Realtime publishers

The Shortcut Guide™ To

Business Security Measures Using SSL

sponsored by

VeriSign®

Dan Sullivan

Introduction to Realtime Publishers

by Don Jones, Series Editor

For several years now, Realtime has produced dozens and dozens of high-quality books that just happen to be delivered in electronic format—at no cost to you, the reader. We've made this unique publishing model work through the generous support and cooperation of our sponsors, who agree to bear each book's production expenses for the benefit of our readers.

Although we've always offered our publications to you for free, don't think for a moment that quality is anything less than our top priority. My job is to make sure that our books are as good as—and in most cases better than—any printed book that would cost you $40 or more. Our electronic publishing model offers several advantages over printed books: You receive chapters literally as fast as our authors produce them (hence the "realtime" aspect of our model), and we can update chapters to reflect the latest changes in technology.

I want to point out that our books are by no means paid advertisements or white papers. We're an independent publishing company, and an important aspect of my job is to make sure that our authors are free to voice their expertise and opinions without reservation or restriction. We maintain complete editorial control of our publications, and I'm proud that we've produced so many quality books over the past years.

I want to extend an invitation to visit us at http://nexus.realtimepublishers.com, especially if you've received this publication from a friend or colleague. We have a wide variety of additional books on a range of topics, and you're sure to find something that's of interest to you—and it won't cost you a thing. We hope you'll continue to come to Realtime for your educational needs far into the future.

Until then, enjoy.

Don Jones

Copyright Statement

This sponsored eBook is valid until June 30, 2011.

Chapter 1: Security Threats to IT Operations in the Age of Cybercrime

Over the past decade, businesses have had to adapt to an array of technical changes, including an increasingly hostile cyber environment. We saw the early precursors of cybercrime decades ago when computer use was limited to a relatively small group of specialists and electronics enthusiasts. Innovative programmers, some still in high school, would find ways to display annoying messages on their friends' computers and from there spread to other devices via shared floppy disks. This kind of part practical joke-part vandalism form of malware has been overshadowed by the more serious, technically complex, and financially lucrative form of today's cybercrime.

In this guide, we will examine major types of threats to information security that businesses face today as well as techniques for mitigating those threats. One of the most important tools available to us is SSL technology.

> **Note**
> This is actually something of a misnomer. Secure Socket Layer (SSL) protocols have largely been replaced with Transport Layer Security (TLS) protocols but by convention, we continue to use the term "SSL."

With SSL technology, we enable secure communication, identity verification, and ultimately trust between businesses. SSL technology does not exist in a vacuum, though. Information security is a multifaceted challenge that requires coordination of a variety of security measures, so this guide will examine the business and technical practices that weaken security as well as best practices for improving information security. This guide is organized into four chapters:

- Chapter 1 describes the evolving nature of security threats, including the development of an underground economy for cybercrimes. It also covers the business resources targeted by criminals and the impact of poor security on business operations and innovation.
- Chapter 2 examines common vulnerabilities in IT systems and business practices that undermine information security.

- Chapter 3 focuses on developing and maintaining a high-impact security strategy. In this chapter, the emphasis is on reviewing business practices and workflows, assessing technical infrastructure, and refining security policies and procedures.
- Chapter 4 concludes this guide with a discussion of practices for implementing a business-centric security management strategy. Topics range from protecting infrastructure to securing desktops and other endpoint devices. Special attention is paid to end-user security-awareness training. A checklist of practices and technologies is included to help you begin implementing the measures important to your environment.

Taking the adage "know thy enemy" to heart, we start with a look at the nature of cybercrime.

Evolving Information Security Threats

Before delving into the details of today's cybercrime environment, let's dispel any last semblance of malware, hacking, and related activities as simply mischievous pranks or technical vandalism. Those days are gone.

Minimal Threats: Experimenters and Dabblers

Of course, there are curious, ingenious programmers experimenting with operating systems (OSs), browsers, and application software trying to break them or use them for unintended purposes. There are also less ingenious, less skilled dabblers, known as "script kiddies," who use tools provided by their more technically advanced colleagues. The former group is not a significant threat as long as their work is not let loose into the wild; the latter are not much of a threat because much of their computer-generated malware is easily detected and contained by today's antivirus systems. More significant threats come from attackers with a different set of motives.

Something Old, Something New: Cybercrime Puts a New Twist on Old Crimes

Cybercrime is blanket term that covers a broad range of crimes and malicious activities that can adversely impact a business' operations and even long-term viability. Forms of cybercrime include:

- Fraud, which can occur, for example, because of mistaken identity, poor access controls that allow unauthorized users to tamper with data, or misappropriating software tools to hide unauthorized transactions.
- Identity theft, which is facilitated by poor identity management, insufficient access controls, unencrypted communications, or other sloppy data protection measures.

- Embezzlement is a classic insider threat; computer technology can help enable as well as prevent this crime. Proper authentication, such as with digital signatures implemented with SSL technologies, can help mitigate this threat through non-repudiation. Extortion with a high-tech twist can come in the form of Denial of Service (DoS) attacks that effectively render network devices inaccessible because of an overload of malicious traffic. Many businesses in Estonia were affected by the widespread DoS attack on that country in 2007. In that case, the attack was prompted by political tensions between Estonia and Russia rather than immediate financial gain.

- Intellectual property theft is not a new problem, but like other forms of crime, it can take on new dimensions when business systems are interconnected. Take, for example, the case of a former Intel employee charged with stealing more than $1 billion in trade secrets from the company (Source: Press Release, U.S. Department of Justice. "Former Intel Employee Indicted for Stealing More than $1 Billion of Trade Secrets," available at http://www.cybercrime.gov/paniIndict.pdf). The man had received a job offer from competitor AMD and he spent the last several days at Intel downloading confidential and proprietary information, including 13 documents designated as "top secret" by the company's data classification standard.

Information technology (IT) has radically changed the way criminals can commit crimes and this exposes businesses to new types of threats. Of course, employees could steal trade secrets in the past by stuffing copies of documents in their brief cases. It is difficult to imagine one man stealing $1 billion worth of secrets using only a copier and a briefcase.

One thing to keep in mind about cybercrime is that the same IT that makes businesses more efficient and able to do more with less is the same technology that allows cybercriminals to do the same. IT professionals, fortunately, have the tools and practices to mitigate these risks. The purpose of this guide is to provide some guidance on which tools, such as SSL certificates, and practices, such as identity management, are appropriate for specific circumstances. Another thing to keep in mind about cybercrime is that the patterns of organization that have helped businesses, industries, and even global markets grow and succeed are now used to extend the reach and impact of cybercrime.

Cybercrime as a Global Industry

Several things that have made modern markets so successful—such as division of labor, specialization, brokers, and exchanges that bring buyers and sellers together—are emerging in the world of cybercrime as well. In 2006, Assistant Director Brian Nagel of the U.S. Secret Service's Office of Investigations observed:

Cyber crime has evolved significantly over the last two years, from dumpster diving and credit card skimming to full-fledged online bazaars full of stolen personal and financial information (Source: Press Release, U.S. Secret Service, "United States Secret Service's Operation Rolling Stone Nets Multiple Arrests," March 28, 2006, available at http://www.secretservice.gov/press/pub0906.pdf).

More recently, Kilian Strauss, of the Organisation for Security and Cooperation in Europe (OSCE) observed how difficult it is to keep up with the pace of innovation in cybercrime:

> These criminals, they outsmart us ten, or a hundred to one (Source: Sarah Marsh, "Cybercrime Could Be as Bad as the Credit Crisis," Reuters, November 29, 2008, available at http://www.itpro.co.uk/608466/cybercrime-could-be-as-bad-as-the-credit-crisis).

Cybercrime is now functioning like an industry. Like other industries, this one is profit driven, so patterns that work for businesses, such as outsourcing specialized services, forming markets to exchange goods and services, and countering competitive threats, will be found in cybercrime. As a first step to understanding this "industry," we need to understand the specialists that constitute the major actors, such as:

- Malware developers
- Bot herders
- Spammers and phishers
- Hackers and data thieves
- Brokers

Each of these actors plays a critical role in current-day cybercrime. Without any one of them, the nature of today's cybercrime would be significantly altered.

