

7

Discretion

Chapter Outline

Trusted Computing Base
Security Through Obscurity
Information Sharing
Information Reconnaissance
Obscurity Layers
Organizational Compartments
National Discretion Program
Top-Down and Bottom-Up Sharing of Sensitive Information
Summary
Chapter Review Questions/Exercises

> The British spook said it on the way to the pub—a seemingly random confession that stood out in contrast to the polite evasions that were Ellis's standard form of reply. Public key cryptography? "You did a lot more with it than we did," he said.
>
> **Steven Levy**[1]

A belief found occasionally in the hacking community is that all information should be free and that anyone trying to suppress information flow is evil. The problem with this view is that it suggests that sensitive personal data should be exposed to the world. As such, this extreme view is commonly modified by hackers as follows: All information associated with *organizations*, especially government, should be free, but private data about individuals should never be disclosed. From a logical perspective, this is a curious distinction, because large organizations are comprised of individuals, but in practice the view makes perfect sense. Hackers are almost universally concerned with protecting the rights of the individual; this view of information establishes a charter for the hacking community to make public anything that might degrade individual rights.

The result is a hacking culture where it is considered acceptable to expose proprietary information from government and industry in hacking magazines, on websites, at conferences, and across the Internet. Hackers often claim that reporting commercial and national vulnerabilities is a useful public service that prompts a more rapid security fix. This certainly does not justify leaking proprietary information that has nothing to do with vulnerabilities, but it does offer some value—albeit in an overly forceful manner. Regardless of the motivation, the fact is that proprietary information in companies and agencies will most definitely be widely exposed if discovered by hackers. Perhaps worse, terrorists and information warriors are also interested in this information, but for more malicious purposes—and they will rarely make their intentions public in advance.

The result is that national infrastructure protection initiatives must include means for protecting sensitive information from being leaked. The best approach is to avoid vulnerabilities in the first place, as this information is the most urgently sought and valuable for public disclosure. More practically, however, national infrastructure includes a wide spectrum of information ranging from innocuous tidbits and gossip to critically sensitive data about infrastructure. This spectrum requires a customized protection program focused primarily on the most critical information. Any practical implementation should therefore combine mandatory, functional security controls with programs that dictate the use of *discretion* by individuals possessing important information. Mandatory controls can be implemented centrally, but discretion must be embedded in the local culture and followed in a distributed and individualized manner.

Exposure of vulnerabilities can force a quick response, but that same exposure might lead adversaries directly to private data.

Trusted Computing Base

The nearest the computer security community has come to recognizing the importance of human discretion lies in an architectural construct introduced in the 1980s called a *trusted computing base* (TCB). The definition of TCB is the totality of hardware, software, processes, and individuals whose correct operation and decision-making are considered essential to the overall security of the system. In an operating system, this would include the system files and processes in the underlying kernel. In an organization, this would include the system and security administrators who operate the critical protection systems. For an organization, it would also include all constructs for managing and storing personally identifiable information (PII) about employees and customers. Candidates for exclusion from a TCB include anything whose malfunction or public disclosure would not create a significant or cascading problem. In modern infrastructure, the TCB generally extends to the systems and networks of partner and supplier groups. This greatly complicates the protection of TCB assets because it extends the TCB perimeter to an environment that is more difficult to control.

A modern TCB extends beyond a single organization, making protection all the more difficult.

The primary goal of any program of discretion in national infrastructure protection should be to ensure that information about TCB functionality, operations, and processes is not exposed inappropriately to anyone not properly authorized and to avoid disclosure to anyone who does not possess a clear business need for that information. Such a program will combine two distinct components:

- *Mandatory controls*—These are the functional and procedural mechanisms that are put in place to ensure that information is protected from unauthorized access. Other than key administrators within the TCB, no individual in any organization should be able to bypass mandatory controls, which will typically include firewalls, intrusion detection systems, and honey pots.
- *Discretionary policy*—These are the rules, recommendations, and guidelines that are put in place by an organization to protect its information, especially with respect to the TCB. The discretion here is generally driven by practical concerns; for example, no functional mechanism can control what people mention informally to colleagues or customers. The only way to ensure protection here is the discretionary guidance afforded by the local culture. This can certainly be complemented with severe punishments if someone clearly violates the spirit of protection for TCB-related information.

As one might expect, the TCB is easiest to protect if its size and complexity are minimized. Having fewer people that must be trusted to support security, for example, is better than having to trust many different people and groups. Similarly, the fewer the systems one must trust in some base, and the less complex these systems are, the better off an organization will be from a security perspective. So, the minimization of a TCB is an excellent goal, albeit one that is often ignored in practice. Security practice has all too often involved the introduction of some new security system that is large and complex and requires full trust (see Figure 7.1).

A smaller, less complex TCB is much easier to protect.

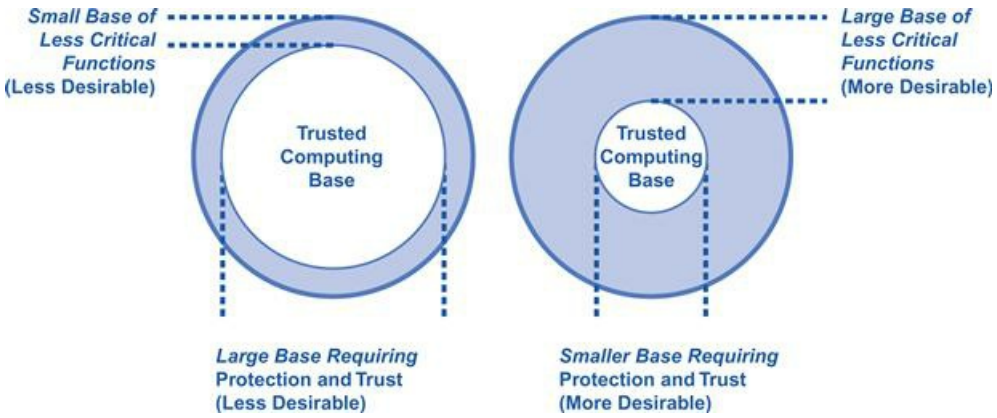


Figure 7.1 Size comparison issues in a trusted computing base.

A major consideration in the protection of national infrastructure thus becomes how to manage, promote, and ensure proper human discretion around critical information related to TCB assets. This requires that policies, procedures, and even functional controls be put in place to assist in exercising such discretion. The idea is that, before any TCB-related information is disclosed that could have an impact on the security of some national asset, the following types of questions must be considered:

Asking the right questions can help determine the impact of TCB security-related disclosures.

- *Assistance*—Could this information assist an adversary in attacking some aspect of national infrastructure? For example, if terrorists or country-sponsored information warriors had this information, could they mount a malicious campaign against services such as emergency 911?
- *Fixes*—Does disclosure of this information assist in identifying a timelier or more effective security fix? For example, will this disclosure provide someone with information that can reduce the time required to fix the problem?
- *Limits*—Can the information disclosure be limited to those in a position to design a security fix? More specifically, can the disclosure be done quietly and in private to a targeted group such as the vendor or service provider that can directly solve the problem?
- *Legality*—Is disclosure of this information a legal or contractual requirement in the local environment? Or, is this disclosure being done for some other reason—perhaps personal gain or pent-up anger with some organization for moving too slowly?
- *Damage*—Is any individual or group harmed or damaged by protection and nondisclosure of this information?
- *Need*—Do others need this information to protect their own systems or infrastructure?

As suggested, proper human discretion in the interpretation of these questions, along with subsequent decision-making, is critical to protecting national assets. In many cases, government organizations will demand information related to some national infrastructure component or service, especially if the

information relates to some trusted computing base. This is fine, as long as the purpose of sharing is reasonable and focused on improving the situation. When such information is demanded by a government group for unspecified purposes (or, at worst, for the purpose of power or gossip), then such sharing is not recommended.

Before sharing critical information, consider who is requesting it and what the purpose is behind their request.

In any event, regardless of the security process, architectures, and systems put in place to protect assets, humans will remain a critical link in the chain. In fact, in many environments, they may be the weakest link. This is why the exercising of discretion in sharing information is such an important principle.

Security Through Obscurity

A barrier to proper discretion is the much maligned and poorly understood notion of *security through obscurity*. Ask any security expert what they think of this concept, and you will receive a religious argument, especially from cryptographers, that deliberately hiding information to ensure security will not work. Their claim is that anyone trying to hide design, implementation, or operational detail is probably just trying to conceal flaws. Furthermore, all information presumably finds its way public, they will argue, and any dependencies on suppression will eventually topple. The most objectionable applications of security through obscurity can be described in the following two scenarios:

There are many opponents of security through obscurity as a meaningful protection strategy.

- *Long-term hiding of vulnerabilities*—This involves the operators of a system concealing the existence of some exploitable flaw as their primary, long-term means for securing the system, as opposed to the more desirable approach in which the vulnerability would be removed.
- *Long-term suppression of information*—This involves the operators of a target system deliberately suppressing general information about a system to make things more difficult for adversaries, hackers, and third parties to discover potential flaws in a system.

In each of these scenarios, the primary control involves hiding information. Most would agree that this is not a reliable long-term method, because suppressed information has a tendency to eventually become public. The situation can be depicted as a knowledge time line, where zero information is initially made public about some system. With time, a gradual increase will occur in available public knowledge. If this increase reaches the point where sufficient information is available to mount an exploit, then the security through obscurity scheme has failed. Obviously, disruptive events such as hacker announcements can create abrupt increases in knowledge (see Figure 7.2).

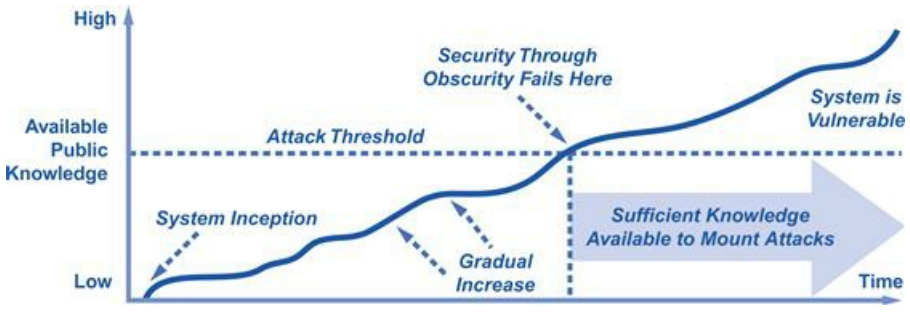


Figure 7.2 Knowledge lifecycle for security through obscurity.

Although security through obscurity is not recommended for long-term protection as a primary control, it remains an excellent complementary control in many cases, as well as being an essential requirement in the short term for many types of security problems in infrastructure. For example, there are no compelling reasons for information about some organization's security architecture to be made public. As long as the security

design receives expert local treatment, it is best left not publicized. Certainly, no one should recommend this as a primary control, and it should not be used to hide flaws, but such discretion raises the bar against adversaries and might be the difference between an attack that succeeds and one that fails.

Security through obscurity should not be a primary protective strategy but can certainly be part of a defense package.

Correspondingly, when some exploitable flaw is discovered locally that requires immediate attention, the worst thing that can happen is for that information to be shared broadly. When this occurs, perhaps as a result of a posting to the Internet, the local response becomes distorted by concerns related to public relations, imminent threat, and legal concerns. Engineering solutions would be much improved if the flaw can be analyzed carefully and embedded into proper development and operations lifecycles. In addition, suppose that the steady state for some system is that sufficient security exists to ensure proper operation, and any vulnerability that might exist is sufficiently obscure as to make the technology reasonably dependable. If a severe vulnerability is then found, the result is that the new steady state could jump to an unacceptably high risk state, and the integrity and dependability of the operation could be in jeopardy. This is simply not acceptable, even for short periods of time, for essential national services.

Essential national services cannot afford to be in a high risk state, even for a short period of time.

The familiar argument that hackers often make here is that by exposing the vulnerability, a fix is rushed into place. In addition, when the fix is embedded into the original system, the integrity of that system has, by definition, been increased, simply because an existing flaw has been removed. This is a powerful argument and is in fact a correct one. The problem is that for essential services, the vulnerability period—during which risk grows higher than some tolerable threshold—must be avoided. Cold logic generally goes out the window when a service must be in place to ensure that a heart attack victim receives aid, or that tenants in an inner city receive electricity and heat, or that operators of a nuclear power plant can avoid dangerous emergency situations that could create serious health disasters (see Figure 7.3).

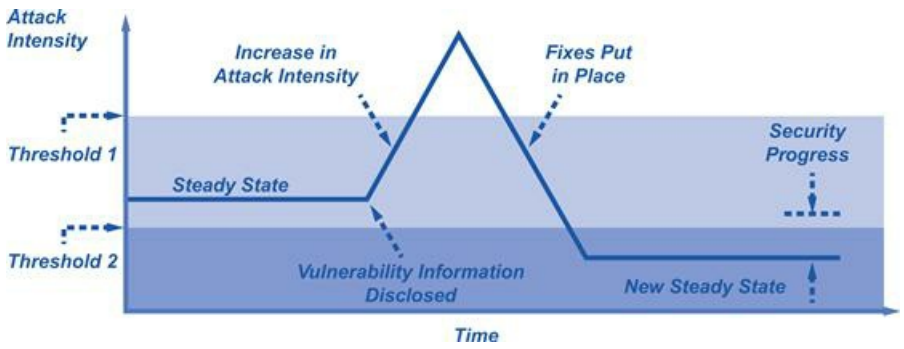


Figure 7.3 Vulnerability disclosure lifecycle.

Regardless of the specific steady-state attack intensities and acceptable thresholds, the requirement is that

the individuals charged with protecting vulnerability information must exercise proper discretion to ensure a level of obscurity for their systems. Without such discretion and obscurity, the chances are great that attack intensities can exceed desired levels, thus leading to serious problems. In general, the practice should be to avoid public disclosure of vulnerabilities until a responsible fix has been put in place. This suggests that disclosure of vulnerability information must be minimized and confined to those in a position to design and embed a proper solution.

Information Sharing

Sensitive information can be exposed in different ways, including deliberate leaks, stray comments, document theft, and hacker disclosure. Each of these occurrences can be jolting for a security team, and their potential creates a general feeling of unease, especially in national infrastructure settings. An additional path for the exposure of sensitive information involves willful information sharing with some controlled, authoritative group. While this is a predictable event, and the recipients are usually delighted with the information, the group doing the sharing is rarely pleased with the overall process.

Information sharing may be inadvertent (stray comments), secretive (document theft), or willful (federal regulations or audits).

Government agencies are the most aggressive in promoting information sharing. Obviously, where legal requirements dictate reporting of data, there is no reason for debate. Law enforcement groups and federal regulators, for example, regularly demand information, but this is done under extremely controlled conditions and rarely, if ever, results in vulnerability-related data being disclosed to an adversary. For cyber security, however, government agencies request that industry share sensitive information for the following reasons:

- *Government assistance to industry*—In theory, attack signatures and related security data could be provided by government to industry, as long as government is fully aware of the vulnerabilities that might reside in commercial infrastructure. This requires information sharing from companies to government.
- *Government situational awareness*—For government to properly assess cyber security risk at the national level, information sharing from industry is required, as such a large portion of national infrastructure resides in industry.
- *Politics*—Government groups are political by nature, and sensitive information provided by industry serves as a type of "power currency" that is used to push political objectives within government. This is rarely stated, but no government official would deny its validity.

In practice, information sharing between industry and government tends to provide spotty results for both parties. The idea of government providing direct cyber security assistance to industry, for example, is mostly theoretical. Valid scenarios can easily be imagined, especially for attack signatures that might be known by a military or intelligence group, but the practical realization of this is rarely seen. Similarly, the idea of government using shared information to form an aggregate view of national cyber security risk sounds great, but has never been done—at least in any public way. In contrast, the political objective has been the primary driver for most information sharing initiatives, which helps explain the enthusiasm that remains in government for this activity. This is a shame, because of all the motivations this one is the least important to the operator sharing data. In fact, an inverse relationship seems to exist between the respective measures of value to the sharing and receiving parties (see Figure 7.4).

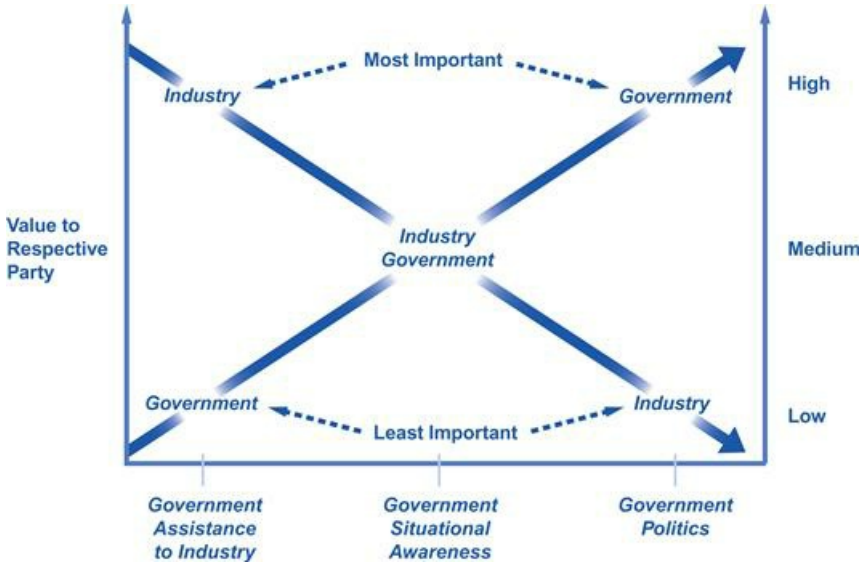


Figure 7.4 Inverse value of information sharing for government and industry.

Government and industry are not mutually invested in information sharing for the same reason.

The relationship illustrated in Figure 7.4 shows that whereas government primarily seeks political power with information, industry cares the least about this; correspondingly, where industry would benefit most from government assistance, this is an area where government is in the weakest position to help. Both government and industry would agree that it is moderately important that government maintain situation awareness of vulnerabilities, but neither would list this as their primary objective. It is this inverse relationship that helps one understand why information sharing initiatives have rarely worked. It also goes without saying that any cases where information has been shared with government and is then sloppily handled, perhaps even leaked to the press, just makes matters worse.

Certainly, poor handling of sensitive or private information lessens industry's trust in government when sharing information on vulnerabilities.

The recommendation here is that any energy available for expenditure in this area should focus on flattening the two curves somewhat. Government should be less focused on politics, and industry should be less concerned with getting something in return for sharing. The end result is that sharing objectives will naturally converge to an agreed-upon situational awareness objective, which is important but certainly not so important as to warrant all the attention this issue brings to the cyber security discussion.

Information Reconnaissance

Reconnaissance activity performed by an adversary is another means by which sensitive information can be exposed. This is important to recognize because attacks on national infrastructure will always include some form of reconnaissance. It can be done at arm's length using remote access over the Internet; it can be done using compromised or planted insiders with access to critical local data; it can be done using social engineering techniques; it can be done via deliberate theft, remote hacking, or quiet sabotage, and so on. Regardless of the technique or vantage point, reconnaissance is used to plan and prepare for attacks on infrastructure.

Adversarial attacks are rarely spontaneous; some amount of planning goes into each attack.

This three-stage model suggests that at each layer of information collection by an adversary the opportunity exists for security engineers to introduce information obscurity. The purpose of the obscurity would be to try to prevent a given type of information from being disclosed through the reconnaissance activity. The specific types of security-related national infrastructure information that should be obscured are as follows:

Reconnaissance Planning Levels

Three levels of reconnaissance are followed in most instances of cyber attack planning:

1. The first level involves broad, wide-reaching collection from a variety of possible sources. This can include web searches, personal contact, and business interaction.
2. The second level of reconnaissance involves targeted collection, often involving automation to provide assistance. Network scanning is the most common functional support for this second level of reconnaissance.
3. The third level involves direct access to the target. A successful hacking break-in to some system, followed by the collection of targeted data, is an example.

One possible scenario that strings the three phases together might involve broad reconnaissance, where something found on the Internet would prompt more targeted reconnaissance, which would involve the scanning activity to find something that could then be used in the third phase for direct access to a target (see Figure 7.5).

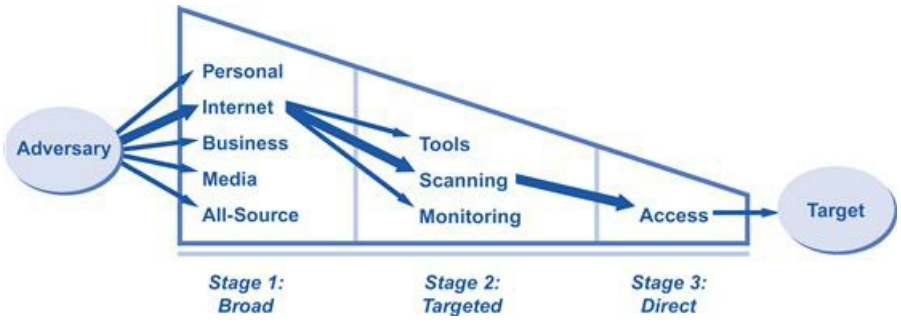


Figure 7.5 Three stages of reconnaissance for cyber security.

- *Attributes*—This is information about seemingly nonsecurity-related features, functions, and characteristics of the computing, networking, applications, and software associated with national infrastructure. It could include equipment type, vendor name, size and capacity, and supported functionality. Adversaries often covet this type of information because it helps provide context for a given attack.
- *Protections*—This is information related to the security protection of a national asset. It can range from technical configuration or setup data about systems to nontechnical contact information for key security administrative staff. The value of this information should be obvious; when obtained, it provides a roadmap for the type of countermeasures an adversary must consider in planning a successful attack.
- *Vulnerabilities*—This is information related to exploitable holes in national infrastructure. It can range from well-known bugs in commercial operating systems to severe vulnerabilities in some national asset. Adversaries will seek this type of information from any possible source. This can include the national infrastructure management team, relevant technology or service vendors, or even the general public. The hacking community is also a rich source of vulnerability information, especially as it relates to national assets.

Of these three attributes, vulnerability information tends to dominate most discussions about the types of information an adversary might desire. Go to the technical section of any bookstore, for example, and you can find thick tomes chronicling the exploitable holes in virtually any technology you can imagine. This gives you some idea of how difficult it really is to obscure vulnerability information. This should not discourage the operators of national infrastructure; when serious problems are discovered that can degrade essential services, the only responsible action is to work toward some sort of fix with the responsible parties before the information is shared to the rest of the world, which obviously includes the adversary.

Although truly obscuring vulnerability information is likely an impossibility, security managers should strive for discretion and privacy on this point whenever possible.

Obscurity Layers

One conceptual approach to managing discretion in protecting national infrastructure information involves *obscurity layers*. These layers are intended to reduce the likelihood that critical information is disclosed to unauthorized individuals. Techniques for introducing layers of obscurity range from common-sense human discretion to more structured processes for controlling information flow. If designed properly, obscurity layers should make unauthorized disclosure possible only if multiple, diverse obscurity techniques are somehow bypassed. In this sense, obscurity layers can be viewed as an instance of defense in depth.

Layering the methods of obscurity and discretion adds depth to a defensive security program.

In the best case, obscurity layers provide diverse, complementary, and efficient coverage around national asset information. That is, an asset might first be protected by an obscurity layer that includes data markings to remind individuals of their obligation to use discretion. A second obscurity layer might involve some mandate that no technical information about local networks, software, or computing platforms be shared beyond the team of trusted administrators. A third layer of obscurity might then involve the mandate that, if information does somehow leak out about critical infrastructure, the organization will never comment publicly on any aspect of the leak.

These three example layers are complementary and provide guidance to individuals on how to exercise discretion in what information to share and what information to suppress. As such, they can be viewed as an effective discretionary tool for protecting assets (see Figure 7.6).

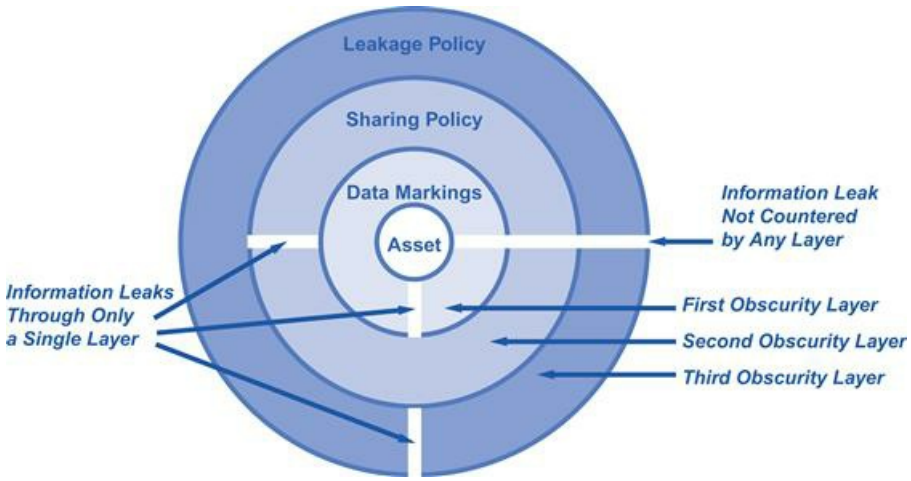


Figure 7.6 Obscurity layers to protect asset information.

Leaks through obscurity layers might make their way through each layer or might be countered by one or more layers. For example, in Figure 7.6, an information leak that would not be countered by any layer might involve someone exercising poor discretion by ignoring data markings (through the first layer), violating information sharing policies (through the second layer), and being ignorant of policies for disclosure after an incident (through the third layer). This demonstrates the human element in the use of discretion to protect

critical infrastructure information. Additional examples of obscurity layers in national infrastructure protection include the following:

Even with layered obscurity, asset information may leak through to an adversary.

- *Public speaking*—A policy might be in place that would deliberately prevent anyone with responsibility for national infrastructure from speaking publicly without explicit public relations preparation and planning.
- *Approved external site*—A ubiquitous mechanism, such as a website, might be in place to constantly and consistently provide organizationally approved information about infrastructure that might be desired by external entities.
- *Search for leakage*—Search engines might be used via ethical hacking techniques to determine the degree and scope of inappropriate information that might already be located on websites or in a cache. This can be complemented by modern data leakage protection (DLP) tools.

As suggested above, the purpose of these discretionary controls is not to suppress information for the purposes of hiding incompetence or inappropriate behavior. The purpose is to responsibly control the type of information made available to a malicious adversary.

Organizational Compartments

An information protection technique used successfully by the U.S. federal government, especially in the military and intelligence communities, involves the compartmentalization of individuals and information. These compartments can be thought of as groups for which some set of policy rules uniformly apply. Typically, individuals are put through a background check to determine their level of trustworthiness. They are then given a designated security *clearance*. Information is similarly put through an analysis to determine its level of criticality; it is then given a designated security *classification*.

Government clearance levels and information classification are techniques used to protect data by limiting accessibility.

The specifics of how clearances and classifications work are beyond the scope of this book, but a key notion is that each combines some notion of hierarchical level (e.g., Top Secret, Secret, Confidential, Unclassified) with a corresponding notion of "need to know" categories (e.g., Navy, Air Force). The cross-product of some set of classified information with the corresponding individuals cleared to access that information is called a *compartment*. Policy rules for accessing data, such as classified documents, from a compartment can then be implemented (see Figure 7.7).

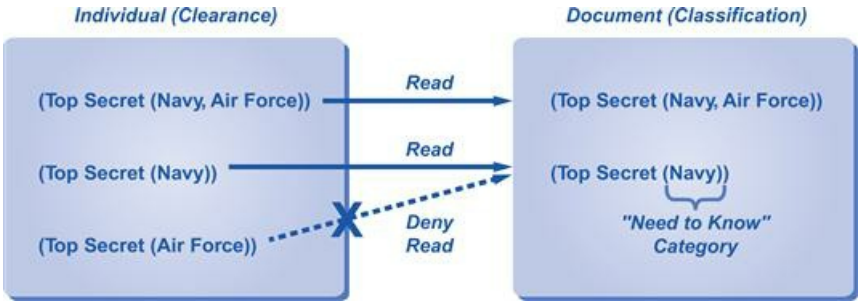


Figure 7.7 Using clearances and classifications to control information disclosure.

The examples in Figure 7.7 show an individual cleared to Top Secret in categories Navy and Air Force being successful in reading a document that is classified to the same level and categories. In addition, an individual cleared to Top Secret in category Navy is successful reading a document cleared to the same level and categories. On the other hand, an individual cleared to Top Secret in category Air Force is denied access to a document whose category is only Navy. This type of approach is especially powerful in an actual government setting, because information leaks can be interpreted as violations of federal law. In the most intense case, such violations could be interpreted as espionage, with all the associated punishment that comes with such action. The result is a mature environment in most government settings for reducing the chances that national security-related information will be leaked.

Certain secure government data can only be accessed by a few top-level officials.

Clearly, the protection of national services is not just the responsibility of government. Thus, industry needs a corresponding approach to policy-based access control. The good news is that translation of government compartments to a corporate setting is relatively straightforward. Clearance and classification levels can be mapped to company-defined organizational levels such as "supervisor" and "senior manager." Categories can be mapped to specific projects in a company. Thus, a compartment in some company might correspond to the senior manager level, within some project A and project B (see Figure 7.8).

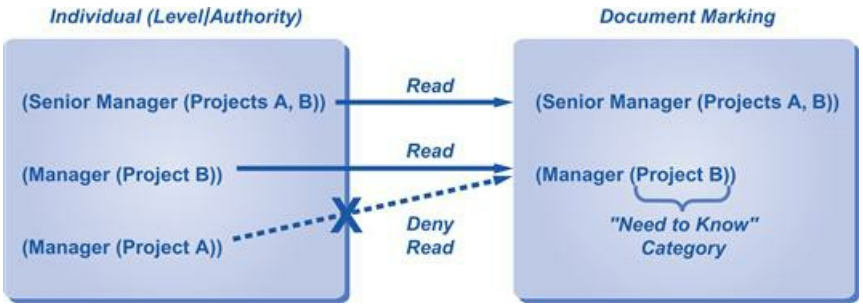


Figure 7.8 Example commercial mapping of clearances and classifications.

The bottom line with compartmentalization is that it should be used to help define boundaries around which information can or cannot be accessed. This helps guide decisions that require human discretion. Too often, in computer security settings today, the underlying goal of many projects and in the management of many critical systems is to avoid the use of information boundaries, often in the interest of openness and sharing. These concepts are valuable for many types of standards, information, data, software, and services, but unfortunately openness and sharing are not always consistent with protecting security-related information about national infrastructure.

Private companies can mirror government clearance levels by classifying data and limiting access.

National Discretion Program

To implement a national program of information obscurity and discretion, several management and security engineering tasks will be required:

- *TCB definition*—Although it could be difficult to do so, effort should be directed by suitable national authorities toward trying to define a nationwide trusted computing base. This will require coordination between government and industry, but the resulting construct will help direct security management decisions.
- *Reduced emphasis on information sharing*—Government must immediately reduce its emphasis on demanding that information be shared by industry. Any information sharing initiatives that do maintain such an emphasis should focus only on providing government with situation status.
- *Coexistence with hacking community*—The national infrastructure community in government and industry would benefit by creating an improved spirit of cooperation with the hacking industry. This could come in the form of financial support from government for hacking groups and forums, or it could be more explicit in terms of actual tasking on real programs.
- *Obscurity layered model*—A national obscurity layer should also be put in place to guide decisions about human discretion in protecting sensitive national infrastructure-related information.
- *Commercial information protection models*—Industry should be provided with incentives and rewards for demonstrating some degree of embedded policy-based access control similar to the military model.

Certainly, to increase the chances that these tasks are successful, a culture of human discretion around sensitive information must be created. Senior managers must reinforce this culture by not exercising their right to bypass discretionary controls; for example, all documents, even those created by senior managers, should be marked appropriately. Similarly, if violations of basic information discretion do occur, the consequences should be similarly applied, regardless of organizational position or level.

Finally, let's briefly look at some practical ways to protect sensitive information in critical national infrastructure environments. During cyber attack situations, the DHS monitors risk management activities and status at the functional/operations level, the local law enforcement level, and the cross-sector level. Sensitive information sharing may also incorporate information that comes from pre- and postevent natural disaster warnings and reports. While sensitive information sharing is multidirectional within the networked model, there are two primary approaches to information sharing during or in response to a threat or incident: top-down and bottom-up.

Top-Down and Bottom-Up Sharing of Sensitive Information

Under the top-down sharing of sensitive information approach, information regarding a potential terrorist threat originates at the national level through domestic and/or overseas collection and fused analysis, and is subsequently routed to state and local governments, critical national infrastructure owners and operators, and other federal agencies for immediate attention and/or action. This type of sensitive information is generally assessed against DHS analysis reports and integrated with critical national infrastructure-related information and data from a variety of government and private sector sources. The result of this integration is the development of timely sensitive information products, often produced within hours, that are available for appropriate dissemination to critical national infrastructure partners based on previously specified reporting processes and data formats.

On the other hand, under the bottom-up sharing of sensitive information approach, state, local, tribal, private sector, and nongovernmental organizations report a variety of security- and cyber attack-related information from the field using established communications and reporting channels. This bottom-up sensitive information is assessed by DHS and its partners in the intelligence and law enforcement communities in the context of threat, vulnerability, consequence, and other information to illustrate a comprehensive risk landscape.

Cyber attack threat sensitive information that is received from local law enforcement or private sector suspicious activity reporting is routed to DHS. The sensitive information is then routed to intelligence and operations personnel to support further analysis or action as required. In the context of evolving cyber attack threats or incidents, further national-level analyses may result in the development and dissemination of a variety of products. Further sensitive information-sharing and cyber attack management activities are based on the results of the integrated national analysis and the needs of key decision makers.

DHS also monitors operational sensitive information such as changes in local risk management measures, pre- and postincident disaster or emergency response information, and local law enforcement activities. Monitoring local incidents contributes to a comprehensive picture that supports incident-related damage assessment, recovery prioritization, and other national- or regional-level planning or resource allocation efforts. Written products and reports that result from the ongoing monitoring are shared with relevant critical national infrastructure partners according to appropriate sensitive information protection protocols. The establishment and use of a risk management process to assess and manage the sharing of sensitive information, cyber attack threats, vulnerabilities, and consequences (see "An Agenda for Action for Balancing the Sharing and Protection of Sensitive Information Through Risk Management") are extremely vital for the intelligence information gathering process.

An Agenda for Action for Balancing the Sharing and Protection of Sensitive Information Through Risk Management

The critical national infrastructure encompasses a number of protocols/actions that facilitate the flow of sensitive information, mitigate obstacles to voluntary information sharing by critical national infrastructure owners and operators, and provide feedback and continuous improvement for sensitive information-sharing

structures and processes. When completing the sharing and protection of sensitive information protocols checklist, an organization should adhere to the provisional list of actions to increase risk (of detection) for the adversary, while reducing his or her chances of success or making successful cyber attacks unaffordable. The order is not significant; however, these are the activities for which the research would want to provide a detailed description of procedures, review, and assessment for ease of use and admissibility. As a minimum, the following risk management and information sharing protocols/actions should include (check all tasks completed):

1. Establishing and using a risk management process to assess and manage threats, vulnerabilities, and consequences.
2. Establishing a risk management process that is based on a system-wide assessment of risks and obtaining management approval of this process.
3. Updating the system-wide risk assessment whenever a new asset/facility is added or modified, and when conditions warrant (changes in threats or intelligence).
4. Using the risk assessment process to prioritize security investments.
5. Coordinating with regional security partners, including federal, state, and local governments and entities with a shared critical national infrastructure (other transit agencies or rail systems) to leverage resources and experience for conducting risk assessments (leverage resources such as the Security Analysis and Action Program operated by TSA's Surface Transportation Security Inspectors).
6. Participating in a sensitive information-sharing process for cyber attack threat and intelligence information.
7. Participating in sensitive information-sharing networks or arrangements with:

a. State and local law enforcement and homeland security officials.
b. DHS's Homeland Security Information Network (HSIN) and its mass transit portal (the HSIN portal enables secure information sharing among transit agencies and passenger rail systems at no cost to users).
c. FBI Joint Terrorism Task Force (JTTF) and/or other regional antiterrorism task force (Terrorism Early Warning Group [TEW], U.S. Attorney's Office).
d. TSA Surface Transportation Security Inspectors (STSI) and Public Transportation Information Sharing and Analysis Center (PT-ISAC).

8. Establishing and using a reporting process for suspicious activity (internal and external).
9. Through training and awareness programs, ensuring transit agency employees understand the what, how, and when to report observed suspicious activity or items.
10. Using exercises to test employee awareness and the effectiveness of reporting and response procedures.
11. Ensuring public awareness materials and announcements provide clear direction to the public on reporting of suspicious activity.
12. Maintaining protocols to ensure that designated security coordinator(s) report threats and significant security concerns to appropriate law enforcement authorities and TSA's Transportation Security Operations Center (TSOC).

13. Maintaining protocols that ensure actionable security events are included in reports to the FTA's National Transit Database (NTD).

Summary

This chapter focused on the critical national infrastructure protection initiatives that include means for protecting sensitive information from being leaked. The best approach is to avoid vulnerabilities in the first place, as this information is the most urgently sought and valuable for public disclosure. More practically, however, the critical national infrastructure includes a wide spectrum of information ranging from innocuous tidbits and gossip to critically sensitive data about infrastructure.

The chapter also covered how the federal government is working with state and local partners and the private sector to create the sensitive information-sharing environment for terrorism and homeland security information, in which access to such information is matched to the roles, responsibilities, and missions of all organizations engaged in countering terrorism and is timely and relevant to their needs. It is important to note that most of the sensitive information shared daily with the critical national infrastructure sensitive information-sharing environment is necessary for coordination and management of risks resulting from natural hazards and accidents. Consequently, for sensitive information sharing to be efficient and sustainable for critical national infrastructure owners and operators, the same environment needs to be used to share terrorism information.

Furthermore, with its breadth of participants and the complexity of the critical national infrastructure protection mission, the critical national infrastructure sensitive information sharing breaks new ground. It also creates business risks for the owners and operators. Significant questions are raised, such as the following: What sensitive information is required for a productive two-way exchange? How is information most efficiently delivered and to whom to elicit effective action? How is sensitive information (both proprietary and government) appropriately protected? How will the sectors take appropriate action in coordination with all levels of government? How can business risks be mitigated when an exchange takes place?

Of particular criticality is the coordination of the national infrastructure sensitive information sharing at the national level with that at the local level, where most decisions are made and actions are taken to support the critical national information protection mission. The integration of the critical national infrastructure sensitive information-sharing environment, into the national information-sharing environment, as its private sector component, in recognition of its comprehensiveness and engagement between critical national infrastructure owners and operators and all levels of government, strengthens the foundation for effective coordination.

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

Chapter Review Questions/Exercises

True/False

1. True or False? A belief found occasionally in the hacking community is that all information should be free and that anyone trying to suppress information flow is evil.
2. True or False? A modern TCB extends beyond a single organization, making protection all the more difficult.
3. True or False? A barrier to proper discretion is the much maligned and much understood notion of *security through obscurity*.
4. True or False? Sensitive information can be exposed the same way, including deliberate leaks, stray comments, document theft, and hacker disclosure.
5. True or False? Reconnaissance activity performed by an adversary is another means by which sensitive information can be exposed.

Multiple Choice

1. The primary goal of any program of discretion in national infrastructure protection should be to ensure that information about TCB functionality, operations, and processes is not exposed inappropriately to anyone not properly authorized and to avoid disclosure to anyone who does not possess a clear business need for that information. Such a program will combine which two distinct components:

A. Mandatory controls
B. Informal subjective reasoning
C. Discretionary policy
D. Use-case studies
E. Testing and simulation

2. The idea is that before any TCB-related information is disclosed that could have an impact on the security of some national asset, which one of these types of questions should be considered?

A. Does nondisclosure of this information assist in identifying a timelier or more effective security fix?
B. Could this information assist an adversary in attacking some aspect of the national infrastructure?
C. Can the information nondisclosure be limited to those in a position to design a security fix?
D. Is nondisclosure of this information a legal or contractual requirement in the local environment?
E. Is any individual or group harmed or damaged by protection and disclosure of this information?

3. The most objectionable applications of security through obscurity can be described in which two of the following scenarios?

A. Long-term hiding of vulnerabilities
B. Long-term diversity of proof factors
C. Long-term small factors
D. Long-term online factors

E. Long-term suppression of information

4. For cyber security, however, government agencies request that industry share sensitive information for the following reasons, except which two:

A. Government assistance to industry
B. Government situational awareness
C. Attributes
D. Protections
E. Politics

5. Additional examples of obscurity layers in national infrastructure protection include the following, except which two?