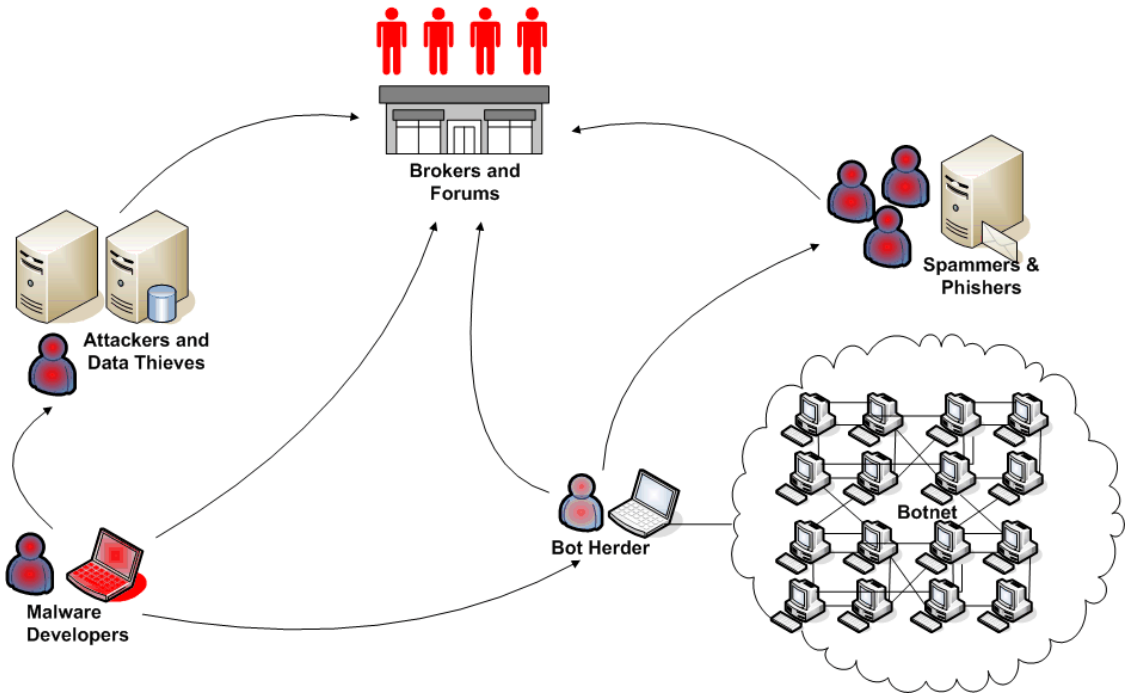

Figure 1.1: Cybercrime has evolved to support a complex mix of different skills and services much like legitimate businesses.

Malware Developers

Malware developers are the innovators that produce the new tools for the cybercrime industry. These software creators are the source of viruses, worms, Trojan horses, bots, rootkits, and other exploits. Given the financial motivation of cybercriminals, the malware that is in greatest demand is that which can lead to financial gain, including the ability to steal:

- Credit card data sufficient to successfully commit fraud
- Personal information that would allow someone to steal another person's identity
- Intellectual property, such as trade secrets, that can provide a competitive advantage to the ultimate recipient of the stolen goods
- Authentication credentials, such as usernames and passwords, that would allow an attacker to gain access to those kinds of data listed previously
- Computing and network resources that allow others to generate spam or launch DoS attacks at low or no cost

There is a specialization of labor in cybercrime, so it is not surprising that malware developers are not necessarily using their own software. That is left to others, such as bot herders and spammers.

Bot Herders

A *bot*, aka a *zombie*, is a computer under the control of someone other than its legitimate user. Put a group of bots together and you have a *botnet*.

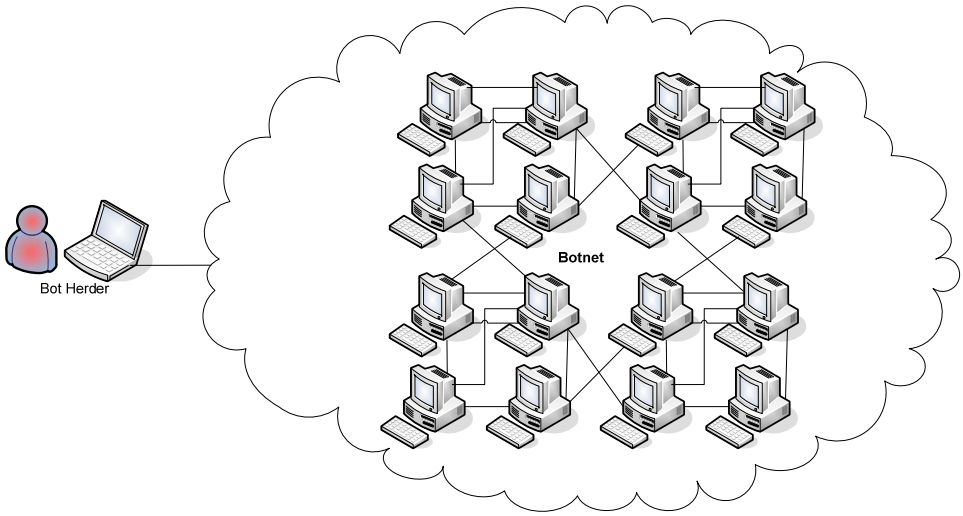

Figure 1.2: A botnet is a collection of compromised computers controlled by a bot herder. The most resilient botnets do not depend on a single server for command and control structure; rather, they use more distributed communications methods and employ recovery techniques to work with different bots should other bots they had been working with become unavailable.

From a purely disinterested point of view, botnets are highly useful distributed systems. They provide on-demand computing and networking services to the people that control them. They can generate phishing lures and send those lures to millions of email recipients or launch DoS attacks to disrupt business or government operations. The legitimate business world has an analog of botnets in the form of cloud computing.

Cloud services provide (legitimately) on-demand computing resources, storage, and networking for specialized projects or ongoing business operations. Amazon's S3 storage service and EC3 computing services are probably the best known examples of cloud services. The reason botnets are popular in cybercrime is the same reason cloud computing is of growing interest to business: little or no capital investment is required, the ongoing operational costs are minimized, and you can scale rapidly to meet peak demand without having to maintain peak capacity during less-demanding periods.

The resiliency of botnets became clear recently. In a well-publicized counterattack against spammers in November 2008, the Internet service provider (ISP) that had been hosting command and control servers for the 450,000-bot Srizbi botnet cut off service to the bot herder. For several days, there was guarded hope that this might put a dent into the amount of spam generated, but that hope was short lived. The botnet developers had planned for such a contingency and the bots were able to re-establish communication with new command and control servers.

Spam is not the only potential way to make money with botnets; launching DoS attacks is another revenue stream. In one case, a Michigan business man was sentenced to 30 months in prison for conspiring with a bot herder to disrupt competitors' business by launching DoS attacks against their Web sites and online sales servers (Source: U.S. Department of Justice Press Release, "Michigan Man Gets 30 Months for Conspiracy to Order Destructive Computer Attacks on Business Competitors," August 25, 2006, available at http://www.cybercrime.gov/araboSent.htm). Other businesses using the same ISP hosting the victim were also adversely affected. These included a major online retailer, banks, and a communications and data services company.

How big is the botnet problem? In 2007, 10% of online computers were infected by malware and by the end of 2008, that number is expected to have grown to 15%, according to researchers at the Georgia Tech Information Security Center (Source: David Stevenson "Profit from the Fight Against Cyber-Crime," Money Week, December 19, 2008,available at http://www.moneyweek.com/investment-advice/profit-from-the-fight-against-cyber-crime-14304.aspx).

Spammers and Phishers

Although most of us will not have much direct contact with malware developers and bot herders, we are all too familiar with the products of spammers and phishers. If we can say anything positive about these purveyors of unwanted and unsolicited email, it is that they are persistent, they are efficient, and they are effective.

The constant deluge of junk email we get in our email and content-filtering systems is a testament to spammer's persistence. The problem shows no signs of abating and, given the resiliency of botnets like Srizbi and the expected increase in the size of botnets, it is prudent to assume that spamming and phishing are with us for the long term.

We can deduce the efficiency and effectiveness of spammers by the fact that they choose to continue to operate. The low cost of spamming means that minutely small response rates can still yield a profitable business model. In the case of phishing, we can deduce that extra time and effort to create smaller targeted attacks, known as "spear phishing," pay off as well.

Hackers and Data Thieves

Some attacks are launched at a broad pool of potential victims; the attackers are trolling with wide nets to catch as much as possible. Other attacks are more targeted and seek to victimize a single business. Some examples of this include:

- The largest breach to date occurred at TJX Companies which operates T.J. Maxx and Marshalls stores in the US as well as T.K. Maxx stores in the U.K. and Ireland. The cost was more than $100 million to the company itself with other costs to banks who had to re-issue credit cards.
- The supermarket chain Hannaford Bros. Co. suffered a data breach from December 2007 to March 2008 when attackers were able to capture data in transit.
- In 2008, extortionists tried to compel Express Scripts, a pharmacy benefits management company, to pay or else risk having personal information about millions of customers exposed. In a noteworthy twist, the company refused to pay and instead offered a $1 million reward for information leading to the arrest and conviction of the perpetrator(s).

Hackers and data thieves can use many different techniques to compromise corporate computers. Vulnerability scanners can probe networks and devices on networks looking for unpatched software that can be exploited to gain elevated privileges or access otherwise restricted data. Information sent over wireless networks that is not encrypted may be picked up by eavesdroppers. Poorly designed Web applications may expose databases to SQL injection attacks that can leak private and confidential data. Weak passwords and default passwords can leave servers and network devices vulnerable to dictionary attacks. With so much valuable data within business systems and so many ways to launch targeted attacks, it is not surprising that criminals have taken to this opportunity.

Brokers and Exchanges

Markets depend on buyers and sellers being able to efficiently find each other. Brokers facilitate this process in many markets and cybercrime is once again following tried and true patterns of business. Cybercriminals who have managed to steal valuable data can sell it through collaborative systems such as underground forums.