A. Network-based firewalls
B. Public speaking
C. Approved external site
D. Internal firewalls
E. Search for leakage

Exercise

Problem

In this scenario, hackers launch cyber attacks that affect several parts of the nation's financial infrastructure over the course of several weeks. Specifically, sensitive credit card processing facilities are hacked and numbers are released to the Internet, causing 120 million cards to be cancelled; automated teller machines (ATMs) fail nearly simultaneously across the nation; major companies report payroll checks are not being received by workers; and several large pension and mutual fund companies have computer malfunctions so severe that they are unable to operate for more than a week. Identify the countermeasures that need to be implemented to prevent these cyber attacks from occurring in the future.

Hands-On Projects

Project

At five o'clock in the morning, John Fringe tried to sign in to the computer at his workstation. Each time he tried to sign in, the computer did not respond and halted. After three unsuccessful attempts, John went into the next room and learned that the attending nurse was having the same problem. John called the Help Desk, but there was no tech-support agent available. He went around the department. He found that everyone was facing trouble signing onto their computers. The hospital's staff members were anxious and complaining to their shift supervisor. Patients were being admitted, but no running computer system was available. So, what would be some of the first actions that the hospital's IT technical team would take to control the situation,

and what countermeasures should the team take to prevent future cyber attacks in order to protect sensitive hospital records at the organizational and end-user level?

Case Projects

Problem

With the anonymity of the Internet, there is greater access to sensitive information online; and, there are large profits to be made by reselling hot products like the iPhone, that subscription fraud is becoming a common problem for telcos. As a result, having insight into "who is who," "who knows who," and "who does what" is essential in stopping fraud before financial losses occur. Please explain how you would reduce potential vulnerabilities, protect against fraud, and better anticipate future threats.

Optional Team Case Project

Problem

An IT group of a utility company faces several major challenges in supporting more than 700 employees nationwide with a staff of only 5. For example, many of the mobile sales representatives and field technicians use laptops and work away from office locations. These employees sometimes lose track of their laptops (which contain sensitive company information), either through misplacing them or through theft. One of the main concerns of IT was to protect sensitive business data on laptops and prevent that data from falling into the wrong hands. In addition, IT wanted to be able to rapidly replicate lost data onto a new laptop so that an employee could quickly continue working after a loss. Keeping the preceding in mind, identify how the company would go about tackling this type of problem.

[1] S. Levy, The open secret: public key cryptography—the breakthrough that revolutionized email and ecommerce—was first discovered by American geeks. Right? Wrong, *Wired*, 7(4), 1999.

8

Collection

Chapter Outline

Collecting Network Data
Collecting System Data
Security Information and Event Management
Large-Scale Trending
Tracking a Worm
National Collection Program
Data Collection Efforts: Systems and Assets
Summary
Chapter Review Questions/Exercises

It is important to have a fairly clear understanding of what you are looking for and what events you are interested in, because you cannot collect or detect everything.

Stephen Northcutt[1]

A basic tenet of computer security is that diligent and ongoing observation of computing and networking behavior can highlight malicious activity. This works best when the observer has a good frame of reference for what constitutes normal behavior. Algorithms and human judgment can then be used to compare profiles with observations to identify activity that might be suspicious. Follow-up analysis can then be used to partition suspicious activity into benign and malicious categories. All this processing and analysis can only be done in the context of an existing program of *data collection*.

At the national level, security-relevant data must first be collected at the local or regional level by individual asset managers and a subset then selected for broader aggregation into a national collection system. In some cases, local and regional collection can be directly connected to national programs. Larger-scale collection points on wide-area networks, perhaps run by carriers or government agencies, can also be embedded into the collection scheme and combined with local, regional, and aggregated data (see Figure 8.1).



Figure 8.1 Local, regional, and national data collection with aggregation.

Such a national collection process does not exist today in any organized manner. To build one will require considerable resolve. From a technical perspective, each collection point requires that decisions be made about which data is gathered, what methods will be used for collection, how it will be used, and how it will be protected. It is not reasonable for any organization to collect any sort of data without having specific answers to these simple questions. Improper collection of data, where no clear justification exists for its aggregation, could lead to serious legal, policy, or even operational problems for organizations charged with protecting some national asset.

Data collection should not be attempted until an organized plan is in place to analyze and protect the data.

As an illustration, many government groups have done a terrible job in the past protecting data once it has been aggregated. Several years ago, for example, sensitive information collected from chemical companies in the New York area was published by a government agency on its website. This information was then collected by reporters and reproduced as an article in a New York City newspaper, replete with a map showing which types of dangerous chemicals were present and their exact location, as well as noting the health and safety implications of these chemicals. This type of information is of great interest, obviously, to terrorists. Dissemination of this information could also have a negative impact on business operations and the reputations of these companies.

At both local and national levels, data collection decisions for national infrastructure should be based on the following three security goals:

- *Preventing an attack*—Will the data collected help stop a present or future attack? This implies that the recipient of collected data must justify its role in stopping the attack. If the recipient manages some critical infrastructure component, such as a backbone network, that can be used to throttle or stop the attack, then the justification is obvious. If, however, the recipient is a government agency, then the justification might be more difficult.
- *Mitigating an attack*—Will the data collected assist in the response to an ongoing attack? The implication here is that the recipient of data should be able to help interpret what is happening or should be able to direct resources toward a solution. One of the most relevant questions to be answered

about an ongoing attack, for example, is how widespread the attack might be. Collecting information from a broad distribution will help to answer this question.

- *Analyzing an attack*—Will the data collected assist in the forensic analysis of an attack after it has occurred? This goal is important but can be easily abused, because it could justify collection of any sort of data available. Forensic analysts generally maintain that their task is made easier in the presence of large volumes of data. Care must therefore be taken to ensure that inappropriate data collection does not occur simply because a forensic analyst might claim to need the information.

These three requirements should direct the scope, coverage, and degree of detail associated with a data collection program for every national infrastructure component. In fact, they provide a suitable template for determining exactly what sort of data should be collected and aggregated. At the local, regional, wide area, and national levels, data collection should only proceed if affirmative answers to these questions can be made (see Figure 8.2).

Data collection must be justified as to who is collecting the data and why.

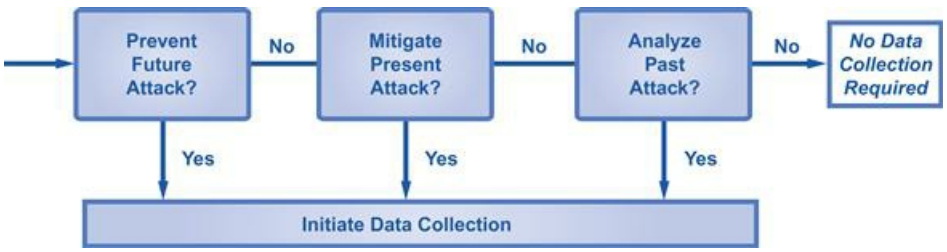


Figure 8.2 Justification-based decision analysis template for data collection.

The decision to *not* collect data might be among the most difficult for any organization, especially a government agency. One of the great axioms of government computer security has been that more data is always better, especially if a path exists to perform such collection. The reality, however, is that improper data collection not only is unnecessary but could also actually weaken national infrastructure.

Beware the "more is better" axiom regarding data collection; focus on quality, not quantity.

Collecting Network Data

Perhaps the most useful type of data for collection in national infrastructure is network *metadata*. Also referred to as *netflow*, metadata provides many security-relevant details about network activity. In a Transmission Control Protocol (TCP)/Internet Protocol (IP) environment, metadata allows the security analyst to identify source address, destination address, source port, destination port, protocol, and various header flags in a given session. This information is security relevant because it provides a basis for analyzing activity. A nontechnical analogy would be that metadata is akin to the information that postal workers can see in the mail they process. The size, weight, color, texture, and addressing information on the envelopes and wrappers are apparent, whereas the contents are not.

Metadata is information *about* the data, not what the data is about.

The collection of metadata involves the placement of equipment or software into the target network for the purpose of producing metadata records. These records are collected and stored for analysis. Obviously, to make this collection feasible, certain functional considerations must be made. There must be legal justification for collecting the data, there must be sufficient storage capacity for maintaining collecting data, and there must be analysts with proper capability to make effective interpretations about the data. Perhaps the most important consideration, however, is whether the collection functionality is sufficiently powerful to keep up with the target network bandwidth capacity (see Figure 8.3).

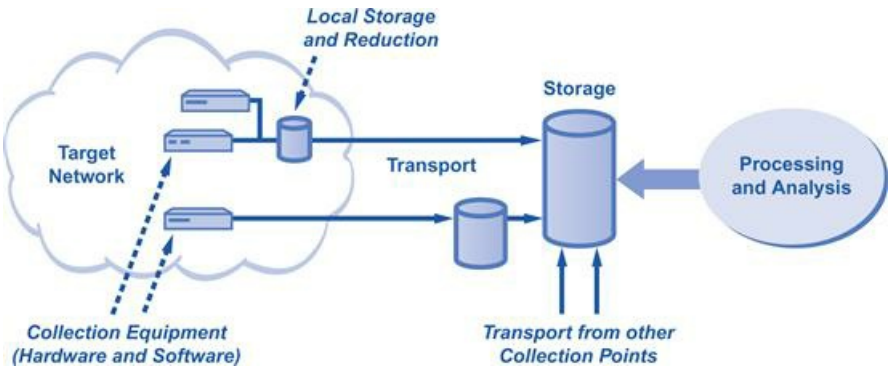


Figure 8.3 Generic data collection schematic.

One issue with large-scale versions of this collection approach is that many metadata collection systems were deployed in carrier backbones during the early part of the century, with the intention of pulling security data from 10-Gbps backbones. The challenge is that carrier backbones have begun to grow to 40- and even 100-Gbps capacities. If the collection systems are not increased at a commensurate rate, then the ability to collect metadata could decrease by as much as a factor of ten.

Data collection systems need to keep pace with growth of carrier backbones.

One solution many security analysts use to deal with increasing network capacity is to *sample* the data. This technique involves grabbing some of the data at predetermined intervals so the inbound flow matches the ability to process. Sampled data is generally acceptable for broad analysis of network activity, but it is not as effective for detailed forensics as unsampled metadata. In an unsampled environment, analysts can often detect tiny anomalies in massive amounts of data. This design consideration affects the overall collection process.

Sampling data is less time consuming, yet unsampled data may reveal more vulnerabilities in the system.

As an example, several years ago unsampled metadata on an IP backbone allowed analysts in a global carrier environment to detect that a small number of packets of an unusual protocol type were beginning to show up. Packets of this type had not been seen on the backbone for years, so this was clearly an anomaly to be investigated. Suspicious packets from this unusual event were collected and observed for four days, until a key equipment vendor contacted the carrier to report a serious security flaw in their operating system software. Interestingly, exploits of this vulnerability involved traffic being sent over precisely the protocol type being observed. The collection point thus detected network activity evidence of a security issue that had not even been publicly reported (see Figure 8.4).

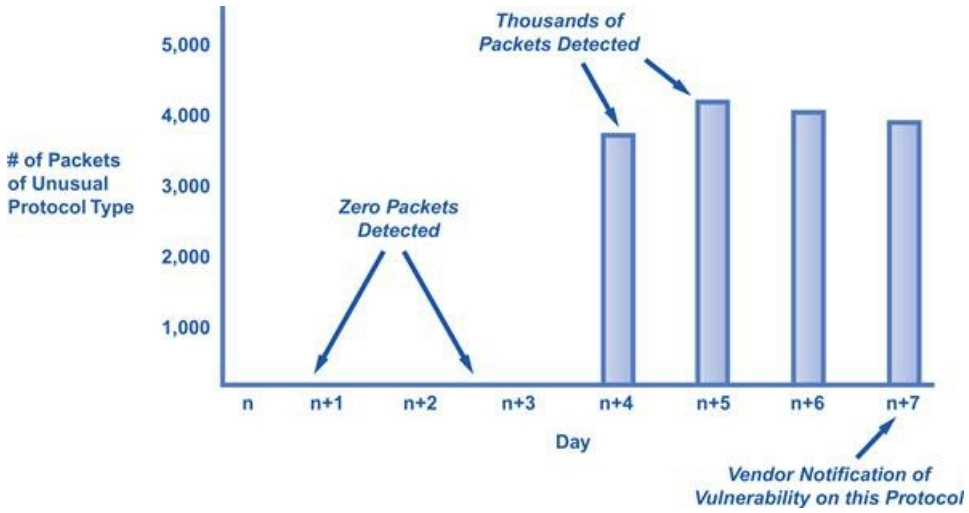


Figure 8.4 Collection detects evidence of vulnerability in advance of notification.

The key observation here is that, under normal conditions, no instances of this type of protocol packet were observed on the carrier backbone. When suddenly the unusual protocol type essentially came alive, there was no easy way to determine why this was the case other than that some sort of anomaly was taking place. When the vendor reported the problem on this protocol, analysts were able to put together this information to solve the riddle of why this anomaly had occurred. This illustrates the importance of integrating all-source information into any data collection environment. National infrastructure protection must include this type of collection and associated analysis to be fully effective in protecting essential services.

Analysis of unsampled metadata can reveal concerning data traffic patterns that would otherwise go

unnoticed.

Collecting System Data

National infrastructure protection initiatives have not traditionally included provision for collecting data from mainframes, servers, and PCs. The justification for such omission is usually based on the scaling and sizing issues inherent in the massive amount of data that would have to be processed from these computers, along with the common view that such systems probably do not provide much security-relevant data. An additional consideration is the potential for privacy abuses, an issue that the citizens of most nations have come to recognize as being important to their lives. As a result, no serious national infrastructure protection initiative to date has included a proposal or plan for this type of functionality.

Regarding scaling and sizing issues, the computing infrastructure required for collection of data from every mainframe, server, and PC deemed part of national infrastructure services would certainly be complex. That said, computing historians know that it is not unprecedented for the complex requirements of one generation to become routine features in another. Furthermore, the tactical approach of identifying a workable subset of the relevant computers in a nation is possible. For example, the mainframes, servers, and PCs running in companies and agencies charged with national infrastructure could be targeted for collection, and this is a tractably sized challenge.

We may not currently have the capacity to collect data from all relevant computers, but it is an important goal to try to reach.

On the issue of whether mainframes, servers, and PCs provide suitable security-relevant information for national infrastructure protection, many critical incidents are best identified through collection of data at this level. Operating system logs, mainframe event summaries, and PC history records provide excellent evidence that malicious activity might be ongoing. Engineering metrics such as memory utilization or processor load can also provide valuable signals about security issues. For example, when a server shows increases in processor usage as a result of an attack, this condition is often easiest to identify using monitoring tools embedded in the operating system of the computer.

System monitoring provides an overview of activity that may reveal troubling patterns.

System monitoring is important to national infrastructure protection because it is often the *only* indicator that some security event is under way—even in the presence of firewalls, intrusion detection systems, and other security tools. As a result, national infrastructure protection initiatives will have to include provision for the gathering and processing of data from mainframes, servers, and PCs. This data will have to be selected, collected, transmitted with suitable protection, stored in an environment properly sized for large amounts of data, and processed in real time. Four specific types of information that should be collected include those listed in the box below.

Top Four Data Collection Areas

1. *Utilization*—One of the most important metrics in determining whether an attack is ongoing is the

utilization profile of servers in the local environment. National asset managers must identify which servers are relevant for monitoring and should instrument an associated program of data collection. This will require cooperation between government and industry, as well as the inclusion of appropriate functional requirements in infrastructure support contracts.

2. *Usage*—Patterns of usage on the mainframes, servers, and PCs in a given nation are important to establish for protection of infrastructure. If certain mainframes are never touched after hours, for example, then this will help to identify smaller attacks during unusual times. Detecting small, active usage events is often easier in a quiet environment than in a noisy environment; however, detecting usage drops is often easier in a noisy environment than in a quieter one.
3. *Applications*—Collecting data about the applications resident on system infrastructure provides useful hints about possible cyber attacks. A common metric is a "top ten" list of most commonly used applications. If the mix changes in some meaningful way, then this could signal an attack. Network gateway systems including proxies are excellent candidates for collecting this type of data for an enterprise. Carriers could provide this type of data in a wide area network or across a given region.
4. *Outages*—Information about outages is important for security, because events that are presumed to have been benign might actually be part of a cyber attack. It is not uncommon for system managers to ignore this possibility; hence, data collection around outages is important. As an example, root-cause analyses after serious outages should be viewed as important information for gathering and analysis.

Two techniques are useful at embedding system management data into cyber security infrastructure. First, an inventory process is required to identify the systems that are considered critical in a given environment. This process might require engineering analysis across relevant government and industrial infrastructure to determine if a given system resides in the critical path of some national service. Alternatively, the decision might be made to try to collect information from every system that is available for collection. Second, for those systems deemed worthy of data collection, a process of either instrumenting or reusing data collection facilities must be identified. This could involve the use of operating system audit trails or it could involve the installation of some sort of application-level logging program.

Regardless of the approach, data would flow from the target computers of interest across some network medium to various aggregation points. Regional and enterprise networks would probably have to introduce an aggregation function for their organization before the data is shared externally. One would expect that network carriers could easily step into this role of providing different types of aggregations; that is, customers of DSL and cable services could agree, under suitable incentives, to allow for collection of data related to the presence of malware, viruses, and other compromising software. Encryption could be used to help protect the confidentiality of the data in transit and storage.

Aggregation points would allow for regional collection of data.

There would also have to be some sort of filtering or data reduction to focus the collection on specific systems of interest and to limit data to only that which is likely to be useful. For example, if a nation tried to collect security-related data from hundreds of thousands or millions of PCs every day, the resultant daily

dataflow would be measured in the multiple terabyte range. Commercial databases would probably be insufficient for storing this volume, so customized databases would be required. The volume of collected data would ultimately be made available to a security processing and interpretive system that could be used for national infrastructure purposes.

Although more creative overall architectures could be imagined, such as peer-to-peer, the centralized collection approach would be more likely to be implemented in practice. It also lends itself quite well to the establishment and operation of a national security operations center (see Figure 8.5).

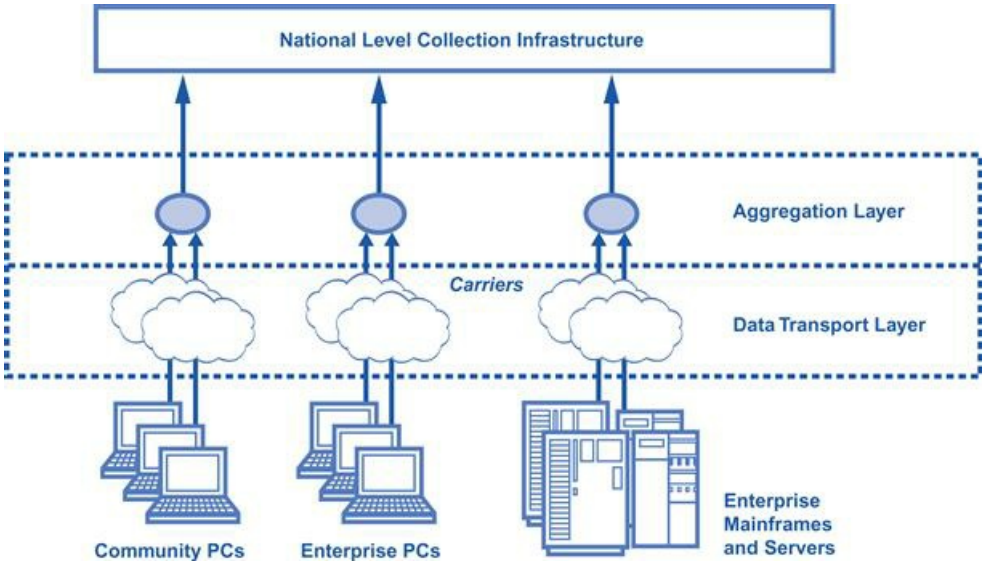


Figure 8.5 Collecting data from mainframes, servers, and PCs.

Readers might cringe at the idea of collecting data in this manner, especially from end-user PCs scattered across a nation, but this practice is more common than one might think. Every large enterprise and government agency, for example, routinely embeds integrity management software, such as tripwire functionality, into their mainframes and servers. Furthermore, almost every enterprise and agency uses software agents on PCs to collect relevant security and management data. Perhaps ironically, botnet operators have also perfected the idea of collecting data from massive numbers of end-user computers for the purpose of attack. The idea that this general schema would be extended to benevolent national infrastructure protection seems straightforward.

A national data collection center may not differ much from current enterprise and agency data collection.

The challenge is that this sort of scheme can be abused. Computer scientists lament software running with high privilege on their systems, and citizens resist the notion of an unknown monitor pulling data from their system to some unknown collection facility, possibly violating privacy principles. Both concerns are valid and need to be debated publicly. If an acceptable compromise is reached between government and its businesses and citizenry, then the result can be incorporated into the design of an appropriate national system. At minimum, such a compromise would have to include demonstrable evidence that mainframes, servers, and

PCs provide only harmless computer security-related information such as scan data, security state, and signature-based malware detection. Anything more penetrating that might allow, for example, remote access and execution from a centralized control station would probably be unacceptable, even though organizations do this routinely with their employee base.

A national data collection program would have to be sensitive to citizens' concerns for privacy.

Another possibility might be some sort of citizen-sponsored, citizen-run, grassroots data collection effort for PCs and servers, where participants agree to provide security information to a massive distributed system of peers. Such a system would not perfectly match the geographic or political perimeter of a nation, and many citizens would refuse to participate based on principle. Few members, however, of massive peer-to-peer networks for music or video complain about the privacy implications of running such software, often questionable or illegal, on their local machine. They just enjoy getting free content. The idea that a similar construct could be used to help secure national infrastructure would require demonstrating some sort of benefit to participants. This may not be possible, but the effort is worthwhile from a security perspective because data collected from a massive deployment of computers across a given nation would provide a valuable and unmatched window into the security posture of national infrastructure.

Citizens who see the benefit of a national data collection system would likely be willing to participate voluntarily.

Security Information and Event Management

The process of aggregating system data from multiple sources for the purposes of protection is referred to in the computer security community as *security information and event management* (SIEM). Today, SIEM tools can be purchased that allow collection of a diverse set of technologies from different vendors. This typically includes firewalls, intrusion detection systems (IDS), servers, routers, and applications. Just about every commercial enterprise and government agency today includes some sort of SIEM deployment. One could easily imagine this functionality being extended to include collection feeds from mainframes, servers, and PCs (see Figure 8.6).

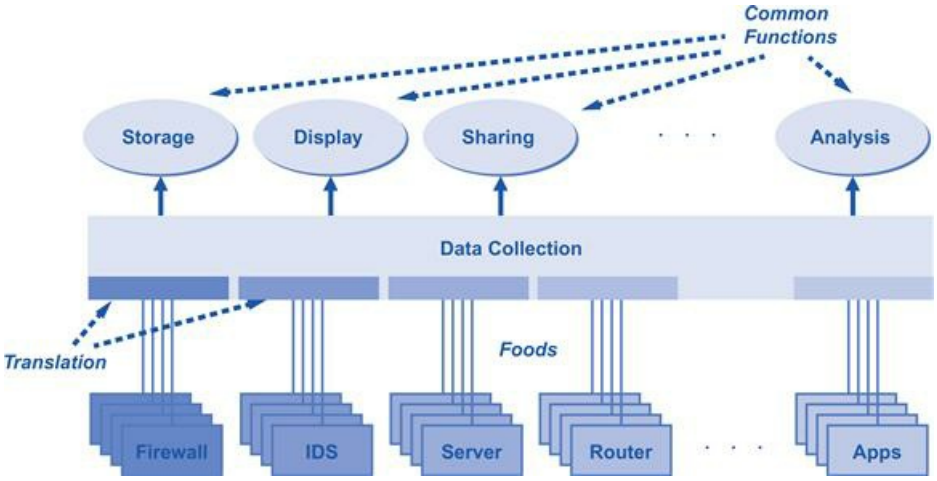


Figure 8.6 Generic SIEM architecture.

The SIEM system will include translation functions to take proprietary outputs, logs, and alarm streams from the different vendors into a common format. From this common collection format, a set of common functions can thus be performed, including data storage, display, sharing, and analysis. National infrastructure protection must include rational means for interpreting SIEM data from components, if only because many organizations will already have a SIEM system in place for processing their locally collected data. This interpretation of SIEM data from multiple feeds will be complicated by the fact that most existing SIEM deployments in different companies, sectors, and government agencies are mutually compatible. A more critical problem, however, is the reluctance among most security managers to instrument a real-time connection from their SIEM system to a national collection system. A comparable problem is that service providers do not currently feed the output of their consumer customers' data into a regional SIEM system.

Security managers will be reluctant to link their SIEM system to a national collection system.

In any event, the architecture for a national system of data collection using SIEM functionality is not hard to imagine. Functionally, each SIEM system could be set up to collect, filter, and process locally collected data for what would be considered nationally relevant data for sharing. This filtered data could then be sent encrypted over a network to an aggregation point, which would have its own SIEM functionality.

Ultimately, SIEM functions would reside at the national level for processing data from regional and enterprise aggregation points. In this type of architecture, local SIEM systems can be viewed as data sources, much as the firewalls, intrusion detection systems, and the like are viewed in a local SIEM environment (see Figure 8.7).

Local and regional SIEM systems would work as filters to feed only relevant data to a national collection point.

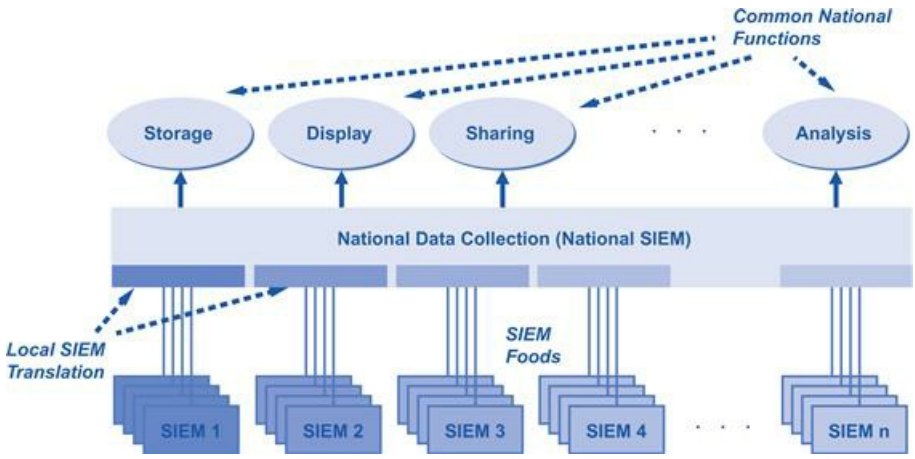


Figure 8.7 Generic national SIEM architecture.

Unfortunately, most local infrastructure managers have not been comfortable with the architecture shown in Figure 8.7 for several reasons. First, there are obviously costs involved in setting up this sort of architecture, and generally these funds have not been made available by government groups. Second, it is possible that embedded SIEM functionality could introduce functional problems in the local environment. It can increase processor utilization on systems with embedded SIEM hooks, and it can clog up network environments, especially gateway choke points, with data that might emanate from the collection probes.

Will a national data collection system put an increased financial burden on private agencies and enterprises?

A much more critical problem with the idea of national SIEM deployment is that most enterprise and government agency security managers will never be comfortable with their sensitive security data leaving local enterprise premises. Certainly, a managed security service provider might be already accepting and processing security data in a remote location, but this is a virtual private arrangement between a business and its supplier. The data is not intended for analysis other than to directly benefit the originating environment. Furthermore, a service level agreement generally dictates the terms of the engagement and can be terminated by the enterprise or agency at any time. No good solutions exist for national SIEM implementation, other than the generally agreed-upon view that national collection leads to better national security, which in turn benefits everyone.

There are still too many unanswered questions about the security of sensitive data leaving private enterprises.

Large-Scale Trending

The most fundamental processing technique used for data that is collected across national infrastructure involves the identification of *trends*. In many cases, trends in collected data are obvious, as in simple aggregate volume increases, such as packets delivered on a network. In other cases, however, trends might not be so obvious. For instance, when the collection process or monitored systems are experiencing change, the trend identification might not be easy. Suppose, for example, that a monitored network is growing, but the collection system is not. The result is that critical data might be missed, which could be misleading. Similarly, if a change is made to the underlying collection system, perhaps involving a new technology or vendor, then this could influence the trends presumably being observed.

At the simplest level, a trend involves some quantitative attribute going up (growth), going down (reduction), staying the same (leveling), or doing none of the above (unpredictability). When data jumps around, for example, it might not be easy to draw a conclusion; however, the fact that it is jumping around might itself be an important and useful conclusion. Perhaps the most common question infrastructure managers ask with respect to security is whether attacks are increasing, decreasing, or staying the same with respect to some component in question. This question about attack trends is a favorite among CEOs and national legislators. It can only be answered accurately in the context of collected data.

Tracking trends may tell us whether adversarial attacks are increasing, decreasing, or staying the same.

As a concrete example, over a nine-month period from June 2006 to March 2007, a stable collection system embedded in a global service provider's backbone detected an increase in behavior consistent with malicious bots. As was outlined in the first chapter, a bot is a piece of software inserted into a target system, usually a broadband-connected PC, for malicious or questionable purposes. The bot might be used to attack some target, it might be used to send spam, or it might be used to steal personal information. The detection of bot behavior comes from collecting traffic information for the purpose of identifying communication between a number of end-user PCs and a smaller number of servers on the Internet.

By collecting evidence of bot behavior and rendering the results in a simple histogram, the growth of bots can be seen clearly, and local management decisions can be made accordingly (see Figure 8.8).

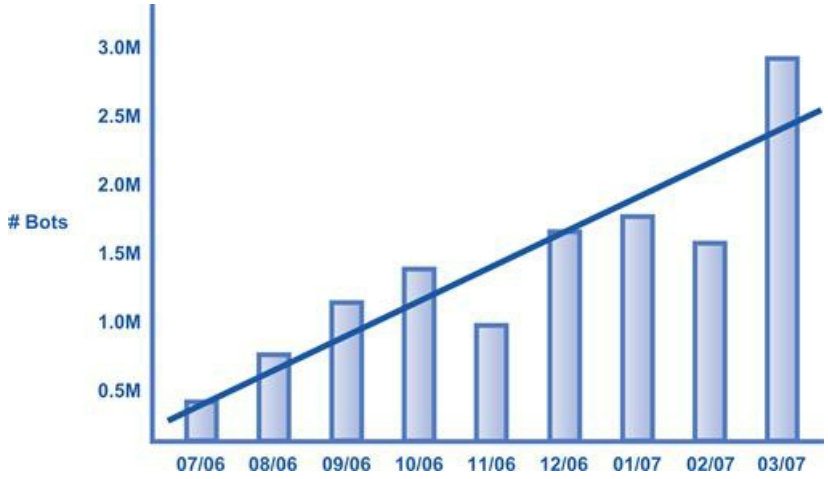


Figure 8.8 Growth trend in botnet behavior over 9-month period (2006–2007).

Most managers shown the growth trend in Figure 8.8 would conclude that bots represented an increasing threat during this time period; however, proper national infrastructure protection requires a more thorough analysis before any real conclusions are drawn. The following are some basic practical considerations that must be made by security analysts before the trend in any data collection chart can be trusted:

Collected data must be analyzed to determine what it can accurately tell us about trends.

- *Underlying collection*—Amazingly, trend data such as that shown in Figure 8.8 is often provided in the context of a collection architecture that might be changing. For example, if a collection system for bots is getting more accurate through algorithmic improvements or better coverage, then the observed growth in bots might simply reflect a more effective use of detection tools.
- *Volunteered data*—It is common for government organizations to use data volunteered from commercial entities as the basis for drawing conclusions about trends. This can be dangerous, because weak or nonexistent controls are in place regarding how the information is collected and managed. It is also possible that data might be volunteered that is incorrect or tampered with for some malicious or mischievous purpose.
- *Relevant coverage*—The amount of coverage across the environment in which the data is collected will affect the validity of an observed trend. Suppose, for example, that a small organization with an Internet connection uses that connection to draw conclusions about traffic trends. This certainly would be a less attractive vantage point than a global Internet carrier making the same determination.

These issues highlight the importance of national infrastructure managers taking a mature approach to the interpretation of collected data. This is especially important because trend information so often drives the allocation of critical resources and funding. At the national level, for example, experienced security experts can point to dozens of cases where some sort of trend is used to advance the case for the funding of an initiative. This often involves hype about the rise of some virus or worm.

Trends must be interpreted carefully before they are used to justify changes in funding levels.

The Conficker worm, for example, reportedly included some sort of embedded attack that would occur on April 1, 2009. Conficker was especially relevant—and still is—because its operation involved several million bots. This makes it one of the more potentially powerful botnets known to the security community. Most security experts understood that there was nothing in the Conficker code to suggest such an event on that particular date, but predicted attack dates are convenient for attracting attention and are thus common. National infrastructure protection begs a more mature approach to the public interpretation of collected data.

Tracking a Worm

Data collection provides an excellent means for tracking a worm. Recall that a worm is a program that does three things: (1) it finds network-visible computers that can accept a copy of the worm program, (2) it sends a copy of itself to one of the identified network-visible machines, and (3) it initiates remote execution of the new remote instance of the program on the network-visible target. This starts a chain reaction in which the identifying, copying, and remote execution continue indefinitely. By collecting network metadata while this is all happening, security analysts can generally determine what the worm is doing and how serious the event might be. In the best possible cases, the collection might even provide hints that can be used to stop a worm from developing, which is obviously attractive for national infrastructure security.

Collecting network metadata allows security analysts to track a worm's progress and predict its course.

In 2003 and 2004, the Internet experienced an unusually large number of worm events. This was due primarily to the poor processes that were in place at the time for operating system and application-level software patching. This patching problem was true for both enterprise systems and home broadband users. During this time period, one worm after another seemed to rage across the Internet, and most observers viewed these events as largely spontaneous; that is, the general consensus was that worms would spread in just a few minutes, and that data collection was useless. If a worm was going to get you, the thinking went, it would get you fast, and there was nothing you could do in advance to stop the event.

The reality of the situation was actually more subtle. The SQL/Slammer worm of January 2003, for example, was one that appeared to have a spontaneous impact on traffic. In the minutes during which the worm appeared to have spread significantly, packets of User Datagram Protocol (UDP) traffic went from small, predictable volumes with few anomalies to an immediately spiked upward volume. On first glance, this happened in a manner that suggested no warnings, no time for preparation, and no time for incident response (see Figure 8.9).

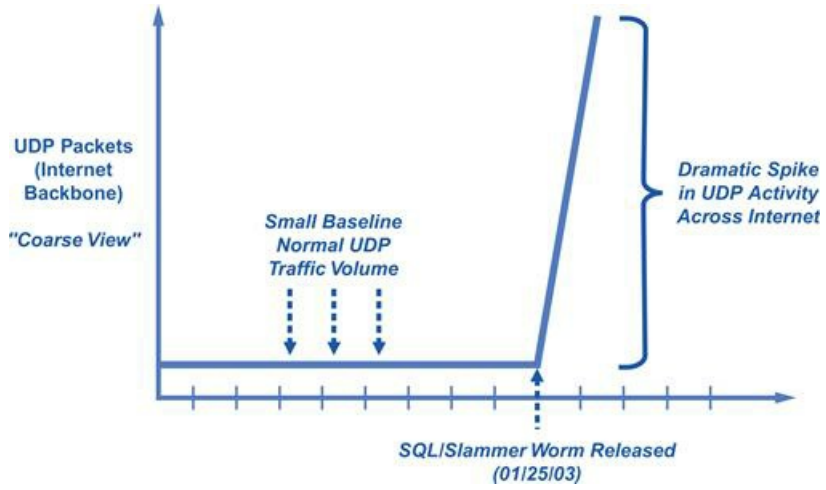


Figure 8.9 Coarse view of UDP traffic spike from SQL/Slammer worm.(Figure courtesy of Dave Gross and

Brian Rexroad.)

The spike in packet volume due to the SQL/Slammer worm certainly appeared to be immediate and without warning. Upon much closer examination, however, one finds that the UDP data leading up to this event might have carried some indications and warning value from a security perspective. In particular, starting in early January 2003, UDP volumes on the specific SQL port used by the worm were displaying anomalous behavior. On January 2, 2003, the first spike occurred, and this was followed by three weeks of similarly odd behavior. While it might be a stretch to absolutely conclude that these odd spikes were early attempts at producing a worm, no one can argue that they suggested a serious change in UDP behavior on the Internet (see Figure 8.10).

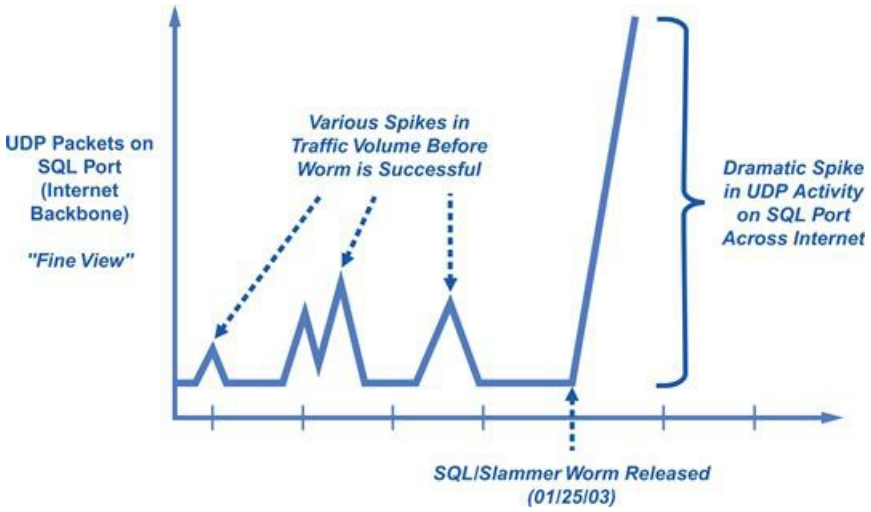


Figure 8.10 Fine view of UDP traffic spike due to SQL/Slammer worm.(Figure courtesy of Dave Gross and Brian Rexroad.)

The suggestion here is that a more detailed inspection of UDP behavior on the SQL port before the SQL/Slammer worm achieved its aim could have given valuable data to security engineers. In particular, the vulnerability exploited by the SQL/Slammer worm was known at the time, although most security managers were lax to install the patch. If the information in Figure 8.10 had been widely disseminated at the time, then anyone wise enough to heed the warning and install the patch would have been immune from the SQL/Slammer worm. The implications of this situation should be obvious from the perspective of national infrastructure protection.

Collecting and analyzing data are important steps; the next is acting on the data in a timely manner.

National Collection Program

Implementing a program of national data collection for infrastructure security will require a combination of public outreach initiatives before any large-scale structures can be put in place. The citizenry and business community must fully understand the purpose, usage, and controls associated with a collection system. Mechanisms for preventing privacy abuses must be paramount in the discussion and embedded into any architecture that might be proposed. The specifics of how this debate might be influenced are beyond the scope of this book, but it goes without saying that no national collection program can be successful without this requisite step.

A successful national data collection program must address the concerns of citizens and the business community regarding protection of private data.

Once general acceptance has been obtained for the creation of a national data collection program, the following technical and architectural issues must be addressed:

- *Data sources*—National attention is required to define which data sources are deemed valuable for providing security information to the broad collection function. Important mainframes and servers in organizations and agencies charged with infrastructure protection would seem the most obvious to include. End-user PCs owned and operated by private citizens would probably be the most difficult to include.
- *Protected transit*—Security-relevant data collected from identified sources would need to be transmitted via suitable networks with sufficient encryption. Sizing consideration could dictate limits on the amount of information that could be pulled from a particular source.
- *Storage considerations*—The amount of information collected is obviously controllable, but the appetite for data from security analysts is usually unlimited. As such, pressure would exist to maximize the amount of information stored, as well as the length of time the data is available for analysis.
- *Data reduction emphasis*—Across the entire national initiative for data collection, time and energy should be directed toward reducing the amount of data being handled. Obviously, this is critical if a given collection method should inadvertently grab more information than is needed or might include information that has no relevance to the security challenge.

A planned, stepwise approach to national data collection could create a system that would be of immense value in the quest to protect our national infrastructure.

While each of these issues represents a technical challenge, particularly in terms of sizing and scaling, they can be combined into a reasonable system if engineered properly. The overall approach will benefit from stepwise refinement methods that start with a tractable subset of data sources initially which gradually increases with time.

Finally, let's briefly look at some practical ways to collect data in the critical national infrastructure through the systems risk view. The Transportation Security Administration (TSA) is responsible for

developing an understanding of data collection for asset dependencies, interdependencies, and critical functionality beyond what is required for the National Asset Database (NADB), including collecting and storing system-level data. In conjunction with the Government Coordinating Council (GCC) and the Sector Coordinating Council (SCC) members, the transportation systems sector will work to identify targeted data sets, based on strategic risk objectives (SROs), that are required to accomplish risk-informed security activities. While the NADB is currently asset-focused, the transportation systems sector will seek to build a systems perspective into the existing NADB. This will not result in a secondary repository for information, but rather enhance the existing NADB.

In collecting cyber asset transportation data, the TSA will use previous data collection efforts (the NADB); current TSA data collection approaches (corporate security reviews, risk assessments, rail inspections, commercial site vulnerability checklist for cyber assets); and publicly available information, such as the Securities and Exchange Commission filings. The GCC/SCC construct will serve as the primary vehicle for sharing cyber asset data within the sector (the logical collection of assets, systems, or networks that provide a common function to the economy, government, or society). The transportation systems sector is one of 17 critical infrastructure and key resources (CI/KR) sectors. Cyber asset information will also be shared on an as-needed basis with other sector lead agencies, such as the National Cyber Security Division (NCSD) communications sector and the Department of Energy (DOE).

The data gathered will be used in a variety of ways throughout the risk assessment and prioritization processes. Uses of the information will include, but are not limited to, risk assessments on systems, interdependency analyses, critical national infrastructure modeling, infrastructure prioritization, and reporting. The transportation systems sector will ensure that information protection mechanisms are in place to protect against misuse, unauthorized disclosure, or theft.

Data Collection Efforts: Systems and Assets

Collecting critical national infrastructure data through the systems risk view focuses on multiple, heterogeneous, geographically distributed systems that are embedded in networks at multiple levels. The four views capture multiple ways of addressing systems and add to a more robust assessment:

- Modal view
- Geographic view
- Functional view
- Ownership view

Modal View

The modal view treats all classes of assets within a mode collectively as a system.

Critical national infrastructure information in the modal view is categorized by interdependencies and supply chain implications that are specific to a particular mode of transportation. In addition to focusing on individual assets, nodes, and links, information specific to the modal view includes how those assets, nodes, and links interact within the mode and with other modes, their emergent properties and governing principles, or legislative information with specific modal impact. The sector will collect data through existing mode-specific data lists and readily available databases. Sector partners, in cooperation with other federal agencies, state and local governments, the GCC and SCC, trade associations, nongovernmental organizations, and industry subject matter experts, will work to build a complete data set to best understand the risks to these modes.

Geographic View

The geographic risk view compiles transportation infrastructure data within specific regions of the nation. The boundaries of those regions may vary based on the purpose and necessary parameters of an assessment. Regions may contain markedly different assets and systems, and thus the risks to those systems and the types of data collected from those regions will differ as well. Data collection in this view will allow an information set to be defined by what is physically located within that region and the processes or policies that impact that specific region. Therefore, assets, links, nodes, and emergent properties within a defined geographic area are evaluated as an integrated system.

Functional View

The functional view of data collection looks at the function a system fulfills within the supply chain. Examples of a functional view of systems include all of the assets, links, nodes, processes, policies, and emergent properties associated with:

- Delivery of critical medicines.
- Delivery of chlorine for drinking water or other purposes.
- Delivery of heating oil to the Northeast.

By examining the function a system plays in society, the critical aspects of the system can be measured. This view also will have value in identifying interdependencies with other critical national infrastructures. Collection efforts in the functional view are in the early stages and will be expanded over time.

Ownership View

According to the GAO, the private sector owns approximately 90% of the nation's assets. The ownership view examines information on ownership of assets, including the owner's/operator's decision structure, policies, and procedures, and recognizes those assets owned by the same entity as an integrated system. Any data requested from owner/operators by the federal government for risk analysis need not be all-encompassing. Rather, critical national infrastructure information required from owners by the federal government will be targeted and based on SROs.

Finally, the asset data are segmented by transportation modes. Data collection efforts by the transportation systems sector will not attempt to be all-encompassing. In addition to using asset data collected in the NADB, the sector security partners will establish SROs through the systems-based risk management (SBRM) approach, and only targeted data related to those SROs will need to be collected. The transportation systems sector plans to employ the GCC/SCC framework to aid in the process of identifying and acquiring that targeted asset data. Specific information concerning the data collection efforts of individual modes can be found in the respective modal implementation plan annexes. This results in a data collection strategy being guided by a set of core principles (see "An Agenda for Action for Data Collection Strategy Guiding Principles for the Critical National Infrastructure").

An Agenda for Action for Data Collection Strategy Guiding Principles for the Critical National Infrastructure

To provide timely, comprehensive, relevant, and accurate data that can guide and improve policymaking, program development, and performance monitoring in support of a data collection strategy, the following set of core principles must be adhered to (check all tasks completed):

1. The data collected are timely, accurate, relevant, and cost-effective.
2. Data efforts are cost-efficient and purposeful, and minimize redundancy and respondent burden.
3. Data are used to inform, monitor, and continuously improve policies and programs.
4. Data activities seek the highest quality of data and data collection methodologies and utilization.
5. Data activities are coordinated within the agency, maximizing the standardization of data and sharing across programs.
6. Partnerships and collaboration with federal and nonfederal stakeholders will be cultivated to support common goals and objectives around data activities.
7. Activities related to the collection and use of data will be consistent with applicable confidentiality, privacy and other laws, regulations, and relevant authorities.
8. Data activities will adhere to appropriate government-wide guidance issued by OMB, its advisory bodies, and other relevant authorities.

Summary

This chapter focused on data collection in the critical national infrastructure. At the national level, security-relevant data must first be collected at the local or regional level by individual asset managers and a subset then selected for broader aggregation into a national collection system. In some cases, local and regional collection can be directly connected to national programs. Larger scale collection points on wide-area networks, perhaps run by carriers or government agencies, can also be embedded into the collection scheme and combined with local, regional, and aggregated data.

The critical national infrastructure collects data in minutes from measurement sites and delivers it to a data management facility (DMF) for processing within 2 minutes. The underlying network infrastructure is critical to this ability. The architecture of the entire data flow provides reliability and performance through a division of labor from data collection to the doorstep, and back to the DMF. The ingests can process data in minutes, making them available for many uses that leads to overall improved data quality. A data archive provides this high-quality data to a global audience of data users.

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

Chapter Review Questions/Exercises

True/False

1. True or False? A basic tenet of computer security is that diligent and ongoing observation of computing and networking behavior can downgrade malicious activity.
2. True or False? Perhaps the most wasteful type of data for collection in the national infrastructure is network *metadata*.
3. True or False? National infrastructure protection initiatives have traditionally included provision for collecting data from mainframes, servers, and PCs.
4. True or False? The process of aggregating system data from single sources for the purposes of protection is referred to in the computer security community as *security information and event management* (SIEM).
5. True or False? The least fundamental processing technique used for data that is collected across national infrastructure involves the identification of *trends*.

Multiple Choice

1. At both local and national levels, data collection decisions for national infrastructure should be based on the following three security goals:

A. Planning an attack
B. Preventing an attack
C. Strategizing an attack
D. Mitigating an attack
E. Analyzing an attack

2. The following are some basic practical considerations that must be made by security analysts before the trend in any data collection chart can be trusted, except which two:

A. Effective security
B. Underlying collection
C. Information nondisclosure
D. Volunteered data
E. Relevant coverage

3. Data collection provides an excellent means for tracking a worm. Recall that a worm is a program that does three things:

A. It finds network-visible computers that can accept a copy of the worm program.
B. It focuses on long-term diversity of proof factors.
C. It sends a copy of itself to one of the identified network-visible machines.
D. It directs long-term online factors.
E. It initiates remote execution of the new remote instance of the program on the network-visible target.

4. Once general acceptance has been obtained for the creation of a national data collection program, the following technical and architectural issues must be addressed, except which one:

A. Data sources

B. Situational awareness

C. Protected transit

D. Storage considerations

E. Data reduction emphasis

5. The four views capture multiple ways of addressing systems and add to a more robust assessment, except which one:

A. Modal view

B. Geographic view

C. Functional view

D. Internal view

E. Ownership view

Exercise

Problem

This scenario covers an agency that is developing a comprehensive critical national infrastructure asset management system using mostly internal resources. It has tried several approaches for critical infrastructure data collection and has used both the agency's personnel and consultants. Please identify what type of critical infrastructure data collection is needed for supporting decisions at the network level.

Hands-On Projects

Project

An agency has focused its system development and critical infrastructure data collection efforts on separate engineering management systems for different types of assets and is working on the integration of these systems. In this case, the agency focused on the data collection for two types of assets: pavement and storm water management facilities. Please identify what type of critical infrastructure data collection is needed for pavement and storm water management facilities.

Case Projects

Problem

This case study illustrates a different approach for asset management that relies heavily on the private sector support. The agency outsources most of the maintenance of its assets through performance-based contracts.

Although consultants perform most of the data collection, the agency has also emphasized incorporation of citizen input on the asset evaluation process. Please identify what type of critical infrastructure data collection is needed for asset management through private sector support.

Optional Team Case Project

Problem

This agency focused on critical infrastructure data collection practices that support one of the components of the agency's asset management system and the maintenance management system (MMS). The agency has developed the system and conducted the initial critical infrastructure data collection by using a consulting firm that specializes in asset management. Please identify what type of critical infrastructure data collection is needed for the agency's asset management system and maintenance management system (MMS).

[1] S. Northcutt, *Network Intrusion Detection: An Analyst's Handbook*, New Riders Publishing, Berkeley, CA, 1999, p. 34.

9

Correlation

Chapter Outline

Conventional Security Correlation Methods
Quality and Reliability Issues in Data Correlation
Correlating Data to Detect a Worm
Correlating Data to Detect a Botnet
Large-Scale Correlation Process
National Correlation Program
Correlation Rules for Critical National Infrastructure Cyber Security
Summary
Chapter Review Questions/Exercises

A benefit of anomaly detection is that it can potentially recognize unforeseen attacks. A limitation is that it can be hard to distinguish normal from abnormal behavior.

Dorothy Denning[1]

Computer and network security experts understand that correlation is one of the most powerful analytic methods available for threat investigation. Intrusion detection systems, for example, are only useful when the alarm streams that result from signature or profile-based processing can be correlated with data from other areas. When alarms are viewed in isolation, they are of only limited use. This limitation in processing alarms is directly related to the complexity of the target environment; that is, decision makers in more complex environments will be more reliant on correlating collected data than in more limited environments. Proper national infrastructure protection is therefore highly dependent upon a coordinated program of information correlation from all available sources.

Data in a vacuum is irrelevant; it must be compared with other data to determine its relevance and importance.

From a foundational perspective, four distinct analytic methods are available for correlating cyber security information: *profile-based*, *signature-based*, *domain-based*, and *time-based correlation*. Profile-based correlation involves comparison of a normal profile of target activity with observed patterns of activity. Presumably, if a substantive difference exists between normal and observed, this could signal a possible intrusion. Obviously, many situations exist where observed activity is not normal but does not signal an intrusion. Websites running specials or supporting some limited-time engagement, for example, will see traffic spikes during these periods

that do not match normal patterns. Nevertheless, anomalies with activity profiles are worthy of attention from a security perspective (see Figure 9.1).

Comparing data determines what is normal and what is an anomaly.

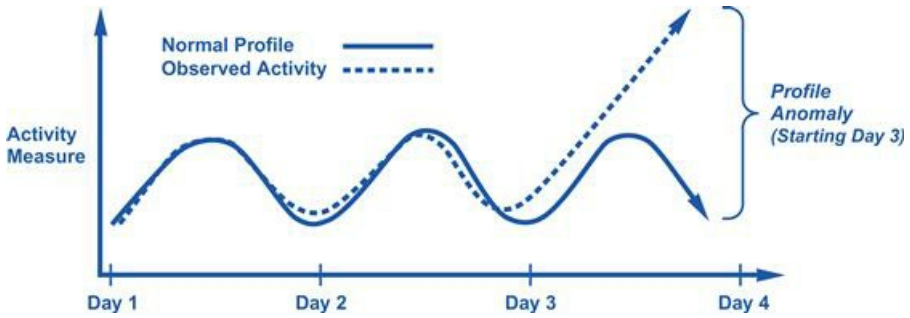


Figure 9.1 Profile-based activity anomaly.

Signature-based correlation involves comparing a signature pattern of some known malicious condition to observed activity. If the two match, then high confidence exists that an intrusion is under way. The challenge is when observed activity shares characteristics with a signature but does not exactly match. This requires diligence from the security team to stay focused. Most signature-based correlation patterns involve some sequence of events, such as commands, which are defined as a discrete signature, and comparison against logs of observed activity. For example, antivirus software, antispam algorithms, and intrusion detection systems all operate in this manner (see Figure 9.2).

Data comparison, especially from different domains, creates a clearer picture of adversary activity.

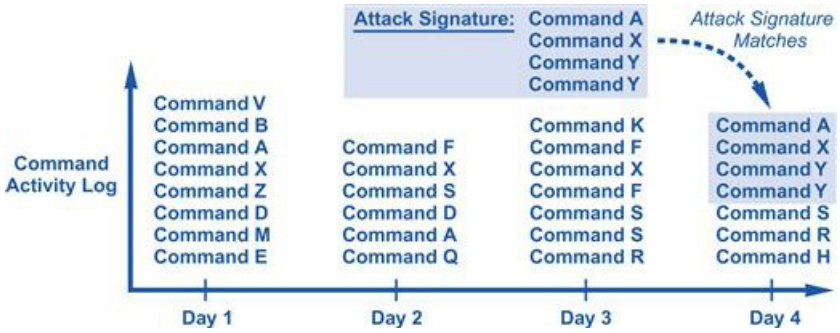


Figure 9.2 Signature-based activity match.

Domain-based correlation involves comparing data from one domain with data collected in an entirely different context. Relevant differences in the data collection environments include computing environment, software architecture, networking technology, application profiles, and type of business being supported. For example, data collected by a power company about an attack could easily differ from data collected by a federal civilian agency on the same incident. Similarly, two targets of a botnet attack could report different isolated

views that could be correlated into a single common view. This requires a prearranged transport, collection, and analysis approach leading to a common correlated output (see Figure 9.3).

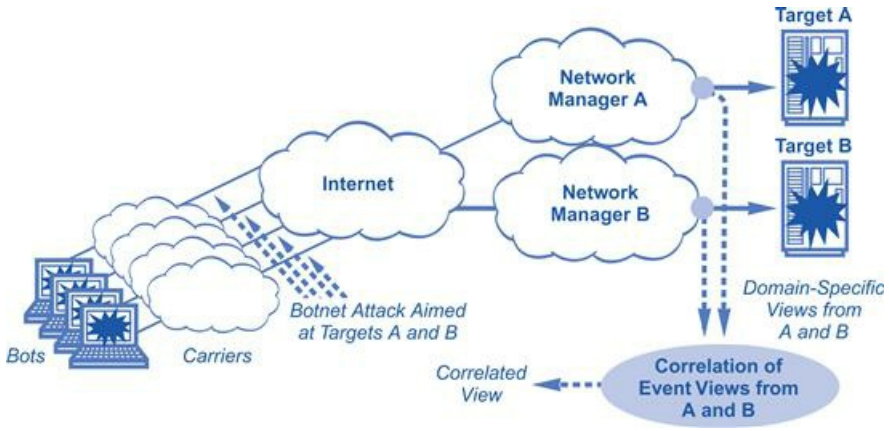


Figure 9.3 Domain-based correlation of a botnet attack at two targets.

Time-based correlation compares data collected during one time period with data collected at a different time. This can involve the same or different data source but is obviously more effective if the data source is the same, because this removes one variable from the correlative analysis. Many types of attacks will not be time sensitive and are thus not well suited to this type of correlation; for example, a single break-in, during which malware is embedded in a target system, might not be a good candidate for time-based correlation. Attacks that are multistage, however, such as many "low and slow" approaches, are quite well suited to the approach. Botnet attacks are increasingly being designed by adversaries in this manner, with the distributed program attacking its target in a slower and more deliberate manner than via a single bombardment. Detection of such an event is well suited to time-based correlation, because potentially significant time periods could exist between successive steps in an attack. Time-based correlation would be required to connect relevant steps and to filter out noisy, irrelevant activity (see Figure 9.4).

Changes that appear over time may indicate a slowly building, deliberate attack.

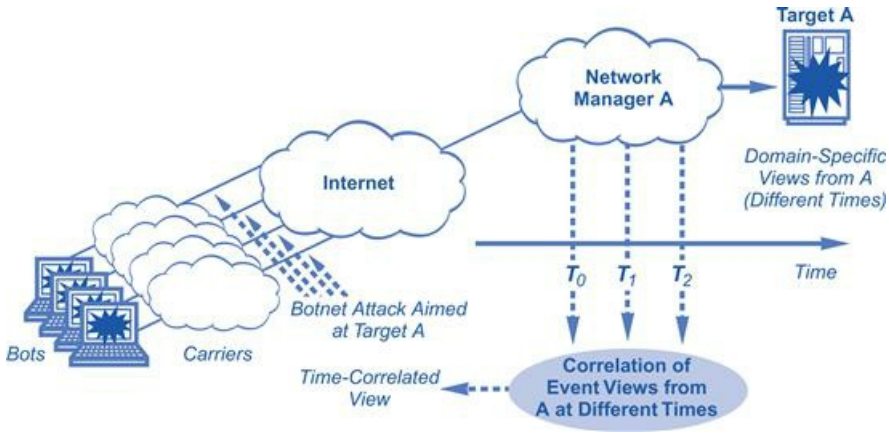


Figure 9.4 Time-based correlation of a botnet attack.

The essence of correlation in cyber security involves comparison of various pieces of data to determine whether an intrusion is under way. In the most desirable set of circumstances, this involves comparing two pieces of data for which every associated, relevant attribute is the same *except for one*. Such a scenario allows the analyst to focus in on that one attribute. Time-based correlation works nicely when the collection environment is exactly the same but the data is collected at different times. The analyst does not have to worry about whether changes in other factors are affecting the data, as only the time changes. In the most complex case, however, multiple pieces of data are collected from environments where the associated, relevant attributes differ. The analyst thus must juggle concerns about which attributes in which environments might be affecting the data. This greatly complicates the correlation task (see Figure 9.5).

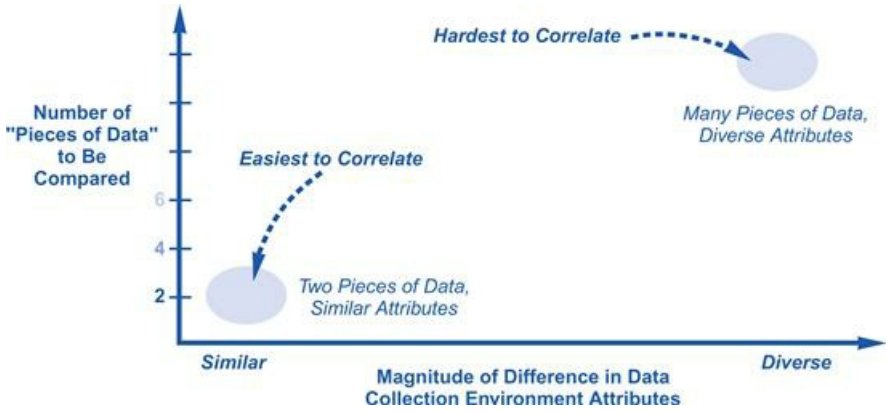


Figure 9.5 Taxonomy of correlation scenarios.

This data collection attribute taxonomy is important to national infrastructure, because most practical cases tend to be very complex cases that are difficult to correlate. Information that becomes available during an incident usually originates from wildly different sources with diverse collection methods, processing tools, network views, and so on. Worm events on the Internet, for example, are often observed with considerable scrutiny by some organizations (perhaps with bad consequences), whereas other groups might not even notice that a security event is ongoing. Only the most mature correlation analysts will have the ability to factor these differences in viewpoint into an accurate broad conclusion about security. To date, this has required experienced human beings with considerable training. Additional research is required before dependable tools will be available to perform accurate correlation on multiple, diverse inputs.

We currently rely on human analysis of data across different domains and during different time periods; no software or program can factor in all relevant elements.

Conventional Security Correlation Methods

The current state of the practice in day-to-day network security correlation in existing national infrastructure settings is based on a technique known as *threat management*. In this approach, data aggregated from multiple sources is correlated to identify patterns, trends, and relationships. The overall approach relies on a *security information and event management* (SIEM) system for the underlying collection and aggregation of relevant data. A SIEM system does the best it can in terms of identifying correlation instances, using the best available algorithms for profile, signature, domain, and time-based analysis, subject to the practical limitations mentioned above. Four of the primary feeds into a typical commercially available SIEM tool for threat management are listed in the box.

Information Feeds for SIEM Threat Management

- *Firewall audit trails*—Firewalls generate audit records when certain types of security-relevant events occur such as denied connection requests. These records are of limited use in isolation but are often useful for correlation with other data. Other static information about a firewall, such as its inbound and outbound policy, is also important for correlation.
- *Intrusion detection and prevention system alarms*—Intrusion detection and prevention systems are designed specifically to generate alarm data when suspicious activity is observed. The problem is that it is not always easy to determine if something suspicious is truly malicious. Generally, correlation with other data is required to make this determination.
- *Operating system or application logs*—Output log files generated by activity on an operating system or software application can provide useful indications and warnings for security. The first step in forensics, for example, involves examination of log files for evidence. (Good hackers know not to leave such obvious tracks, of course.) In addition to logs, the specific attributes of the operating system and application are also important for correlation. This can include version, vendor, and configuration data.
- *Network device metadata*—Information about network behavior is quickly becoming recognized by cyber security experts as possibly being the most powerful tool available for threat management. Metadata showing source and destination information about addresses and ports, as well as information about protocol, direction of flow, and status of protocol flags and settings, gives security analysts a view into network activity unavailable through any other means.

The interplay between the various security devices in a local threat management system is sometimes straightforward. If an intrusion detection system generates an alarm signaling some sort of problem involving a given Internet protocol (IP) source address and corresponding destination port, and if the local environment also allows inbound traffic to this destination port, then the correlation process could generate a recommendation that the local firewall block either this source address or this port. Many commercial firewalls and intrusion detection systems provide this capability today, although the reality is that many network managers do not make use of this type of protection. This is usually due to a lack of familiarity with

the process, as well as a common lack of local knowledge about the egress and ingress traffic through an enterprise gateway or perimeter. This is a shame, because when it is done properly the protection achieved can be quite powerful (see Figure 9.6).

Many security managers underutilize the commercial firewalls at their disposal.

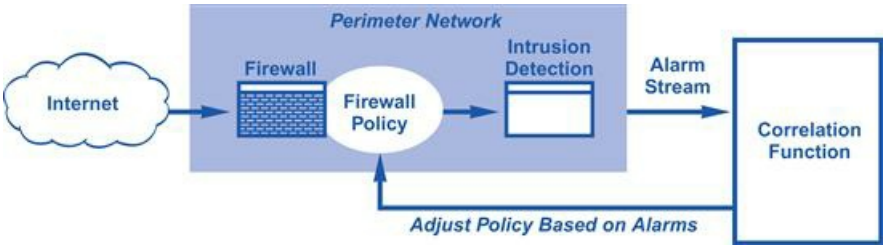


Figure 9.6 Correlating intrusion detection alarms with firewall policy rules.

The example shown above demonstrates the natural feedback loop that can occur when data is correlated—that is, as interpretation resulting from the correlation task is fed back to the firewall as a new data input. This in turn affects processing and will eventually change the correlation function output. This feedback loop will cease when the resultant interpretations are no longer new and have no changes to report back to the firewall. Security managers often configure their intrusion detection systems to suppress output when this steady-state condition occurs. This reduces operator burden but great care must be taken to ensure that valid indicators are not being missed.

Exercise caution in suppressing output once a steady-state condition has been achieved; otherwise, valid indicators may be missed.

The correlation function can extend to different parts of the same organization with different networks, servers, applications, and management groups. Surprisingly, many correlation activities are complicated by such decentralization. To illustrate, suppose that two groups in a company experience similar security problems. The root-cause data from each group should be correlated toward the optimal interpretation. If, for example, each group found similar malware in their systems, then this observation could signal the source of the attack, such as a common software vendor. This fact might not be easy to determine by either group in isolation.

Quality and Reliability Issues in Data Correlation

To create a proper security correlation process for national infrastructure protection in an environment of large, cross-organizational size, scope, and scale, several technical and operational factors must be considered. The most important such considerations involve the *quality* and *reliability* of the data sources. This calls into question any national-level initiative for which these attributes cannot be controlled.

Regarding the quality of data, the best situation involves a service level agreement between the data source and correlation function. Managed security services are useful, because the provider will ensure that data quality exists within well-defined parameters. When data originates from a mix of organizations with no service level agreements, the potential exists for inaccurate, misleading, or invalid data to be made available. This can only be dealt with by automated or human filtering in which the data source and attributes are factored into the analysis. This is troublesome when correlation relies on information *volunteered* across the Internet. Grass roots efforts to collect volunteered data will always have an issue with guaranteed data quality.

Service level agreements help guarantee quality of data.

A similar concern exists with the reliability of a data source, especially for volunteered feeds. When data is important for regular analysis, perhaps based on a profile, its continued reliability is essential; for example, if a data stream experiences gaps or changes, perhaps at the whim of the feed source owner, this could easily confuse the correlation process. Gaps, in particular, make it tough to match observed activity against the desired patterns. This issue is especially difficult to manage when data is being volunteered by varied sources. Thus, in addition to quality issues, correlation based on any imperfect collection process, including the use of volunteered data, will also face inherent challenges related to reliability (see Figure 9.7).

Due to limited oversight of volunteered data, its quality and reliability cannot be guaranteed.

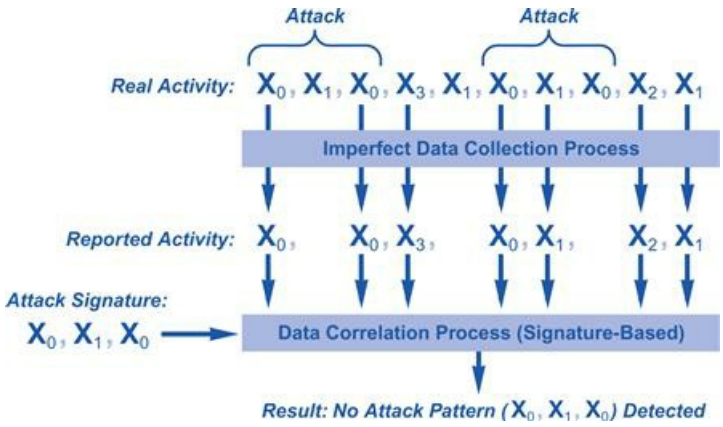


Figure 9.7 Incorrect correlation result due to imperfect collection.

Many national initiatives today rely on data sources agreeing to provide data on a best effort basis. These initiatives must be viewed with great suspicion, because the conclusions being drawn will be based on a subset

of relevant data. This includes initiatives where participants send intrusion detection alarms, highly sampled netflow summaries, and log files. If the delivery is not consistent, predictable, and guaranteed, then the correlation result is suspect; for example, attack signature patterns can be missed, profiles can be incorrectly matched or missed, and so on. National infrastructure managers should thus only collect data that is associated with a consistent service level agreement.

Without consistent, predictable, and guaranteed data delivery, correlations are likely to be incorrect and data is likely missing.

Correlating Data to Detect a Worm

Network service providers have a particularly good vantage point for correlating data across multiple companies, agencies, groups, individuals, and regions. All government, business, and consumer traffic must traverse a provider backbone at some point, so this becomes an excellent source of correlation information. Obviously, if this is done, great care must be taken to ensure full compliance with applicable laws with a deep respect for privacy considerations. The effort is worth the time, because service providers collecting netflow information on a broad scale can generally correlate observed activity with known patterns to detect large-scale events such as worms. This is typically done with greater accuracy than existing computer and network security techniques using firewalls and intrusion detection systems.

Network service providers are in a unique position to collect information across multiple venues.

As an illustration, consider that security devices such as intrusion detection systems are put in place to detect worms and viruses. Unfortunately, many worms and viruses are not so easy for an intrusion detection system to detect. The Nachi worm is such an example; it raged across the Internet during the summer of 2003, using the Internet Control Messaging Protocol (ICMP) as one of its mechanisms for operation. Some speculate that the worm was actually intended to find infected systems on the Internet and go fix them. What happened instead was that the ICMP packet flows got out of hand, which is the main reason why this worm caused more damage than perhaps had been intended by its designer.

Most intrusion detection systems were not set up well to detect this problem, because an intrusion detection system is typically not interested in changes to some service port. In contrast, any network system that was monitoring ICMP flows would see that something was amiss. On one service provider's backbone this increase was evident as the Nachi worm began to operate. By simply counting ICMP packets crossing gateways on its backbone, the provider could quickly see the spike in traffic flows due to the worm across several key network gateways. The resultant time-based correlation of collected ICMP data over several hours revealed the impending worm event (see Figure 9.8).

Network service providers have unique views of network activity that allow them to see when something is amiss.

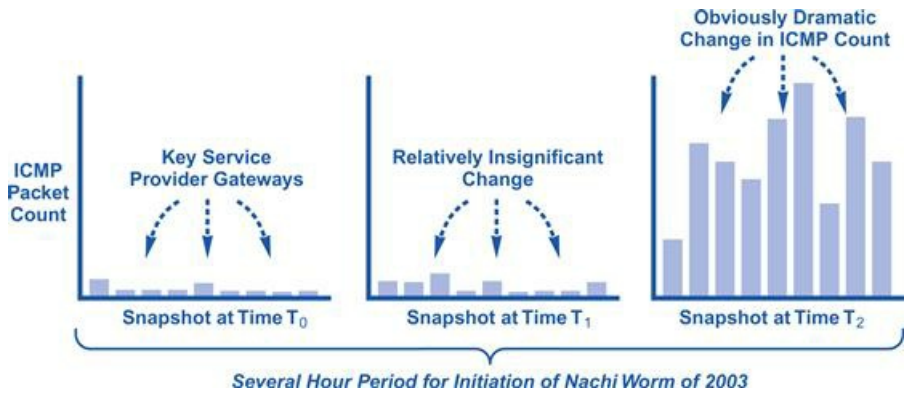


Figure 9.8 Time-based correlation to detect worm.

One might conclude from the above example that by monitoring broad network traffic collected across organizations a much more accurate security picture can be drawn. A complementary conclusion that can be drawn from this example is that the network service provider clearly plays a key role in the detection of large-scale attacks. Over the past decade, so much security responsibility has been distributed to end users and organizational managers that no common strategy exists for infrastructure protection. Instead, when a problem occurs, all vulnerable endpoints must scramble to determine a suitable means of addressing it, and this can involve conflicting approaches. One group might choose to ignore and drop all packets associated with an attack, whereas another group might choose to collect, process, and send responses to the sources of attack packets. This distribution of security implies that national infrastructure protection should include some degree of centralized operations. For large-scale network service, this can only be reasonably managed by the service provider.

Correlating Data to Detect a Botnet

The most insidious type of attack one finds today in any large-scale, distributed, Internet-connected network environment is the botnet. The way a botnet works is that an attacker rounds up a collection of Internet-connected computers to be used as bots; these computers are generally PCs attached to some home broadband service and are generally poorly administered by the home user. Such improper system administration allows for easy insertion of malware, perhaps through fishing or other social engineering means.

Once the bots have been configured with suitable malware, they are commanded by a series of bot controllers located around the Internet. These controllers generally utilize some familiar protocol such as Internet Relay Chat (IRC) simply for convenience, although they could certainly use any sort of communication protocol to interact with their bots. The idea is that the controller commands the bots to perform an attack task aimed at a target predetermined by the botnet operator. This works to the advantage of the attacker, because the bots are generally distributed across a broad geographic spectrum, and their bandwidth capacity might be substantive when viewed as a collective capability.

A botnet uses home-based PCs to distribute an attack.

If two bots can generate 1 Mbps of attack traffic, then a target with a 1-Gbps inbound connection can be filled up by 2000 bots, which turns out to be a modestly sized botnet. Following this logic, a much larger botnet, perhaps with hundreds of thousands or even millions of bots, can be viewed as a particularly substantive problem for national infrastructure that requires attention. The correlation issue in this case is that no single endpoint will have a suitable vantage point to determine the size, scope, or intensity of a given botnet. One might suggest that the only reasonable chance one has of actually performing the proper correlation relative to a botnet is in the context of carrier infrastructure.

Steps for Botnet Detection

The steps involved in the detection of a botnet via correlative analysis by a network carrier are roughly as follows:

1. *Broad data collection*—The detection of a botnet requires a broad enough vantage point for collecting data from both broadband-connected PCs as well as enterprise servers visible to the Internet. The type of information needed is essentially netflow-type metadata, including source, destination, and traffic types.
2. *One-to-many communication correlation*—From the collected data, the correlative analysis must focus on identifying the typical one-to-many fan-out pattern found in a distributed botnet. This pattern can include several botnet controllers, so multiple one-to-many relations typically overlap in a botnet.
3. *Geographic location correlation*—It is helpful to match up the bots and controllers to a geographic location using the associated IP address. This does not provide pinpoint accuracy, but it offers a general sense of where the bots and controllers are located.
4. *Vigilant activity watch*—The security analysis should include close, vigilant watch of activity from the

bots and servers. The most important activity to be identified would be a distributed attack from the bots to some target.

The steps in the box above allow for the construction of a logical map of a botnet, showing the geographic locations of the bots, their associated service provider (usually a local broadband carrier), the set of servers used as botnet controllers, and a general depiction of any relevant activity. Typical activity found in a botnet includes recruitment of new bots, as well as attacks from the bots toward some designated target (see Figure 9.9)

Botnets can have a far-reaching geographic distribution.



Figure 9.9 Correlative depiction of a typical botnet.

The botnet diagram demonstrates some of the conclusions that can be drawn immediately from such an analysis. The typical pattern of bot clumping that one finds in a botnet might give hints as to the type of social engineering or lure used to drop malware onto the target PCs. Useful hints might also be gathered from regions where the botnet seems to have gathered no bots. One area where correlative analysis is often not useful is trying to determine correlations between the geographic locations of botnet controllers. This generally results in no useful information, as botnet controllers tend to be scattered across the globe, driven by opportunistic hacking.

It goes without saying that national infrastructure protection requires the real-time capability to monitor botnet configuration and activity. The risk of botnets has grown so much in recent years partly because they have been able to exist under the radar of most government and commercial organizations. The first step in reducing this risk involves the creation of a national capability to collect information about botnets and to advise the participants on how best to avoid being either duped into hacking someone else or directly targeted for an attack.

Disseminating information about botnet tactics may help consumers avoid future lures.

Large-Scale Correlation Process

For national infrastructure protection, large-scale correlation of all-source data by organizations with a broad vantage point is complicated by several technical, operational, and business factors, including the following:

- *Data formats*—Individual national asset environments will most likely collect data in incompatible formats due to a lack of standards in security data collection tools. As a result, almost all security-relevant data is collected in a proprietary or locally defined format. This represents a significant challenge for any large-scale collection from multiple sources.
- *Collection targets*—Individual asset environments will likely collect data from different types of events and triggers. Some, for example, might collect detailed information about networks and only limited information from systems, whereas others might do the opposite. This obviously complicates the comparison of aggregated data from multiple sources.
- *Competition*—Various commercial groups collecting relevant data might be in direct business competition. (Most government groups will admit to their share of mutual competition as well.) This competitive profile implies that any aggregated information and any interpretation that would result from correlative analysis must be carefully protected and associated with suitable anonymity.

To deal with these challenges on a large scale, a deliberate correlation process must be employed. The process must break down each component of the correlation task into discrete entities with well-defined inputs and outputs. This process is best viewed in aggregate as consisting of five different passes leading from collected data to actionable information (see Figure 9.10).

Large-scale data correlation initiatives must overcome challenges posed by competition, incompatible data formats, and differing collection targets.

Five Passes Leading to Actionable Information

1. The first pass in this process schema involves resolution of all incompatible data formats from the different sources. In addition to the data generated by familiar security devices, these inputs can also include human-generated data that could be obtained through telephony or even social processes. The resolution must be automated via filters that produce a common output. Amazingly, very little work has been done in the computer security community to standardize relevant formats.
2. The second pass in the schema involves a leveling of the various types of data collected. The most common task in this pass is to categorize similar data into the appropriate set of categories. This must be done because different organizations routinely refer to the same security-relevant events by different names. Commercial tools also tend to refer to the same attacks by different names and alarm types. Large-scale correlation thus requires a common understanding of the semantics associated with activity of interest. Small-scale analysis methodologies using a common threat management tool from one vendor can skip this step; large-scale analysis from multiple, diverse sources cannot.
3. The third pass involves the actual comparison of collected data to relevant attributes. Computer

security experts often refer to this pass itself as correlation. This pass is where security algorithms for profile, signature, domain, and time-based correlation are incorporated into the analysis. It typically involves a combination of automated processing using tools, with the interpretation of human experts. In the best case, this pass in the process occurs rapidly, almost in real time, but the reality is that the analysis step can take considerable time in the most complex scenarios. This pass, along with the first two passes, can be viewed collectively as the *correlation engine*.

4. The fourth pass involves storage and protection of the output. This is likely to include interpretation of the data once it has been aggregated and compared. Insights are often evident at this stage of the process, and these can be represented as either deliberately stored information in a database or simply as information known to security analysts involved in the overall process. In either case, the information must be protected. For large-scale applications, the size of the information collected can be massive, which implies that special database technology with the ability to scale might be required.
5. The fifth and last pass in the process involves filtering and dissemination of the information. This might result in a feedback loop where output recommendations become input to a new series of five correlation passes. Alternatively, it can be used by appropriate parties for immediate action such as real-time incident response. This pass, along with the storage pass, can be viewed collectively as the *correlation back end*.

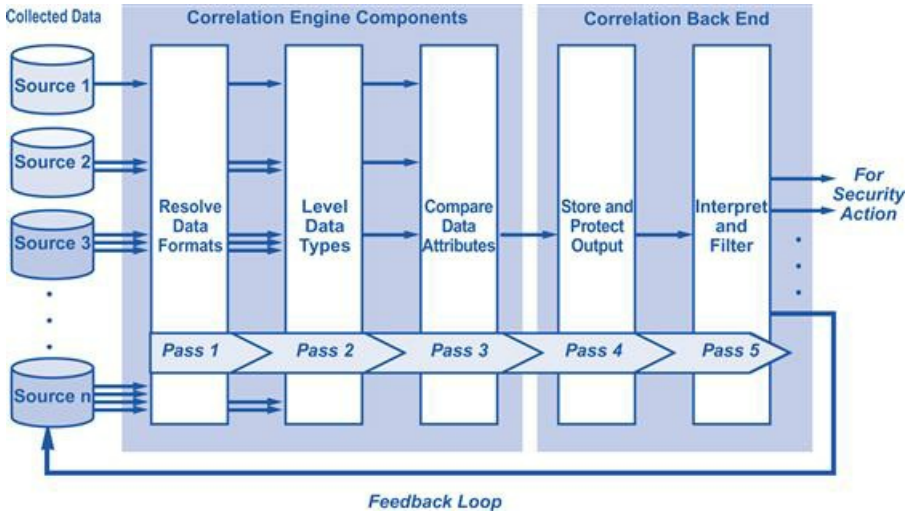


Figure 9.10 Large-scale, multipass correlation process with feedback.

National Correlation Program

Implementation of a national correlation program is likely to follow two specific directions. First, steps might be taken to encourage individual organizations with national infrastructure responsibility to create and follow a local program of data correlation. This can be done by embedding correlation requirements into standard audit and certification standards, as well as within any program solicitations for government-related infrastructure work. The likelihood of success of this approach is high and is thus recommended for immediate adoption at the national policy level.

Data collection can be encouraged by making it a requirement of contracted government-related projects.

Second, national-level programs might be created to try to correlate collected data at the highest level from all available sources. This approach is much more challenging and requires addressing the following technical and operational issues:

- *Transparent operations*—The analysis approach used for correlation should be fully known to all participants. Thus, whether profiles, signatures, or the like are used, the process should be clearly explained and demonstrated. This will allow participants to help improve such aspects of the process as data feed provision, data reduction algorithms, and back-end interpretation.
- *Guaranteed data feeds*—Any participant providing data to the correlation process must be held to a guaranteed service level. Obviously, this level can change but only under controlled conditions that can be factored into the analysis. Without such guarantees, correlation algorithms will not work.
- *Clearly defined value proposition*—Participants should recognize a clearly defined value proposition for their provision of data. The worst situation involves a "black hole" collection process where the output recommendations from the correlation activity are not generally shared.
- *Focus on situational awareness*—The output of the process should certainly be action oriented but should also recognize the limitations inherent in broad correlation. It is unlikely that any national-level correlation function will be able to give a real-time silver bullet to any participant. More likely, the output will provide situational awareness that will help in the interpretation or response to an event.