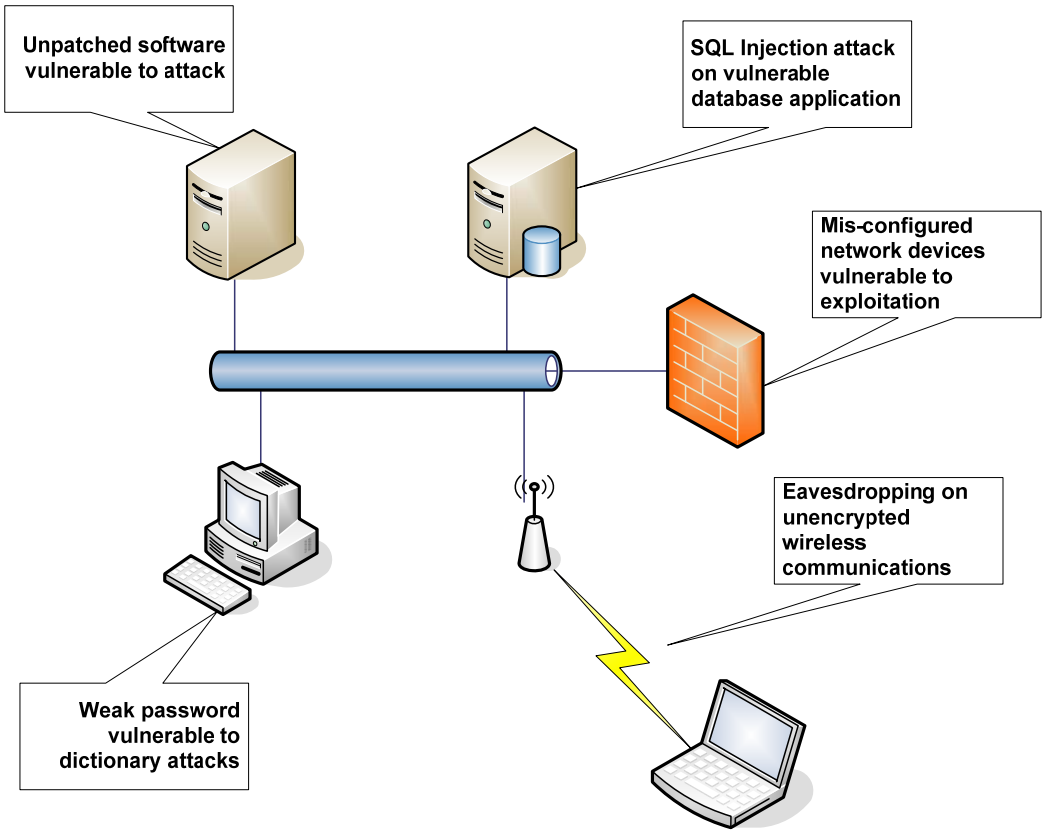

Figure 1.3: Attackers can exploit multiple types of vulnerabilities on desktops, servers, databases, applications, and networks to steal private and confidential business data.

Increasing Numbers and Sophistication of Attacks

Some security researchers monitor communication channels as well as other indicators of overall cybercrime activity, and have observed patterns that indicate an upturn in cybercrime activity. For example:

- In one study spanning a one-year period, 69,130 advertisers sought to sell stolen information in underground forums; the top-10 sellers offered $16.3 million in credit card data and $2 million in bank account data (Source: Symantec Press Release, "New Symantec Report Reveals Booming Underground Economy," November 24, 2008, available at http://www.symantec.com/about/news/release/article.jsp?prid=20081123_01).
- One security service provider observed a 30% increase in network and Web-based security events over a 4-month period among their clients; the number of events rose from 1.8 billion to 2.5 billion events per day (Source: IBM Press Release, "Citing a Surge in Online Cybercrime, IBM Bolsters Security Service," December 4, 2008, available at http://www-03.ibm.com/press/us/en/pressrelease/26232.wss).
- The price of stolen information is dropping. Credit card numbers now sell for $2 to $3 and full victim profiles, with credit card number, mother's maiden name, Social Security number, and so on are selling for $10 (Source: Taylor Buley, "Crime Still Pays for Identity Thieves—Just a Little Less than It Once Did," Forbes, October 27, 2008, available at http://www.forbes.com/security/2008/10/25/credit-card-theft-tech-security-cz_tb1024theft.html).

Clearly, cybercriminals are adapting to new opportunities presented by the changing economic landscape. There are likely multiple reasons for the increase, on both the supply and the demand side. The global downturn leaves fewer legitimate opportunities for computer professionals, some of whom may be turning to cybercrime. Victims looking to make up for lost income can be easy prey for phishers and other scammers. Along with the increase in volume of attacks, there is an increase in sophistication of attacks.

Case Study in Credit Card Theft

From late 2007 to early 2008, a major supermarket chain was subject to a sophisticated data breach that netted more than 4 million credit and debit card numbers for the attackers. 300 stores in the Hannaford Bros. chain had servers infected with malware that intercepted credit card data and sent it to servers outside the country. Unlike other well-publicized data breaches, Hannaford Bros. was not storing more data than allowed under industry rules and the company was in compliance with Payment Card Industry (PCI) standards.

The problem was that data was captured as it was transmitted from the point-of-sale device to transaction processing service. This example shows that even when in compliance with industry standards, data breaches can still occur. Even when data is transmitted on trusted networks, encrypting data in transit using SSL technologies can mitigate the risk of this type of attack.

Doing Business in the Age of Cybercrime

Cybercrime is evolving and becoming more dangerous. It is useful to think of cybercrime as an industry with similar division of labor, service provider models, and drives to efficiency and revenue growth seen in legitimate businesses. We also need to keep in mind that compliance with regulations is a minimal set of requirements for securing business information. Malware developers, bot herders, spammers, phishers, and other attackers have demonstrated that they can and will develop new techniques to bypass security countermeasures.

Now that we have highlighted some of the structural characteristics of the cybercrime industry, let's turn our attention to business targets of their attacks.

Business Resources Targeted by Cybercrime

Businesses have primarily two assets of value to cybercriminals: information and computing resources. Both are actively sought after in the cybercrime underground economy.

Targeted Information Assets

Information is in many ways an ideal target for criminals. It is intangible, so you do not need to be in physical proximity of the information to steal it. There are many ways to hide your identity and eliminate traces of malicious activity. Perhaps best of all, large amounts of valuable information tend to be stored in centralized repositories, such as databases, or are transmitted across common paths, such as from a point-of-sales system to a transaction processing server. In such cases, it takes only marginally more effort to steal thousands or even millions of credit cards than it does to steal one or two.

Three types of information of value to cybercriminals are:

- Identity information
- Credit card and financial account data
- Proprietary information and intellectual property

Identity Information

Identity information is the key to successfully committing identity theft. The object of identity theft is to commit fraud using the credit profile of the victim. Identity theft victims may find fraudulent bank withdrawals, new accounts opened in their names, and even bankruptcy filed in their names. Specialized forms of identity theft can wreak even more havoc on victims. Medical identity theft, for example, occurs when someone uses another person's identity to receive payment for medical treatment or provide medical goods. In addition to the usual credit problems that follow for identity theft, these victims may have to correct inaccurate medical records. The ripple effects of identity theft can include complications with taxpayer records that need to be resolved with the Internal Revenue Service (IRS).

Credit Card and Bank Account Data

Credit card and bank account fraud is big business. One study found that almost one third of all advertisements in a cybercrime forum were for credit card data. In 2008, the FBI and other international law enforcement agencies shut down one forum, known as Dark Market, that had at its peak 2500 registered members (Source: FBI Press Release, "FBI Coordinates Global Effort to Nab 'Dark Market' Cyber Criminals," October 16, 2008, available at http://www.fbi.gov/pressrel/pressrel08/darkmarket101608.htm). The forum was notorious as a market for credit card data, login credentials, and even some equipment used in financial crimes. Breaking up that one forum resulted in 56 arrests and prevented $70 million in losses due to fraud.

Identity theft and credit card fraud are well-publicized aspects of cybercrime, so much so, that one might think cybercrime is primarily a problem for banks, retailers, and others with high volumes of consumer financial transactions. That is certainly not the case.

Proprietary Information and Intellectual Property

Trade secrets and other intellectual property are not the commodity products of cybercrime the way credit card and bank account data are, but it can still be a highly valued target. Consider some examples of cybercrime involving proprietary information:

- A former Netgear engineer was indicted for theft, misappropriation, and unauthorized downloading of trade secrets. It is alleged that the engineer used access to a semiconductor supplier's technical documentation to download trade secret information about the supplier's switches and transceiver products. He then took those documents with him when he went to work for one of the supplier's competitors (Source: U.S. Department of Justice Press Release, "Silicon Valley Engineer Indicted for Stealing Trade Secrets and Computer Fraud," December 22, 2005, available at http://www.cybercrime.gov/zhangIndict.htm).
- Two former employees of NetLogics Microsystems stole chip design trade secrets from their then employer as well as other companies and then started their own company in the hopes of obtaining venture capital funding for their efforts (Source: U.S. Department of Justice Press Release, "Two Bay Area Men Indicted on Charges of Economic Espionage, "Sept. 26, 2007).
- Three chemical company executives were indicted for conspiring with an employee of another chemical company to steal trade secrets. The indictment alleges one of the conspirators would download trade secret data to an external storage device prior to meetings. The conspiracy appears to have continued for more than 6 years (Source: U.S. Department of Justice Press Release, "Trade Secret Charges Filed Against Company Executives and South Korean Nationals," November 12, 2008, available at http://www.cybercrime.gov/shinIndict.pdf).