By addressing these issues, the technical and operational feasibility of a successful, national-level correlation function increases dramatically. Unfortunately, many legal, social, and political issues—considered outside the general scope of this book—will complicate the creation of such a function.

Finally, let's briefly look at some practical ways to correlate data in the critical national infrastructure through a real-time event and multisource correlation. Correlation is based on the application of threshold and scenario-based rules against multisource and real-time event streams. It can easily be distributed to support scalable parsing processes for large deployments and has virtually no limit on event rate or volume. While real-time correlation performs dynamic parsing, normalization, filtering, analysis, and alerting, a separate data fork of the same unparsed event logs and subsequent alerts is sent to a long-term data repository in a tamper-resistant, raw format.

Correlating data from multiple log sources and assessing multiple events using a set of universal

attack/event sequences is what is known as multisource correlation. Multisource correlation provides for greatly improved incident coverage and alert accuracy. Critical infrastructure organizations can use combinations of preincident reconnaissance activity, postincident activity, and thresholds of events in order to describe scenarios which indicate a serious risk, attack, and/or successful compromise of systems or applications. As such, there is not a 1:1 ratio of alert coverage to rule. One rule can cover dozens of threats, resulting in less time managing and creating dozens of repetitive rules.

Correlation Rules for Critical National Infrastructure Cyber Security

To be effective in critical national infrastructure network defense, and not just for forensic analysis, the network and cyber security event data must also be analyzed and correlated in real time. This information needs to be manageable and actionable as well. Forensics are not enough. Detecting and stopping today's zero-day, multivector, and blended cyber threats requires real-time, in-memory analytics that can capture, correlate, and respond to network attacks and insider abuse at network speed. There are numerous obstacles to performing this task efficiently, securely, and with minimal personnel resources.

The first significant obstacle to real-time event correlation is the fact is that none of the core defense technologies deployed in the classic defense in depth and best-of-breed model are designed to communicate with each other. They are simply point solutions and represent silos of information. The data from these disparate systems must be aggregated and normalized to a common taxonomy—effectively, a universal translator is required to map the French, German, Russian, and Chinese of the various technologies into English.

Another major obstacle to real-time event correlation is the construction of the correlation rules. Few organizations think in terms of correlation rules, but they are certainly familiar with network policies and they can describe business rules and objectives. The challenge is to find a way to bridge their knowledge and objectives with the construction of correlation rules, without requiring IT personnel to become system programmers. It is also critical that there be a mechanism to build the correlation rules quickly because the need for targeted monitoring or network assessment can change quite rapidly.

Traditional event-modeling techniques make it tedious and time consuming to build multiple event correlation systems. To minimize complexity, these systems often place arbitrary limits on the number and type of data elements or fields that can be used in the correlation rules, and rigidly enforce linear or static evaluation paths. In addition to the ease with which new rules can be created, organizations should adhere to prebuilt correlation rules that cover the critical network infrastructure and change management and network security functions (see "An Agenda for Action for Maximizing the Benefits of Correlation Rules").

An Agenda for Action for Maximizing the Benefits of Correlation Rules

Correlation enables system users to take the audit data analysis to the next level. Rule-based and statistical correlation allows the user to (check all tasks completed):

1. Dramatically decrease the response times for routine cyber attacks and incidents by using the centralized and correlated evidence storage.
2. Completely automate the response to certain cyber threats that can be detected reliably by correlation rules.
3. Identify malicious and suspicious activities on the network even without having any preexisting knowledge of what to look for.
4. Increase awareness of the network via baselining and trending and effectively "take back your network."
5. Fuse data from various information sources to gain cross-device business risk view of the organization.

6. Use the statistical correlation to learn the threats and then deploy new rules for site-specific and newly discovered violations. Overall, combining rules and algorithms provides the best value for managing an organization's cyber security risks.
7. Uniquely identify steps or vectors of the critical cyber attack scenarios, such as traversal from one network segment to another.
8. Enforce common process control networks policies. Since the process control networks are typically quite static compared with business networks, violations that can be alerted upon include rogue systems, configuration changes, and port scans.
9. Implement a data dictionary for process control-specific cyber security events, mapping proprietary logged process control system events to standardized cyber security events.

Summary

This chapter focused on four distinct analytic methods that are available for correlating cyber security information: *profile-based*, *signature-based*, *domain-based*, and *time-based correlation*. Profile-based correlation involves comparison of a normal profile of target activity with observed patterns of activity. Signature-based correlation involves comparing a signature pattern of some known malicious condition to observed activity. Domain-based correlation involves comparing data from one domain with data collected in an entirely different context. Time-based correlation compares data collected during one time period with data collected at a different time.

Furthermore, the chapter also covered how organizations can actively defend the critical national infrastructure network through highly targeted correlation rules, behavior analysis, and integration with network infrastructure. The defensive arsenal includes the ability to quarantine, block, route, and control services, processes, accounts, privileges, and more. Real-time analysis, event correlation, and active response are the basis for next-generation technology that provides organizations with visibility into their networks and a defense against insider abuse and cyber attacks.

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

Chapter Review Questions/Exercises

True/False

1. True or False? Computer and network security experts understand that correlation is one of the most powerful analytic methods available for threat investigation.
2. True or False? The current state of the practice in day-to-day network security correlation in existing national infrastructure settings is based on a technique known as *threat management*.
3. True or False? To create a proper security correlation process for national infrastructure protection in an environment of large, cross-organizational size, scope, and scale, several technical and operational factors must be considered.
4. True or False? Network service providers have a particularly good vantage point for correlating data across multiple companies, agencies, groups, individuals, and regions.
5. True or False? The most insidious type of attack one finds today in any large-scale, distributed, Internet-connected network environment is the botnet.

Multiple Choice

1. From a foundational perspective, four distinct analytic methods are available for correlating cyber security information, except which one:

A. Profile-based correlation
B. Attack-based correlation
C. Signature-based correlation
D. Domain-based correlation
E. Time-based correlation

2. Firewalls that generate audit records when certain types of security-relevant events occur are called:

A. Firewall audit trails
B. Firewall collection
C. Firewall information
D. Firewall data
E. Firewall coverage

3. One of the steps that is involved in the detection of a botnet, via a correlative analysis by a network carrier, is called:

A. One-to-many communication
B. Geographic location
C. Vigilant activity watch
D. Long-term online factors
E. Broad data collection

4. For national infrastructure protection, large-scale correlation of all-source data by organizations with a broad vantage point is complicated by several technical, operational, and business factors, including the following, except which two:

A. Data formats
B. Collection targets
C. Competition
D. Storage considerations
E. Data reduction

5. National-level programs might be created to try to correlate collected data at the highest level from all available sources. This approach is much more challenging and requires addressing the following technical and operational issues, except which one:

A. Transparent operations
B. Guaranteed data feeds
C. Functional view
D. Clearly defined value proposition
E. Focus on situational awareness

Exercise

Problem

At a banking customer's site, log-on failures were being generated from a branch workstation after 10 p.m. The bank's IT security staff correlated the log-on failures, the source IP address, the rapid succession of events, and the activity that was occurring outside of business hours. Please explain what the IT security staff was able to do to solve this problem.

Hands-On Projects

Project

In this case study, hacker probes were followed by a cyber attack. The correlation rule watches for the general cyber attack pattern consisting of a reconnaissance activity followed by the exploit attempt. Cyber attackers often use activities such as port scanning, application querying to scope the environment, finding targets for exploitation, and getting an initial picture of system vulnerabilities. After the initial information gathering is performed, the cyber attacker returns with exploit code or automated attack tools to get to the actual system penetration. The correlation enriches the information reported by the intrusion detection systems and serves to validate the cyber attack and suppress false alarms. Please explain how the cyber security administrator was able to solve this problem.

Case Projects

Problem

This case study has to do with login guessing. The correlation rule watches for multiple attempts of failed authentication to network and host services followed by a successful login attempt. Please explain how you would go about solving this problem.

Optional Team Case Project

Problem

In this case project, a simulation was created at a test-bed site in order to counter potential threats to the oil and gas industry, based on hypothetical cyber attack scenarios. One cyber attack scenario highlighted the increased risk that control systems are exposed to as they get connected to Internet-enabled business networks. It showed how an outside intruder can hack into the business network and then, once inside, gain access to other networks, like a SCADA system, and actually tamper with a piece of equipment in the field. Please explain how you would go about solving this case project.

[1] D. Denning, *Information Warfare and Security*, Addison-Wesley, New York, 1999, p. 362.

10

Awareness

Chapter Outline

Detecting Infrastructure Attacks
Managing Vulnerability Information
Cyber Security Intelligence Reports
Risk Management Process
Security Operations Centers
National Awareness Program
Connecting Current Cyber Security Operation Centers to Enhance Situational Awareness
Summary
Chapter Review Questions/Exercises

Intelligence, the information and knowledge about an adversary obtained through observation, investigation, analysis, or understanding, is the product that provides battlespace awareness.

Edward Waltz[1]

Situational awareness refers to the collective real-time understanding within an organization of its security risk posture. Security risk measures the likelihood that an attack might produce significant consequences to some set of locally valued assets. A major challenge is that the factors affecting security risk are often not locally controlled and are often deliberately obscured by an adversary. To optimize situation awareness, considerable time, effort, and even creativity must be expended. Sadly, most existing companies and agencies with responsibility for national infrastructure have little or no discipline in this area. This is surprising, as a common question asked by senior leadership is whether the organization is experiencing a security risk or is "under attack" at a given time.

Awareness of security posture requires consideration of several technical, operational, business, and external or global factors. These include the following:

- *Known vulnerabilities*—Detailed knowledge of relevant vulnerabilities from vendors, service providers, government, academia, and the hacking community is essential to effective situational awareness. Specific events such as prominent hacking conferences are often a rich source of new vulnerability data.

Consider attending a hacking conference to learn more about potential vulnerabilities.

- *Security infrastructure*—Understanding the state of all active security components in the local

environment is required for proper situational awareness. This includes knowledge of security software versions for integrity management and anti-malware processing, signature deployments for security devices such as intrusion detection systems, and monitoring status for any types of security collection and processing systems.

- *Network and computing architecture*—Knowledge of network and computing architecture is also important to understanding an organization's situational security posture. An accurate catalog of all inbound and outbound services through external gateways is particularly important during an incident that might be exploiting specific ports or protocols.
- *Business environment*—Security posture is directly related to business activities such as new product launches, new project initiation, public relations press releases, executive action involving anything even mildly controversial, and especially any business failures. Any types of contract negotiations between management and employee bases have a direct impact on the local situational security status.
- *Global threats*—Any political or global threats that might be present at a given time will certainly have an impact on an organization's situational security posture. This must be monitored carefully in regions where an organization might have created a partnership or outsourcing arrangement. Because outsourcing tends to occur in regions that are remote to the organization, a global threat posture has become more significant.

The increase in global outsourcing requires awareness of how international political events may impact your vendors.

- *Hardware and software profiles*—An accurate view of all hardware and software currently in place in the organization is also essential to situational awareness. A common problem involves running some product version that is too old to properly secure through a program of patching or security enhancement. A corresponding problem involves systems that are too new to properly characterize their robustness against attack. In practice, an optimal period of product operation emerges between the earliest installation period, when a product or system is brand new, and the latter stages of deployment, when formal support from a vendor might have lapsed (see Figure 10.1).

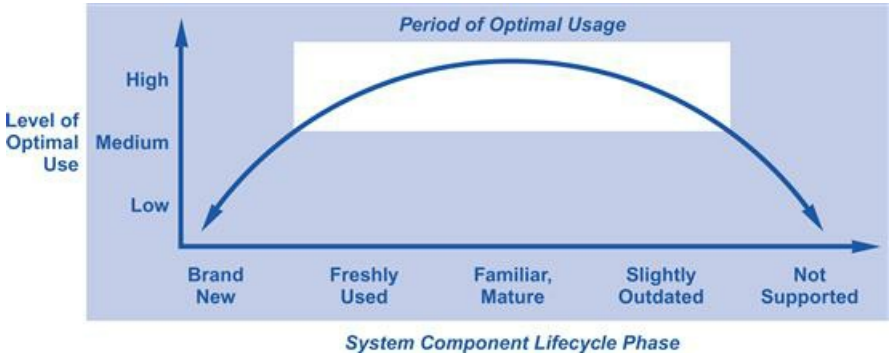


Figure 10.1 Optimal period of system usage for cyber security.

Each of these factors presents a set of unique challenges for security teams. An emerging global conflict, for example, will probably have nothing to do with the vulnerability profile of software running locally in an enterprise. There are, however, clear dependencies that arise between factors in practice and will improve situational awareness. For example, when vulnerabilities are reported by a hacking group, the organization's security posture will depend on its local hardware, software, and security infrastructure profile. As a result, it is generally reasonable for an organization to combine the value of all situational status factors into one generic measure of its security posture. This measure should be able to provide a rough estimate of the broad, organizational security risk at a given time. It should then weigh the likelihood and potential consequences of serious attack against the normal, everyday level of risk that an organization lives with every day. Presumably, risk on a day-to-day basis should be lower than during a serious incident, so it stands to reason that a rough metric could capture this status, perhaps as a high, medium, and low risk characterization (see Figure 10.2).

Factoring in all elements of situational awareness and any related challenges should create an overview of an organization's current security risk.

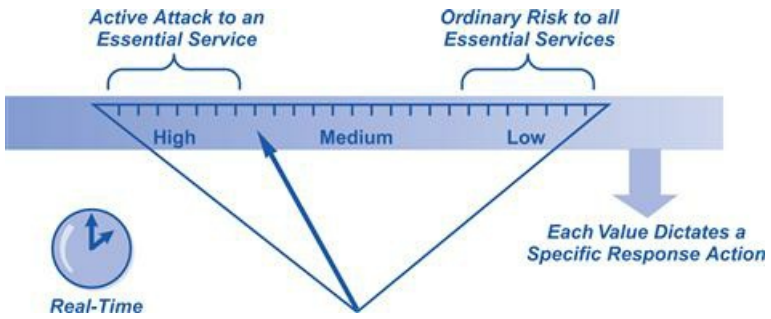


Figure 10.2 Rough dashboard estimate of cyber security posture.

Descriptors such as high, medium, and low to describe security risk are too vague to be helpful.

Unfortunately, the public perception of categorizing high, medium, and low security risks is that it does not provide useful information. This is certainly true for such measures as the public threat metric, which was used previously by the U.S. Department of Homeland Security to characterize risk. The problem with this metric was that it dictated no concrete actions to be taken by citizens. If risk was characterized as low, citizens were warned to remain vigilant and on guard; if risk was characterized as medium or even high, the advice was essentially the same. Citizens were told to go on with their normal lives, but to be *somehow* more careful. Obviously, this type of advice causes confusion and is to be avoided in national infrastructure protection.

The only way a posture metric can be useful is if it is driven by real-time events and is connected directly to an explicit incident response program. When this is done, an ongoing rhythm develops where the situational status helps direct security management activity. This could involve some serious flaw being detected in an organization (which would drive the threat level upward), followed by detection of a real exploit in the wild (which would drive the threat level further upward), followed by a patch activity that fixes the

problem (which would drive the threat level back down) (see Figure 10.3).

Security risk levels should be set to correlate with actionable items.

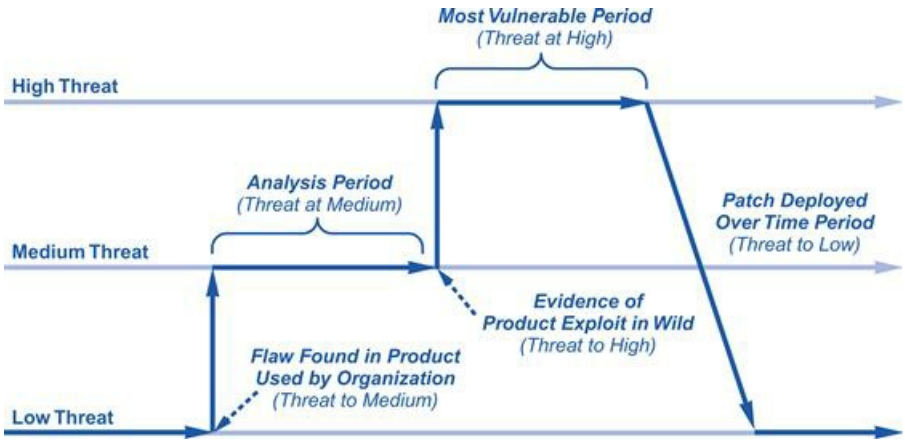


Figure 10.3 Security posture changes based on activity and response.

Regardless of public perception with respect to previous government threat metrics, any program of situational awareness for cyber security must include a broad characterization of real-time risk. The attributes of this broad characterization will be based on a much more detailed understanding of the real-time posture. Collectively, this posture is referred to as situational awareness and is based on an understanding of whether or not the infrastructure is under attack, which vulnerabilities are relevant to the local infrastructure, what sort of intelligence is available, the output of a risk management process, and information being generated by a real-time security operations center. These elements are described in the sections that follow.

Detecting Infrastructure Attacks

The process of determining whether an attack on national infrastructure is under way is much more difficult than it sounds. On the surface, one would expect that, by observing key indicators, making the determination that an attack has begun or is ongoing would seem straightforward. Correlating observed activity with profiles, signatures, and the like can provide a strong algorithmic basis, and products such as intrusion detection systems offer a means for implementation. These factors are misleading, however, and the truth is that no security task is more difficult and complex than the detection of an ongoing attack, especially if the adversary is skilled.

There are many tools for detecting attacks, yet no single tool is comprehensive a or foolproof.

To illustrate this challenge, suppose you notice that an important server is running in a somewhat sluggish manner, but you cannot diagnose the problem or explain why it is occurring. Obviously, this is suspicious and could be an indicator that your server has been attacked, but you cannot state this with any certainty. There could be a million reasons why a server is running slowly, and the vast majority of them have nothing to do with security. Suppose, however, that you discover a recently installed directory on the server that is filled with unfamiliar, strange-looking files. This will clearly raise your suspicion higher, but there are still numerous explanations that do not signal an attack. Perhaps, finally, someone in the enterprise steps forward and admits to running some sort of benign test on the server, thus explaining all of the errant conditions. The point is that confidence that a target is under attack will rise and fall, depending on the specifics of what is being observed. Obviously, there is a threshold at which the confidence level is sufficiently high in either direction to make a sound determination. In many practical cases, analysis never leads to such a confidence threshold, especially in complex national infrastructure environments (see Figure 10.4).

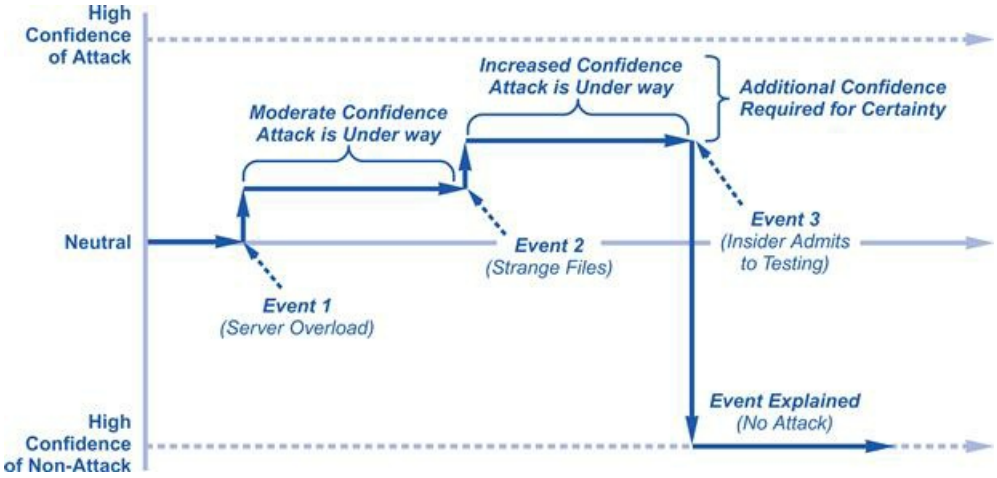


Figure 10.4 Attack confidence changes based on events.

In our example, you eventually became confident that no attack was under way, but many scenarios are not terminated so cleanly. Instead, events expose a continuing stream of ongoing information that can have a

positive, negative, or neutral effect on determining what is actually going on. In many cases, information that is incorrect or improperly interpreted has the effect of confusing the process. Relatively new technologies, such as mobile wireless services, tend to exhibit this property, especially in cases where a particular incident has never been seen before. The primary disadvantage of never determining the root cause of an attack is that the security posture cannot be accurately measured. This is especially troublesome when the attack is severe and targets essential national infrastructure services.

Determination of security risk level is a fluid process; it changes as new information is revealed or as situations change.

T.me/library_Sec

Managing Vulnerability Information

A common cynical view of computer security is that its experts are engaged in nothing more than a game of trivial pursuit around attack and vulnerability information. Support for this view is evident in the security books published to date, most of which contain page after page of esoteric attack specifics that are often long-since irrelevant. It is also evident in social circles at security and hacking conferences, where the discussion rarely addresses foundational topics of software engineering or system design but instead focuses on such trivia as which systems have which bugs in which versions on which hardware. Some security experts and hackers have become walking encyclopedias of such knowledge, even viewing information as the driver of power and skill. Anyone not possessing sufficiently detailed knowledge is thus tagged a newbie, lamer, or perhaps worse —a *manager*.

In spite of this odd phenomenon, situational awareness for national infrastructure protection does require a degree of attentiveness to daily trivia around vulnerability information. We refer to the information as trivia simply because, once addressed and fixed, the value of the information drops very close to zero. Nevertheless, it is important information to collect, and most national infrastructure teams use the default approach of *active opportunism*, where a set amount of effort is expended to gather as much data as possible and anything else that comes in is welcomed. The problem with active opportunism is that it will never be complete and cannot be depended upon for accurate management decisions. For example, the question of whether a given vulnerability has been coded into an exploit and made available on the Internet can be researched by one, two, or 50 people. If no evidence of such an exploit is found, then the weak conclusion can be drawn that it does not exist. Obviously, information about the vulnerability could be tucked away in some IRC discussion or on an obscure hacking site, but unless it is found or volunteered the security team will never know for sure.

Collecting daily trivia around vulnerability information should not be dismissed as unimportant but should be considered one of many methods of achieving situational awareness.

The best one can hope for is to create as active and complete a vulnerability information-gathering process as possible. See the box for practical heuristics that have been useful for infrastructure protection in the past.

Practical Heuristics for Managing Vulnerability Information

- *Structured collection*—The root of all vulnerability management processes must be some sort of structured collection approach with means for assuring proper delivery of information, validating the source, cataloguing the information in a suitable taxonomy, and maintaining a useful database for real-time reference with provision for indexing and crawling vulnerability data in real-time. This structured approach should be integrated into all day-to-day cyber security activities so that accurate vulnerability information is available across the entire security infrastructure and team. Filters should exist to assure incoming data, as well as to ensure that external entities only obtain appropriate information (see Figure 10.5).
- *Worst case assumptions*—Many situations arise where a security team cannot determine whether some

important piece of vulnerability-related information has actually been disclosed or has become known to an adversary group. The most mature and healthy approach in such scenarios is to assume the worst possible case. Most experts would agree that if the possibility arises that some vulnerability *might* be known externally, then it probably *is* known.

- *Nondefinitive conclusions*—Making definitive statements about national infrastructure security is not recommended. Too many cases exist where a security team draws the confident conclusion that a system is secure only to later obtain vulnerability-related information to the contrary. Experienced managers understand, for example, that they should always include caveats in security posture reports given to senior leaders in government or industry.
- *Connection to all sources*—Managing vulnerability information should include connections to all possible sources such as industry groups, vulnerability-reporting services, hacking conferences, internal employee reports, and customer data. Sometimes the most critical piece of vulnerability information comes from the most unlikely source.

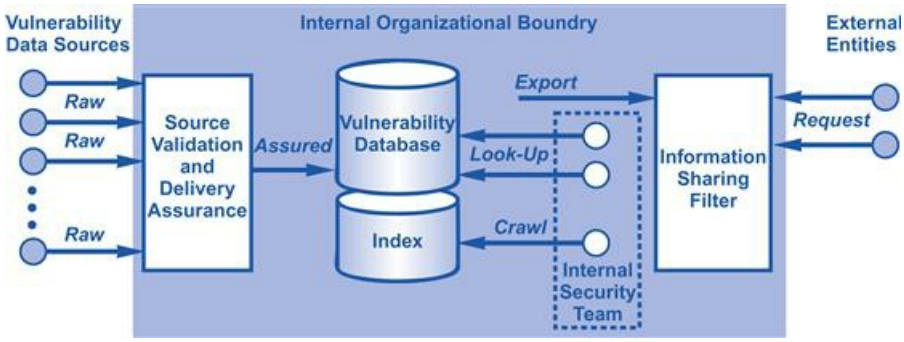


Figure 10.5 Vulnerability management structure.

Following the heuristics listed in the box above will help to ensure that the best available data is collected, stored, and used, but these heuristics can never provide assurance that the vulnerability management process is perfect. Instead, managers are strongly advised to follow three basic rules: (1) always assume that the adversary knows as much or *more* about your infrastructure than you do, (2) assume that the adversary is always keeping vulnerability-related secrets from you, and (3) never assume that you know everything relevant to the security of your infrastructure. Such complete knowledge is unattainable in large, complex national infrastructure settings.

Cyber Security Intelligence Reports

A technique commonly used in government intelligence community environments, but almost never in most enterprise settings, involves the creation and use of a regularly published (usually daily) intelligence report. For cyber security, such a report generally includes security-related metrics, indicators, attack-related information, root-cause analysis, and so on for a designated period. It is typically provided to senior management, as well as all decision-makers on the security and infrastructure teams. The report should also be indexed for searches on current and previous information, although this is not a common practice.

Daily cyber security intelligence reports that are standard in government agencies would be equally useful in enterprise settings.

Although the frequency and content of intelligence reports should be tailored to the needs of the local environment, some types of information that one would expect in any daily intelligence report include the following:

- *Current security posture*—The situational status of the current security risk would be required in any intelligence report, especially one issued over a daily or weekly interval (monthly intervals create too long a gap for information to be considered "intelligence").
- *Top and new security risks*—Characterization of the top risks, as well as any new risks, is also important to include in an intelligence report. Visualization and other techniques are often helpful to highlight changes in risk posture.
- *Automated metrics*—Security systems that generate metrics should provide input to the intelligence report, but care must be taken to avoid the creation of a voluminous document that no one will read. Also, raw output from some devices is indiscernible and should be either summarized or avoided in the report.
- *Human interpretation*—Ultimately, the most useful cyber security intelligence includes analysis by experienced and expert human beings who can interpret available security data and recommend suitable action plans. It is unlikely that this interpretation function will be automated in the near future.

Human interpretation is bound to catch vulnerabilities that automated algorithms will miss.

The activity associated with the realization of a cyber security intelligence report can be viewed as an ongoing and iterative process made up of three tasks (see box).

Tasks for Creating a Cyber Security Intelligence Report

1. The first task involves *intelligence gathering* of available vulnerability and security posture data. This can be automated but should allow for manual submission from people who might have useful information to share. Many organizations do this gathering in the early morning hours, before the bulk of the business activity begins (a luxury that does not exist for global companies).

2. The second task involves *interpretation* and *publication* of the gathered data, not unlike similar processes in daily news publications. The interpretation should focus on the audience, never assuming too much or too little knowledge on the part of the reader. It is during this task that the human interpretive summary of the collected data is written.
3. The third task involves protected *dissemination* and *archiving* of the report for use by end users with a need to know. Report transmission is generally protected by encryption, and report archives and storage are protected by access controls (see Figure 10.6).

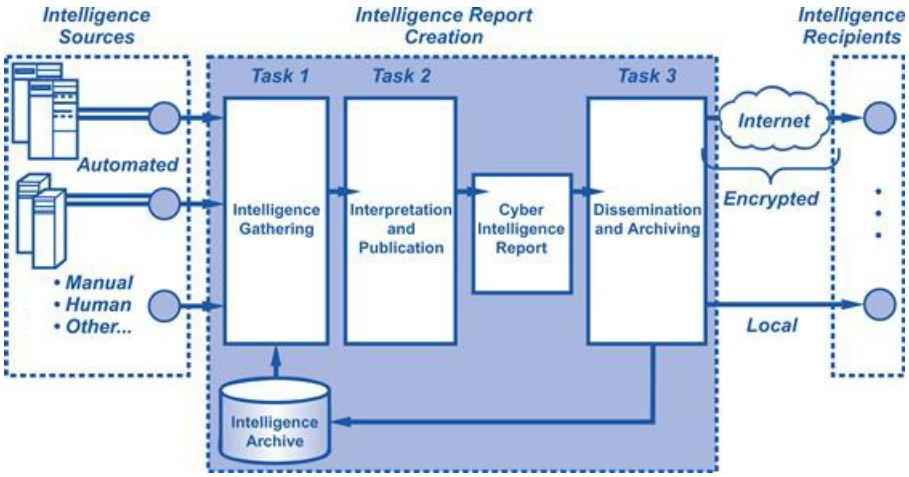


Figure 10.6 Cyber security intelligence report creation and dissemination.

One byproduct of creating an intelligence report is that it helps guide the local culture toward greater attentiveness to real-time security considerations. Everyone knows that, during an incident, response activity summaries will find their way to senior managers which tends to heighten concentration on the accuracy and completeness of the report. In addition, when an incident occurs that does not find its way into the report, managers can justifiably question the completeness of reporting around the incident.

Risk Management Process

Managers of essential national services must understand the security risks associated with their underlying infrastructure. Although this can be done using all sorts of fancy risk taxonomies, tools, and methodologies, the recommended approach is to simply maintain a prioritized list. Depending on the severity of the risks in the list, managers can decide to focus on a subset of the top ones, perhaps the top 10 or 20. Funding and resource allocation decisions for cyber security can then be driven by the security risk profile of the organization, keeping in mind that the list of risks will change with any adjustments in threat environment, technology deployment, or reported vulnerabilities.

Security risks must be tracked (listed) and prioritized to drive appropriate funding and resource allocation.

The generally agreed-upon approach to measuring the security risk associated with a specific component begins with two estimations:

- *Likelihood*—This is an estimate of the chances an attack might be successfully carried out against the specific component of interest.
- *Consequences*—This is an estimate of how serious the result might be if an attack were carried out successfully.

These two estimates must be performed in the context of an agreed-upon numeric range. The actual values in the range matter less than the relative values as the estimates increase and decrease. The simplest and most common values used are 1, 2, and 3, corresponding to low, medium, and high for both estimates. Once the likelihood and consequences have been estimated, risk is obtained by multiplying the values. Thus, if some component has a high likelihood of attack (value 3) and medium consequences resulting from an attack (value 2), then the associated risk is 3 times 2, or 6. If security measures are put in place to reduce the likelihood of an attack to medium (value 2), then the risk is now 2 times 2, or 4. Again, the absolute value of risk is less important than the relative value based on security decisions that might be made.

The actual numeric value of a security risk is less important than its overall relative risk.

A useful construct for analyzing security decisions in infrastructures compares relative security risk against the costs associated with the recommended action. The construct allows managers to consider decision paths that might increase, decrease, or leave unaffected the security risk, with the balancing consideration of increased, decreased, or unaffected associated costs (see Figure 10.7).

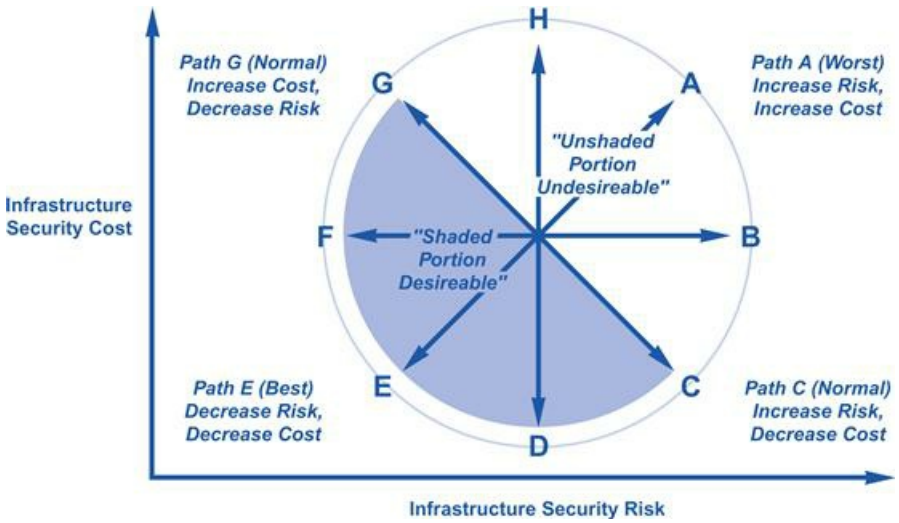


Figure 10.7 Risk versus cost decision path structure.

To interpret the choices in the decision path structure, start at the middle of the diagram and consider the effects of each path labeled A through H. The path labeled G shows a security decision that increases costs in order to reduce risk. This is a normal management decision that is generally considered defensible as long as sufficient budget is available. Similarly, the path labeled C is also normal, as it accepts increased risk in order to reduce costs, which is unfortunately a common enough decision.

Interestingly, any decision path in the area shaded on the figure will be generally acceptable in most cases because the relationship between cost and risk is reasonable. The decision paths in the unshaded portion of the graph, however, are generally considered unacceptable because of the odd balance between the two factors. Decision path H, for example, increases costs with no impact on security risk. This case corresponds to the situation encountered all too often where a security safeguard is put in place that actually has zero impact on the risk profile.

Increasing risks likely incur increased costs; assessing relative risk will help determine the value of investing in risk reduction.

To summarize, all decisions about national infrastructure protection should be made in the context of two explicit management considerations: (1) maintaining a prioritized list of security risks to the system of interest, and (2) justifying all decisions as corresponding to paths in the shaded portion of the decision path structure shown in Figure 10.7. If these two simple considerations were mandatory, considerable time, effort, and money would be immediately saved for many infrastructure management teams.

Security Operations Centers

The most tangible and visible realization of real-time security situational awareness is the *security operations center* (SOC), also referred to as a *fusion center*. The most basic model of SOC operations involves multiple data, information, and intelligence inputs being fed into a repository used by human analysts for the purpose of operations such as interpretation, correlation, display, storage, archival, and decision-making. The SOC repository is constructed by active solicitation or passive acceptance of input information, and information processing combines human analysis with automated processing and visual display (see Figure 10.8).

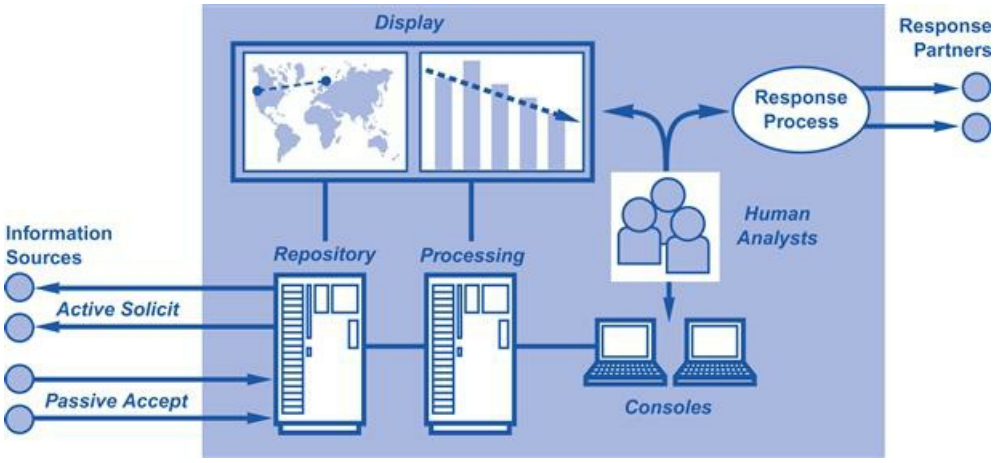


Figure 10.8 Security operations center (SOC) high-level design.

Most SOC designs begin with a traditional centralized model where the *facility* is tied closely to the operations of the center. That is, methods and procedures are created that presume SOC resources, including all personnel, are located in one place with no need for remote coordination. All data is stored in a local repository that can be physically protected in one location. This approach has its advantages, because it removes so many coordination-related variables from the management equation. That said, an SOC can be created from distributed resources in geographically dispersed locations. Repositories can be distributed, and analysis can be performed using remote coordination tools. Generally speaking, this approach requires more work, but the main benefit is that more expert analysts can be recruited to such an approach, especially if the requirement is that 24/7 operations be supported. Experts can be hired across the globe in a "follow-the-sun" support arrangement.

The advantage to global dispersal of SOC resources is an around-the-clock real-time analysis of security threats.

Typical operational functions supported in an SOC include all human interpretation of data by experts, management of specific incidents as they arise, support for 24/7 contact services in case individuals have security-relevant information to share, and processing of any alarms or tickets connected to a threat management or intrusion detection system. The 24/7 aspect of SOC operation is particularly useful to

national-level situational awareness, because key infrastructure protection managers will know that they can obtain a security posture status at any time from a human being on call in the SOC. Government procurement efforts for national services should include requirements for this type of coverage in the SOC.

National Awareness Program

The goal of supporting a national-level view of security posture should not be controversial to most security and infrastructure managers. Everyone will agree that such a view is necessary and useful for supporting national infrastructure protection-related management decisions. The challenge, however, lies with the following important practical considerations:

- *Commercial versus government information*—To achieve full situational awareness at the national level will require considerable support from both commercial and government entities. Groups supplying security status information must be provided with incentives and motivations for such action. Patriotic justification helps, but global companies must be more deliberate in their sharing of information with any government.
- *Information classification*—When information becomes classified, obviously the associated handling requirements will increase. This can cause problems for data fusion. In fact, the essence of data compartmentalization for classified information is to prevent and avoid any type of fusion, especially with unclassified data. The result is that situational awareness at the national level will probably include two views: one unclassified and public, the other based on more sensitive views of classified information.
- *Agency politics*—Government agencies are famous for using *information* as a basis for political agendas, including support for project funding, hiring plans, and facility expansion. This tendency is counter to the goal of information sharing for situation awareness and must therefore be managed carefully.
- *SOC responsibility*—If a national SOC is to be realized, then some organization must be designated to run it. The decision as to whether this should be a defense- or civilian-related initiative is beyond the scope of this book, but most security experts agree that current defense-related awareness initiatives provide many of the elements required in a fully functioning SOC.

If these challenges are not addressed properly, the risk is that inaccurate views of situational awareness could arise. If an agency, for example, finds out about a vulnerability but decides to *not* share this information, then a hole emerges in any national-level risk estimation. Similarly, if a commercial organization is unable to receive and process classified information, then their view of current security risk posture will not be accurate. Attentiveness to managing these issues on a case-by-case basis, perhaps as part of a national SOC, would seem the best approach.

Finally, let's briefly look at some practical ways to connect current cyber operation centers to enhance situational awareness. There is a pressing need to ensure that government information cyber security offices and strategic operations centers share data regarding malicious activities against federal systems. This would be consistent with privacy protections for personally identifiable and other protected information as legally appropriate in order to have a better understanding of the entire threat to government systems and take maximum advantage of each organization's unique capabilities to produce the best overall national cyber defense possible. This initiative provides the key means necessary to enable and support shared situational awareness and collaboration across centers that are responsible for carrying out U.S. cyber activities. This effort focuses on key aspects necessary to enable practical mission bridging across the elements of U.S. cyber

activities: foundational capabilities; investments such as upgraded infrastructure; increased bandwidth; integrated operational capabilities; enhanced collaboration, including common technology, tools, and procedures; and enhanced shared situational awareness through shared analytic and collaborative technologies. The National Cybersecurity Center (NCSC) within the Department of Homeland Security will play a key role in securing U.S. government networks and systems under this initiative by coordinating and integrating information from centers to provide cross-domain situational awareness, analyzing and reporting on the state of U.S. networks and systems, and fostering interagency collaboration and coordination.

Connecting Current Cyber Security Operation Centers to Enhance Situational Awareness

The importance of enhancing cyber security situational awareness in the critical national infrastructure is well understood by the Department of Defense. Critical national infrastructure information gained may not be readily actionable if it remains in the minds of just a few people or is not fused with other key information. Technology can be an enabler in enhancing cyber security situational awareness while ensuring that it does not have to be built from the ground up.

Available commercial collaboration tools allow for rapid and secure sharing of information with key team members while integrating with existing applications already in use. Operational teams can collaborate with voice, video, instant messaging, blogs, and wikis, in addition to shared documents and workspaces. These solutions allow for rapid training, further enhancing leaders' awareness of the operational environment. Operational teams, network defenders, and leaders can function virtually rather than abide by the parochial concept of large operations centers, where cyber operators are locked away watching banks of monitors displaying volumes of trivial events. Technology-enabled sharing and collaboration solutions can bring together disparate and geographically dispersed individuals and teams, creating a better cyber operational environment.

Nevertheless, the Department of Defense needs to improve real-time cyber security situational awareness in the cyber domain rather than rely on after-the-fact forensics. Cyber indications and warning information must be a result of smartly aggregated and correlated data sets that have been fused from all sources. Cyber defense and offense operations must be synchronized across all Department of Defense operating entities. The Department of Defense needs real-time cyber security situational awareness and synchronized cyber operations across the vast cyber domain that is critical to all its missions across all domains. Getting there requires an optimized approach to governance and an efficient model for cyber security situational awareness. Commercial technology solutions can facilitate the intra- and interagency, as well as public-private sector information sharing and collaboration that is required. By devising a model for cyber security situational awareness in the critical national infrastructure (see "An Agenda for Action for Enhancing Cyber Security Situational Awareness in the Critical National Infrastructure") that focuses on aggregation and correlation of information as well as building a layered monitoring framework that roots out truly anomalous activity for human interaction, the Department of Defense can efficiently utilize human capital and conduct dynamic cyber security operations.

An Agenda for Action for Enhancing Cyber Security Situational Awareness in the Critical National Infrastructure

The enhancement of cyber security situational awareness is the focal point for the critical national infrastructure for receiving, tracking, monitoring, and reporting of cyber security incidents (check all tasks completed):

1. Monitoring the critical national infrastructure cyber security vulnerabilities, maintaining an awareness of the threat to the critical national infrastructure, and providing appropriate information to senior

critical national infrastructure officials, so they can maintain an up-to-date awareness of the threat and vulnerability to that threat.

2. Providing a centralized capability for reporting of cyber-related security incidents against the critical national infrastructure's internal information technology infrastructure.
3. Monitoring the critical national infrastructure intrusion detection and intrusion prevention systems.
4. Maintaining an information cyber security incident response report database, conduct trending analysis of events, and recommend actions to minimize or prevent releases.
5. Reviewing actions and conduct root cause analysis of critical national infrastructure information cyber security incidents.
6. Interfacing with the critical national infrastructure on patch review and applicability of patches to ensure prioritization.
7. Coordinating activities and responses to internal critical national infrastructure cyber-related security incidents with appropriate offices.
8. Communicating relevant cyber security information such as security alerts, advisories and bulletins, software vulnerability data and reports, vendor patch notifications, virus alerts, and other relevant cyber security information.
9. Providing an electronic clearinghouse for information assurance tools, antivirus software, and recommended or best practice cyber security guidelines.
10. Serving as the primary reporting authority to the U.S. Computer Emergency Readiness Team, OMB, law enforcement and criminal investigative groups in the reporting of cyber-related attacks against the critical national infrastructure.
11. Serve as the critical national infrastructure observer to the Committee on National Security Systems.
12. Participating in relevant federal cyber security groups such as the National Cyber Response Coordination Group and Government Forum of Incident Response and Security Teams.
13. Conducting penetration testing and vulnerability scanning of the critical national infrastructure's network.

Summary

This chapter focused on situational awareness, which refers to the collective real-time understanding within an organization of its security risk posture. Awareness of security posture requires consideration of several technical, operational, business, and external or global factors.

Furthermore, the chapter also covered the enhancement of situational awareness through the understanding of the current environment and being able to accurately anticipate future cyber security problems to enable effective actions. This was approached through the context of sensemaking, in contrast with the traditional situational awareness approach, recognizing that both are valid and necessary approaches.

For example, sensemaking is the ability to make sense of an ambiguous situation. It is the process of creating situational awareness and understanding to support decision making under uncertainty—an effort to understand connections among people, places, and events in order to anticipate their trajectories and act effectively.

The traditional situational awareness approach involves a human's mental representation of the world in terms of perception and comprehension of elements in the environment. Traditional situational awareness research tends to focus on user interface issues in displays and visualizations. In contrast, sensemaking situational awareness research addresses not only the user interface design issues, but also the underlying goal-directed behaviors such as the cyber security problem-solving context, goals, assumptions, expectations, and biases that affect human performance.

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

Chapter Review Questions/Exercises

True/False

1. True or False? *Situational awareness* refers to the collective real-time understanding within an organization of its security risk posture.
2. True or False? The process of determining whether an attack on national infrastructure is under way is much less difficult than it sounds.
3. True or False? A common cynical view of computer security is that its experts are engaged in nothing more than a game of trivial pursuit around attack and vulnerability information.
4. True or False? A technique not commonly used in government intelligence community environments, but almost never in most enterprise settings, involves the creation and use of a regularly published (usually daily) intelligence report.
5. True or False? Managers of essential national services must understand the security risks associated with their underlying infrastructure.

Multiple Choice

1. Awareness of security posture requires consideration of several technical, operational, business, and external or global factors, including one of the following:

A. Infrastructure vulnerabilities
B. Attack vulnerabilities
C. Known vulnerabilities
D. Domain-based vulnerabilities
E. Time-based vulnerabilities

2. Although the frequency and content of intelligence reports should be tailored to the needs of the local environment, some types of information that one would expect in any daily intelligence report include the following, except which one:

A. Current security posture
B. Intelligence gathering
C. Top and new security risks
D. Automated metrics
E. Human interpretation

3. The generally agreed-upon approach to measuring the security risk associated with a specific component begins with two estimations:

A. Likelihood
B. Location
C. Activity
D. Consequences

E. Data

4. The goal of supporting a national-level view of security posture should not be controversial to most security and infrastructure managers. Everyone will agree that such a view is necessary and useful for supporting national infrastructure protection-related management decisions. The challenge, however, lies with the following important practical considerations, except which one:

A. Commercial versus government information
B. Collection targets
C. Information classification
D. Agency politics
E. SOC responsibility

5. To learn more about potential vulnerabilities, consider attending a:

A. Network conference
B. Security conference
C. Database conference
D. Hacking conference
E. Storage conference

Exercise

Problem

Recently, a large-scale functional exercise on cyber security situational awareness was coordinated among international and U.S. federal and state governments, and private sector organizations. Planners were integral to the exercise design process and were organized to help in the management and development of a situational cyber attack scenario to meet their objectives. Expand on what the general findings were from this exercise.

Hands-On Projects

Project

In this case study, a small company provides a subscription service to a specialized database, and its network consists of 40 workstations, two SQL servers, two exchange servers and two dedicated website servers, all linked together via a broadband connection. The company did not have a trained cyber security team, just one person serving part-time in a cyber security administrator role. When the company's webserver suddenly started experiencing much higher levels of traffic from countries where they did not conduct business, they suspected cyber criminals had broken into their network. Explain how the cyber security administrator was able to solve this problem.

Case Projects

Problem

This case study has to do with a new generation of malware which has evaded existing cyber security products, which were engineered by highly skilled programmers and showed that traditional cyber security approaches are no longer sufficient for effective protection. Tasked with protecting highly sensitive data assets, government cyber security teams must defend against these threats on a daily basis. Explain how government cyber security teams would go about solving this new malware problem.

Optional Team Case Project

Problem

In this case project, the current paradigm for cyber security is based on protection. Protection depends on identifying vulnerabilities and applying countermeasures to neutralize their effects. These are complex human-based activities whose results are uncertain and not capable of according 100% assurance. While used with some effect for components, applications, and stand-alone systems, the paradigm of protection is insufficient for ensuring systems, such as the nation's critical infrastructure and DOD's Global Information Grid. Explain how you would go about anticipating and avoiding the effects of adversity in solving this case project.

[1] E. Waltz, *Information Warfare: Principles and Operations*, Artech House, Norwood, MA, 1998.

11

Response

Chapter Outline

Pre- Versus Post-Attack Response
Indications and Warning
Incident Response Teams
Forensic Analysis
Law Enforcement Issues
Disaster Recovery
National Response Program
The Critical National Infrastructure Incident Response Framework
Transitioning from NIPP Steady State to Incident Response Management
Summary
Chapter Review Questions/Exercises

Incident response is a vital part of any successful IT program and is frequently overlooked until a major security emergency has occurred, resulting in untold amounts of unnecessary time and money spent, not to mention the stress associated with responding to a crisis.

Kenneth van Wyk and Richard Forno[1]

The most familiar component of any cyber security program is the *incident response* process. This process includes all security-related activities that are initiated as a result of an attack that is imminent, suspected, under way, or completed. Incident response will generally be optimized to the local environment in an organization, but in most cases it will include at least the following four distinct process phases:

1. *Incident trigger*—Some warning or event must trigger the incident response process to be initiated. Obviously, if the trigger involves a system that has already been maliciously attacked, then the response must be focused on reconstitution and disaster recovery. If the trigger involves an early warning, then it is possible that the incident response process could avoid visibly negative effects.
2. *Expert gathering*—This involves a gathering together of the appropriate experts to analyze the situation and make recommendations. Most organizations have a base set of incident response staff that work all incidents and manage a repository of information related to all previous incidents. In addition, each incident will dictate that certain subject matter experts be brought into the process to work the details. These experts will also provide a local information base relevant to the incident at hand.
3. *Incident analysis*—Analysis of the incident is the primary task for the experts gathered during incident response. This can include detailed technical forensics, network data analysis, and even business

process examination. Generally, the most difficult part of any analysis involves figuring out the underlying cause of the incident. Once this has been determined, developing the best solution is the key goal.

4. *Response activities*—The output of any incident response process will be a set of management recommendations on how to deal with the incident. These often include rebuilding systems, working around problems, informing customers, and the like. Providing this information to the correct individuals and organizations requires that the incident response teams be properly plugged into the specifics of which groups are responsible for which relevant functions.

Specific incident response processes will vary from organization to organization, but virtually every company and agency process is based on some version of these four elements and includes incident response processes local to an organization or that might exist as a special response resource for citizens, businesses, or government groups (see Figure 11.1)

Most organizations have some form of incident response process in place that generally incorporates the same elements.

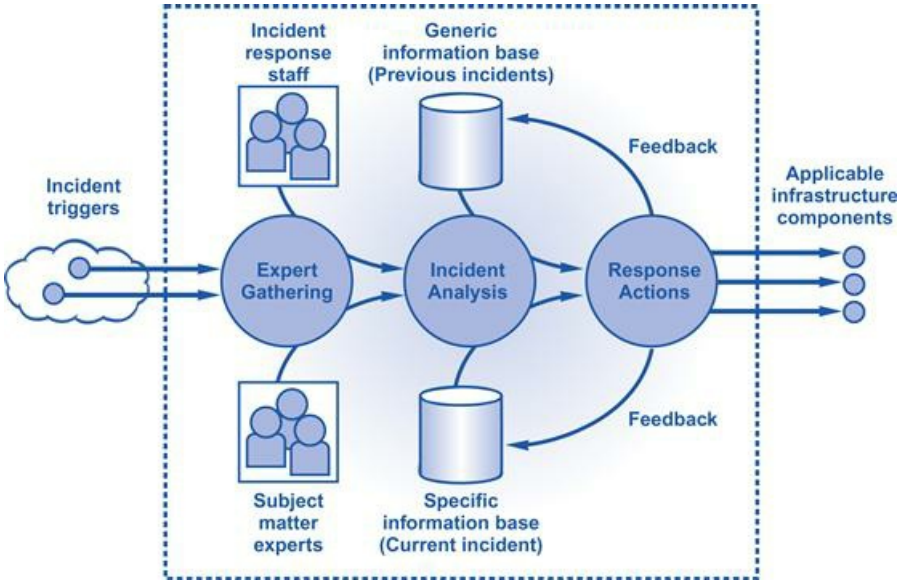


Figure 11.1 General incident response process schema.

In spite of the commonality inherent in the incident response processes found in various companies and agencies, great differences exist in their respective success patterns. The biggest differences reside in the relative effectiveness of incident response in avoiding, rather than simply responding to, serious infrastructure problems. To optimize the early-warning aspect of incident response, certain key considerations must be well understood. These include a focus on pre- versus post-attack responses, detailed understanding of what constitutes a valid indication or warning, proper construction of how an incident response team should be managed, best practices in forensic analysis, optimal interactions with law enforcement, and good processes for recovering from disasters. These elements are explained in more detail below, with an emphasis on how

national infrastructure response processes must be constructed and operated.

Effective incident response is critical, but avoiding infrastructure problems in the first place will reduce the work required of the incident response team.

Pre- Versus Post-Attack Response

The most critical differentiating factor between incident response processes involves the two fundamental types of triggers that initiate response. The first type involves tangible, visible effects of a malicious attack or incident. These effects are usually noticed by end users in the form of slow application performance, clogged gateway performance, inability to get e-mail, slow or unavailable Internet access, and so on. Incident response in this case is usually urgent and is affected by the often vocal complaints of the user base. The second type of trigger involves early warning and indications information, usually embedded in some system or network management information. These triggers are usually not visible to end users but are prone to high levels of false positive responses, where the warning really does not connect to a malicious action.

Early warning triggers are generally not visible to end users and are prone to high levels of false positives.

Incident response processes can thus be categorized into two specific approaches, based on the degree to which these triggers are addressed:

- *Front-loaded prevention*—This includes incident response processes that are designed specifically to collect indications and warning information for the purpose of early prevention of security attacks. The advantage is that some attacks might be thwarted by the early focus, but the disadvantage is that the high rate of false positive responses can raise the costs of incident response dramatically.
- *Back-loaded recovery*—This includes incident response processes that are designed to collect information from various sources that can supply tangible, visible information about attacks that might be under way or completed. This approach reduces the false positive rates but is not effective in stopping attacks based on early warning data.

Hybrid incident response processes that attempt to do both front-end and back-end processing of available information are certainly possible, but the real decision point is whether to invest the time, resources, and money necessary for front-loaded prevention. These two types of processes can be illustrated on the time line of information that becomes available to the security team as an attack proceeds. For front-loaded prevention, the associated response costs and false positive rates are high, but the associated risk of missing information that could signal an attack is lower; for a back-loaded response, these respective values are the opposite (see Figure 11.2).

Combining front-loaded prevention with back-loaded recovery creates a comprehensive response picture; however, an emphasis on front-loaded prevention may be worth the increased cost.

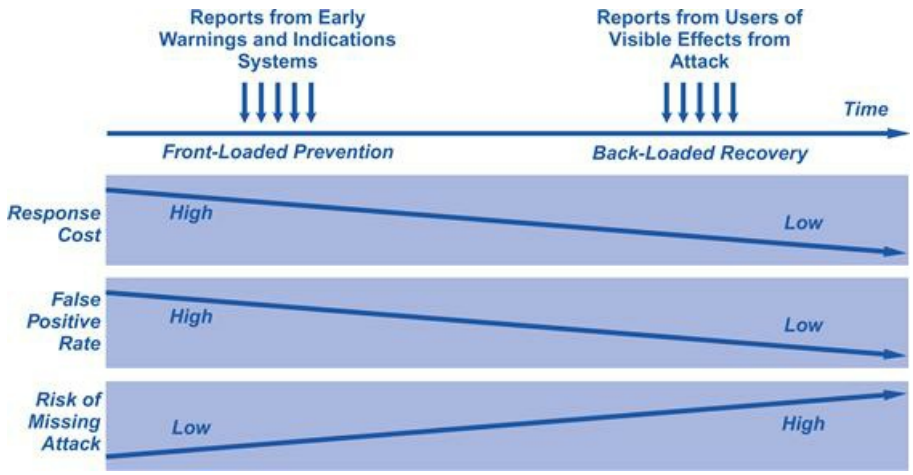


Figure 11.2 Comparison of front-loaded and back-loaded response processes.

Back-loaded incident response might be acceptable for smaller, less-critical infrastructure components, but for the protection of essential national services from cyber attack the only reasonable option is to focus on front-end prevention of problems. By definition, national infrastructure supports *essential* services; hence, any process that is designed specifically to degrade these services misses their essential nature. The first implication is that costs associated with incident response for national infrastructure prevention will tend to be higher than for typical enterprise situations. The second implication is that the familiar false positive metric, found so often in enterprise settings as a cost-cutting measure, must be removed from the vocabulary of national infrastructure protection managers.

It is worth suffering through a higher number of false positives to ensure protection of essential national assets.

Indications and Warning

Given the importance in national infrastructure protection of front-loaded prevention based on early indications and warning information, it becomes urgent to clarify the types of early triggers that should be used to initiate response processes. Because these triggers will vary between organizations due to obvious differences between their respective environments, the best that can be done is to categorize the various types of triggers into a broad taxonomy. Some of the elements of the taxonomy will be obvious and consistent with most current methodologies, whereas others will be quite different from current practice and will require process enhancements.

Taxonomy of Early Warning Process Triggers

The taxonomy of early warning process triggers includes:

- *Vulnerability information*—Knowledge of any new vulnerability is an obvious trigger for front-loaded prevention. The vulnerability might never be exploited, but response teams should still analyze the possibilities and work toward developing proactive steps to ensure that an exploit cannot occur. In many cases, the vulnerability will be reported by a vendor, which implies that they will have to become part of the local incident response process.
- *Changes in profiled behavioral metrics*—Incident response teams should use meaningful changes in any measured behavioral metric as a trigger for process initiation. This can include changes in network behavior, changes in processor utilization, or changes in some application profile. Initiation of incident response as a result of behavioral change represents a dramatic departure from current incident response processes in most organizations.
- *Match on attack metric pattern*—Similarly, if a signature or attack metric pattern is detected on some application, system, or network, then preventive incident response dictates that analysis be performed on the data for security implications. This is also a departure from current incident response approaches.
- *Component anomalies*—Any anomalous behavior detected in an infrastructure component is a candidate trigger for incident response. More intense behavioral anomalies found on more critical components will clearly trigger greater response processes.
- *External attack information*—Information that comes from external sources about attacks that might be locally relevant could trigger an incident response process. For national infrastructure protection, this is even more important if the information comes from a credible source regarding systems or technology having some local significance.

One way to view the difference between the front-loaded and back-loaded methods is in the context of the trigger intensity required to initiate a response process. For the trigger approaches listed above, the information should be sufficient to cause the incident response team to take immediate action. In more conventional and familiar contexts, these triggers would not be sufficient for such action (see Figure 11.3).

Front-loaded prevention responses have a high sensitivity to triggers; that is, response is initiated more often than with a back-loaded recovery response.

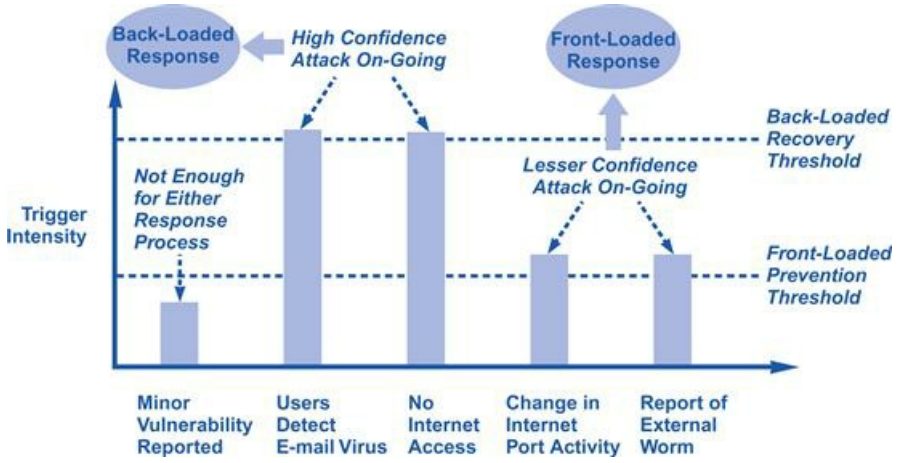


Figure 11.3 Comparison of trigger intensity thresholds for response.

The triggers for front-loaded response share one important aspect—namely, they provide partial information that may signal a possible attack but that also could be explained in a non-security context. Thus, the major obligation of the incident response team in front-loaded prevention is to piece together all partial information into as complete a view as possible; from this view, for national infrastructure protection, the most conservative recommendation should be made. That is, it should be presumed that an attack is ongoing even if the team is not sure. This increases costs and decreases convenience to the local staff, but it errs on the side of caution and is thus appropriate for protecting essential services.

Erring on the side of caution is worth the extra time and expense when it comes to protecting our national assets.

Incident Response Teams

The optimal incident response team for national infrastructure protection includes two different components. First, a core set of individuals will manage the incident response process, maintain relevant repository information, document all incident-related data, provide briefings to anyone interested in the process (including senior management), and interact with other incident response teams. Second, a more dynamically allocated set of subject matter experts will be brought into the incident response activity when an attack is targeting systems they understand best.

In complex settings, the core incident response team is likely to be working multiple incidents simultaneously, generally with different sets of subject matter experts. Thus, response triggers will spawn new cases, which are worked in parallel to successful completion. In smaller environments, it is rare for multiple cases to be ongoing, but for larger, more complex critical infrastructure it is unusual to find times when multiple incident response cases are not being worked simultaneously. This leads to the unique incident response obligation for national infrastructure protection of ensuring that concurrent response activities do not mutually conflict (see Figure 11.4).

Individuals on incident response teams need to ensure they are not working at cross-purposes with their colleagues.

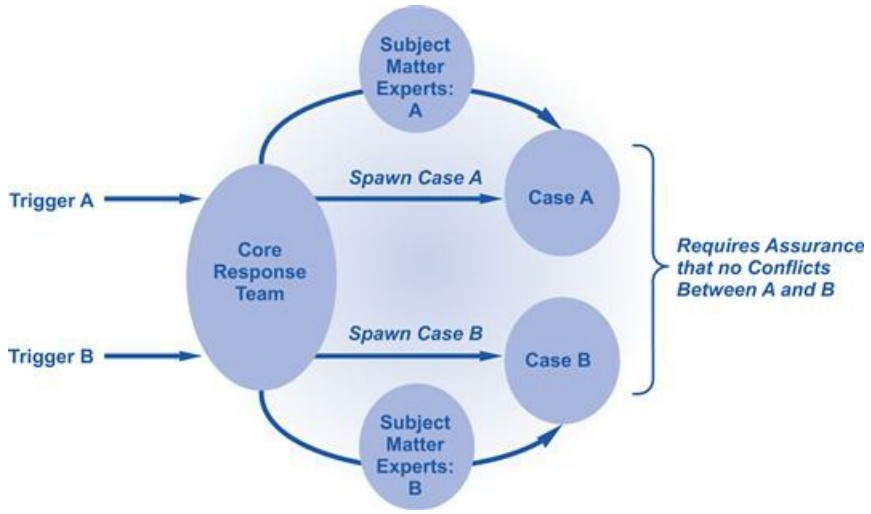


Figure 11.4 Management of simultaneous response cases.

The notion of managing simultaneous response cases is largely unexplored in conventional computer security. This is unfortunate, because every large organization eventually comes to the realization that this is not only possible but is generally the norm. Furthermore, those national attack scenarios with the most serious potential consequences to infrastructure routinely include multiple concurrent attacks aimed at the same company or agency. Response teams in a national setting must therefore plan for the possibility of multiple, simultaneous management of different incident response cases. Some considerations that help plan properly

for this possibly include the following:

It is unlikely that a large organization would not have simultaneous attack scenarios to face.

- *Avoidance of a single point of contact individual*—If a single individual holds the job of managing incident response processes, then the risk of case management overload emerges. This might seem like a minor management detail, but given the importance of response, especially in a recovery scenario, avoidance of such weaknesses is a requirement.
- *Case management automation*—The use of automation to manage, log, and archive incident response cases will improve the productivity of the core incident response team and can lead to streamlined analysis, especially if previous case information is available for online, automated query and search.
- *Organizational support for expert involvement*—The entire organization must readily agree to provide experts for incident response when requested. This is not controversial when the process follows a back-loaded recovery method, because everyone is visually aware of the consequences of the incident. It is more challenging, however, when a front-loaded prevention approach is used and the triggers that initiate incident response are more subtle.
- *24/7 operational support*—Without full 24/7 coverage every day of every year, the success likelihood of managing multiple, concurrent incident response cases drops considerably. Most organizations integrate their incident response function into an SOC to ensure proper management coverage.

An interesting recent trend in infrastructure management involves the outsourcing of certain security operations to a third party. For status monitoring of security devices such as firewalls and intrusion detection systems, this is a reasonably mature activity and will have no materially negative effect on local security protection efforts (unless the outsourcing firm is incompetent). Even for certain SOC operations, outsourcing is often an excellent idea, especially because collection and correlation are always more effective if the vantage point is large. Outsourced SOC operations can also provide the security team with access to technical skills that may not reside locally.

Outsourcing some aspects of security operations may make good business sense.

Incident response processes, however, can easily become awkward for full outsourcing because of the embedded nature of prevention and recovery efforts for local infrastructure. Certainly, an outsourcing provider or vendor can and should be of assistance, and third-party SOC experts might offer excellent guidance and advice. Ultimately, however, incident response must be a local management function, and the organization will have no choice but to expend time, energy, and resources to ensure that the correct local management decisions are made. Third parties can never prioritize actions or tailor recovery procedures to the local environment as well as the organization itself. Instead, they should be used to augment local functions, to provide expert guidance, to automate processes, to manage equipment and networks, to support data collection and correlation, and to assist in recovery.

Companies cannot avoid complete responsibility for incident response by outsourcing the entire process;

prioritizing and tailoring recovery procedures must be done locally.

Forensic Analysis

Forensic analysis involves those activities required to investigate, at both a high level and a detailed lower level, the root cause and underpinnings of some event. Typical questions addressed during the forensic analysis process include:

- *Root cause*—How specifically was the target system attacked?
- *Exploits*—What vulnerabilities or exploits were used in the attack?
- *State*—Is the system still under an active state of attack by an adversary?
- *Consequences*—What components of the system were read, stolen, changed, or blocked?
- *Action*—What actions will stop this attack (if ongoing) or prevent one in the future?

To answer these difficult questions during incident response, forensic analysis requires the ability to drive deeply into a target system of interest, gathering relevant information but doing so in a manner than never destroys, affects, or changes key evidence. This is a critical requirement, because clumsy forensic analysis might overwrite important files, change important stamped dates on system resources, or overwrite portions of memory that include critical evidence. Forensic analysis is a difficult activity requiring great skill and competency, as well as the ability to investigate a system both manually and with the assistance of special tools (see Figure 11.5).

Great care must be taken during forensic analysis not to change or destroy files or other critical evidence.



Figure 11.5 Generic high-level forensic process schema.

The forensic process is performed on a computer to determine how, when, and where some event on that computer might have occurred as the result of hardware, software, human, or network action. Corporate security groups, for example, often perform forensic analysis on a computer when the owner is suspected of violating some guideline or requirement. Law enforcement groups perform similar actions on computers seized from suspected criminals. Forensics can, however, be performed on a target much broader than a computer. Specifically, for the protection of essential national services, the organization must have the ability to perform forensic analysis on the entire supporting infrastructure.

Forensic analysis can be specific (one computer) or broad based (entire supporting infrastructure).

The individual technical skills required to perform such broad forensic analysis are easy to write down, but qualified personnel are not always so easy to recruit and hire. This problem is so severe for most large organizations that it is not uncommon for a company or agency to have no local expert with sufficient skills to properly lead the investigation of a widespread infrastructure attack. This is unacceptable, because the only options for that organization are to locate such talent externally, and this will result in a less intimate evaluation process. Long-term employees who are committed to a career in an organization will always be more knowledgeable than consultants or third parties; furthermore, they will be suitably trusted to investigate an incident into the deep recesses of the local environment.

An internal expert will be the one most likely to properly lead a company investigation, but few company employees have the requisite skills.

As such, the irony of forensic analysis is that most businesses and agencies would be wise to begin building and nurturing a base of talent with these skills. Typically, to maintain and satisfy forensic experts requires several things:

- *Culture of relative freedom*—Most good forensic analysts are creative individuals who learned their craft by exploring. They tend to maintain their skills by continuing to explore, so organizations must give them the freedom to seek and analyze systems, networks, applications, and other elements of interest. When they are working an incident, the target is obvious, but when they are not then managers must offer them the freedom to explore as they see fit. This is not easy for some managers, especially in relatively mature organizations with (ahem) long legacies of tight employee controls.
- *Access to interesting technology*—A related aspect of the local environment required to keep forensic analysts happy is constant access to interesting, changing, and emerging technology. What this means is that assigning your best forensic analysts to day-to-day operations around a single technology might not be the best idea.
- *Ability to interact externally*—Forensic analysts will also need the freedom to interact with their peer community and to learn from experts outside the organization. This must be permitted and encouraged.

These environmental elements are not unique to forensic experts, but of all the skill sets required in a national infrastructure protection setting forensic analysis is the one that is the most difficult for an organization to obtain. Good forensic analysts can command the highest premium on the market and are thus difficult to keep, especially in a relatively low-paying government job. As such, attention to these quality-of-work-life attributes becomes more than just a good idea; instead, it becomes a requirement if the organization chooses to have the ability to perform forensic analysis as part of the overall incident response process.

Investing in a good forensic analyst will be expensive but worthwhile for the protection of national security assets.

Law Enforcement Issues

A common issue faced by response teams is whether a given incident should be turned over to law enforcement for support. Most countries have laws that obligate response teams to contact law enforcement groups in the event of certain crimes; incident response teams must be familiar with these laws and must obey them without question. They must, in fact, be burned into incident response processes with full review by legal council in the organization. The issue of law enforcement involvement is also driven, however, by emotional considerations, especially when great time and effort have been directed toward dealing with some incident. The team often wishes to see tangible retribution, perhaps involving the bad guys actually going to jail.

Carefully review local, regional, and national laws regarding when law enforcement must be contacted during a security incident.

In the end, however, interaction with law enforcement for infrastructure protection should follow a more deliberate and routine process. National infrastructure protection has a singular goal—namely, to ensure the continued and accurate delivery of essential services to the citizenry and businesses of a nation. This does not include the goal of catching bad guys and throwing them in jail, as much as security teams might like this result. The result is that discretionary law enforcement involvement should only be considered when the local security team believes that such enforcement could help with a current incident, perhaps through offering some relevant data or hints, or could help prevent a future incident by putting away some group that appears to be a repeat offender. A decision process for law enforcement involvement emerges as shown in Figure 11.6.



Figure 11.6 Decision process for law enforcement involvement in forensics.

This decision process does recognize and support the clear requirement that crimes must be reported, but the figure also highlights a particularly fuzzy aspect of cyber security—namely, detecting suspicious behavior on a computer network usually does not constitute sufficient evidence of a crime being committed. Even if evidence of a break-in to a given system is observed, the argument could be made that no crime has occurred, especially if the break-in is the result of some automated process as one finds in a botnet attack.

Incident response teams should report relevant information to law enforcement, even if it does not result in arrest.

The result is that national infrastructure protection teams will need to understand the decision process

for law enforcement and follow it carefully during every incident. They will also need to create a local process for determining whether a crime has been committed in the context of their infrastructure. The result not only will optimize the interface between an organization and law enforcement but will also minimize the inevitable resource demands that will arise for the local team if law enforcement gets involved.

Disaster Recovery

The process of disaster recovery after a security attack is more mature than other aspects of incident response. This stems from the commonality that exists between recovery from attack and recovery from natural disasters such as floods, tornados, fires, and the like. Unfortunately, many large organizations charged with responsibility for national infrastructure do not properly address their obligation to include disaster recovery in their planning. Specifically, disaster recovery programs have three fundamental components, whether they are driven by concerns of malicious attack or natural disaster (see box).

Three Components of a Disaster Recovery Program

- *Preparation*—The decision to prepare in advance for disaster recovery is easy to make but much more difficult to support in practice. Operational funding is usually the stumbling block, because the process of preparing for disaster in advance involves more than just writing down a list of potential actions. Instead, it often requires architectural changes to avoid single points of potential failure. It could require installation of safe, redundant means for communication between recovery teams, and it could even require upgrades to cyber security systems to ensure proper protection through a disaster.
- *Planning*—An essential element in a disaster recovery program is an explicit plan that is written down and incorporated into all operational methods and procedures. The plan can be continually improved as the organization deals with real disasters. For example, many organizations who relied on the use of commercial airplanes to shuttle equipment to disaster sites found that this did not work well in the aftermath of 9/11.
- *Practice*—The decision to practice for disasters is also an expensive one, requiring that teams of experts be funded to support mock drills. The best way to practice for a disaster is to create a realistic scenario and work through the specifics of the written plan. Usually, this will involve the use of spare computing or networking capacity that is set aside in a hot configuration (see Figure 11.7).

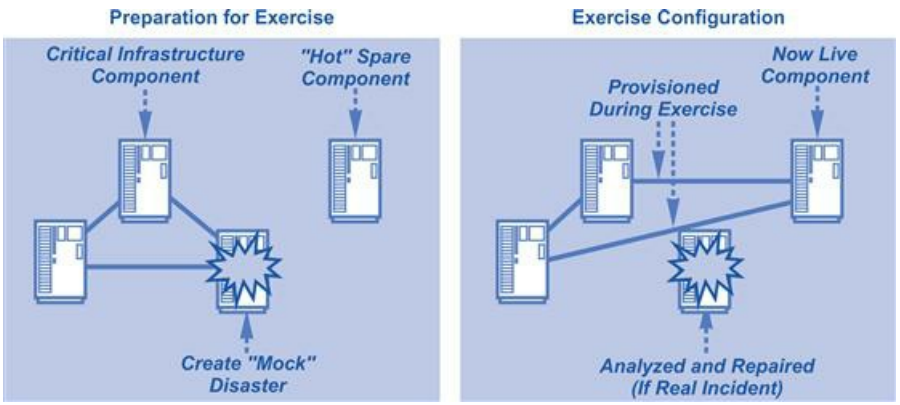


Figure 11.7 Disaster recovery exercise configurations.

Realistically, very few organizations actually practice for disasters. It requires a discipline that is generally

missing from most enterprise system and network teams and can only work if the senior leadership team makes it a priority. Sadly, the only time disasters are considered is after they occur, especially after they have some impact on the local environment. This familiar process of taking disasters seriously only after they occur is something we have all witnessed in our society, especially as it relates to natural disasters and terrorism. For proper protection of national infrastructure from cyber attack, this attitude must be adjusted.

Proper planning for disaster response and recovery requires time and discipline, but the outcome is well worth the effort.

National Response Program

The most important function in any national response program involves emergency coordination among government, business, citizens, and other nations during a cyber attack incident. The respective interfaces must be identified and managed as part of response planning. National programs can provide centralized coordination, but intrasector coordination should also be encouraged (see Figure 11.8).



Figure 11.8 National response program coordination interfaces.

This coordination function would seem obvious, but most existing national emergency response programs and computer emergency response team (CERT) programs tend to focus on dissemination of vulnerability-related information. This is useful, especially for smaller organizations that have no security team, but this focus tends to leave a gap in national-level coordination should a major national incident occur. Amazingly, at the time of this writing, such a major national incident has yet to occur, but if one should happen soon then national coordination in the United States is unlikely to be smooth. This is unacceptable and requires immediate attention to properly protect national infrastructure from the effects of cyber attack.

Finally, let's very briefly look at some practical ways the federal government plays a significant role in managing intergovernmental (federal, state, local, and tribal) and, where appropriate, public-private coordination in response to cyber incidents of national significance. Federal government responsibilities include:

- Providing indications and warning of potential threats, incidents, and cyber attacks.
- Sharing information both inside and outside the government, including best practices, investigative information, coordination of incident response, and incident mitigation.
- Analyzing cyber vulnerabilities, exploits, and attack methodologies.
- Providing technical assistance.
- Conducting investigations, forensics analysis, and prosecution.
- Attributing the source of cyber attacks.
- Defending against the cyber attack.
- Leading national-level recovery efforts.

The preceding activities require a concerted effort by federal, state, local, and tribal governments, and nongovernmental entities such as private industry and academia. Also, together, the National Infrastructure

Protection Plan (NIPP) and the National Response Framework (NRF) provide a comprehensive, integrated approach to addressing these key activities/elements of the nation's homeland security mission to prevent terrorist attacks, reduce cyber attack vulnerabilities, and respond to incidents in an all-hazards context. The NIPP establishes the overall risk-informed approach that defines the nation's steady-state posture with respect to critical infrastructure and key resources (CIKR) protection and resiliency, while the NRF and National Incident Management System (NIMS) provide the overarching framework, mechanisms, and protocols required for effective and efficient domestic incident response management. The NIPP risk management framework, information-sharing network, and partnership model provide vital functions that, in turn, inform and enable incident response management decisions and activities.

The Critical National Infrastructure Incident Response Framework

The National Response Framework (NRF) provides an all-hazards approach that incorporates best practices from a wide variety of disciplines, including fire, rescue, law enforcement, public works, and emergency medical services. The operational and resource coordinating structures described in the NRF are designed to support decision making during the response to a specific cyber threat or incident and serve to unify and enhance the incident response management capabilities and resources of individual agencies and organizations acting under their own authority. The NRF applies to a wide array of natural disasters, terrorist threats and cyber security incidents, and other emergencies.

The NRF specifies incident response management roles and responsibilities, including emergency support functions designed to expedite the flow of resources and program support to the incident area. Federal agencies have roles within the NRF structure that are distinct from, yet complementary to, their responsibilities under the NIPP. Ongoing implementation of the NIPP risk management framework, partnerships, and information-sharing networks sets the stage for CIKR cyber security and restoration activities within the NRF by providing mechanisms to quickly assess the impact of the incident on both local and national CIKR, assist in establishing priorities for CIKR restoration, and augment incident-related information sharing.

Collaborative Efforts for Cyber Attack Watch, Warning, and Incident Response

The federal government is working strategically with key allies on cyber security policy and operational cooperation. For example, the Department of Homeland Security (DHS) is leveraging preexisting relationships among computer security incident response teams (CSIRTs). DHS also has established a preliminary framework for cooperation on cyber security policy, watch, warning, and incident response with key allies. The framework also incorporates efforts related to key strategic issues as agreed on by these allies. An International Watch and Warning Network (IWWN) is being established among cyber security policy, computer emergency response, and law enforcement participants representing 18 countries. The IWWN will provide a mechanism through which the participating countries can share information in order to build global cyber security situational awareness and coordinate incident response.

Transitioning from NIPP Steady State to Incident Response Management

The variety of alert and warning systems that exist for natural hazards, technological or industrial accidents, and cyber security and terrorist incidents provide the bridge between steady-state operations using the NIPP risk management framework and incident response management activities using the NRF concept of operations. These all-hazards alert and warning mechanisms include programs such as National Weather Service hurricane and tornado warnings, and alert and warning systems established around nuclear power plants and chemical stockpiles. In the context of terrorist incidents, Homeland Security Advisory System (HSAS) provides a progressive and systematic approach that is used to match protective measures to the nation's overall cyber attack threat environment. This link between the current cyber attack threat environment and the corresponding protective actions related to specific threat vectors or scenarios, and to each HSAS threat level, provides the indicators used to transition from the steady-state processes detailed in the NIPP to the incident response management processes described in the NRF.

DHS and CIKR partners are developing and implementing stepped-up protective actions to match the increased terrorist threat conditions specified by HSAS and address various other all-hazards alerts and warning requirements. As warnings or threat levels increase, NRF coordinating structures are activated to enable incident response management. DHS and CIKR partners carry out their NRF responsibilities and also use the NIPP risk management framework to provide the CIKR protection dimension of cyber security incident operations (see "An Agenda for Action for Integrating CIKR Protection with Cyber Security Incident Response Management").

An Agenda for Action for Integrating CIKR Protection with Cyber Security Incident Response Management

The process for integrating CIKR protection with cyber security incident response management and transitioning from NIPP steady-state processes to NRF incident response management coordination includes the following actions by DHS and other CIKR partners (check all tasks completed):

1. Increasing protection levels to correlate with the specific cyber security threat vectors or threat level communicated through HSAS or other relevant all-hazards alert and warning systems, or in accordance with sector-specific warnings using the NIPP information-sharing networks.
2. Using the NIPP information-sharing networks and risk management framework to review and establish national priorities for CIKR protection, facilitating communications between CIKR partners, and informing the NRF processes regarding priorities for response and recovery of CIKR within the cyber security incident response area as well as on a national scale.
3. Fulfilling roles and responsibilities as defined in the NRF for cyber security incident response management activities.
4. Working with sector-level information-sharing entities and owners and operators on information-sharing issues during the active cyber security incident response mode.
5. Establishing a communications protocol to facilitate timely information exchange and necessary coordination with the CIKR sectors and their federal, state, local, and private sector partners during

those national-level cyber security incidents that involve a coordinated federal response.

Summary

This chapter focused on the most familiar component of any cyber security program: the *incident response* process. This process includes all security-related activities that are initiated as a result of an attack that is imminent, suspected, under way, or completed. Incident response will generally be optimized to the local environment in an organization.