Cybercrime provides the means to avoid the high cost of research and development in intellectual-capital intense industries. It is not surprising that even within legitimate businesses, there are those that will turn to cybercrime or use IT systems in the course of their intellectual property theft.

Targeted Computing Assets

When you consider the cost businesses incur to purchase and maintain IT infrastructure, it becomes clear why cybercriminals would have an interest in stealing computing assets. Just as in legitimate businesses, cybercrime operators need to be able to ensure:

- They have adequate computing, storage, and network resources to meet demands
- Failover and disaster recovery procedures are in place
- Costs are minimized without adversely affecting performance
- Ironically, malware and attackers do not gain control of their infrastructure

Botnet malware and bot herders are integral parts of acquiring and maintaining a cybercrime infrastructure. As noted earlier, botnets are designed to avoid single points of failure and to gracefully degrade and ultimately recover in response to isolated failures. The more sophisticated botnets also use blended threats to detect bots in competitors' botnets, disable the alternate bot software, and add the bot to their own botnet. The benefit of well-designed bot software is a virtually free IT infrastructure; there are none of the typical support costs including power, hardware maintenance, service support, rent, software licensing, and so on.

As a baseline for the value of botnets, we can look to a legitimate provider of on-demand computing and storage: Amazon. The Amazon Simple Storage Service (S3) and Elastic Compute Cloud (EC2) provide customers with long-term storage and computing services for costs often below the charges small organizations, such as business IT departments, can offer. Nonetheless, there are costs.

Businesses are attractive targets for cybercriminals. They have valuable commodity data, such as credit card and bank account information, identity information sufficient to enable identity theft, as well as proprietary information that may be of value to less scrupulous competitors. Businesses also have well-managed computing infrastructures with the computing, storage, and networking services needed in the cybercrime economy. The business consequences of cybercrime include the immediate effects of data breaches and related attacks as well as subtler and sometimes underappreciated impact on business.

Poor Security's Impact on Business

Headlines about security breaches and data losses at major retailers, banks, and government agencies certainly do get attention, especially when costs are mentioned. The full cost of poor security is not captured even in these attention-grabbing incidents. They are more like the proverbial tip of the iceberg than a reflection of the full impact of weak security measures. To understand the full extent of cybercrime's adverse impact on business, we should consider the obvious as well as the less obvious consequences.

Damage in Plain Sight

The cost of poor security is apparent after a security breach. Consider a fictional but representative example. Suppose a disgruntled employee has decided that he has been underpaid and mistreated by his employer. To compensate himself, he decides to capture customer credit card data as it moves across the network. This employee has access to internal systems, so this task is not a problem, especially because this type of data is only encrypted when it is sent outside the trusted network. After the employee collects a sufficient amount of credit card data, he copies the data to his iPod, heads home, and posts an advertisement on a cybercrime forum. If he is successful, he will earn a couple of dollars for each account.

Now it is time to tally up the costs to the business:

- The cost of violating any of the many state and federal privacy regulations protecting consumer data
- The cost of possible industry regulation violations, such as PCI data protection standards
- The cost of litigation associated with lawsuits
- The cost of notifying customers of the breach and possibly paying for credit monitoring services for victims
- The soft cost of brand damage and loss of customer loyalty

These costs could have been avoided with the use of SSL technologies to encrypt communication between servers and endpoint devices.

Hidden Costs of Poor Security

Not all costs are as obvious as those related to data breaches and associated regulation violations. The less obvious costs come in the form of reduced effectiveness of business operations, and in particular:

- Reduced innovation
- Costly ad hoc responses to incidents
- Opportunity costs to other IT initiatives

Imagine a strategy session with executives and business managers planning to overhaul a business process with partners. Someone suggests working with suppliers to offer drop shipping from their facilities rather than maintain high levels of inventory within the company's warehouses. The company could work with the suppliers to leverage their shipping and order processing systems and rebrand the supplier's Web site to look like the company's when its customers are checking shipping information. A software development manager makes some suggestions about using Web services, passing customer data to the supplier, and receiving shipping details in return. So far, so good. Then one of the more security-conscious members of the meeting chimes in with questions such as:

- How do we ensure order information is not tampered with during transmission?
- How do we know protected customer information is not leaked?
- How will the company's application verify it is working with the supplier's Web service and not a fake Web service set up to capture customer information?

Without proper security measures, such as SSL technologies for encrypting data and verifying digital identities, innovative business processes such as these might be left on the drawing board. Ultimately, if we do not protect information assets, we can expose our businesses, partners, and customers to compromise.

Day-to-day operations can be adversely affected by poor security practices. Ad hoc responses to incidents such as malware infections and the need to patch applications can ultimately cost more than a more methodical approach. With proper asset management applications, patch management tools, and an incident response plan, businesses can more effectively and efficiently respond to adverse events.

Overall, the true cost of poor security is reflected in a combination of costs from data breaches and other security incidents and the opportunity cost of not implementing innovative procedures and processes because of fear of potential security problems. It is worth emphasizing that such fear is not unfounded; there may be significant risks to changing workflows and opening systems to work with business partners' applications when proper security measures are not in place. One of the goals of this guide is to provide you with information about techniques such as using SSL for encryption and digital identity verification to help control some of these risks.

Summary

Viruses and hacking are no longer just electronic forms of vandalism carried out by programmers demonstrating their technical prowess. Cybercrime has evolved into an industry-like phenomenon complete with markets, specialization of services, and multiple business models for turning stolen information and computing resources into cash. For businesses to succeed and thrive in such an environment, they must manage security processes and leverage technologies such as SSL for encryption and digital identity verification. The remaining chapters of this shortcut guide will delve into details of how to accomplish this.

Chapter 2: Common Vulnerabilities in Business IT Systems

Businesses, governments, and other organizations face a wide array of information security risks. Some threaten the confidentiality of private information, some threaten the integrity of data and operations, and still others threaten to disrupt availability of critical systems. Chapter 1 examined the role of organized cybercrime, the prevalence of malicious software and the underground marketplaces that facilitate the exchange of stolen information, and tools of the cybercrime trade. In this chapter we turn our attention inside the organization. Although the external threats are considerable, they are not the only component in the risk equation. Another important set of factors are the vulnerabilities that lie within an organization.

For our purposes, we will broadly organize these vulnerabilities into two categories: technical weaknesses and organizational weaknesses. This specification is to draw attention to the fact that information security is not just about technology, although that is an obvious component. How we perform business operations, how we attend to information systems management, and how we train and help others understand the nature of security risks can make a critical difference in the overall effectiveness of an information security strategy. Perhaps more importantly, it is crucial to understand that technical controls will not compensate for poor organizational practices, and the best trained staff and most well intentioned IT professionals will not be able to protect information assets without proper technical controls. An overall security posture is a combination of technical and organizational controls.

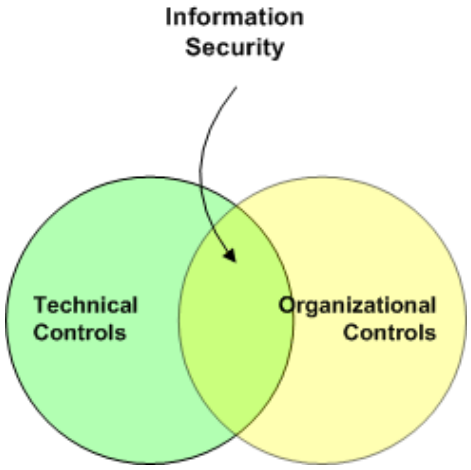

Figure 2.1: Technical and organizational controls overlap and are both essential to information security.

This chapter will examine common weaknesses in technical and organizational controls and then discuss options for addressing those weaknesses.

Technical Weaknesses

Technical weaknesses are vulnerabilities that can be mitigated using technical controls, such as the implementation of new firewall rules or an update of antivirus signatures on a client device. There are many different types of such vulnerabilities; we will concentrate on several that are all too common:

- Unencrypted communications
- Insufficiently patched operating systems (OSs) and applications
- Insufficient use of antivirus and personal firewalls
- Weak boundary security
- Poor application security

For each of these, let's consider types of attacks enabled by these vulnerabilities and their cost to business.

Unencrypted Communications

Rapid, reliable, and trustworthy communications are essential in today's business world. Although postal mail and telephones are still used widely, some of the most cost-effective communications take place online. We routinely email colleagues, customers, clients, and other professional and personal contacts. Instant messaging is especially useful for geographically distributed teams who need an electronic equivalent of talking across the room or over the top of a cubicle partition. Many have taken to social networking services, from LinkedIn and Facebook to Twitter, to keep up to date with large groups of individuals. All these communication mechanisms have their advantages and few would want to ban them from the office, but with their convenience and efficiency comes security risks.

When communications are transmitted in unencrypted forms—such as plain text—there is the potential for someone to intercept the message to learn the contents or tamper with the contents before they arrive at the intended recipient's inbox. We will consider two examples of such attacks: the man-in-the-middle (MITM) attack and the replay attack.

Man-in-the-Middle Attack

An MITM attack injects a malicious third party into a communication between two presumably unsuspecting victims. The purpose of the attack is to control the communications between the two victims and alter messages between them. Several conditions must be in place for an MITM attack to succeed:

- The attacker must have access to the communication channel between the two parties
- The attacker must be able to impersonate each of the victims sufficiently to overcome technical controls and potential suspicions on the part of either victim.
- The attacker must be able to alter, inject, or remove messages sent on the communication channel without detection

Accessing communication channels used to require access to wired network equipment, such as routers or hubs, but the prevalence of wireless networks allows attackers to gain access to a communication channel from a distance.

> **Note**
>
> Using just any encryption for wireless communication is not sufficient to protect communications. The Wired Equivalent Privacy (WEP) protocol was defined in the late 1990s for encrypting wireless communications. Within several years, flaws were found in the algorithm, and tools are available today to break WEP encryption in minutes. Wireless networks should use the Wi-Fi Protected Access (WPA) or WPA version 2 (WPA2) encryption, both of which are stronger than WEP.

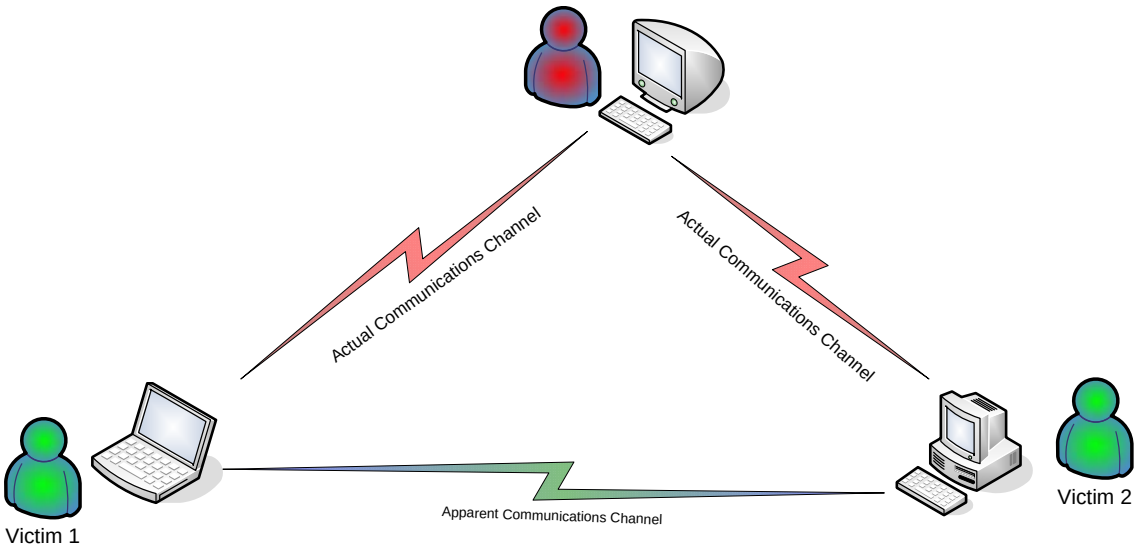

Figure 2.2: MITM attacks inject a malicious third party into a communications channel with the intent of reading and tampering with messages sent between victims.

To impersonate both victims, the attacker needs to overcome any technical controls in place. For example, unless authentication mechanisms are in place, such as those used in SSL-based communications, it is possible for an attacker to spoof, or impersonate, the victims. SSL communications can use a combination of public and private pieces of information known as keys to authenticate the parties in communication, so an attacker would need access to the private keys of both victims to carry out a successful MITM attack.

Note

SSL and Transport Layer Security (TLS) use both symmetric and asymmetric cryptography. Asymmetric encryption is used for authentication while symmetric encryption is used for large data transfers, as it is computationally more efficient. It is conceivable that an MITM attack could occur by breaking the symmetric key encryption after authentication has occurred. The use of strong symmetric encryption algorithms, such as the Advanced Encryption Standard (AES), makes that highly unlikely.

In addition to overcoming technical controls, the content of the messages injected by the attacker must be believable enough to convince the victim they are authentic. This is not difficult, especially in business communications where many exchanges are standardized. For example, it would not be difficult to change quantities on an order or replace a credit card number with another legitimate credit card number without raising suspicion.

Replay Attack

A replay attack is a type of MITM attack in which a message is captured by a malicious third party and resent or replayed for the target victim. For example, if Alice was to send a message to Bob saying "Send 100 widgets to Charles and charge to my account" and that message was captured by an attacker, the attacker could then resend the message to Bob. Bob in turn would then have orders to send a total of 200 widgets to Charles and charge them all to Alice. In a more realistic example, the message would be a structured transaction following a well-defined protocol, but the point is that unprotected messages can be captured and used again in unintended ways.

One way to protect against replay attacks is to use some type of session variable. For example, each message from Alice to Bob would include a message counter. The message counter is incremented after each transaction is sent. If this technique were used, Bob would recognize the second message sent by the attacker was a repeat of the first message and could safely ignore it. However, if the message transmission is in plain text (that is, unencrypted), the attacker could simply change the value of the message counter.

Alice and Bob might try to outwit eavesdroppers by having a non-obvious pattern in the way they increment the counter. Instead of incrementing the message counter by one, they might increment by 2, 101, the number of the day of the month of the transaction, or any other pattern. Unless the attacker knows the proper increment, the expected message counter would be incorrect and the recipient would recognize the message as invalid. Attackers could solve this problem by monitoring traffic between Alice and Bob until they have enough sample transactions to determine the rule for incrementing the message counter.

This simple example illustrates how "homegrown" solutions to protecting confidentiality can break down. Cryptography, the study and development of encryption algorithms, is a science as is cryptanalysis, the study of code breaking. It is highly unlikely that someone other than a specialist in cryptography could develop a sufficiently difficult-to-crack algorithm to warrant attempts at such development. A better solution is to use public algorithms, such as AES. Confidentiality is assured by a combination of the strength of the algorithm, which is publicly known, and the keys, which are kept private, used to encrypt messages.

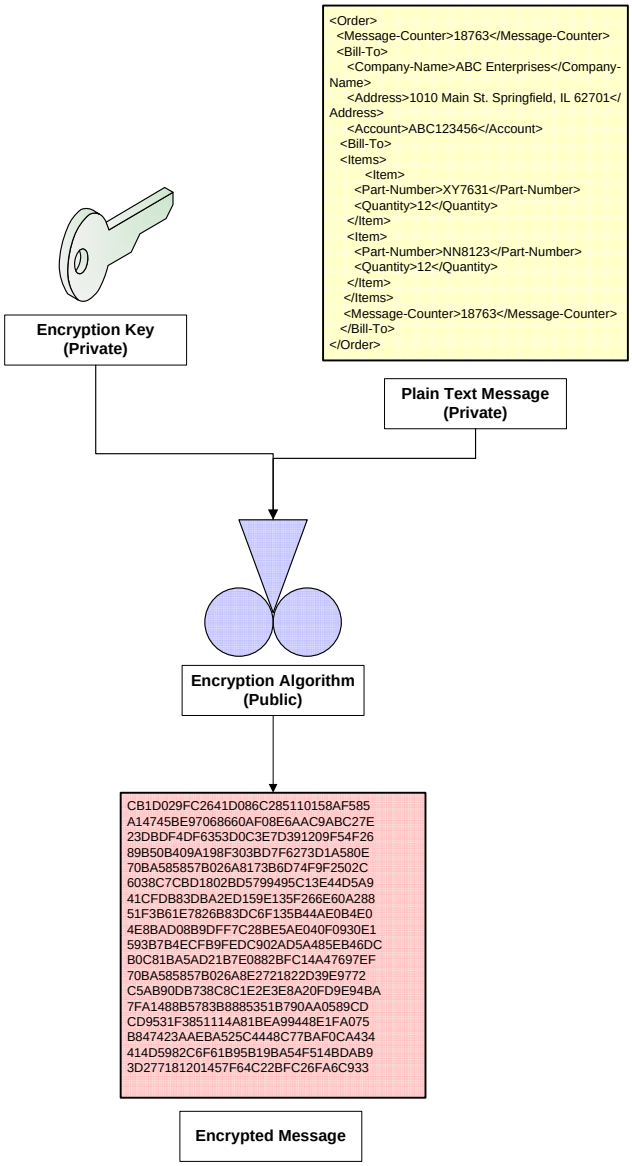

Figure 2.3: Confidentiality is ensured if a message and the encryption key are kept secret; there is no need to use a secret or home grown algorithm. In fact, public algorithms are subject to a great deal of cryptanalysis scrutiny and are more likely to provide codes that cannot be broken in a reasonable amount of time with reasonable resources.