Furthermore, the chapter also covered why specific incident response processes vary from organization to organization. However, virtually every company and agency process is based on some incident response processes that are local to an organization or that might exist as a special response resource for citizens, businesses, or government groups.

In addition, the chapter presented recommendations to help those facilities that use control systems to be better prepared for and respond to a cyber security incident regardless of source. The chapter also suggested ways to learn from cyber security incidents and to strengthen the system against potential cyber attacks.

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

Chapter Review Questions/Exercises

True/False

1. True or False? The least familiar component of any cyber security program is the *incident response* process.
2. True or False? Early warning triggers are generally not visible to end users and are prone to high levels of false positives.
3. True or False? Given the importance in national infrastructure protection of front-loaded prevention based on early indications and warning information, it becomes less urgent to clarify the types of early triggers that should be used to initiate response processes.
4. True or False? Individuals on incident response teams need to ensure they are not working at cross-purposes with their colleagues.
5. True or False? Forensic analysis involves those activities required to investigate, at both a low level and a detailed higher level, the root cause and underpinnings of some event.

Multiple Choice

1. The most familiar component of any cyber security program is the *incident response* process. This process includes all security-related activities that are initiated as a result of an attack that is imminent, suspected, under way, or completed. Incident response will generally be optimized to the local environment in an organization, but in most cases it will include at least the following four distinct process phases, except which one:

A. Incident trigger
B. Expert gathering
C. Incident analysis
D. Domain-based vulnerabilities
E. Response activities

2. Incident response processes can thus be categorized into two specific approaches, based on the degree to which these triggers are addressed:

A. Current security posture
B. Intelligence gathering
C. Front-loaded prevention
D. Back-loaded recovery
E. Human interpretation

3. The taxonomy of early warning process triggers includes one of the following:

A. Likelihood
B. Location
C. Activity

D. Consequences
E. Vulnerability information

4. Response teams in a national setting must therefore plan for the possibility of multiple, simultaneous management of different incident response cases. Some considerations that help plan properly for this possibly include the following, except which one:
A. Avoidance of a single point of contact individual
B. Case management automation
C. Information classification
D. Organizational support for expert involvement
E. 24/7 operational support

5. A typical question addressed during the forensic analysis process includes which of the following?
A. What nonexploits were used in the attack?
B. Is the system still under an nonactive state of attack by an adversary?
C. What components of the system were blocked?
D. What actions will stop this attack (if ongoing) or prevent one in the past?
E. How specifically was the target system attacked?

Exercise

Problem

An organization with a global presence discovered a portion of its network was exposed to the Internet without adequate cyber security protection. The company's cyber security personnel discovered that the security lapse began several months earlier. Expand on what the general solutions were from this exercise.

Hands-On Projects

Project

In this case study, one of the largest electric utilities in the United States, serving 24 million customers and 1,670 communities, needed to secure a wide range of IT systems, including a nuclear power generation station and California's electric power grid, with limited internal security resources. Explain how the utility's cyber security administrator was able to secure a wide range of IT systems.

Case Projects

Problem

This case study has to do with how financial institutions need to balance a high level of security with convenient access for their diverse set of users. So, when the legacy network access control (NAC) system of a

financial institution failed to respond to an internal penetration test during an audit, their cyber security team began looking for a new solution immediately. Explain how the cyber security team went about solving this incident response problem.

Optional Team Case Project

Problem

In this case project, a major university realized cyber attackers had figured out that operating systems (OS) are becoming harder to penetrate and noticed them focusing their attacks on applications within the OS, such as MS Office, Firefox, and so forth. So, the cyber security manager at the university decided it was time to upgrade the school's intrusion prevention and incident response solution. Please explain how the cyber security manager went about solving this case project.

[1] K. van Wyk and R. Forno, *Incident Response*, O'Reilly Media, Sebastopol, CA, 2001.

APPENDIX A

National Infrastructure Protection Criteria

Any discussion of computer security necessarily starts with a statement of requirements.

DOD 5200.28-STD "Orange Book"

As mentioned earlier in the book, during the early 1980s, the U.S. government created a set of functional and assurance computer security criterion requirements entitled the *Trusted Computer System Evaluation Criteria* (TCSEC), also known informally as the Orange Book. For roughly a decade, the Orange Book provided a useful set of criteria for evaluating the security of a given computer or network system. This allowed for a common reference point in evaluating the relative security strengths and weaknesses in a target system, usually a computer operating system. Procurements by government agencies, as well as commercial entities, would simply specify one of the criterion categories as a requirement, and prospective bidders, suppliers, and vendors understood exactly what needed to be done.

Perhaps the greatest strength of the Orange Book was its simplicity; almost two-and-a-half decades after its inception, this author can still recite most of the requirements from memory. By focusing on simplicity, rather than completeness, the Orange Book was a spectacular success, and had a largely positive impact on the development of computer security as a legitimate technical discipline. Sadly, as the "gold rush" associated with the World Wide Web lured many computing professionals (including security experts) to dot-com start-ups, the formal discipline associated with computer security and the Orange Book died off. This is unfortunate, because the requirements embedded in the Orange Book would have helped to reduce the risk of viruses, worms, and botnets during the past decade.

In this appendix, and in the spirit of the Orange Book, we offer a simple set of criterion requirements for national infrastructure protection. The requirements are written in the context of an organization, rather than an operating system, and each requirement is binary, that is, a given requirement is either met or not in a given company or agency. The criteria should be straightforward for incorporation into a local security policy, or into local procurement requirements (especially in government), or into a security audit plan—perhaps even as part of a self-audit discipline in an organization. As a further simplification, the set of criteria includes only two classes—compliant (all requirements met) and not compliant (some requirement not met). Interested readers should have no trouble extending the criterion requirements to their own context, or to tailor them for whatever purpose.

Finally, the requirements are simplified to the point where it should be relatively straightforward to interpret whether a local team is actually compliant with a given requirement. Great care is taken to optimize

this point, albeit subject to comments in the chapter on commonality about how important elements in a security program are often impossible to grade, score, and evaluate objectively. As one would expect, the categories of requirements correspond to the chapters of this book, and each requirement is written as a definitive statement of some condition, capability, or function that the organization should be able to support with a compliance argument using concrete evidence.

Deception Requirements

The organization must …

(DEC-1) … operate deceptive honeypot functionality that is attractive and locally accessible to malicious insiders.

This requirement ensures that effort has been made to operate trap functionality focused on insiders, including employees, consultants, visitors, and contractors. The deployment does not have to be extensive, but should operate somewhere in the enterprise. The decision to run in stealth or nonstealth mode can be a local decision.

(DEC-2) … operate deceptive honeypot functionality that is attractive and externally accessible to malicious outsiders.

This requirement ensures that effort has been made to operate trap functionality focused on outsiders who might target organizational resources via the Internet. The deployment also does not have to be extensive, but must be present and can be done in a stealth or nonstealth manner.

(DEC-3) … operate honeypot management and support systems that can detect exploit attempts at honeypot resources for the purpose of initiating response.

This requirement ensures that the organization has sufficient backend systems for the deception to be effective, especially if it is done in stealth mode. This requirement is best met by a honeypot alarm notification system connected to human beings trained to interpret the results and direct response activity. Without such backend support, the deception is unlikely to operate properly.

Separation Requirements

The organization must …

(SEP-1) … proactively redirect or filter live DDOS traffic before it reaches local network ingress points.

This requirement ensures that the organization is not completely exposed to the crippling effects of inbound DDOS attacks. A service provider operating filters on a large capacity backbone is a good option here. The amount of filtering should be expressed as a multiple of inbound ingress capacity. Obviously, the multiple must be greater than 1—and the greater, the better. The filters must not operate on the ingress gateway, for obvious reasons (do the math).

(SEP-2) … flexibly enforce network access controls (firewalls) between designated groups of insiders.

This requirement ensures that the organization is using internal firewalls to create trusted internal domains. Casual insiders including the majority of a typical employee base should not have the ability to view, change, or block a broad set of internal resources as a result of their special access. Certainly, employees working on a specific component might have the ability to cause local problems, but this effect should be limited to the local component. This can be accomplished with firewalls, access lists on routers and switches, and other types of mechanisms.

SEP-3 … flexibly enforce network access controls (firewalls) between organizational resources and any untrusted external network.

This requirement ensures that firewalls are in place between an organization and any external, untrusted network such as the Internet. Remote access systems must be included as well, and mobility-based access over carrier networks is a new type of access that must be considered as well, particularly with the speeds promised in the 4G infrastructure soon to emerge. It's worth noting that network-based firewalls are particularly efficient in ensuring complete coverage of connections to untrusted networks.

(SEP-4) … stop inbound email and web-based viruses, Spam, and other malware before they reach local network ingress points.

This requirement ensures that inbound network garbage is collected before it hits the ingress point. This greatly reduces the risk of a volume-based attack using these services, as well as simplifying gateway security requirements. Efficiency and cost reduction concerns are a good by-product in this approach, even though they are not the primary motivations for inclusion here.

Commonality Requirements

The organization must …

(COM-1) … have a written security policy, with supporting training for decision-makers, and explicit mechanisms for enforcement and violation consequences.

This requirement ensures that common attention is being placed in the organization to basic security policy considerations and how the attendant requirements are enforced. Clearly, an organization could have a bad security policy, but experience dictates that the effort to actually produce and enforce a policy is usually accompanied by the discipline to ensure that the requirements are reasonable. In contrast, security problems generally arise when an organizational security policy is totally missing.

(COM-2) … demonstrate organization compliance to at least one recognized information security standard attested by an external auditor.

This requirement ensures that the organization has targeted at least one reasonably well-known and accepted security standard for compliance. Although there are some differences between standards, the reality is that the recognized ones all include a basic core set of requirements that dictate essentially the same sort of controls. Thus, it really doesn't matter—in the vast majority of cases—which standard is selected, so long as at least one is being used.

Diversity Requirements

The organization must …

(DIV-1) … provide evidence that no single vendor failure or compromise can produce a cascading effect in critical application, computing, or networking functionality across the entire organization.

This requirement ensures that no single cascading chain of failure exists because of a common vendor thread in some critical technology. Unfortunately, this is a tough requirement for most organizations to meet on the desktop, given the pervasiveness of a single vendor architecture and set of applications. It is nevertheless critical that the cascading problem be addressed through attention to diversity.

(DIV-2) … use at least one live, alternative back-up vendor in a substantive manner for mission-critical software, enterprise PCs, enterprise servers, network devices, and network services.

This requirement implies one possible component of the solution to meeting DIV-1 by dictating a live alternative backup.

Depth Requirements

The organization must …

(DEP-1) … provide evidence that no individual, inside or outside the organization, can directly access systems affecting the integrity or operation of any essential national service without at least two diverse security authentication challenges.

This requirement ensures that no

(DEP-2) … provide a convincing argument that the organization has deployed a network-based mechanism for throttling, diverting, or stopping attack traffic before it reaches the local inbound gateway.

(DEP-3) … provide evidence that failure of any one protection system cannot lead to a direct compromise in any critical application, computing, or networking functionality across the entire organization.

Response Requirements

The organization must …

(RES-1) … provide evidence that the organization has the ability to respond to indicators and warning signals in advance of an attack on any critical resource.

(RES-2) … provide evidence that the organization maintains documentation and metrics on the root cause of past security problems, as well as the effectiveness of response activities for past and present security incidents.

Awareness Requirements

The organization must …

(AWA-1) … provide evidence that cyber security intelligence information is collected on a regular basis and disseminated to decision makers on a timely basis.

(AWA-2) … provide evidence that a real-time security operations function exists that coordinates any preventive or response actions based on collected information and correlative analysis (presumably in an operations center).

Discretion Requirements

The organization must …

(DIS-1) … provide evidence that all organizational information is properly marked and that such markings are suitably enforced.

(DIS-2) … provide evidence that organizational staff is fully trained in local policies for how information is handled and shared externally.

Collection Requirements

The organization must …

(COL-1) … provide evidence that a set of criteria has been established for which types of information in which contexts should be collected and stored by the organization.

(COL-2) … provide evidence that collection systems are in place to gather in real time and store in a secure manner all desired information from applications, systems, and networks.

Correlation Requirements

The organization must …

(COR-1) … provide evidence that algorithms are in place to correlate relevant information in real-time toward actionable results.

(COR-2) … provide evidence that correlative output (presumably in a security operations center) is connected to organizational awareness and response functions.

APPENDIX B

Case Studies

John R. Vacca

For the Case Study Review Questions in this Appendix, please note that the answers to the case studies can be found at: http://www.elsevierdirect.com/companion.jsp?ISBN=9780123918550.

Case Study 1: Cyber Storm

Cyber Storm II (in 2008) was designed to support the strategic vision of the Department of Homeland Security (DHS) and the National Cyber Security Division (NCSD). This a part of the National Protection and Programs Directorate's (NPPD) Office of Cyber Security and Communications (CS&C) and the President's National Strategy to Secure Cyberspace. The primary goal of planning and executing Cyber Storm II was to provide the arena to examine the processes, procedures, tools, and organizations in response to a multi-sector-coordinated attack through, and on, the global cyber infrastructure. The exercise incorporated a wide spectrum of players representing federal, state, and international governments, interagency coordination bodies, and the private sector. The coordinated cyber attacks facilitated incident response from the technical, operational, and strategic perspectives.

In 2008, a cadre of intruders leveraged their collective capabilities to mount a simulated coordinated cyber attack on a global scale. Although primary motives differed among the entities, a sophisticated network of relationships enabled the intruder to degrade Internet connectivity, disrupt industrial functions, and ultimately erode confidence in everyday communications. By generating counterfeit digital certificates, the intruders directed unknowing web users to "spoofed" websites where funds were extorted and personal information was mined. Coordinated attacks on domain name servers and telecommunications router infrastructure resulted in a distributed denial of service and unreliable telephony. Users were intermittently unable to access websites, send e-mail, and make phone calls. Victims of the attack were forced to explore alternative methods of communication during the disruptions. The intruders' intent was to cause cascading disruptions stemming from specific, focused attacks.

As the events unfolded, law enforcement and intelligence agencies gathered information and responded as necessary. In coordination with the impacted private sector entities and other government agencies, law enforcement and the Intelligence Community worked to halt attacks and restore confidence in the Internet. All participating organizations relied on trusted relationships and forged new communications paths to share information and build and pass along situational awareness.

Cyber Storm II objectives were examined through the exercise planning and execution period. A number of findings from the Cyber Storm II exercise were identified. These findings were made through observations by participants and observer/controllers. This part of the case study provides the exercise's significant findings, some solutions, and supporting observations:

- Value of standard operating procedures (SOPs) and established relationships
- Physical and cyber interdependencies
- Importance of reliable and tested crisis communications tools
- Clarification of roles and responsibilities
- Increased noncrisis interaction
- Policies and procedures critical to information flow
- Public affairs influence during large-scale cyber incidents
- Greater familiarity with information sharing processes

The U.S. Department of Homeland Security's (DHS) Cyber Storm exercise series (I, II, and III) is part

of the Department's ongoing efforts to assess and strengthen cyber preparedness, examine incident response processes in response to ever-evolving threats, and enhance information sharing among federal, state, international, and private sector partners. The Cyber Storm III exercise scenario reflected the increased sophistication of intruders, who have moved beyond more familiar webpage defacements and denial of service (DOS) attacks in favor of advanced, targeted attacks that use the Internet's fundamental elements against itself. The goal here was the compromising of trusted transactions and relationships.

Throughout the exercise, the goal of exercise players was to identify, in real time, the ongoing attack and mitigate the compromises and vulnerabilities that allowed it to occur, as well as possible consequences to compromised systems. At its core, the exercise was about resiliency—testing the nation's ability to cope with the loss or damage to basic aspects of modern life. Cyber Storm III was the first opportunity to test the new National Cybersecurity and Communications Integration Center (NCCIC). NCCIC served as the hub of national cybersecurity coordination and was inaugurated in October 2009. Cyber Storm III findings are still being reviewed as of this writing.

Case Study Review Questions

1. What was the primary goal of planning and executing Cyber Storm II?
2. What were some of Cyber Storm II's significant findings, solutions, and supporting observations?
3. Throughout Cyber Storm III, what was the goal of the exercise players?

Case Study 2: Cyber Attacks on Critical Infrastructures—A Risk to the Nation

There has been a great deal of research related to cyber attacks and vulnerabilities and critical infrastructure, but there is an incomplete understanding of the cascading effects a cyber-caused disruption could have on other critical national infrastructures and the ability of the affected infrastructures to deliver services. This case study describes a cyber-attack-consequence assessment process developed to coordinate Sandia National Laboratories' capabilities in assessing the effects of a cyber attack and in assessing the infrastructure impacts and economic consequences of those attacks.

Step 1 of this process identifies a cyber attack, and Step 2 identifies a system vulnerability that will allow a cyber attack to be successful. These two steps may occur simultaneously because a cyber attack is likely to attempt to exploit a system vulnerability to ensure success.

Step 3 of this process is the assessment of the effects of a successful cyber attack on a critical infrastructure control system. This step answers the question, "How does the attack affect the control system and the components that are connected to the system?" Simulators that model control systems can be used to assess how the control system will react to the attack. This step can be informed by general heuristics, or rules-of-thumb, about the structure of the control systems to help inform the assessment.

During Step 4 of the process, the impact of the control system effects to the critical infrastructure being attacked (and possibly other, related infrastructure) is assessed. Infrastructure models are used to determine how the control system effects might spill over to other parts of the infrastructure that are not controlled by the attacked system. The result of this step is an infrastructure-impact scenario, which is a specific scenario of how the infrastructure is affected by the cyber attack. The scenario should specify the particular components of the infrastructure that are affected, as well as the details (time, severity, etc.) of the impacts.

Finally, during Step 5 of the process, the economic consequences of the infrastructure disruptions are found using the infrastructure-impact scenario. If the infrastructure-impact scenario constructed in Step 4 finds that the cyber attack may create disruptions in infrastructure, there will likely be economic ramifications to the loss. Economic models are available that can be used to assess the economic consequences of infrastructure disruptions caused by cyber attacks.

A cyber attack on a control system may have effects beyond those of the attacked infrastructure identified in the infrastructure-impact step of the process. Infrastructures are interdependent, which means that a failure in one component may spill over to other components of the same infrastructure as well as associated infrastructures and industries. This interdependence is clear in the electrical power industry because almost all industries require electrical power in some manner. Disruptions of infrastructure may also spill over to economies. Economic activity depends on the infrastructure. A sustained loss of electric power, for example, may cause economic activity to nearly stop.

The consequences of infrastructure disruptions are complicated and difficult or impossible to measure in many cases and may vary greatly in their effects. An outage at a single generator during a period with adequate reserve capacity is unlikely to disrupt service. Spot prices might be affected by the outage, but there will likely be little change to overall economic activity. The consequences of an outage that results in unserved load are

more difficult to measure. For a short load-shedding event, the economic consequences will likely be light because many short-term economic losses are recoverable. For example, consumer purchases can be delayed to another day or time, and interrupted manufacturers can draw on inventories that can be replenished over time. Many of the losses that do occur may be difficult to quantify. For example, short losses of power chemical plants sometimes cause the release of chemicals and have the potential to cause accidents.

This case study focuses on Sandia's capabilities in carrying out the cyber-attack-consequence assessment process using electric power control systems as an example. The process can be used with other critical infrastructure control systems with modifications to existing capabilities and the addition of infrastructure-impact simulations for new infrastructures.

Of the three steps of the cyber-attack-consequence assessment process focused upon in this case study, the system-effect step and the infrastructure-impact step need to be modified from the electric power walk-through. For the final step (economic consequence assessment), the REAcct tool can continue to use the same type of infrastructure-disruption scenario as an input (specifications of which counties are affected, how long the disruption lasts, what fraction of their area is affected, and what fraction of economic activity is disrupted), provided that the necessary mappings of infrastructure disruptions to economic disruptions are made. Many of the economic assessment tools that filled the gaps of REAcct are similarly flexible or can include new infrastructures by expanding their models.

System Effect

The process walk-through detailed methods and tools that can currently be used to simulate a cyber attack on an electric power control system and assess the impacts to the electric power grid. Although these tools are tailored to the electric power industry, some tools, such as the VCSE can be modified to different infrastructures. Other types of physical infrastructure can be simulated by either interfacing existing tools with the VCSE or creating new tools.

Infrastructure Impact

The infrastructure-impact step of the process maps changes in critical infrastructure control systems that are caused by cyber attacks to overall changes in infrastructure. The tools necessary to assess the infrastructure impact of cyber attacks will vary depending on the infrastructure being simulated, especially for infrastructures that have complex interdependencies among components. Thus, models of the specific infrastructure will be useful for developing a detailed and reliable infrastructure-impact scenario that shows how cyber attacks against a control system affect an infrastructure.

Economic Consequence

As mentioned earlier, the economic consequence tools are very flexible and can accommodate a variety of infrastructures, provided that the infrastructure-impact scenario can be mapped to a specific economic disruption. This mapping may be more difficult in infrastructures other than electric power. Most economic activity is highly dependent on electric power, but the same cannot be said for many other infrastructures. For example, a cyber attack on water treatment that resulted in a boil order would likely be more of an

inconvenience than an event that halts all economic activity. In the extreme case of an infrastructure-impact scenario where all water service was disrupted for a municipality, all economic activity would not be halted; much economic activity does not require water, and there are many common, alternative ways of obtaining water (such as wells).

More detailed economic consequence models, such as the National Infrastructure Simulation and Analysis Center (NISAC-Agent-Based Laboratory for Economics (N-ABLE™)), may be able to better model infrastructure disruptions that lead to more subtle economic disruptions than do interruptions in electric power. Heuristics can be used (or developed) to aid REAcct in mapping an infrastructure disruption to an economic disruption.

Case Study Review Questions

1. What is the five-step cyber-attack-consequence assessment process developed to coordinate Sandia's capabilities in assessing the effects of a cyber attack and in assessing the infrastructure impacts and economic consequences of those attacks?
2. What does it mean when infrastructures are interdependent?
3. What are the two extensions that can better integrate the different steps of the cyber-attack-consequence assessment process?

Case Study 3: Department of Homeland Security Battle Insider Threats and Maintain National Cyber Security

In 2009, the Department of Homeland Security's Office of the Chief Information Officer, Information Technology Services Office, and Risk Management Control Division were faced with the challenge of unifying 21 component agencies. Their challenge was to strengthen the components through the creation of one secure network and reduce the number of data centers. To do this, the DHS needed to coordinate centralized, integrated activities across components that are distinct in their missions and operations. With scores of administrators accessing key critical national infrastructure at these core data centers, the DHS' Risk Management Control Division was tasked with ensuring contained access and monitoring, logging, and tracking all administrative changes to its systems. In addition to stringent security policies, the DHS is subject to compliance regulations including Federal Desktop Core Configuration (FDCC) standards. Launched by the Office of Management and Budget in 2007, the FDCC ensures that federal workstations have standardized, uniform, desktop configurations to enable more consistent and better documented security while reducing costs. The DHS needed a solution that would allow it to support the component consolidation effort, transforming the 21 sites by unifying and controlling access to key servers at those sites while maintaining the separation of duties within and across the component agencies. It also needed a solution that could quickly and easily be dropped into technology already in place. This was a challenging task because the DHS has a wide range of platforms and operating systems, including mainframes, UNIX, LINUX, and Microsoft Windows.

The solution criteria were crystal clear. The DHS needed a solution that supported remote access, desktop virtualization, two-factor authentication, and auditing. It also needed out-of-the-box multi-platform support along with integration with existing cyber security products. As part of the selection process, the DHS vetted several cyber security products from a variety of market leading vendors. The DHS selected a cyber security product that provides access control for privileged users, including company employees, partners, consultants, and IT staff, along with the computing infrastructure. The cyber security product controls, contains, and audits the activity of privileged users, whether they originate from inside or outside of the network. The cyber security product also enforces fine grained access control policy on users, contains them to authorized systems and applications, and monitors, logs, records, and reports their activities for compliance and security risk management. This gives DHS control over its privileged users and high-risk assets. It also allows DHS to enforce access control policies and contains users in a manner that enables them to see only the network resources to which they have access. With an identity-based access control solution, the cyber security product provides the DHS with access control, user containment, and audit-quality logging in a single appliance-based offering. From an operations and risk perspective, this allows the DHS to granularly control who gets access to what servers, when and for how long in an easy-to-manage unified offering. The cyber security product also enables DHS to contain users from its 21 sites to authorized systems and applications without any reconfiguration of its network. The cyber security product's capabilities also addressed the DHS requirement to maintain end-to-end accountability.

Finally, the cyber security product has increased security awareness at the DHS. With the cyber security

product, the DHS has been able to provide privileged users with highly secure access to key servers in its facilities. As a result, the DHS has increased network security while enforcing the cyber security policy. The DHS has used the cyber security product to maintain FDCC compliance. It does this at the desktop level since the secure access is provisioned *via* a web browser without an additional desktop client required. The DHS has also used the cyber security product to streamline operations. This has been possible because the cyber security product provides a single solution for controlling, monitoring, logging, and tracking all administrator changes. Now, DHS can easily determine when a change was made and the implications of that change. The DHS derived several additional benefits from the appliance. First, DHS found the anti-leapfrogging capabilities beneficial, which contain users to authorized resources. Another benefit was being able to add keystroke loggers to administrative accounts and prevent them from doing any intentional or unintentional damage.

Case Study Review Questions

1. What are the cyber security solution criteria for the Department of Homeland Security?
2. How does the cyber security solution for the Department of Homeland Security (DHS) enforce fine grained access control policy on users; contain them to authorized systems and applications; and, monitor, log, record, and report their activities for compliance and security risk management?
3. How has the cyber security solution increased security awareness at the DHS?

Case Study 4: Cyber Security Development Life Cycle

A utility company's website is attacked by a botnet, a program built specifically to replicate malicious software on the web. It was spreading rapidly online by injecting itself into vulnerable websites and then waiting for unsuspecting users to click on the site. When they did, the code copied itself on their computers. In a few months, 360,000 sites had been infected. The botnet was diabolically engineered to sniff out the Achilles heel in SQL. The botnet co-opted an application on the company website and injected itself directly into a company database. The fear was that in the process, it could get past the utility's larger security perimeter and have its way with the company's software portfolio of applications, database tools, and other code. It also had the potential to install itself on the computers of anyone who visited the utility's website. The attack was a legitimate risk to the utility company.

The utility knew it wanted (needed) a new culture for how it engineered, developed, and tested its software. It also knew it wanted that culture grounded in widely accepted standards. That way, coders could learn from one another, and the company would not be reinventing its cultural wheel to make its software more secure. The catch was, no one in staff knew much about how to make applications safer.

The design phase of the cyber security development life cycle (CSDL) requires developers to create something called a cyber threat model. That is, a sense of the cyber attacks an application *might* face. What kind of exploits might a cyber attacker use? How would hackers gain access to an application running on a computer network? What older, existing pieces of code associated with the new application might be vulnerable? This overall feel for the risks an application might come under allows coders to *anticipate* risks. Threat models need not be complex: even high-quality ones can be done on the back of cocktail napkins.

Once the standard was set, critical areas were addressed and basic training was completed, next up was spreading the new cyber security culture inside the utility. Two basic lines of work emerged: remediation on the existing code where needed, and maximizing the cyber security of all new code created from that point on. The company-wide remediation was a copy of the early, high-level work on the website: carefully anticipating threats identified by the utility's version of CSDL, analyzing each threat, and then refactoring code where necessary. This strategic work was buttressed by scanning tools that helped identify high, medium, and low risks. But, despite this automatic assistance, it was immediately clear the work ahead would not be easy.

Time was something the utility's coders had little of. Its IT department was designed to be an internal resource for the coding needs of various departments: providing the company's energy traders with a new way to manage their inventory, helping human resources manage employee benefits, and planning how utilities route their electricity or gas. But under a mandate from the top, they found a way. And, slowly, cyber software security at the utility moved from afterthought to top of mind. Under CSDL, the utility now *started* with cyber security. Step 1 in the process was identifying a well-thought-out set of cyber threats that showed where a piece of software might be weak. How would the code be used? What was at risk? Then, using its new test tools and protocols, the entire development team became responsible for keeping the code within the standard. The utility had even gone so far as to install a last step—a human review to triple check that all new code cleared the cyber security bar before it went live.

Case Study Review Questions

1. What does the design phase of the cyber security development life cycle (CSDL) require developers to create?
2. Once the standard was set (critical areas were addressed and basic training was completed; next up was spreading the new cyber security culture inside the utility), what were the two basic lines of work that emerged?
3. Why is cyber security *not* absolute?

Case Study 5

Cyber security is an essential tool for managing risks in today's increasingly dynamic and capable cyber threat landscape. Yet, the market for cyber security remains small, and organizations are making only tactical investments in cyber security measures—one of the reasons why there has been an increase in cyber attacks. Evidence suggests that this trend will last for some time to come. However, the anticipation of an increasingly open and mobile enterprise should help refocus the spotlight on strategic investments in areas like cyber security. Cyber security professionals who wish to see cyber security move up in IT's priority queue should take immediate steps such as demanding secure software from suppliers and requiring rigorous acceptance tests for third-party code to help promote cyber security in the long run.

Because cyber security has a significant impact on vulnerability management, one could infer that the spotlight is only shifting to a different perspective and that commitment to cyber security may not have declined in the final analysis. Although viewed as a priority by many cyber security professionals, cyber security has not seen the appropriate commitment level reflected in IT's budget allocation.

For example, data breaches resulting from web application hacking are almost always accomplished through the exploitation of application vulnerabilities like SQL injection or cross-site scripting. If cyber security is not improved at a larger scale, the industry will continue to be plagued with security incidents that result in data breaches or other consequences that are even more disastrous. Changing the attitude toward cyber security, however, would require a culture shift, a shift that places importance on proactive risk management rather than immediate return on investment (ROI). This shift won't happen overnight. In the meantime, cyber security professionals should follow the below recommendations to implement a few immediate measures to effect positive changes:

- Demand software quality and security from suppliers.
- Perform stringent acceptance tests for third-party code.
- Disable default accounts from applications.
- Establish a secure operational environment for applications.
- Implement effective bug-reporting and handling.

As the buyer side starts to demand secure cyber software, the power balance will start to shift toward more strategic approaches to managing cyber-level risks. Cyber security professionals can encourage this change by engaging in the following longer term initiatives:

- Work toward an industry certification program for secure development practices.
- Implement a cyber security program.
- Continue to drive awareness of the changing cyber threat landscape.

So, to improve cyber security, companies and cyber security professionals should work in a concerted fashion to cultivate a culture that values and promotes cyber security. To help usher in such a culture, cyber security professionals should:

- Do their part to promote a cyber security ecosystem.
- Use mobile proliferation as a catalyst for cyber security.

Cybercriminals from China have spent more than 6 years cautiously working to obtain data from more than 70 government agencies, corporations, and nonprofit groups. The campaign, named Operation Shady RAT (remote access tool) was discovered by the security firm McAfee.

Although most of the targets have removed the malware, the operation persists. The good news: McAfee gained access to a command-and-control server used by the cyber attackers and has been watching, silently. U.S. law enforcement officials are working to shut down the operation. The Chinese government is denying that it sanctioned the cyber attack operation, although configuration plans for the new DoD F-35 stealth fighter were composed of the cyber attackers. So, with the preceding in mind, the following are five things that came to light:

- Seventy-two organizations were compromised.
- It was just not North America and Europe.
- When the coast was determined to be clear, the cyber attackers struck.
- This was a single operation by a single group (probably the Chinese).
- The only organizations that are exempt from this cyber threat were those that didn't have anything valuable or interesting worth stealing, from a national security point of view.

The loss of this data represents a massive economic cyber threat not just to individual companies and industries, but to entire countries that face the prospect of decreased economic growth in a suddenly more competitive landscape; the loss of jobs in industries that lose out to unscrupulous competitors in another part of the world; and not to mention, the national security impact of the loss of sensitive intelligence or defense information.

Yet, the public (and often the industry) understanding of this significant national cyber security threat is largely minimal due to the very limited number of voluntary disclosures by victims of intrusion activity compared to the actual number of compromises that take place. With the goal of raising the level of public awareness today, this is not a new cyber attack, and the vast majority of the victims have long since remediated these specific infections. Although whether most victims realized the seriousness of the intrusion or simply cleaned up the infected machine without further analysis into the data, loss remains an open question.

The actual intrusion activity may have begun well before 2006, but that is the earliest evidence that was found for the start of the compromises. The compromises themselves were standard procedure for these types of targeted intrusions: a spear-phishing e-mail containing an exploit is sent to an individual with the right level of access at the company, and the exploit when opened on an unpatched system will trigger a download of the implant malware. That malware will execute and initiate a backdoor communication channel to the web server and interpret the instructions encoded in the hidden comments embedded in the webpage code. This will be quickly followed by live intruders jumping on to the infected machine and proceeding to quickly escalate privileges and move laterally within the organization to establish new persistent footholds *via* additional compromised machines running implant malware, as well as targeting for quick exfiltration the key data that the cyber attackers came for. In the end, one very critical question remains unanswered: Where was the DHS all over this cyber breach during the last 6 years when "Operation Shady Rat" was alive and well? After all, isn't DHS supposed to be the security guardians of the cyber world?

If "Operation Shady Rat," wasn't bad enough, hackers are now using outfitted model planes/drones to

hack into your wireless system. Built from an old Air Force target drone, the Wireless Aerial Surveillance Platform (WASP) packs a lot of technological power into a flying high-end cyber endurance package.

Case Study Review Questions

1. To implement a few immediate measures to effect positive changes, what recommendations should cyber security professionals follow?
2. Cyber security professionals can encourage change by engaging in which longer term initiatives?
3. Which five things came to light, after cybercriminals from China spent more than 6 years cautiously working to obtain data from more than 70 government agencies, corporations, and nonprofit groups?

REVIEW

Answers to Review Questions/Exercises, Hands-On Projects, Case Projects, and Optional Team Case Projects by Chapter

Chapter 1: Introduction

Chapter Review Questions/Exercises

True/False

1. True
2. False
3. True
4. True
5. True

Multiple Choice

1. B and D
2. A and C
3. E
4. D
5. E

Exercise

Solution

The following steps will aid the cyber security team (CST) in coming up with incident response and recovery solutions, as a result of the cyber attack:

1. Establishing criteria and procedures to activate an information technology/information system (IT/IS) command center (partial or complete) during emergencies
2. Developing systems and/or procedures to determine what cyber-systems are affected by certain events
3. Establishing procedures to obtain information on a possible entry point for a cyber security violation
4. Developing procedures to evaluate firewall management and containment, and to respond accordingly
5. Establishing policies for the chief information officer (CIO) or IT/IS manager to direct key IT/IS staff in identifying potential problem areas
6. Developing communication methods for the CIO or IT/IS manager to issue organizational alerts regarding cyber-system failures or viruses affecting systems
7. Determining contact lists and communications methods in order for the CIO or IT/IS manager to immediately notify the nursing staff (nursing house supervisor) and/or senior medical staff (chief of staff) regarding affected cyber-systems that will have a direct impact on health care delivery and potential to adversely affect patient safety
8. Establishing procedures for emergency incident notification when affected systems will take greater

than 3 hours to return to full operational status to alert the incident commander and key disaster response personnel

9. Developing procedures for all administrators and key health care delivery staff to use manual documentation systems or nonaffected portable devices and later merging data with recovered systems
10. Establishing procedures to identify medical care, patient records, admissions, financial, supply management, computer-aided facility management (CAFM), and other critical systems and operations directly impacted by the cyber attack
11. Developing a plan to notify patients about any delays in service and the situation
12. Establishing procedures to ensure resources (personnel, equipment, software, and hardware) are obtained as appropriate to provide the fastest and most secure level of cyber-systems recovery
13. Developing procedures to implement regular briefings on cyber-system restoration status for personnel
14. Establishing predeveloped, departmental business continuity plans with clear recovery time objectives (RTOs) in place. Making sure that these plans have practice exercises
15. Developing criteria to restore normal operations
16. Establishing procedures to complete incident documentation and archiving
17. Developing procedures to debrief staff and identify corrective actions
18. Identifying components to include an after action report (AAR), including a cost analysis of time spent on restoration efforts
19. Establishing procedures to revise an emergency operations plan (EOP) as needed, including enhanced staff awareness training

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following partial project solution:

Your cyber security team (CFT) identified the intruder by network analysis. The outbound IP address for C2 (used to define the destination IP address) was flagged. A list of "notable" IPs collected via all source intelligence prepared. This means that a full content internal network collection allowed for the monitoring of the intruder as well as a collection of tools that were being utilized by the intruder. Reverse engineering of the tools disclosed similarities to known intrusion sets. In one instance, administrators had previously installed anti-spyware utilities, but could not rid system of its strange behavior. In summary, the CFT also identified that the e-mail:

- Was sent from a spoofed e-mail address.
- Messages were sent to the Executive Distrolist (a set of software components [open source components] assembled into a working whole and distributed to a user community).
- Contained a Trojan Horse Adobe PDF or MS Office attachment that contained real Adobe or Office documents.

- Contained a malicious injection file.
- Had reverse shell capability.

The recent exploit by the intruder took advantage of a memory corruption vulnerability in the JBIG2 filter in Adobe reader. Eventually, Adobe issued its first ever scheduled quarterly update for its Reader/Acrobat product line, a mega-patch covering 13 documented security vulnerabilities. This update resolves multiple heap overflow vulnerabilities in the JBIG2 filter that could potentially lead to code execution (e.g., CVE-2009-0509, CVE-2009-0510, CVE-2009-0511, CVE-2009-0512, CVE-2009-0888, and CVE-2009-0889).

Case Project

Solution

The following is a partial Department of Homeland Security (DHS) project solution. Students should be able to expand on the DHS project analysis through extensive research. Nevertheless, it becomes clear here that the following points should at least be developed or improved upon:

- Centralized command for cyber security needs to be expanded. In 2009, The Obama administration created the National Cybersecurity and Communications Integration Center (NCCIC). This is a 24-hour, DHS-led coordinated watch and warning center that will improve national efforts to address threats and incidents affecting the nation's critical information technology and cyber infrastructure. The NCCIC provides an integrated incident response facility to mitigate risks that could disrupt or degrade critical information technology functions and services while allowing for flexibility in handling traditional voice and more modern data networks. The new unified operations center combines two of DHS's operational organizations: the U.S. Computer Emergency Readiness Team (US-CERT), which leads a public-private partnership to protect and defend the nation's cyber infrastructure, and the National Coordinating Center (NCC) for Telecommunications, the operational arm of the National Communications System.
- Interests of national security should be integrated with international policy.
- Gathering, analysis, and sharing of information need to be improved.
- National and flexible insertion of cyber security expertise needs to be developed.
- A national policy with regard to defensive capacities needs to be developed.

Optional Team Case Project

Solution

The following is a partial solution to aid the participants and observer/controllers in coming up with their own solution to solve this case: Four overarching objectives were examined through the simulated exercise planning and execution period. A number of additional findings were identified through observations by participants and observers/controllers. The following are the simulated exercise's significant findings and supporting

observations:

1. *The value of standard operating procedures (SOPs) and established relationships*: Preparation and effective response are significantly enhanced by established and coordinated SOPs and existing relationships in the cyber response community. These SOPs and relationships facilitate rapid information sharing among community members.
2. *Physical and cyber interdependencies*: Cyber events have consequences outside the cyber response community, and noncyber events can impact cyber functionality. Fully acknowledging this reality is critical to refining comprehensive contingency plans and response capabilities. It is necessary to continue to converge and integrate response procedures tailored for physical crises and those developed for cyber events. The unique activities related to cyber response activities must be highlighted in cyber response processes and procedures to clearly reflect the inherent differences between cyber response and traditional physical/crisis response activities.
3. *Importance of reliable and tested crisis communications tools*: Tools and related methods developed and deployed for handling crisis communications need further refinement and enhancement. To maximize tools' efficiency and effectiveness during a crisis, the cyber response community needs to examine placement of tools, the impact of tools' capabilities and limitations on response procedures, and identification and authentication protocols used with the tools.
4. *Clarification of roles and responsibilities*: Substantial improvements were observed in the interagency integration and coordination of a cyber event response with senior leadership across interagency boundaries. Continued development and clarification of roles, responsibilities, and communication channels should further enhance capabilities.

Chapter 2: Deception

Chapter Review Questions/Exercises

True/False

1. False
2. True
3. False
4. False
5. False

Multiple Choice

1. A
2. B
3. A, C, and D
4. A, B, and E
5. B, D, and E

Exercise

Solution

The Advanced Persistent Threat (APT) evaluated the extent of the intrusion and identified compromised systems by collecting and analyzing volatile and static host data. They began the investigation by first looking for all of the signatures they were aware of including those collected during previous investigations, those provided by the client (the corporation that hired the APT), and generic indicators of system configurations that could signify system compromise. As data was collected and analyzed, the APT identified several new indicators that were unique to that client's environment. The team added them to the search list and scanned the entire network again to look for the new indicators.