MITM and replay attacks could be quite costly to businesses for two reasons. First, individual transactions could be repeated or tampered with as a means to commit fraud. The potential cost of a single act of fraud may be great enough on its own to justify implementing stronger security measures, such as using SSL for all business-essential communications. Perhaps a greater reason for concern is that without SSL encryption, *any* electronic communications could be called into question. This position is extreme but the lack of trust in communications systems could undermine business operations and efficiencies. Will salespersons call customers on the phone to verify electronically submitted orders? The use of a second means of communication, known as *out-of-channel communications*, is one way to reduce potential fraud, but it is highly inefficient for both parties. Securing communications with SSL-based communications is more efficient and practical for business operations.

Encrypting message transmissions protects data in motion. Data at rest and the servers and other devices used to store and process that data require additional technical controls to provide sufficient security for typical business operations.

Insufficiently Patched OSs and Applications

One of the most memorable malware attacks to broadly impact the Internet hit in January 2003. The SQL Slammer worm spread across the globe and infected tens of thousands of machines in minutes. The worm's Denial of Service (DoS) attack slowed Internet traffic and effectively blocked traffic on some segments. The malware took advantage of a vulnerability in the SQL Server database and the Microsoft SQL Server Desktop Engine. Microsoft had released a patch 6 months before the attack; unfortunately, many users of the affected systems did not patch their systems.

Although the impact of SQL Slammer was quite dramatic, the existence of program vulnerabilities is far from rare. The National Vulnerability Database (http://nvd.nist.gov/home.cfm), which tracks known vulnerabilities, listed 35,142 software vulnerabilities as of early February 2009, publishing on an average of 15 vulnerabilities per day. Vulnerabilities are not limited to popular databases and OSs; consider some of the vulnerabilities discovered over the past few years in widely used applications:

- Internet Explorer—A vulnerability in IE could allow remote code execution (Microsoft Security Advisory 961051).
- Microsoft Access—A vulnerability in an ActiveX Control could allow remote code execution(Microsoft Security Advisory 955179)
- Microsoft Excel—A vulnerability in Excel could allow remote code execution (Microsoft Security Advisory 947563)
- Xterm (Linux)—The default configuration of xterm on Debian GNU/Linux, and possibly Ubuntu, could potentially allow arbitrary code execution (CVE-2006-7236)

The cost of unpatched systems to businesses can be significant. As the example vulnerabilities highlighted in the previously list show, commonly deployed applications can be used to execute arbitrary code. When malicious code can be executed with administrator or root privileges, it is difficult if not impossible to prevent an attacker from gaining control of a device. Unpatched applications can provide attackers with a stepping stone to committing data breaches, tampering with databases, or denying access to mission-critical applications.

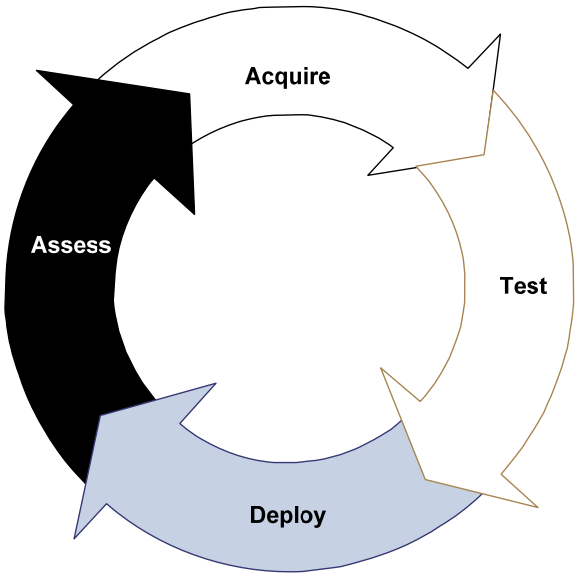

Figure 2.4: The patch management cycle starts with acquiring patches from vendors and other sources, testing them to ensure critical functions are maintained, deployed to devices, and assessed in operations.

Keeping track of applications, versions, configurations, and patch levels is challenging but a set of practices known as the patch management cycle (see Figure 2.4) is designed to address these challenges. The key steps in the patch management cycle are (1) acquiring the patch, (2) testing the patch in a controlled environment, (3) deploying the patch to production systems, and (4) assessing any problems with a patch deployment. Asset management systems can improve the efficiency of patching by automatically pushing patches to devices and providing reports on the status of patch operations. Hardware vendors are improving remote device manageability through offerings such as Intel's vPro and AMD's support for the desktop and mobile architecture for system hardware (DASH); asset management and patch management tools may take advantage of these for additional efficiency improvements.

The same tools that help with patch management can also help with another typical technical weakness.

Insufficient Use of Antivirus and Personal Firewalls

Using antivirus software is like driving with seat belts—we all know we should use the precautionary measure. The analogy quickly breaks down though. Although we rarely need seat belts because most of us have few if any accidents, most users are likely to encounter malicious software. Part of the problem is the prevalence of malware.

Malware can infect devices from multiple points of entry into a system:

- Malware attached to emails
- Malware-infected media files, such as video and music files
- Malware downloaded when visiting a compromised Web site, a technique known as a "drive-by download"
- Malware transmitted from another infected machine on the network using weaknesses in firewall configurations or network vulnerabilities to infect other devices

Anti-malware vendors are constantly updating signature-detection database and behavior-analysis systems used to detect malware. Like application and OS patching, anti-malware software has to be routinely updated to counter new and emerging threats. Malware developers know this, and some malware includes code to block updates. Sometimes the blocking techniques are simple, such as editing a local file used to map domain names to IP addresses, so antivirus software is directed not to the vendor's update site but to another non-functional site (thus, updates are never downloaded).

Personal firewalls can help stem the spread of malware by blocking traffic on ports that are not needed for legitimate purposes. This can, for example, prevent worms from accessing a device via a blocked port; it can also block outbound traffic, such as spam generated by a bot that has already infected the machine. Low-cost and free personal firewalls are readily available for Windows; Mac OS X and most Linux distributions include firewalls. Proper firewall configuration can provide an additional layer of security on devices.

The cost of insufficient use of antivirus and personal firewalls is manifested in poor performance in devices, unnecessary consumption of bandwidth in the case of devices infected with botnet software, increased demand for Help desk service to diagnose performance problems, and the cost of removing malware once it is detected.

Keyloggers and video frame grabbers are particularly dangerous types of malware. These not only compromise the systems they infect but also are designed to steal information, such as login credentials or confidential information, and transmit it to a point where the attacker can retrieve it. One of the reasons passwords and other authentication mechanisms should be updated frequently is because they may be leaked or stolen. Credit cards, drivers' licenses, and digital certificates all use expiration dates because something can go wrong and those artifacts, for whatever reason, cannot be trusted. Credit card and drivers' license issuers cannot go into the field and retrieve the cards (at least in any practical sense). Similarly, we cannot recover stolen passwords. Malware is just one of the reasons to frequently change authentication information.

Weak Boundary Security

As systems become more distributed and we adapt more service-oriented architectures, we find the need to move data further and sometimes across organizational boundaries. This practice is undermining the traditional notion of the network perimeter.

In the past, a company may have had all traffic moving over a firewall between the internal network and the Internet. Traffic across this boundary was restricted to those protocols needed for Web browsing, email, and instant messaging. Today, companies may have

- A database hosted by a third-party site with database protocols used to exchange data between client and server
- Internal applications invoking Web services provided by business partners; confidential data is moved back and forth between these two systems (in which case, digital certificates should be used to authenticate the partner's Web service and SSL should be used for communications)
- Remote users connecting to the corporate network using virtual private networks (VPNs)

Network perimeters today are more porous than they have been in the past. Now rather than depending too heavily on boundary security, we must have multiple layers of overlapping security (known as defense in depth) to protect data and systems. This security includes implementing technical controls to avoid the common weaknesses described in this section as well as securing data at rest and in motion with the use of encryption. Organizations that do not address the boundary security requirements risk well-known problems, including data breaches, compromised devices, and the potential loss of computing and network services.

We must be careful not to confuse information security with just network security; applications are another broad area of concern in information security.

Poor Application Security

It is somewhat ironic that improvements in our ability to protect OSs and network devices have led to a heightened awareness of application vulnerabilities. Like water seeking the lowest level, attackers look for the easiest way to reach their target. Today, the target is often information. Application vulnerabilities include:

- Injection flaws, such as SQL injection attacks in which SQL commands are sent as part of input data
- Cross-site scripting attacks, which allow attackers to execute scripts within the context of a user's browser
- Poorly managed authentication in distributed applications that allow, for example, a victim's username and passwords to be stolen
- Insecure communications, in which private and confidential information is sent in unencrypted or easily decrypted form

All of these and other common application vulnerabilities can be avoided with sound coding and software engineering practices.

> **Note**
>
> For more information about application security, especially Web applications, see the Open Web Application Security Project (OWASP) at http://www.owasp.org.

Automated application vulnerability scanning can help identify vulnerabilities in deployed applications and pre-deployment code. Some scanners work with source code using static analysis to identify weaknesses apparent from the structure of code, such as potential out-of-bounds references; other scanners perform dynamic analysis and probe applications for vulnerabilities while they run. The latter is especially useful when source code is not available.

As noted earlier, even widely used applications can contain vulnerabilities. Businesses, government agencies, and others can mitigate the risk and potentially avoid the cost of having application vulnerabilities exploited if they are detected before the system is moved into production. It is also less disruptive and more cost effective to correct problems as early as possible in the software development life cycle.

It should also be noted that incorrect configurations can lead to application vulnerabilities. Using default configuration and default passwords, for example, provide an easy way for attackers to get started compromising an application. As a general rule, configurations should implement only functions needed by business requirements. The more subsystems enabled in an application, the greater the surface area for an attack. Each unnecessary subsystem may bring with it vulnerabilities that can be leveraged by attackers.

This section has highlighted some of the technical weaknesses that can undermine information security. Not surprisingly, these weaknesses span the breadth of IT infrastructure from network architecture to endpoint devices to the ways we transmit sensitive and confidential information. Weaknesses, however, are not limited to technical issues.

Organizational Weaknesses

In many respects, the challenges of implementing and managing effective technical controls pale in comparison with the difficulties in addressing organizational weaknesses, such as insufficient or ineffective security awareness training. This section will consider how end user security training, security policies governing mobile devices, and the inappropriate use of business computers and networks can result in security vulnerabilities.

End User Training and Security Awareness

Technical controls alone will never constitute a comprehensive security strategy. Humans can override, alter, disconnect, turn off, and ignore technical controls. Technology is a supporting part of security controls; it is not the full picture; thus, it is imperative that employees, contractors, consultants, and business partners understand their role in the information security mosaic that protects business assets and data.

To get a sense of just how difficult it is to mitigate vulnerabilities related to the human factor in IT security, consider some of the findings of a 2008 survey by Cisco and Insight Express on data leaks (Source:http://cisco.com/en/US/solutions/collateral/ns170/ns896/ns895/Cisco_STL_Data_Leakage_2008.pdf). Some of the more telling findings include:

- 10% of surveyed employees have stolen or know of other employees who have stolen data or devices
- 10% of employees have lost or were the victims of theft of a company-issued computer, mobile device, or portable storage device containing corporate data in the past 12 months
- 11% of US IT decision makers indicate their company has suffered a data breach that included the theft of company data
- The top three concerns for data leaks are, in order: portable USB drives, email, and stolen laptops

These statistics demonstrate the widespread ignorance of sound security practices with regards to computer use or a disregard for those practices. Such breadth of weakness is not necessary to create a significant risk. For example, that same survey found that 14% of global respondents had changed security settings on their computers; 2% of US respondents had done so. Of those that did make changes, half of them admitted they did so to visit sites regardless of their company's policy, and more than one-third felt it was not the concern of their company if they did change security settings on company-issued devices! Like a chain that is only as strong as its weakest link, a small number of employees with cavalier attitudes are enough to compromise security.

Statistics such as these and anecdotal evidence about lost laptops, simplistic phishing lures, and irresponsible behaviors have led to a couple of myths about end user security awareness training that need to be dispelled.

End User Training Myths

Unfounded myths about users and their willingness or need to learn undermine an appreciation for what is necessary to improve the human factor components of information security.

Myth #1: "If security training worked, it would have worked by now."

This fatalistic view only rings true if we assume that our training methods are sufficient and we do not need to try other approaches. Widespread public health campaigns, such as anti-smoking efforts, and public safety campaigns, such as promoting the use of seat belts, have largely succeeded and can offer guidance on how to proceed. These successful campaigns are continuous and long running. Anti-smoking efforts that started in the late 1960s and early 1970s continue to some degree today. It is difficult to drive across state lines in the US without seeing signs to buckle up. Successful campaigns use a combination of techniques to get their point across, including humor. Talking crash test dummies taught us about car collisions. The point is that we should not give up on training employees about security because past methods have not worked; we can learn from others' successes.

Myth #2: "Younger workers are more tech savvy and therefore more skeptical of scams and do not need security training."

The idea that one generation will not repeat the mistakes of previous generations is appealing but lacks sufficient evidence to be believed. More importantly, social engineering attacks, malware, hacking techniques, and anti-forensic techniques are constantly changing. Some of us will not be tempted by a phishing scam promising extraordinary returns if we just send money to a foreign national in a temporary bind; that is no reason to assume we are immune to other scams or that we know all the ways attackers can infect a device with malware. Drive-by downloads from compromised Web sites were not known 10 years ago; why should we think that 10 years from now new techniques won't stump today's tech-savvy generation?

The impact on business, including the cost, of insufficient and ineffective end user training could be measured in computers infected with malware from sites users should not have visited, leaked information given in response to phishing lures that should have been ignored, and inaccurate data left after a disgruntled employee gained access to data using someone else's account left open after hours.

Lax Security with Mobile Devices

Mobile devices require both technical and organizational controls. Antivirus, personal firewalls, and vulnerability scanning (at least with tools such as the Microsoft Baseline Security Analyzer available at http://technet.microsoft.com/en-us/security/cc184923.aspx) fall on the technical side of the equation. Once again, the more difficult challenges come on the organizational side of things.

Part of the challenge with lax security with mobile devices is that employees are not aware of risks to mobile devices. The Privacy Clearinghouse Chronology of Data Breaches (http://www.privacyrights.org/ar/ChronDataBreaches.htm) has plenty of examples of stolen laptops containing tens of thousands of database records containing personal information. As more smartphones are used to access and store data, there will be more opportunity for confidential information to be lost or stolen. Employees should be trained in reasonable procedures for protecting mobile devices when they are in cars (a popular target) and in the use of encryption to prevent data from falling into the hands of thieves if a device is stolen.

Another problem, one that gets less attention, is the growing use of personal mobile devices in the workplace. Employees may purchase Blackberry and iPhone smartphones on their own and use them to access corporate data. These devices are not owned by the company, so there are limits to what the company can dictate while still allowing these devices to access corporate repositories. Consider how policies may need to be re-worked to accommodate these devices:

- If these devices were company owned, they could be standardized; however, the company may not want to limit access to only those with a particular device type or OS.
- As these are personal devices, companies may not be able to dictate how they are used when not accessing corporate systems. Sites that may be blocked from corporate networks may be readily accessible from a smartphone also used to access confidential data.
- Companies may have a policy dictating minimum security measures for an employee-owned device used on the corporate network but may not have the means to enforce that policy. For example, a policy may dictate up-to-date antivirus signatures but not be able to verify a configuration before allowing a user to download data to their device.

Here again, we have an example where technical controls are not enough. We need educated and cooperative employees who understand and follow policies. The cost to business, and presumably an employee's career, can be significant if a data breach is traced to a poorly-secured, personally-owned smartphone.

Inappropriate Use of Business Computers and Network Services

A final example or an organizational vulnerability is the improper use of computers and network services. Some might try to look at this from a lost productivity standpoint—if an employee is checking personal email or ordering personal items online, they are not productive from the company's perspective. However, it is equally plausible to argue that use of company systems allows an employee to attend to personal errands more efficiently and therefore leaves them more time to focus on their work. There is no universal formula for finding the proper balance, but we can reasonably conjecture that one exists. A more pressing problem than unproductive time is the potential to introduce malicious software on the network.

If an employee checks a personal email account, there may not be the same filters that are applied to the corporate email system, thus allowing malicious software to enter the network via email. Similarly, employees browsing to non-work-related sites can result in drive by downloading of malware. These sites are not just those considered inappropriate for the workplace; legitimate popular sites, such as news sites, could be compromised because of vulnerabilities in their systems which in turn result in an adverse impact on your network. The service support staff probably has enough to do without having to clean up a botnet infection on the corporate network because an employee surfed somewhere she did not belong.

Organizational weaknesses generally stem from human behavior. Changing human behavior is an art that may never be mastered. Nonetheless, helping employees understand the nature of security threats and their role in protecting the company's assets as well as themselves is the starting point to mitigating organizational weaknesses.

Options for Addressing These Threats

Broadly speaking, there are three approaches to dealing with technical and organizational weaknesses. There is always the option of doing nothing, or more properly, the option of continuing to function as is. At best, one can reasonably presume that the organization would continue with the same levels of risks. If there have been no major breaches, confidential communications have not been intercepted, and malware outbreaks are infrequent, this might seem like a prudent course of action. The problem with this scenario is that it assumes the overall security and business environment will stay the same. We know that is not true. Malware has become more difficult to detect, it spreads by more methods, the size of major data breaches is increasing, and cybercriminals appear to be getting better at covering their tracks during an attack.

At the other end of the spectrum is the spare-no-expense approach. Even in the best of economic conditions, this is not reasonable. We cannot simply buy security systems and deploy end point security applications like buck shot in the hopes of hitting all the weaknesses in our network.

A balanced approach is, not surprisingly, the one that is called for. We cannot let fear of security threats keep us from aligning security strategy with business strategy. One of the hallmarks of this alignment is identifying risks to the business strategy and then implementing a combination of technical and organizational controls.