During this process, the APT investigated 40,000 hosts, searched for over 360 different indicators of compromise, and provided the client a list of affected systems within the first four days on-site. The deceptive intruders had accessed the client's network multiple times over the course of 28 months. They gained initial entry through a phishing attack targeted at several senior executives. The company had responded to this first intrusion by discovering and removing malware from the victim machine. They had investigated a sampling of other machines in the organization and determined that no machines were compromised in the same manner.

At this point, the client believed the attack had been remediated successfully and the threat removed. But, in fact, the attackers had left undiscovered back doors on the network. The back doors allowed the attacker to continue pilfering the information from the company.

Once the list of known and suspected indicators was analyzed and all the suspicious system configurations were reviewed, the APT worked hand in hand with the client to create a remediation plan tailored to their circumstances. The plan included short-, medium-, and long-term strategies to protect their network from further attack while addressing the needs of various business units and senior management.

The client's ability to conduct a comprehensive investigation of the enterprise allowed the APT to identify all of the compromised hosts, all of the compromised user accounts, and all of the compromised data on the network. Moreover, the client didn't have to take any systems off-line and minimized the disruption to daily operations, all while doing so at a revolutionary scale and speed.

Because the problem was solved so quickly and completely, the client did not have to repeat costly remediation work or wonder if they had found the entire problem. The APT helped them respond to the incident on their terms, not the intruder's, thus saving their time, effort, and money.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

Leaders at Defense Information Systems Agency (DISA) used an innovative and pragmatic solution, one that enables the agency to simulate Internet-scale cyber war in a controlled environment by using a single compact device to:

- Generate large amounts of realistic user application traffic.
- Play canned scenarios with minimal configuration and update effort.
- Script simulation data flows.
- Support and emulate advanced networking protocols like MPLS and IPv6.

DISA relies on the device to support the requirements of each exercise, whether for deception threat injection, operating system types, patch levels, enclave machines, or network services. The latest in cyber range innovation empowers DISA to hold vendors accountable for delivering resilient devices that will not weaken the agency's infrastructure, and ensures that DISA's cyber warriors are fully equipped and trained to deal with new cyber attack threats on the critical national infrastructure by:

- Creating personalized and current cyber-warfare scenarios in a matter of minutes, replicating the very latest attack scenarios.
- Eliminating the need for dozens of different hardware and software systems to generate the appropriate levels of traffic and attacks—which saves time, eliminates complexity, and reduces costs by tens of millions of dollars.
- Accessing the very latest deception threat scenarios and global application protocols by using a single compact device.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research. Before certifying equipment and systems as production-ready, the bank's network security team must understand the true breaking point of each piece of equipment and of the network as a whole. To do this, they recreate Internet-scale network conditions to:

- Stress infrastructures to assess capacity and confirm appropriate IT investments.
- Validate functionality and defenses under high-stress load and deceptive cyber attacks.
- Measure the resiliency score of all IT elements prior to purchase, prior to deployment, and following configuration changes.
- Validate accuracy of data loss prevention solutions.

The team also keeps the bank's systems perpetually resilient by validating performance, cyber security, and stability, by rightsizing the overall critical IT infrastructure, while hardening the resiliency of network and data center devices and systems. The rigorous process that the team has put in place allows the bank to:

- Improve rigor in cyber security, performance, stability, and compliance processes.
- Balance security and performance of critical infrastructures to meet business objectives.
- Gain insight in advance as to how all devices, networks, and applications will perform prior to deployment.
- Reduce overall IT costs.

Financial services companies have much on the line when it comes to protecting sensitive customer data and ensuring the cyber security of the financial transactions that traverse their network and data center infrastructures. This major bank is leveraging the realism, ease of use, and scalability to optimize network and data center resiliency, both now and in the future.

Optional Team Case Project

Solution

The following is a partial solution to aid the participants and observer/controllers in coming up with their own solution to solve this case: Yahoo! recreated a realistic Internet-scale network conditions to validate the performance and cyber security of the company's network, data center, and application infrastructures. In particular, the security team used the device to:

- Standardize the validation of Deep Packet Inspection (DPI)-based network and security products prior to purchase and deployment.
- Understand in advance how high-performance, content-aware products will perform when faced with Yahoo!'s unique architecture and network conditions.
- Establish a standardized, deterministic, and vendor-neutral method for evaluating the resiliency of network and data center devices.
- Simulate real application traffic and user load to produce a series of measurements of server farms on the other end of server load balancers.

As part of the solution, Yahoo! was able to:

- Measure and add capacity to withstand peak load by subjecting applications, servers, and network infrastructure to application traffic from millions of users as well as cyber attacks.
- Improve the return on its IT investments.
- Use thorough and comprehensive protocol fuzzing to probe every possible weakness in the critical infrastructure.
- Validate defenses against large-scale distributed denial of service (DDOS) attacks.

Chapter 3: Separation

Chapter Review Questions/Exercises

True/False

1. False
2. True
3. False
4. True
5. False

Multiple Choice

1. A
2. B
3. C
4. D
5. E

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

If a cyber attack produces the death, damage, destruction, or high-level disruption that a traditional military attack would cause, then it would be a candidate for a use of force consideration, which could merit retaliation. The defense department's dependence on information technology is the reason why it must forge partnerships with other nations and private industry to use the separation of critical assets to protect the critical national infrastructure. This strategy will also state the importance of synchronizing U.S. cyber-war doctrine with that of its allies, and will set out principles for new security policies. The most sophisticated computer attacks require the resources of a government. For instance, the weapons used in a major technological assault, such as taking down a power grid, would likely have been developed with state support. Cyber attacks that have a violent effect are the legal equivalent of armed attacks, or what the military calls a use of force. A cyber attack is governed by basically the same rules as any other kind of attack if its effects are essentially the same. The United States would need to show that the cyber weapon used had an effect that was the equivalent of a conventional attack. Pentagon officials are currently figuring out what kind of cyber attack would constitute the use of force. Many military planners believe the trigger for retaliation should be the amount of damage (actual or attempted) caused by the attack. For instance, if computer sabotage shut down as much commerce as would a naval blockade, it could be considered an act of war that justifies retaliation.

Gauges would include death, damage, destruction, or a high level of disruption. Culpability, military planners argue in internal Pentagon debates, depends on the degree to which the attack, or the weapons themselves, can be linked to a foreign government.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

The project found that major issues remain with the communication between public and private sector organizations in the face of cyber attacks on the critical national infrastructure and in that groups' ability to piece together information to understand the scope of distributed threats. Among the findings detailed in the exercise was the conclusion that correlation of multiple incidents across public and private IT infrastructures remains a major challenge.

Although the cyber incident response community was generally effective in addressing single threats, and some distributed attacks, most of the exercises were treated as individual and discrete events, making it less likely for organizations to share data that could help point to widespread events. Exercise leaders indicated that threat response coordination became more challenging as the volume of cyber events increased.

The interagency communication within the government was acceptable, but needs further refinement, specifically the manner in which different bodies, including the federal government's Interagency Incident Management Group (IIMG) and National Cyber Response Coordination Group (NCRCG), work together. The contingency planning, risk assessment, and definition of roles and responsibilities across the entire cyber incident response community must solidify.

On the positive side, Cyber Storm found that the existing framework between international governments operated efficiently in terms of sharing information about domestic and international cyber attacks. The exercise made recommendations for improving performance in future tests, including more cyber threat training and simulation programs, more services to inform the general public about attacks, and new priority planning to deal with threats as they arrive.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

Although efforts to protect U.S. federal network systems from cyber attacks remain a collaborative, government-wide effort, the Department of Homeland Security (DHS) has the lead responsibility for assuring the security, resiliency, and reliability of the nation's information technology (IT) and communications infrastructure. Current measures to prevent future attacks and intrusion attempts should include the following:

- Hiring additional personnel to support the U.S. Computer Emergency Readiness Team (US-CERT), DHS's 24×7 watch and warning center for the federal government's Internet infrastructure

- Expanding the EINSTEIN Program to all federal departments and agencies
- Consolidating the number of external connections including Internet points of presence for the federal government Internet infrastructure, as part of the Office of Management and Budget's (OMB) Trusted Internet Connections Initiative, will more efficiently manage and implement security measures to help bring more comprehensive protection across the federal ".gov" domains
- Creating a national cyber security center to further our progress in addressing cyber threats and increasing cyber security efforts
- Expanding the National Cyber Investigative Joint Task Force (NCIJTF) to include representation from the U.S. Secret Service and several other federal agencies
- Working toward a stronger supply chain defense to reduce the potential for adversaries to manipulate IT and communications products before they are imported into the United States
- Facilitating coordination and information sharing between the federal government and private sector to reduce cyber risk, disseminate threat information, share best practices, and apply appropriate protective actions as outlined within the National Infrastructure Protection Plan (NIPP) framework
- Partnering with academia and industry to expand cyber education for all U.S. Government employees, particularly those who specialize in IT, and enhance worksite development and recruitment strategies to ensure a knowledgeable workforce capable of dealing with the evolving nature of cyber threats
- Partnering with other countries to locate and neutralize all hackers who pose a threat to the critical national infrastructure through the use of traditional military force

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

What all countries need to do is cooperate across international boundaries on their critical national infrastructure protection from cyber attacks by using separation. They also need to share information about what is going on and on how they have dealt with cyber attacks.

Chapter 4: Diversity

Chapter Review Questions/Exercises

True/False

1. True
2. False
3. False
4. True
5. True

Multiple Choice

1. B and E
2. A and C
3. D and E
4. C and D
5. E

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

Department of Defense (DOD) needs to do more to guard the digital storehouses of design innovation. Current cyber attack countermeasures have failed badly to provide companies robust protection for their networks. Classified threat intelligence should be shared with defense contractors or their commercial Internet service providers (ISPs), along with the know-how to employ it in network defense. Such intelligence helps the companies and their ISPs identify and stop malicious activity in their networks. The United States stands at an important juncture in the development of the cyber threat. More destructive tools are being developed, but have not yet been widely used. The defense department needs to develop stronger defenses, before those who mean harm to the United States gain the ability to launch more damaging cyber attacks. The United States has a window of opportunity in which to protect networks against more perilous threats. Through information sharing, intrusions need to be halted by learning more about the diversity of techniques used to perpetrate them.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

IT administrators must create global network user interactions and traffic conditions to validate the resiliency of network and data center equipment. Resiliency is measured for devices and systems while they are under the burdens of traffic from hundreds of real applications, load from millions of users, and comprehensive security attacks, obfuscation, and evasion techniques. The process assesses and validates the resiliency of cyber infrastructure, but beyond that it can be used to certify resiliency on a scientific basis. Resiliency certification provides critical assurance to government agencies tasked with protecting the critical network infrastructure of the United States, and to global enterprises that must secure customer information and valuable intellectual property. By putting systematic resiliency scoring in place, government agencies, enterprises, and service providers can:

- Confidently assess the risk associated with any network element.
- Hold equipment manufacturers accountable for their claims.
- Use the information to harden critical national infrastructures against cyber attacks.
- Identify weak equipment for replacement or reconfiguration.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

Important issues began to show up like the need for executive support and how to get everyone onboard as the company raised security development as a central focus for their internal development group moving forward. The company validates the need to make deep changes when necessary within the software development culture versus performing cyber security around the edges. Other important insights detail how an aggressive timeline created focus and gave everyone a clear goal. The company was able to significantly reduce the number of vulnerabilities and meet its diverse and resilient security goals while setting the company up for long-term success.

What was found of particular interest by the company was that after it went through this experience, it not only created more secure applications, but also found something it hadn't counted on. The company's process requirements and the resultant engineering culture shift had brought together the entire development organization with quality assurance in a way it hadn't seen previously. Together, they engaged in a cyber security development lifecycle process, and as a result there were fewer cyber security bugs that were found and needed to be fixed late in the process—when it is most expensive. The company saw a real productivity gain out of their development organization, not just better application infrastructure security.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

This cyber attack incident was handled by the computer emergency response team (CERT) in Estonia, with different Estonian authorities, in close cooperation with CERT organizations from all over the world. CERT handles security incidents that take place in Estonian computer networks, and takes measures to prevent such incidents and raise the security awareness of users. On the state level, CERT's tasks are performed by the Department for Handling Information Security Incidents of the Estonian Informatics Center. To subdue the cyber attacks and to maintain services within the country, the access of foreign users to the targeted websites had to be restricted for extended periods of time. Consequently, paralysis of e-commerce by spam and denial-of-service attacks on the public sector websites had a particularly crippling effect on the country. During the worst of the cyber attacks, attempts to access a few sites with Estonia's national domain name ".ee" resulted in problems in loading page messages.

Chapter 5: Commonality

Chapter Review Questions/Exercises

True/False

1. False
2. False
3. False
4. True
5. False

Multiple Choice

1. A
2. D
3. B
4. E
5. A

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

During the cyber attacks, the DHS learned that the hypothetical malware originated from a group of servers in Russia. The United States had the offensive capability to shut down those servers. But would Russia see that as an act of war? Currently, the United States does not have a well-developed policy to respond to major cyber attacks. While the DHS was debating how to respond to the cyber attacks, the electrical grid in the eastern United States began shutting down. In what appeared to be an attack coordinated with the smartphone malware, pipe bombs exploded at two energy facilities in the United States, causing a major gas pipeline to shut down. An ongoing heat wave contributed to problems of electrical blackouts.

As the blackouts began to cover much of the northeast and several large midwestern cities, DHS began talking about mobilizing the National Guard and active-duty military members to protect electricity-generating facilities and prevent civil unrest. The DHS suggested the military help for delivering diesel fuel to hospitals in areas where there were blackouts. Hospitals have backup generators that run on diesel, but the generators can only run for 6–12 hours without additional fuel. People, after about 12 hours, were going to start dying in hospitals.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

Certain desirable security attributes must be present in all aspects and areas of the national infrastructure (like banks) to ensure maximal resilience against cyber attacks.

- Raising cyber-threat awareness among business continuity and disaster recovery executives and managers
- Reviewing recent cyber threats and their potential impact to the critical national infrastructure systems
- Discussing how to effectively manage these cyber threats from operational and business perspectives
- Examining existing and necessary commonality of information sharing and incident response processes to address the cyber threats

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

In this project, the European member states will need to cooperate with each other to avoid a simulated total infrastructure network crash. The event was organized by EU member states, with support from the European Network and Information Security Agency (ENISA) and the Joint Research Centre (JRC). This project scenario was followed by more complex scenarios—ultimately going from the European level to the global level. Supporting EU-wide cyber security preparedness exercises is one of the actions foreseen by the Digital Agenda for Europe to enhance online trust and security. In the simulation, citizens, businesses, and public institutions would have difficulties to access critical infrastructure online services (such as e-Government), unless the traffic from affected interconnections were rerouted.

This cyber security project exercise aimed to enhance the understanding of EU member states' of how cyber attacks are handled and test infrastructure communication links and procedures in case of a real large-scale cyber attack. The project exercise tested the commonality of contact points in the participating countries, the communication channels, the type of data exchanges over these channels, and the understanding that EU member states have of the role and mandate of their counterparts in other EU member states.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The utility must find the vulnerabilities of the supervisory control and data acquisition (SCADA) systems and how they can improve their control of such vulnerable cyber-systems. In addition, they must ensure that their staff are adequately trained to manually operate and check all of the utility's systems in the event of a SCADA failure.

Chapter 6: Depth

Chapter Review Questions/Exercises

True/False

1. False
2. True
3. True
4. True
5. True

Multiple Choice

1. B
2. A and E
3. C and D
4. A and B
5. B and C

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

The utility must discuss the vulnerability of SCADA systems and how they can improve their control of such vulnerable cyber-systems. In addition, the utility must work to restore the drinking water supply and cooperate with the fire department to restore fire flow or to establish an alternate water source.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

The cyber attack should have been prevented in the first place, by the introduction of multiple layers of defense, in order to increase the likelihood that a given attack will be stopped or at least slowed down. This likelihood is dependent upon the quality and relative attributes of the various defensive layers to prevent:

- Trojaned e-mail from being sent from spoofed e-mail addresses.
- E-mail messages from being sent to the Executive Distribution list.
- The Trojaned Adobe PDF or MS Office attachment from containing real Adobe or Office documents, malicious injection files, and reversed shell capability.

- The Recent exploit from taking advantage of a memory corruption vulnerability in the JBIG2 filter in the Adobe reader.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

The organization should be able to reduce potential vulnerabilities, protect against intrusion attempts, and better anticipate future threats by preventing USB-delivered malware from:

- Infecting the USB memory stick carrying Trojan.
- Having multiple variants, which caused the malware not to be detected.
- Clearly targeting specific organizations and computing infrastructures.
- Establishing the C2 with communications back to the external locations.
- Relying on the Windows auto-play feature Autorun.inf on infected USB points to the malware.
- Being located in the RECYCLER folder on the device.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The organization should prevent this type of cyber attack, by taking countermeasures like authenticated access, to prevent the following from being stolen:

- All user-generated data
- System files, executables, or other common files
- Personally identifiable information (PII)
- Research documentation, proposals, and proprietary information
- System information that is used to attribute the exfiltrated data
- System configuration
- Network structure
- Mapping of the internal network
- Target lists for lateral movement

Chapter 7: Discretion

Chapter Review Questions/Exercises

True/False

1. True
2. True
3. False
4. False
5. True

Multiple Choice

1. A and C
2. B
3. A and E
4. C and D
5. A and D

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

- The strength of the private sector companies will be tested in regard to prevention and deterrence.
- Although the physical infrastructure was not at great risk, Internet software did deteriorate, and numerous systems had to be repaired.
- All attacks must be recognized in the future. Each attack ended before anyone had enough time to completely diagnose the problem.
- Emergency response will be split between technically bringing systems back online and instituting business continuity process, and controlling the public perception of the situation to restore confidence and prevent panicky behaviors.
- All internet service providers (ISPs), domain name service (DNS) operators, and other organizations will need to evaluate their network topologies, diversity, integrity of backup processes, and other methods of attack prevention.
- Primarily, victim "care" will be based on economic assurance. Citizens will look for government assurances that the Internet is a stable and viable method for conducting business and other financial operations.
- Using intelligence and law enforcement sources and methods, the investigators will need to determine

the likely technical source and the identity of the perpetrators.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

The First Actions

- The shift supervisor immediately contacted the IT manager and systems administrator.
- The IT technical team came in within a few minutes.
- After a very quick investigation, the IT technical team found the hospital's emergency department's computer network had shut down.
- Immediately, the IT technical team declared that the department's network was under attack.

Countermeasures

- Organizational level
- Use up-to-date protection tools such as a firewall, spyware, antivirus, Intrusion Detection System (IDS), and Intrusion Prevention System (IPS) within the entire network.
- Outline and implement effective user policies for the hospital's computing facilities.
- Outline and implement strong user policies for remote users of the hospital's computing facilities.
- Develop guidelines and policies to control internal users' Internet usage activities.
- Implement user access control and monitor systems to analyze, detect, and prevent internal misuse and unauthorized access to sensitive information within the network.
- Implement policies and systems to back up all data so that in case of any unusual event, data can be recovered.
- Have a redundant network running for critical computer systems to continue daily operation in case of a cyber attack on a network or network segment.
- Maintain effective security techniques for wireless and voice-enabled communications systems.
- Periodically scan network resources for vulnerabilities and fix them.
- Educate and train employees about cyber security and best practices.
- Finally, have a cyber security committee.

- End-user level
- Control access to your Computer
- Consider security before making online transactions.
- Update operating systems and software regularly.
- Use strong passwords, and change them often.
- Lock down your system.

• Disable all unnecessary services.
• Use encryption to protect your communications.
• Install (should be installed by IS&T) and regularly update antivirus software.

• Control access to your information
• Think about security before you act.
• Be concerned about "social engineering" tricks.
• Be cautious about the possibility of "spoofed" websites and e-mails.
• Beware of giving out your personal information.
• Know with whom you are working or conducting business.
• Do not open suspicious or unknown e-mails.
• Be cautious about downloading free software and programs.
• Look for privacy and security policies.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

An analysis platform software should be used to help proactively uncover online and in-store fraud by consumers, dealers, and organized crime groups and prevent the subsequent loss of products and revenue. The solution here is to detect fraudulent account activation attempts, enabling the organization to automatically analyze sensitive customer information (including phone number, bank account, postal address, IP address, or any other distinguishing attribute) across disparate data sources. The sensitive information is scored using sophisticated algorithms and compared against historical sensitive information to reveal possible matches, such as if the postal address for an order matches the address of a known fraudster; if orders from several people are placed from the same IP address; if two customers have similar names; or, if dealers circumvent the vetting process by using their own credit cards and taking cash payments. The algorithms take into account the company's risk based on product cost along with data from usage monitoring and payment collections to provide a comprehensive view of customer behavior. With this insight, the staff can identify new patterns fraudsters are using to reduce fraud across the industry. The company also has a comprehensive view of system controls, so it can identify and fix missing controls in the activation system at a lower cost. The Telecom didn't experience the financial losses its competitors did from the iPhone theft. The sensitive information gathered also helped lead to the arrest of nearly a dozen people participating in a subscription scam.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The IT staff faced a number of difficulties in solving their laptop problems. Although they needed to protect the sensitive data on laptops, they didn't want to encrypt the entire disk drive. They also implemented a full-disk encryption before, and had encountered problems with running their applications, as well as with corrupted data. The utility company decided to use off-site tape vaulting to securely store tapes from two locations. The IT staff needed to protect employee laptops properly. The use of off-site tape vaulting protects sensitive and confidential data on all computers, including laptops, whether in or out of the office. Off-site tape vaulting also supports regulatory compliance through intelligent file and folder encryption. If a computer is lost or stolen, off-site tape vaulting can destroy sensitive data automatically to prevent it from falling into the wrong hands. The encryption method is possible at the file level, rather than at the drive level. This encryption method is less intrusive for users than for other solutions. Should a user fail to log on correctly, it is a simple matter for IT staff to use their encryption key to decrypt files. Access to the off-site tape vaulting can be done remotely. This ensures that all sensitive laptop data is always backed up and available for restoration.

Chapter 8: Collection

Chapter Review Questions/Exercises

True/False

1. False
2. False
3. False
4. False
5. False

Multiple Choice

1. B, D, and E
2. A and C
3. A, C, and E
4. B
5. D

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

Data is collected only for supporting decisions at the critical infrastructure network level. Decisions on what data to collect are mainly on the basis of experience. The districts collect data and update information in a central database. The network-level critical infrastructure data collection framework is revised biannually. Project-level data collection policies are quite mature. There are many offices involved with asset management critical infrastructure data collection activities, more specifically, 10 district offices, 56 residences, and 357 area headquarters.

The information is updated annually and stored in a corporate database platform. The pavement and bridge data are kept in separate databases that feed the central database once a year. The accuracy needed for the various data items has not been addressed objectively. However, there are good-quality assurance procedures in place to check the information collected and input to the system. The agency is planning to conduct a sensitivity analysis to determine which data items have the most impact on the decisions.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

The agency started a pilot to inventory hydraulic facilities in 1999. As part of the inventory process, the agency hired contractors to reference each individual node using GPS. These inspectors conduct the initial inspection of the facility. A second inspection is triggered if the performance-based rating is 4 or 5.

Originally, the agency developed the inventory using available plans; however, the process has evolved, and the data is currently verified in the field. Inspectors take the plans to the field, locate the facilities again, and reference their location using GPS. The agency has prepared a detailed inspection manual for engineers to use when they assess these facilities. The condition of the storm water management facilities is evaluated periodically by using two rating systems:

1. *Performance-based rating*: This is a performance evaluation of facilities for functional and structural integrity. Each facility is rated according to 45 items using a subjective 1–5 scale.
2. *Response rating*: This provides an estimate of the level of work required (how the structures will be maintained) and the priority for maintenance and remediation. The facility receives an overall inspection rating from A (no action required) to E (facility failed, hazardous conditions). Facilities with a rating of C or below are candidates for remediation.

Small drainage structures are added when possible. While inspectors collect GPS data (inventory) on the facility, they can easily take a picture and provide a brief rating. Given the number of these facilities (hundreds of thousands), the agency cannot afford to inspect these smaller drainage systems on a regular basis. Most of the storm drainage networks are under the roadways and require video inspections because it is hard to get people into some of these facilities.

The storm water management division spends approximately $2 million on critical infrastructure data collection (inventory and inspection) per year. The critical infrastructure data collection cost was approximately $1 million for the collection of the initial inventory data for one county; however, updating the information the second time was cheaper by approximately half of the initial cost. The most cost effective critical infrastructure data collection procedure uses handheld PDAs and GPSs.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

This agency is conducting a major project to inventory all assets in its transportation system and is developing transition plan. The assets are categorized into 45 groups. A consultant collects the inventory data in the pilot program, and some data also is reduced from the camera information collected for the pavement management system. In addition, the agency collects data for several of its main assets as part of its individual management systems:

- *Pavements*: The agency conducts annual condition assessment through contractors to identify different distresses on pavement for each block segment and computes a pavement condition index (PCI [on a scale of 0–100]) on the basis of the individual distress types. Other performance measures such as the

international roughness index (IRI) also are collected. The agency is expanding its pavement critical infrastructure data collection by changing from one lane in each direction to every traveled lane in both directions. The agency used to collect citywide pavement data every year, but the frequency has been reduced to every other year for local pavements and every year for federal pavements.

- *Bridges*: Consultants conduct annual bridge condition assessments. If the contractor identifies some critical problem that needs to be fixed immediately, the maintenance contractor will receive a task to fix the problem from the agency. Approximately 40% of the bridges are assessed every year.
- *Sidewalks and alleys*: There is an ongoing effort to develop a 7-year sidewalk-and-alley rehabilitation program. Data collected include GPS, pictures, length (extension), condition, and maintenance needs, among other attributes, and sidewalks and alleys are still in the process of inventory. The project selection will be based on both agency assessments and requests from citizens.

Public opinion also was considered in the performance rating process. The agency organized a tour of the transportation system every 3 weeks as part of the assessment for the asset preservation contract to evaluate cleanliness. The evaluators were provided with rating manuals before the tour, and they rated the various management units—three-block segments in the inner city and 2-mile segments on the highway. The results were forwarded to the contractor to help him prepare his work plans. The public also can review the work plans and give comments and suggestions. The process enables the agency to maintain good communication with its neighborhoods and effectively provide the needed services.

Both the central office and individual branch agencies (parking, street lights, signs, and bus and mass transit) collect data on the assets that are their responsibility. The asset management division collects more detailed performance data, whereas individual branch agencies collect data that are more related to the selection of appropriate maintenance treatments. The agency plans to make the asset condition information easily available to the public.

Whereas some of the information is stored in a central database, most of the individual branches maintain their own databases. There is a central repository (web portal) that was developed that includes all documentation related to the project. The agency is investigating the possibility of expanding it for the entire city.

Critical infrastructure data collection methods differ from asset to asset. For sidewalks and alleys, the contractor conducts a walking, visual inspection using laptops and PDAs. Pavement data is collected using automatic data collection vans equipped with cameras. Access equipment and laptops are used to support the bridge data collection.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The road maintenance unit has a maintenance condition assessment program. The program periodically evaluates the condition of certain elements, collects and organizes the data, and analyzes the results to determine the maintenance level of service of the road system. Every 2 years, the road maintenance unit

conducts a condition rating on a statistically representative percentage of the highway systems. The sample size was determined using a statistical analysis that allows determination of the repair cost.

The evaluation is used to estimate percentage of asset at each performance level. The sampling is performed on rural interstate as well as primary and secondary road systems. The urban road systems also are included in the inspections; however, the agency does not plan to continue to include urban roads in future campaigns. The critical infrastructure data collected and the performance measures for different units were determined through committees that have been formed for the different types of assets. The committees will also decide what specific work activities are associated with particular performance measures.

The agency units involved in critical infrastructure asset data condition collection are road maintenance, pavement management, and bridge maintenance. The road maintenance unit employs two or three critical infrastructure data collection teams for each division. The pavement unit uses one team per county, and the bridge unit conducts the assessment using one team per division. Extensive training including distribution of detailed manuals is offered to the critical infrastructure data collection teams before the operations. In addition, three quality assurance teams from the road maintenance unit assure the quality of the critical infrastructure data collected for the entire state by reevaluating the highway samples (randomly selected) covered by the division teams.

Road maintenance data is stored in a central database, but this data is not integrated with the pavement and bridge management data, rather each division maintains its own condition database. In addition, preprocessed data from each unit is periodically fed to a state data warehouse, but no analysis has been done using these data. It is expected that once the new asset management system is implemented, the output could also be used in the maintenance management system. The maintenance management system critical infrastructure data collection is done manually.

Chapter 9: Correlation

Chapter Review Questions/Exercises

True/False

1. True
2. True
3. True
4. True
5. True

Multiple Choice

1. B
2. A
3. E
4. D and E
5. C

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

The bank's IT security staff then immediately alerted the appropriate IT administrator. Since they had also enabled the technology's active response capability, the security staff disconnected the machine at the network interface controller card level without any human intervention. IT administrator couldn't be there at 10:00 p.m., but the IT security staff could (stepping in to protect their network from potential abuse), in this case, it was the janitor.

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

By watching for exploit attempts that follow the reconnaissance activity from the same source IP address against the same destination machine, the cyber security administrator can increase the confidence and accuracy of reporting. After the reconnaissance event is detected by the system, the correlation rule activates and waits for the actual exploit to be reported. If it arrives within a specified interval, the correlated event is generated. The notification functionality can then be used to relay the event to cyber security administrators

by e-mail, pager, and cell phone or invoke appropriate actions.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

Although some intrusion detection systems are able to alert on failed login attempts, the correlation system is able to analyze such activity across all authenticated services, networked (such as telnet, ssh, ftp, Windows Access, etc.) and local (such as UNIX and Windows console logins). This correlation rule is designed to track successful completion of such a cyber attack. The triggering of this correlation rule indicates that a cyber attacker managed to log in to one of your servers. It is well known that system users would often use passwords that are easy to guess from just several tries. Intelligent automated guessing tools, available to hackers, allow them to cut the guessing time to a minimum. The tools use various tricks such as trying to derive a password from a user's login name, last name, etc. In the case that those simple guessing attempts fail, hackers might resort to "brute forcing" the password. The technique uses all possible combinations of characters (such as letters and numbers) to try as a password. After the nonroot (nonadministrator) user password is successfully obtained, the cyber attacker will likely attempt to escalate privileges on the machine in order to achieve higher system privileges. The correlation rule activates after the first failed attempt is detected. The event counter is then incremented until the threshold level is reached. At that point the rule engine would expect a successful login message. In case such a message is received, the correlated event is sent. It is highly suggested to tune the count and the interval for the environment. Up to three failed attempts within several minutes is usually associated with users trying to remember the forgotten password, while higher counts within shorter period of time might be more suspicious and indicate a malicious attempt or a script-based cyber attack.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The goal was to stop intruders before they could cause any damage by correlating and analyzing abnormal events flowing in through the networks. For instance, an event correlation engine could be used to collect messages and log entries from many different devices on the network and infer the relationships among them. For example, if someone repeatedly attempted and failed to log into a workstation, those brute force login attempts were picked up by the system. If somebody changed the IP address of a flow computer, which should rarely be done, that event raised alerts. There were even correlation rules in place that detected rogue users on the network who hadn't been previously identified. By intelligently piecing together the connections among many disparate events coming into the control center, the system could filter out much of the noise, identify significant patterns, and ultimately provide the big security picture to plant operators.

Chapter 10: Awareness

Chapter Review Questions/Exercises

True/False

1. True
2. False
3. True
4. False
5. True

Multiple Choice

1. C
2. B
3. A and D
4. B
5. D

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

The U.S. planners reviewed several of the situational awareness cyber security exercises held in their country and determined four overarching trends from the findings:

- Establish a baseline of participant skills and knowledge, and organizational preparedness capabilities before conducting a situational awareness cyber security exercise.
- Focus coordination across stakeholder groups on reducing perceived barriers to collaboration and information sharing among participants.
- Establish clear triggers and thresholds for identifying and reporting a cyber attack.
- Clearly define federal government involvement in local cyber attack management

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

Upon investigating the situation, it was discovered that one of the workstations had become infected with malware after an employee clicked on a link in a phishing e-mail. All of the servers and a number of

workstations were compromised, giving cyber criminals full access to the network. The company's logs revealed that the webserver was being used to host an illegal music download service, and it was also discovered that the perpetrators had installed hidden rootkits. The disinfection of the company's network required considerable time and expense. The company spent a considerable amount of hours correcting the problems associated with the network breach, including:

- Selecting, ordering, configuring, and installing a quality firewall
- Building a new webserver, uploading digital backups, and bringing it nearline
- Scanning all servers and workstations with several anti-malware tools to locate rootkits
- Wiping and rebuilding Windows on all workstations to ensure removal of all rootkits
- Installing anti-malware software on all servers and workstations
- Bringing a new webserver online and debugging the initial problems
- Repairing things that broke during the rebuilding, installing drivers, bringing printers back online, etc.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

The government cyber security teams used off-the-shelf cyber security software to examine the IP packets in real time by identifying the communications protocols used and extracting flow/session/application content or metadata. This allowed the government cyber security teams to deploy more sophisticated cyber security solutions faster and more affordably than having to develop highly advanced security technology entirely on their own. The off-the-shelf cyber security software was used by the government cyber security teams to develop a custom barrier in order to deal specifically with the abnormal threats that evade most off-the-shelf products. The information retrieved from the network allowed the government cyber security teams to improve their situational awareness so they could take preventive measures and protect sensitive assets. This second line of defense improves detection and mitigation of real and covert cyber attacks that can compromise networks, and hastens response to new threats before they can be implemented by off-the-shelf vendors, enabling government agencies to keep strategies confidential and minimize exposure to cyber threats.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The challenge lies in anticipating and avoiding the effects of adversity, and this depends on a highly refined situational awareness. So, it is in the area of operation sensing and monitoring that a game-changing innovation can be found. What is needed is to obtain a digital situational awareness so as to anticipate cascade triggers in the critical national infrastructure and deploy effective distributed supervisor control protocols that can avoid these triggers. Digital situation awareness can be derived from traffic flow and volume. The method

envisioned to anticipate and avoid cascade triggers in the critical national infrastructure is based on traffic flow and volume and is specified as follows:

- Identify industry sectors of interest to cyber security resiliency.
- Identify each enterprise and organization in each industry sector of interest.
- Identify each computer system of interest in each enterprise and organization.
- Identify each I/O port on each machine of interest.
- Record traffic flow and volume on every port for every second of every day for up to 12 months.
- Determine expected normal operation, using recorded traffic flow and volume, based on derived upper and lower control limits for varying time intervals.
- Derive operating protocols, such as shutdown, switch to backup, and switch to a designated alternate mode, using traffic flow and volume scenarios, for use by intelligent middlemen charged with distributed supervisory control of critical national infrastructure operations.

Chapter 11: Response

Chapter Review Questions/Exercises

True/False

1. False
2. True
3. False
4. True
5. False

Multiple Choice

1. D
2. C and D
3. E
4. C
5. E

Exercise

Solution

The following is a partial exercise solution. The students should be able to expand on the following:

A cyber security team was brought in within 12 hours of notification. The team augmented local staff, providing management and leadership of the cyber security incident, and accomplished the following tasks during 3 weeks of on-site fieldwork:

- Forensic analysis on computer systems
- Live response on computer systems
- Scanning over unknown binaries for malicious content, and manually analyzing binaries of interest
- Providing short-term remediation plan for anonymous proxy software
- Providing a short-term remediation plan for the intrusion
- Providing a long-term remediation plan, documenting recommended countermeasures, process recommendations, and additional cyber security measures to help prevent, detect, and manage future cyber security incidents
- Furnishing a letter providing cyber security team's opinion regarding potential data loss during the intrusion
- Implementing scripts to detect and monitor unauthorized activity on the client networks
- Developing host-based and network-based signatures related to the intrusion set to determine the

scope of the cyber security incident

- Identifying additional hosts requiring remediation
- Developing illustrations that provided visualization of the cyber attack, the timeline of the attack, and the scope of the intrusion

Hands-On Project

Solution

The following is a partial project solution. The students should be able to expand on the following:

The utility's cyber security administrator was able to utilize a combination of people, process, and technology solutions to manage and analyze important, actionable cyber security information and distribute that information to the right people at the right time. The cyber security administrator implemented a computer emergency response team (CERT) across IT and other business units to collectively identify, analyze, and mitigate cyber security threats and vulnerabilities. Additionally, the cyber security administrator implemented a third-party intelligence service product to communicate cyber security threats that are applicable to the utility's computing environment and assisted in prioritizing threats for the CERT based upon threat credibility, severity, and risk.

Case Project

Solution

The following is a partial project solution. Students should be able to expand on the project analysis through extensive research.

The solution uses an out-of-band, policy-driven architecture to deliver centrally managed visibility and access control across wired, wireless, and virtual private network (VPN) environments. Elements in the network environment, including switches, wireless access points, and VPN concentrators, are leveraged to gain visibility of all connected users and devices and to enforce access policies at the edge of the network.

The tool's architecture also allows it to be deployed in phases to meet unique requirements of different organizations. For example, deploying first in *monitor-only* mode provides network-wide visibility of all users and endpoint devices on the network while being completely transparent. This allows the financial institution to *baseline* the network to determine whether users and endpoint devices are compliant with cyber security policies without adversely impacting anyone's network access.

The financial institution can then move on to enforce access policies in later phases of deployment. Advanced capabilities, such as device profiling and securing guest access, can be added in later phases as well, without needing to deploy additional hardware or reconfigure the system. This gives the financial institution customers the ability to adapt a network sentry platform to their own environment.

Optional Team Case Project

Solution

The following is a partial solution to aid students in coming up with their own solution to solve this case.

The university sought to upgrade its intrusion prevention solution (IPS) from the outdated device sitting on its network to a third-generation, purpose-built IPS appliance. They realized that there were fundamental issues with relying too heavily on signature-based protection for their network.

The cyber security manager and his team engaged in an extensive evaluation of the current solutions in the market that were most viable over the long term. The cyber security team examined the actual protocol of the cyber attack, as well as traffic behavior and characteristics.

The cyber security team looked at off-the-shelf cyber security products that went beyond their strong distributed denial-of-service (DDOS) protection history. They were looking for technology that was protecting customers from all types of malicious content, undesired access, and botnet-based attacks. As new threats beyond simple worms and rate-based threats had emerged, the cyber security team knew that they had to keep up, evolving and enhancing their technology to help their customers face these new challenges.

The cyber security team then adopted a phased approach in testing and mimicking cyber attacks and chose the best off-the-shelf security product out of other solutions. They compared their network's traffic and behavior with the entire cyber security rule set active, including signatures, protocol anomaly protection, application usage awareness, deep-packet inspection, and firewalls.

Finally, the cyber security team chose an off-the-shelf cyber security product that maximizes protection for critical IT assets, while allowing full access to legitimate users and applications, and protects against malicious content, undesired access, and rate-based attacks such as DDOS. The cyber security product can be deployed at the network perimeter, on internal network segments, on remote site locations, or at the network core to protect assets and stop attacks, delivering high performance, low network latency, reliability, and high availability.

Index

A

Access control lists (ACLs)
LAN controls, 14
layered access controls, 152

Access controls
functional separation, 65
layered, 152–154, 153f
national infrastructure SSO, 148
remote e-mail, 149–151
separation principle, 14

Access paths, national depth program, 159

Access techniques, separation principle, 13–14

Accuracy
data collection, 225
intelligence reports, 248
national infrastructure firewalls, 69
SCADA systems, 74

ACLs, *see* Access control lists (ACLs)

Actionable information, 230

Active opportunism, vulnerability information management, 245

Actual assets
and honey pot, 41
separation techniques, 66

Administrators, as infrastructure decision-makers, 129

Adversary separation, 65–66

Adversary types, 5, 5f, 9–11

Aggregation

data collection, 192f
system data collection, 198

Air gapping, 75

Alarm feed, intrusion detection information, 158f

Alarms
and actionable information, 230
false, 49–50
intrusion detection, 222
layered intrusion detection, 156–157
SIEM threat management, 222
SOC, 252

Alarm streams
and correlation, 217, 223f, 224
SIEM, 201

Algorithms
and actionable information, 230
antispam, 218
attack detection, 243
and awareness, 25
and collection, 23, 191, 204
and correlation, 7, 24, 221–222, 231
DDOS filtering, 72
and discretion, 21
vs. human interpretation, 247

Amplification approach, DDOS filtering, 73

Analysis objective, as deception, 38

Antivirus software
botnet detection, 152
and botnets, 8
and correlation, 218, 225
relevance, 152
as safeguard, 12
and separation, 15, 63
system data collection, 198

Apple®, PC diversity, 95–96

Application interoperability, and desktop diversity, 97

Application strengthening, encryption methods, 155

Approach factor, separation techniques, 66

Approved external site, and obscurity layers, 178–179

Architectural separation
  layered e-mail filtering, 151f
  trusted networks, 15–16

Architecture issues, and awareness, 240

Asset identification, national depth program, 159

Asset separation
  DDOS and CDN-hosted content, 81f
  overview, 80–81

Assistance questions, TCB, 170

Attack confidence, event dependence, 244f

Attack entry points, and deception, 39–40

Attack metric pattern, early warning process, 265

Attention objective, as deception, 38

Attribute differences, and diversity, 92, 92f

Attributes, information reconnaissance, 169

Audits
  and best practices, 18
  and collection principle, 23–24
  definition, 116
  importance, 18
  meaningful vs. measurable best practices, 118
  national infrastructure simplification, 128
  organization examples, 116f
  purpose, 116
  security, *see* Security audit

SIEM threat management, 222

Authentication
layered, 147–151, 149f
remote e-mail access, 149–151
separation principle, 14

Authorized services, deception scanning stage, 42

Automated attacks
botnets, 53, 55, 272
propagation, 92, 98–99
worms, 95

Automated control systems, 2, 2f

Automated metrics, intelligence reports, 247

Automation
and actionable information, 230
case management, 268
data collection, 21–22
data correlation, 223–224
in deception, 43
fusion center, 23–24
vs. human interpretation, 247
incident response cases, 268
intelligence gathering, 248f
reconnaissance, 176
situation awareness, 25f
SOC, 247
statistics generation, 22

Availability
deception exploitation stage, 49
as security concern, 4, 18

Awareness principle, *see also* Situational awareness
cyber security methodology, 11
implementation, 25–26
large-scale infrastructure protection, 26
process flow, 25f
sample requirements, 296

and security policy, 129

B

Back-end insiders, deception exploitation stage, 50

Back-loaded response, 264, 264f

Backup centers, diversity issues, 103

Behavioral metrics, early warning process, 265

Bell–La Padula disclosure, and MLS, 82–83

Best practices, *see also* Commonality principle
common standards, 117
consistency principle, 17–18
examples, 115
meaningful vs. measurable, 118, 118f
national commonality program, 134
national infrastructure protection, 121
vs. security audit, 120

Biba integrity model, 82–83

Blaster worm, 94–95

Bogus vulnerabilities, and deception
discovery stage, 45–46
exploitation stage, 49
open ports, 44f
and real assets, 49
scanning stage, 42

Botnets
and antivirus software, 152
bot definition, 7
correlation-based detection, 226–228, 228f
and correlation principle, 24
data collection trends, 203–204, 204f
DDOS attack, 9f
and deception, 12–13
domain-based correlation, 218–219, 219f
and PC diversity, 96–97

real-time analysis, 53
as security concern, 4–5
and separation techniques, 68
system data collection, 199
threat, 7–9
time-based correlation, 219–220, 220f

Boundary scanning, air-gapped networks, 76

British Security Standard, best practices, 117

BS-7799, best practices standards, 117

Business environment, and awareness, 240

C

Career path
basic considerations, 131–132
security teams, 121

Carrier-centric network-based firewalls, 71f

Cascade modeling, national diversity program, 106

Case studies
deception discovery stage, 46
depth effectiveness, 144

CDNs, *see* Content distribution networks (CDNs)

Centralized mediation
functional separation, 68, 68f
smart device security, 70

CERT, *see* Computer emergency response team (CERT)

Certification/education programs
best practice recommendations, 128–131
national infrastructure protection, 121
ROI trends, 130–131, 130f

Certified Information Systems Security Professional (CISSP), 131

Circuit-switched technology, 100

CIKR protection, with Cyber Security Incident Response Management, 277–278

Citizen-based data collection, 200

Clark–Wilson integrity model, 79

Classification
  commercial vs. government information, 252
  and information disclosure, 180f
  organizational compartments, 179

Clearance
  commercial mapping, 181f
  organizational compartments, 179

Clear policy, air-gapped networks, 76

Cloud computing
  diversity paradox of, 98–100
  firewalls, 71
  layered e-mail virus/spam protection, 151–152
  network-based firewalls, 15

Clutter
  engineering chart example, 126f
  national infrastructure simplification, 128

Collection principle, *see also* Data collection
  and awareness, 25
  cyber security methodology, 11
  definition, 191
  implementation, 21–23
  large-scale trending, 203–205
  national infrastructure, 22f, 212
  national program, 208–209
  sample requirements, 298
  security goals, 193
  SIEM, 200–203, 201f, 202f

Commercial databases
  data collection, 198–199
  system data collection, 198–199

Commercial firewalls

and correlation, 222
national infrastructure, 69
and SCADA, 75
tailored separation, 65

Commercial operating systems
Mac OS®, 98–99
PC diversity, 95
UNIX®, 152
vulnerabilities, 177
Windows®, 92, 95–96, 98–99, 152

Commercial organizations, *see also* Industry environments
botnet attacks, 228
clearances/classifications, 181f
competition issues, 229
e-mail protection, 151
government assistance, 174
information sharing, 28
insider separation, 77–78
intrusion detection, 156–157
national awareness program, 252
national services, 1–2
and PII, 20–21
security audits, 134
security policy, 122–123
SIEM, 200–201
volunteered data, 204–205
vulnerability reports, 168

Commercial tools
actionable information, 230
and deception exposing stage, 52–53
large-scale offerings, 68
national infrastructure protection, 15f
satellite data services, 104
SIEM, 221–222
threat management systems, 3

Commissions and boards
cyber security principle implementation, 28

national commonality program, 121

Commonality principle
career path, 131–132
certification/education, 128–131
culture of security, 125f
engineering chart
cluttered, 126f
simplified, 127f
infrastructure simplification, 126–128
national infrastructure protection best practices, 121, 135–137
national program, 134–135
overview, 115
past security practice, 132–134
reward structure, 131–132
sample requirements, 293–296
security education ROI trends, 130f
security policy, 122–123, 123f
security protection culture, 123–126
world-class infrastructure protection, 120f

Compartmentalization
and discretion principle, 179–181
information classification, 252
separation techniques, 67f

Competition
large-scale correlation, 229
PC diversity, 97

Complex environments, simplification, 128

Complex networks, firewalls, 64f

Component distribution, as separation objective, 66

Computer emergency response team (CERT), 274

Computer security incident response teams (CSIRTs), 276

Computer-to-human interface, and deception, 53

Conficker worm, 205

Confidentiality

Consistency principle
cyber security methodology, 10
implementation, 17–19

Consumer entertainment systems, 1–2

Content condition, deceptive documents, 47

Content distribution networks (CDNs), 81, 81f

Control Systems Vulnerability Assessment Tool (CSVAT), 31

Correlation principle
actionable information, 230
analytical methods, 217–218
and awareness, 25
basic considerations, 220
botnet detection, 226–228, 228f
collection issues, 224f
conventional methods, 221–223
for critical national infrastructure cyber security, 232–233
cyber security methodology, 11
domain-based example, 219f
implementation, 23–24, 24f
improvement steps, 14–15
intrusion detection and firewalls, 223f
large-scale, 228–230, 229f
national program, 230–232
network service providers, 225
overview, 217
profile-based example, 218f
quality/reliability issues, 223–224
sample requirements, 298–300
scenario taxonomy, 220–221, 221f
signature-based example, 219f
time-based example, 220f
worm detection, 225–226

Cost issues
cyber security principle implementation, 29
and diversity principle, 93

platform diversity, 97
vs. risks, 250f

Coverage
data collection trends, 205
national infrastructure firewalls, 69
SCADA systems, 74

Critical applications, as security concern, 4–5

Critical infrastructure resilience and diversity initiative, 106–108

Critical National Infrastructure Incident Response Framework, 275–276

Critical path analysis, national diversity program, 105

Culture of security
basic considerations, 121, 123–126
implementation, 125, 126
options, 125f

Cyber attacks
on critical infrastructures, 286–287
system effect, 287
infrastructure impact, 288
economic consequence, 288
deception principle and, 55–57
example, 2f
information assurance and, 160
prevention methods, 30–31
protecting critical national infrastructure against, 29–32

Cyber-reliant critical national infrastructures, security of, 136–137

Cyber Security Development Life Cycle, 289, 298

Cyber security methodology
awareness principle, 25–26
collection principle, 21–23, 22f
components, 9–11
consistency principle, 17–19
correlation principle, 14–15, 23–24, 24f
deception, 11–13
depth principle, 19–20

discretion principle, 20–21
diversity principle, 16–17
intelligence reports, 246–248
national principle implementation, 28–29
response principle, 26–28, 27f
separation, 13–16

Cyber security scale, large vs. small, 4f

Cyber security situational awareness, 254–256

Cyber Storm III, 31, 284–285

D

Databases
actionable information, 230
asset separation, 80
as national infrastructure, 1–2
past security practices, 132–133
system data collection, 198–199
vulnerability information management, 245, 246f

Data collection, *see also* Collection principle
with aggregation, 192f
botnet behavior, 204f
botnet detection, 227
and correlation issues, 224f
decision analysis template, 193f
examples, 199f
functional view, 211
generic example, 195f
geographic view, 210
modal view, 210
network metadata, 194–196
network service providers, 225
ownership view, 211–212
system data, 196–200
systems and assets, 209–212
vulnerability detection, 196f
vulnerability information management, 245
worm tracking, 205–207, 206f, 207f

Data feeds
and correlation principle, 7
national correlation program, 231
quality issues, 224

Data formats, large-scale correlation, 228–229

Data leakage protection (DLP) systems, 78–79, 179

Data marking enforcement, 78

Data sampling technique, 195

Data services, 100, 104

Data sources, national collection program, 208

Data storage
encryption methods, 155
national collection program, 208

DDOS, *see* Distributed denial of service attack (DDOS)

Deception principle
and botnets, 12–13
and cyber attacks, 55–57
cyber security methodology, 9–10
deliberately open ports, 43–45
discovery stage, 45–46
documents, 46–48, 48f
example use, 38f
exploitation stage, 48–50
exposing stage, 51–53, 52f
honey pots and software bugs, 42
human–computer interfaces, 53–54, 54f
implementation, 11–13
interface components, 12f
national asset service interface, 43f
national program, 54–55
objectives, 38
overview, 37–38
procurement tricks, 50–51, 51f
sample requirements, 291–293

scanning stage, 42–43
stages, 40
stages for national infrastructure, 40f

Decision-makers
back-end insiders, 50
and certification/education, 121, 128–129, 130f, 131
and data correlation, 217
and intelligence reports, 246–247
and TCB, 171

Decision process
data collection, 193f
forensic analysis, 272f
risk vs. cost, 250f
risk management, 249
security policy, 123f

Decomposition, asset separation, 80

Defense in depth
cyber security methodology, 10
effectiveness, 143–147, 146f, 150f
end-user authentication, 149f
general schema, 142f
implementation, 19–20, 19f, 142–143
intrusion detection information sharing, 158f
layered access controls, 152–154
layered authentication, 147–151
layered e-mail filtering, 151f
layered e-mail virus/spam protection, 151–152
layered encryption, 154–155, 156f
layered intrusion detection, 156–158
national infrastructure, 19f
national program, 158–159
overview, 141
remote e-mail access, 141
sample requirements, 295

Department of Homeland Security (DHS), 276, 289

Depth principle, *see* Defense in depth

Designers, as decision-makers, 129

Desktop computer systems, *see* Personal computers (PCs)

Developers, as decision-makers, 129

Digital rights management (DRM), and worms, 95

Disaster recovery process
- exercise configurations, 273f
- process, 272–274
- program components, 273

Disclosure
- clearance/classification control, 180f
- deception exploitation stage, 49
- as security concern, 4

Discovery phase
- definition, 41
- overview, 45–46

Discretion principle
- clearance/classification commercial mapping, 181f
- cyber security methodology, 10–11
- implementation, 20–21
- information disclosure control, 180f
- information reconnaissance, 176–177
- information sharing, 174–175, 175f
- national program, 181–182
- obscurity layers, 178–179, 179f
- organizational compartments, 179–181
- overview, 167
- sample requirements, 296–298
- "security through obscurity", 171–173
- and TCB, 169
- top-secret information, 20
- trusted computing base, 168–171, 170f
- vulnerability disclosure lifecycle, 173f

Distributed denial of service attack (DDOS)
- and authentication and identity management, 14
- botnets, 9f

CDN-hosted content, 81, 81f
national separation program, 83
network technology diversity, 101–102
separation principle, 10

Distributed denial of service attack (DDOS) filter
challenges, 73
inbound attacks, 72f
separation techniques, 71–73

Distributed mediation, functional separation, 68, 68f

Diversity principle
and attribute differences, 92, 92f
cloud computing, 98–100
cyber security methodology, 10
desktop computer systems, 95–98, 99f
enforcing, 17
implementation, 16–17
national infrastructure, 17f
national program, 105–106
network technology, 100–103
overview, 91
PC nondiversity example, 96f
physical diversity, 103–104
proof factors, 148
sample requirements, 295
satellite data services, 104, 105f
with SSO, 148
and worms, 93–95, 94f, 101f

DLP systems, *see* Data leakage protection (DLP) systems

DNS, *see* Domain name system (DNS)

Domain-based correlation, 218–219, 219f

Domain name system (DNS)
and CDNs, 81
cyber security principle implementation, 28
deceptive open ports, 43–44

DRM, *see* Digital rights management (DRM)

Dual-homing, 76, 77f

Duplication
deception discovery stage, 46
honey pot design, 46, 47f

Duty controls, and best practices, 18

E

Einstein Cyber Shield, expanding, 30

E-mail
layered filtering, 151f
layered virus protection, 151–152
remote access authentication, 149–151

Emergency response
as national infrastructure, 1–2
national program, 274

Encryption
best practices, 117
data collection, 198
deceptive documents, 48
intelligence reports, 248
layered, 154–155, 156f
national infrastructure, 12, 146–147, 154–155
past security practice, 133–134
protected transit, 208

End user education, 131

Energy objective, as deception, 38

Enforceability, security policy, 122

Engineering analysis, depth effectiveness, 143–144

Engineering chart
cluttered, 126f
simplified example, 127f

Engineering standards, quality levels, 19

Enterprise security
and deception principle, 12–13
desktop diversity options, 99f
layered authentication, 148
and PC diversity, 97–98
separation principle, 13–14
well-known techniques, 3

Expert gathering, 262

Exploitation points
deceptive open ports, 45
definition, 5
forensic analysis, 269
national infrastructure, 5f
scanning stage, 42
and "security through obscurity", 172–173

Exploitation stage
definition, 41
overview, 48–50
pre- and post-attack stages, 49f

Exposing stage
definition, 41
overview, 51–53, 52f

External adversary, 5

F

False positives
deception exploitation stage, 49
rate, 27
and response principle, 27
response to, 46

Federal Information Security Management Act (FISMA), 18, 117

Fiber routes
network technology diversity, 101–102
worm propagation, 102f

Field control systems, 74

Filtering
DDOS, 71–73
e-mail, layered, 151f, 152
packet filtering routers, 15–16
system data collection, 198–199
vulnerability information management, 245

Financial applications, as national infrastructure, 1–2

Financial networks, insider separation, 77–78

Firewalls, *see also* Separation principle
approaches, 64–65
carrier-centric network-based, 71f
cloud computing, 71
intrusion detection correlation, 223f
large-scale networks, 65
layered access controls, 153–154, 153f
national infrastructure, 69–71
network-based, *see* Network-based firewalls
SCADA architecture, 75f
separation enhancements, 15f
separation principle, 14
SIEM threat management, 222
simple/complex networks, 64f
and worms, 94

FISMA, *see* Federal Information Security Management Act (FISMA)

Fix questions, TCB, 170

Flaws
and defense in depth, 150
and security posture, 242f

Forensic analysis
decision process, 272f
incident response, 269–271, 270f

Front-loaded response, 263–264, 264f, 266

Functional controls, and defense layers, 19–20

Functional separation

distributed vs. centralized mediation, 68f
overview, 67–69

Fusion centers, *see also* Security operations centers (SOC)
and correlation principle, 8, 23–24
and response principle, 27–28
situational awareness, 250–251

G

Generalization, and infrastructure simplification, 127, 128

Geographic location, botnet detection, 227

Global threats, and awareness, 240

Google®, 95–96

Government agencies/environments
audits, 18
botnets, 228
cloud computing, 15, 15f
commissions/boards, 134
competition, 229
data collection, 22f, 191–192, 199
data markings, 78
and deception, 12–13
discretion issues, 20, 167
firewalls, 15
information sharing, 28, 174, 175f
infrastructure best practices, 119
insider separation, 77–78
intelligence reports, 246–247
intrusion detection, 157, 158
known vulnerabilities, 240
layered authentication, 147
layered intrusion detection, 157
MLS, 82
national awareness program, 252
national commonality program, 134–135
national discretion program, 181
national diversity program, 105–106
national response program, 274, 274f

national separation program, 83
national services, 1–2
network violations, 76
organizational compartments, 179
PC diversity, 95–96
physical diversity, 100
politics, 252
response issues, 262
security policy, 122–123
separation program, 83
SIEM, 200–201
SOC, 252
system data collection, 199
system size issues, 3
TCB, 171
volunteered data, 204–205
vulnerability information management, 246
vulnerability reporting, 6–7
worm detection, 225

H

Hacking
and awareness, 241
and discretion, 167
national discretion program, 182
and "security through obscurity", 173

Hardening (servers), and deliberately open ports, 43

Hardware profiles, and awareness, 240–241

Health Insurance Portability and Accountability Act (HIPAA), 117

Hidden probes, deception exposing stage, 52

HIPAA, *see* Health Insurance Portability and Accountability Act (HIPAA)

HMI, *see* Human–machine interface (HMI)

Homeland Security Information Network (HSIN), 184

Honey pot
and actual assets, 41

and deception, 38
definition, 11–12
duplication, 46, 47f
insider separation, 78
in normal server complex, 45f
and real assets, 13
testing, 12–13
vulnerability mimics, 46

HTTP, *see* Hypertext transfer protocol (HTTP)

Human–computer interfaces, and deception, 53–54

Human–human interfaces, and deception, 53, 54f

Human interpretation, automated metrics, 247

Human–machine interface (HMI), 74

Hypertext transfer protocol (HTTP), deceptive open ports, 43–44, 44f

I

ICMP, *see* Internet Control Messaging Protocol (ICMP)

Identity management, separation principle, 14

Identity theft, as security concern, 4–5

IEC, *see* International Electrotechnical Commission (IEC)

Implied statement, and deception principle, 13

In-band detection, 157

Incident analysis, 262

Incident response, *see also* Response principle
definition, 262
disaster recovery, 272–274, 273f
early warning triggers, 265
forensic analysis, 269–271, 270f
front- vs. back-loaded, 264f
indications/warnings, 265–266
law enforcement issues, 271–272, 272f
national program, 274–275, 274f

phases, 261–262
pre- vs. post-attack, 263–265
process, 263f
security teams, 266–269
simultaneous cases, 267f
trigger intensity thresholds, 266f

Incident trigger
definition, 261
early warning, 265

Inclusiveness, security policy, 122

Indications and warnings
definition, 48–49
early triggers, 265
incident response, 265–266
response principle, 26–27

Industry environments, *see also* Commercial organizations
access control, 180–181
authentication issues, 147
career path/salary, 131–132
data collection, 22f
and hackers, 168
information sharing, 28, 174, 175f
intrusion detection, 157
national discretion program, 181
physical diversity, 100
system size issues, 3
vulnerability information management, 246

Information assurance, in infrastructure networked environments, 160–161

Information management, vulnerabilities, 244–246, 246f

Information reconnaissance
information types, 177
overview, 176–177
planning levels, 176

Information sharing
commercial vs. government, 252

cyber security principle implementation, 28
and discretion, 174–175
by government agencies, 174, 175f
hacker perspective, 167
and incident response, 265
and intrusion detection, 158f
national discretion program, 181
occurrences, 174
and "security through obscurity", 171

Infrastructure networked environments, information assurance in, 160–161

Infrastructure protection
and awareness, 240
and meaningful best practices, 119–121

Infrastructure simplification
commitment, 121, 126–128
national infrastructure, 128

Insider separation, basic considerations, 77–80

Integrity
and best practices, 18
deception exploitation stage, 49
as security concern, 4–5

Intelligence community
daily briefs, 26
and discretion principle, 20
intelligence reports, 246–247

Intelligence reports
creation, 247–248
creation/dissemination, 248f
for cyber security, 246–248

Interfaces
and deception principle, 12, 12f
human–computer, 53–54
national infrastructure simplification, 128
national response program, 274f

Internal adversary, 5

Internal firewalls
- insider separation, 78
- layered access controls, 154

Internal separation
- as firewall approach, 65
- national separation program, 83

International Electrotechnical Commission (IEC), 117

International Organization for Standardization (ISO), 117

International Watch and Warning Network (IWWN), 276

Internet Control Messaging Protocol (ICMP), worm detection, 225

Internet Explorer®, PC diversity, 95–96

Internet Protocol (IP)
- intrusion detection, 222
- layered access controls, 153
- packet-switched technology, 100
- separation principle, 14

Internet Protocol over Satellite (IPoS), 104

Internet Relay Chat (IRC), and botnets, 8, 226–227

Intrusion detection
- with data security, 157
- firewall policy correlation, 223f
- information sharing, 158f
- layered, 156–158
- SIEM threat management, 222

Intrusion prevention, 156–157

Inventory processes, system data collection, 194

Investment considerations
- and centralized security, 71
- ROI and security education, 130–131, 130f

IP, *see* Internet Protocol (IP)

iPhone®, 95

iPod®, 95

IPoS, *see* Internet Protocol over Satellite (IPoS)

IRC, *see* Internet Relay Chat (IRC)

ISO, *see* International Organization for Standardization (ISO)

J

Joint Terrorism Task Force (JTTF), 184

L

LAN controls, *see* Local area network (LAN) controls

Large-scale correlation
- example, 229f
- factors, 228–230

Large-scale cyber security
- firewall protection, 64, 68
- vs. small-scale, 3, 4f

Large-scale trending, data collection, 203–205

Law enforcement issues
- databases as infrastructure, 1–2
- incident response, 271–272

Layered access controls, 152–154

Layered authentication
- end-user, 149f
- overview, 147–151
- remote e-mail access, 149–151

Layered encryption
- multiple layers, 156f
- national infrastructure, 154–155
- overview, 154–155

Layered intrusion detection, 156–158

Layer of protection

defense in depth overview, 141
effectiveness, 144f, 146f

Legality questions, TCB, 170

Likelihood, risk management, 249

Limits questions, TCB, 170

Local area firewall aggregation
example, 70f
technique, 69

Local area network (LAN) controls, 14

Log files
and collection principle, 22
SIEM threat management, 222

Logical access controls, separation principle, 14

Logical diversity, network technology, 100

Low radar actions, 48–49

M

MAC, *see* Media access control (MAC)

Mac OS®-based operating systems, 98–99

Mainframe data storage
encryption methods, 155
system data collection, 197, 199f, 200

Malware
and awareness, 240
botnets, 8, 226
and cloud computing, 99–100
and correlation, 219–220, 223
and data collection, 198
and depth, 146–147, 152
and open ports, 43
and PC diversity, 96–97
and separation, 68, 76–77

Mandatory controls, and TCB, 169

Master terminal unit (MTU), 74

Meaningful best practices
for infrastructure protection, 119–121
vs. measurable, 118, 118f

Measurable best practices vs. meaningful, 118, 118f

Media access control (MAC), separation principle, 14

Metadata
data collection, 194
sampling, 195
SIEM threat management, 222

Military support services, as national infrastructure, 1–2

Misinformation, in deception, 11–12

MLS, *see* Multilevel security (MLS)

Mobile devices
encryption methods, 154
layered authentication issues, 149
virus/spam issues, 151

Mobile telecommunications, as national infrastructure, 1–2

Monitoring
deception exposing stage, 52–53
by network service providers, 225

MTU, *see* Master terminal unit (MTU)

Multilevel security (MLS)
example, 82f
for separation of assets, 82–84

Multiple access control systems, management, 154

N

Nachi worm, 225

National Cyber Incident Response Plan (NCIRP), 31

National Cyber Investigative Joint Task Force (NCIJTF), 30–31

National Cyber Security Center (NCSC), 30

National infrastructure
adversaries and exploitation points, 5, 5f
attack detection, 243–244
and awareness, 25–26, 296
and collection, 21–23, 22f, 298
and commonality, 135–137, 293–296
and consistency, 17–19
and correlation, 14–15, 23–24, 24f, 228–230, 232–233, 298–300
cyber attack vulnerability, 2, 2f
cyber security methodology components, 9–11
cyber threats, vulnerabilities, attacks, 4–7
data collection, 191–192, 212
data correlation issues, 224
DDOS filtering, 72–73
and deception, 11–12, 40f, 49, 51, 291–293
deceptive documents, 47–48
and depth, 19–20, 19f, 295
disaster recovery, 272–273
and discretion, 20–21, 296–298
and diversity, 16–17, 17f, 91, 295
exploitation points, 5f
firewalls, 15f, 69–71
functional separation techniques, 67–68
insider separation, 77–80
layered access controls, 153–154
layered authentication, 147
layered encryption, 154–155
network technology diversity, 102–103
obscurity layers, 178
overview, 1–2
and past security practice, 133
PC diversity, 96
physical attack vulnerability, 2, 2f
and protection, 121, 283
protection against against cyber attacks, 29–32
and response, 26–28, 27f, 264–265, 271–272, 295–296
and separation, 13–16, 63–64, 84–86, 293

service interface with deception, 43f
simplification, 128
situational awareness, 255–256
small- vs. large-scale security, 2f, 3
smart device management, 70
SSO access system, 148
system data collection, 196
TCB assets, 169–171
well-known computer security techniques, 12

National Infrastructure Protection Plan, 135–137

National programs
awareness, 252–253
collection, 208–209
commonality, 134–135
correlation, 230–232
deception, 54–55
depth, 158–159
discretion, 181–182
diversity, 105–106
implementation of principles, 28–29
response, 274–275, 274f
separation, 83

Need questions, TCB, 170

National Response Framework (NRF), 275–276

National Security Presidential Directive 54/Homeland Security Presidential Directive, 23, 29

National Transit Database (NTD), 185

Netflow, 194

Network-based firewalls
and cloud computing, 15
DDOS filtering, 71–72
as firewall approach, 64–65
layered access controls, 153–154
national separation program, 83
simple/complex, 64f

Network data
collection, 194–196
SIEM threat management, 222

Network perimeter, and defense layers, 20

Network routes, diversity issues, 104

Network service providers
data collection, 225
network monitoring, 225

Network technology
diversity, 100–103
and worms, 101f

Network transmission, encryption methods, 155

Nondefinitive conclusions, vulnerability information management, 246

Nonuniformity, and infrastructure simplification, 127

O

Obscurity layers
discretion principle, 178–179
examples, 179f
leaks, 178–179
national discretion program, 182

Obviousness, and infrastructure simplification, 127

One-to-many communication, botnet detection, 227

Online access, security policy, 122

Open solicitations, deception discovery stage, 46

Operational challenges
and collection principle, 22
and deception principle, 51
incident response, 268
smart device security, 70

Operational configurations, and best practices, 18

Operational costs, cyber security principle implementation, 29

Organizational culture
incident response, 268
security implementation, 125, 126
security options, 125f
security protection, 121, 123–126

Out-of-band correlation, 157

Outsourcing
and global threats, 240
incident response, 268
insider separation, 77
security operations, 268
security team members, 130
supply chains, 5

P

Packet filtering routers, separation principle, 15–16

Packet-switched technology, 100

Past security practice, responsible, 121, 132–134

Patching (software and systems), and best practices, 18

Patterns, national infrastructure simplification, 128

Payment Card Industry Data Security Standard (PCI DSS), best practices standards, 117

PCI DSS, *see* Payment Card Industry Data Security Standard (PCI DSS)

PCs, *see* Personal computers (PCs)

Permissions vectors, UNIX®, 152

Personal computers (PCs)
botnet attacks, 7
botnet detection, 226
DDOS attacks, 8
diversity considerations, 95–98, 99f
and diversity principle, 16–17, 95–98
nondiversity example, 96f

system data collection, 197, 199f, 200

Personally identifiable information (PII)
and discretion principle, 20–21
and TCB, 168–169

Physical attacks, national infrastructure vulnerability, 2, 2f

Physical diversity
issues, 103–104
network technology, 100
satellite data services, 104

Physical security, layered access controls, 153, 154

Physical separation
dual-homing example, 77f
technique, 75–77

PII, *see* Personally identifiable information (PII)

PKI tools, *see* Public key infrastructure (PKI) tools

Plain old telephone services (POTS), 100

Planning, disaster recovery program, 273

Platforms, diversity costs, 97

Politics
and awareness, 240
and information sharing, 174
national awareness program, 252

Port scanners, deceptive open ports, 44

Post-attack vs. pre-attack response, 263–265

POTS, *see* Plain old telephone services (POTS)

Power control networks, as national infrastructure, 1–2

Practical experience, depth effectiveness, 143

Practice, disaster recovery program, 273

Pre-attack vs. post-attack response, 263–265

Preparation, disaster recovery program, 273

Prevention
data collection security, 193
front-loaded, 263–264
past security practice, 133

Privacy policy, and collection principle, 22

Procedural controls, and defense layers, 19–20

Process allowance, deception exploitation stage, 50

Process coordination, deception exploitation stage, 50

Procurement discipline
and deception principle, 50–51
national diversity program, 106

Profile-based correlation
definition, 217–218
example, 218f

Proof factors, diversity, 148

Proprietary information
and discretion, 168
national depth program, 159

Protected transit, national collection program, 208

Protection condition, deceptive documents, 47

Protections, information reconnaissance, 177

PSTN, *see* Public switched telephone network (PSTN)

Public key infrastructure (PKI) tools, encryption methods, 155

Public speaking, and obscurity layers, 178

Public switched telephone network (PSTN), 100

Published case studies, deception discovery stage, 46

Q

Quality issues

data collection, 194
data correlation, 223–224
defense in depth, 141
engineering standards, 19

R

Real assets
and bogus assets, 44f, 49
and deception, 37–38, 38f, 49
honey pot connection, 269–270
interfaces and deception, 12f

Real-time analysis
botnet attacks, 53
honey pots, 39

Real-time awareness
implementation, 25
process flow, 25f

Real-time observations, deception exposing stage, 52

Real-time risk, situational awareness, 243

Real vulnerabilities, deception scanning stage, 42

Reliability issues, data correlation, 223–224

Remote terminal unit (RTU), 74

Removal option, PC diversity, 98–99, 99f

Replication, asset separation, 80

Requests for Information (RFIs), deception discovery stage, 46

Requests for Proposals (RFPs), deception discovery stage, 46

Resilience against cyber attacks, 106–108

Response principle, *see also* Incident response
cyber security methodology, 11
implementation, 26–28, 27f
past security practice, 133
sample requirements, 295–296

Return on investment (ROI), security education, 130–131, 130f

Reward structure
basic considerations, 131–132
security teams, 121

RFIs, *see* Requests for Information (RFIs)

RFPs, *see* Requests for Proposals (RFPs)

Right-of-way routes, network technology diversity, 101–102

Risk management process, 248–250

Risk reduction
adversary separation, 65–66
by asset separation, 80–81
and botnet detection, 228
and cloud computing, 98–99
cyber security methodology, 9–11
DDOS attacks, 72–73, 81f
and deception, 13, 38–39
and depth, 19, 145–146
by insider separation, 78–79
national separation program, 83
by network technology diversity, 102–103
by physical diversity, 103
by physical separation, 77
principles, national implementation, 28

Root cause, forensic analysis, 269

RTU, *see* Remote terminal unit (RTU)

S

Salary, 131–132

Sarbanes–Oxley controls
consistency principle, 17–18
diversity principle, 92–93
internal separation, 83

Sasser worm, 94–95

Satellite data services
physical diversity, 104
SCADA configurations, 105f

SCADA, *see* Supervisory control and data acquisition (SCADA) systems

Scaling issues, system data collection, 196–197

Scanning stage
definition, 40–41
overview, 42–43

Search for leakage, and obscurity layers, 179

"Secret," and MLS, 82

Secure commerce, encryption methods, 155

Secure Sockets Layer (SSL), encryption methods, 155

Security audit
vs. best practices, 120
and certification/education, 129
definition, 116
infrastructure protection relationship, 119
meaningful best practices, 119
meaningful vs. measurable best practices, 118
national commonality program, 134
organization examples, 116f
purpose, 116

Security information and event management (SIEM)
definition, 200–203
generic architecture, 201, 201f
generic national architecture, 202f
threat management, 221–222

Security information management system (SIMS), and collection principle, 21–22

Security operations centers (SOC), *see also* Fusion centers
high-level design, 251f
incident response, 268
responsibility, 253
situational awareness, 250–252, 254–256

Security policy
awareness, 129
and certification/education, 128–129
decision process, 123f
intrusion detection correlation, 223f
locally relevant and appropriate, 121, 122–123

Security posture
and activity/response, 242f
estimation, 242f
intelligence reports, 247

Security standard
definition, 116
national commonality program, 134

Security teams
career path/reward structure, 121
incident response, 266–269
as infrastructure decision-makers, 130

"Security through obscurity"
and asset vulnerability, 21
definition, 171–173
and discretion principle, 21
exploitable flaws, 172–173
knowledge lifecycle, 172f
objectionable applications, 171
primary vs. complementary control, 172

Segregation, asset separation, 80–81

Segregation of duties
definition, 79
work functions, 80f

Senior managers
and career path, 132
as infrastructure decision-makers, 129

Sensitive information
as security concern, 4–5
top-down and bottom-up sharing of, 182–185

Separation principle, *see also* Firewalls
asset separation, 80–81
carrier-centric network-based firewalls, 71f
cyber security methodology, 10
DDOS filtering, 71–73
distributed vs. centralized mediation, 68f
firewall approaches, 64–65
firewall enhancements, 15f
functional separation, 67–69
implementation, 13–16
insider separation, 77–80
MLS, 82–84, 82f
national infrastructure firewalls, 69–71
national infrastructure protection, 84–86
national program, 83
objectives, 65–67
overview, 63
physical separation, 75–77, 77f
sample requirements, 293
SCADA architecture, 73–75, 75f
techniques, 66

Separation vs. segregation of duties, 79

Server complex, honey pots, 45f

Server data storage
encryption methods, 155
system data collection, 197, 199f, 200

Service level agreements (SLAs)
and data quality, 224
national infrastructure firewalls, 71

Service ports
bogus assets, 44f
and deception, 39–40, 43–45, 43f

SIEM, *see* Security information and event management (SIEM)

Signature-based correlation
definition, 217–218
example, 219f

Signature sharing, 157

Simple networks, firewalls, 64f

Simplification, *see* Infrastructure simplification

SIMS, *see* Security information management system (SIMS)

Single sign-on (SSO) initiatives
- and diversity, 148
- layered authentication, 147
- national infrastructure, 148

Situational awareness
- attack confidence, 244f
- cyber security posture, 242f
- definition, 239
- implementation, 25
- and information sharing, 174
- infrastructure attack detection, 243–244
- intelligence reports, 246–248, 248f
- national correlation program, 231
- national infrastructure, 255–256
- national program, 252–253
- optimal system usage, 241f
- real-time risk, 243
- risk categorization, 241–242
- risk vs. cost decision paths, 250f
- risk management process, 248–250
- security operations centers, 250–252, 251f, 254–256
- vulnerability information management, 244–246, 246f

Sizing issues
- national infrastructure simplification, 128
- security policy, 122
- system data collection, 196–197

SLAs, *see* Service level agreements (SLAs)

Small-scale vs. large-scale cyber security, 3, 4f

Smart devices
- firewall issues, 70

national infrastructure protection, 70

SMTP, deceptive open ports, 43–44

SOC, *see* Security operations centers (SOC)

Software engineering standards, 19

Software lifecycle, and best practices, 18

Software profiles, and awareness, 240–241

Spam, layered protection, 151–152

Sponsored research, deception discovery stage, 46

SQL/Slammer worm, 94–95
tracking, 206, 206f, 207f

SSL, *see* Secure Sockets Layer (SSL)

SSO, *see* Single sign-on (SSO) initiatives

Standard audit, infrastructure protection, 119

State, forensic analysis, 269

Stream-of-consciousness design, and infrastructure simplification, 127

Subjective estimations, national depth program, 159

Sufficient detail, deception exposing stage, 52

Suitability, and defense in depth, 145

Supervisory control and data acquisition (SCADA) systems
architecture, 73–75, 75f
insider separation, 77–78
IPoS, 104, 105f
layered access controls, 153
as national infrastructure, 1–2
national infrastructure firewalls, 69
separation principle, 10
tailored separation, 65, 83

Supplier adversary
deception techniques, 50, 51f

definition, 5

Suppliers, diversity issues, 17, 103–104

Supply chain, 5

Support and training, and desktop diversity, 97

Surface Transportation Security Inspectors (STSI), 184

System administration, and best practices, 18

System administration and normal usage, 5

System data, collection, 196–200

T

Tailored separation
as firewall approach, 65
national separation program, 83

Target factor
large-scale correlation, 229
separation techniques, 66

Tarpit, 55

TCP/IP, *see* Transmission Control Protocol/Internet Protocol (TCP/IP)

TDM, *see* Time-division multiplexed (TDM) services

Telecommunications
collection systems, 22
insider separation, 77–78
as national infrastructure, 1–2

Terrorist attacks
9/11, physical attack vulnerability, 2

Testing and simulation, depth effectiveness, 144

Theft
deception exploitation stage, 49
as security concern, 4–5

Threat factor

insider separation, 77
separation techniques, 66

Threat management
and best practices, 18
conventional security correlation, 221–222
SIEM, 222

Time-based correlation
definition, 219–220
example, 220f
worm detection, 226f

Time-division multiplexed (TDM) services, diversity, 100–101

Tools and methods, national deception program, 55

Top-secret information
disclosure control, 180, 180f
and discretion principle, 20
and MLS, 82

Transmission Control Protocol/Internet Protocol (TCP/IP), metadata collection, 194

Transparency
and deception, 50
national correlation program, 231

Transportation infrastructure, insider separation, 77–78

Transportation Security Operations Center (TSOC), 185

Trap functionality, and deception principle, 12

Trap isolation, deception exploitation stage, 50

Trends, data collection, 203–205

Trusted computing base (TCB)
basic questions, 169–171
definition, 168–171
discretion program goals, 169
national discretion program, 181
size issues, 169, 170f

Trusted Internet Connections Initiative (TICI), 30

U

UDP, *see* User Datagram Protocol (UDP)

Uncertainty objective, as deception, 38

UNIX®-based operating systems, 152

Usage metric
optimal for security, 241f

Use-case studies, depth effectiveness, 144

User Datagram Protocol (UDP), worm tracking, 206, 206f, 207f

V

Value proposition, national correlation program, 231

Vantage point, centralized security, 70

Vendors
diversity issues, 103–104
and diversity principle, 17

Vigilant watch, botnet detection, 227

Violation issues
access policies, 20
air-gapped networks, 76
and depth, 142
information leaks as, 180
infrastructure protection best practices, 120

Viruses
attack initiation, 2
layered protection, 2f, 151–152
past security practice, 133
and response, 266f
and trending, 205
and voice services, 102–103

Voice services, 102

Volunteered data, 7, 204–205, 223–224, 245

Vulnerability issues
and awareness, 240
and culture of security, 123–124
and data collection, 196f
and deception, 38–39, 42, 48–50
and defense in depth, 19
disclosure lifecycle, 173f
early warning process, 265
honey pot mimics, 46
information management, 244–246, 246f
information reconnaissance, 177
national infrastructure, 4–7
and "security through obscurity", 171, 173

W

Well-known computer security techniques
and exploitation points, 6–7
and national infrastructure, 12

Wide area firewall aggregation
example, 70f
technique, 69

Windows®-based operating systems
access control lists, 152
and diversity principle, 92
PC diversity, 95–96, 98–99

Work functions, segregation of duties, 79, 80f

World-class focus
infrastructure protection, 119–120
methodology, 120f
national commonality program, 134

Worms
attack initiation, 2, 2f
cloud computing, 98
and correlation, 220–221, 225–226, 226f
and diversity, 93–95, 94f

functionality, 93
Microsoft® Windows® target, 96
and network diversity, 100–101, 101f
past security practice, 133
propagation, 93–95, 94f, 102f
protection against, 4–5
and response, 266f
as security concern, 4–5
tracking, 205–207, 206f, 207f
and trending, 205

Worst case assumptions, vulnerability information management, 246