The amount we spend on security should not exceed the value of the assets we are trying to protect and the costs incurred by the organization in the event of a breach. Losing a patient record may not directly cost a hospital, but it may have significant cost to a patient whose identity is stolen and could have detrimental impact on the trustworthiness of the hospital and its brand reputation. Regulations internalize some of those costs which were previously borne by those outside the organization. A risk assessment can help illuminate the assets we need to protect, the threats to those assets, and various combinations of technical and organizational controls that can help mitigate threats to those assets.

Summary

Sometimes we can be our own worst enemy. How we address technical and organizational weaknesses inside the organization can help or hinder our overall goals. Security is a function of technical controls, such as SSL for secure communications and disk encryption for reducing the risk of data compromise, and organizational controls, such as sufficient and effective training and realistic policies that account for changing ways employees access and use data. A balanced approach is based on risk management practices and incorporates both technical and organizational controls; this method can help mitigate risks while accounting for limited resources.

Chapter 3: Developing a High-Impact Security Management Strategy

Effective information security requires a combination of technical and organizational controls; however, running down a generic checklist is rarely sufficient. Instead, a high-impact security management strategy is driven by the particular needs of a business, and these needs span the breadth of business and technical operations within an organization. For example, consider some of the questions one should pose when developing a security strategy:

- What business processes and workflows are vulnerable to attack?
- If a particular server were compromised, what would be the impact on day-to-day operations to users or customers?
- How can we ensure that our networked applications communicate only with trusted, verified partner applications?
- Can exchange of digital documents be as secure, trustworthy, and enforceable as the exchange of paper documents?
- How can we ensure that confidential information can be exchanged over email and online with reasonable assurance that it won't be intercepted and disclosed to an unauthorized party?

The solution to address the answers to these questions will entail a combination of technical measures, such as hardening servers and deploying SSL certificates for secure communications and authentication, as well as organizational measures, such as developing and enforcing security policies, auditing and monitoring network activities, and providing security awareness training. In Chapters 1 and 2, we examined security threats, technical vulnerabilities, and organizational weaknesses that can directly impact the overall security posture of an organization. In this chapter, we build on those discussions and describe a framework for creating a high-impact security strategy. This task entails a number of steps that are divided into three broad categories:

- Review of business processes and workflows
- Review of technical infrastructure
- Definition of security policies and procedures

Each of these steps addresses both technical and organizational aspects of security, which are tightly coupled. We will not have effective security over the long term without appropriate attention to both.

Review of Business Processes and Workflows

Business processes range from the relatively simple, such as processing time cards, to complex multi-organization operations, such as order processing that entails just-in-time delivery. The flow of information is common to virtually all business processes and information security practices have to take into account those workflows.

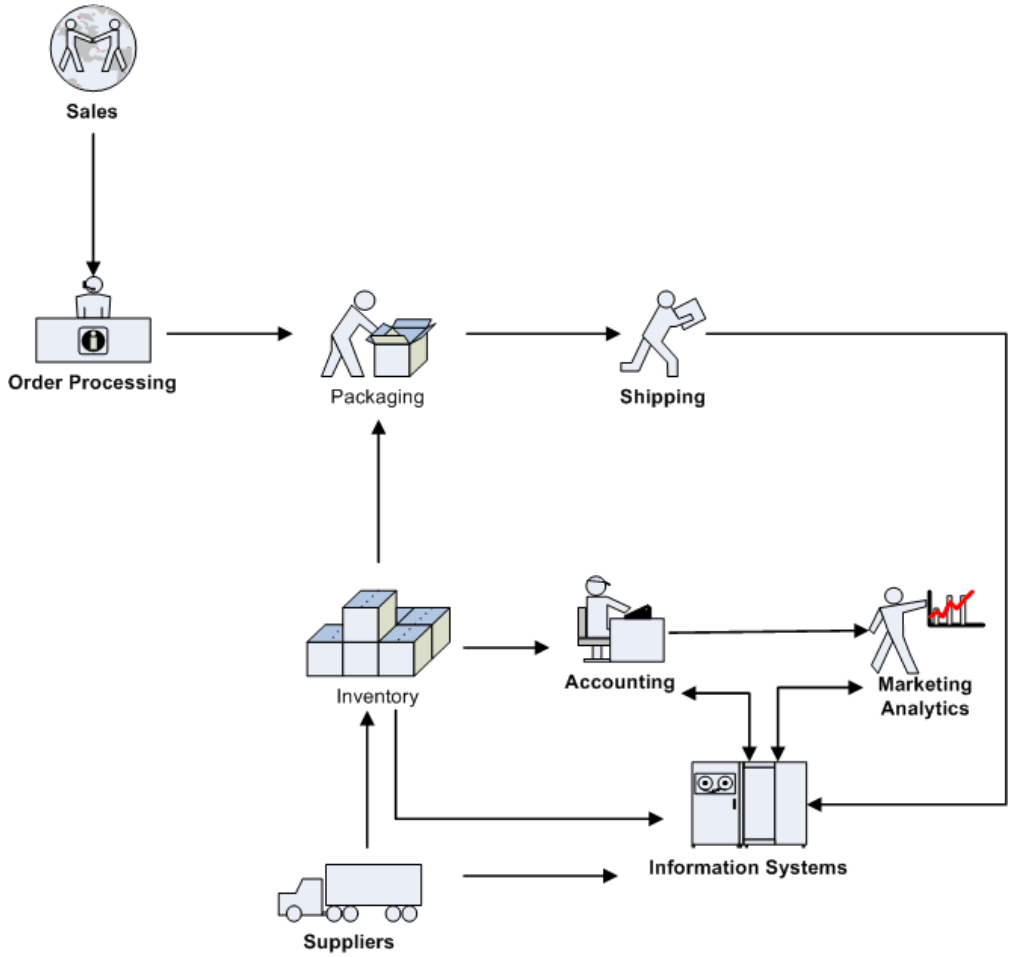

Figure 3.1: Information systems and the flow of information are fundamental aspects of virtually any business process. Securing these information flows begins with understanding the details of the flow and identifying risks to the information.

It is not sufficient to simply protect information at one point in a business flow because, like a chain, a business process is only as secure as its weakest link. For example, a retailer might lock down a database so securely that the time and effort required to break in and steal credit card data is not worth it. However, if credit card data is then sent from a point of sales system to the database using a wireless network encrypted with the weak WEP protocol, attackers will simply target that point in the business process. When we think of protecting information, we need to think in terms of the full life cycle of that information. How and where is it created? How is it transmitted? Where is it stored? How is it backed up and archived? If data is deleted from a production system, how long will it remain in backups? How is data protected when it is moved on physical media, such as backup tapes and disks managed by third-party service providers? These types of questions can be addressed by considering three elements of workflows:

- Data in motion
- Data at rest
- Access to information

When we have a solid understanding of these three elements, we can properly design security measures and implement controls to protect business process and information flows.

Data Classification and Security Measures

When considering information flows, remember that not all information is equally valuable or in need of the same levels of protection. A data classification scheme is a means of defining levels of protection appropriate for different types of information. Public data, such as press releases, do not require special controls because the purpose of this type of data is to share information outside the organization. Prior to release, however, a press release with time-sensitive data may be categorized as sensitive or even confidential if its early release could harm the business. A business' trade secrets or private customer financial data should be treated as confidential and provided with appropriate levels of protection when the data is being transmitted and when it is stored on business systems.

Data in Motion: Identifying Unencrypted Communications

Once we have identified core business processes, we can begin to look into the details of how information moves between servers, workstations, mobile devices, point of sale systems, and other kinds of devices. Key questions to consider are:

- Is the data sensitive, private, or confidential and therefore warrant additional attention to protect the privacy and integrity?
- Does the data move through systems or networks that might be vulnerable to attack?

For the purposes of discussion, we will concentrate on sensitive, private, and confidential information; that is, information, which if disclosed or tampered with, could adversely harm the business, its customers, business partners, or other stakeholders. Sensitive information is information that should not be released for general access, but if were made available, would not have serious impacts on the organization. Private and confidential information, in contrast, is information that if accessed in unauthorized ways would have severe impact on the organization. Private information pertains to individuals, such as customers and employees, while confidential information is related to the business itself, such as trade secrets. With regards to where the information flows, there are so many specific possibilities that it makes sense only to categorize the general range of networks and systems in terms of the level of additional security required.

Movement Within Secured Network Segments

One possibility is that information moves only within a controlled network environment that is already hardened (that is, secured beyond normal default configurations to reduce vulnerabilities). For example, suppose information from a transaction processing database is being copied every night to a data warehouse server on the same network segment. Given the high value of the transaction processing system and the data warehouse, we can assume network security staff has configured servers to run the minimal software needed to complete business operations, keeps the servers patched, and uses network firewalls, intrusion prevention systems (IPS), application firewalls, and database activity monitoring systems. In short, this network segment is made as secure as the risk warrants within the constraints of existing technologies and budgets.

Adding a layer of security with the use of encryption would add another level to a defense-in-depth strategy but at a cost. If the data warehouse required large volumes of data to be transferred within a relatively short window of operation, adding time to encrypt and decrypt data moving over a well-secured network could jeopardize finishing the operation in the time allotted while not significantly reducing the remain risks.

Movement Across Enterprise Networks

Next, consider the case of data moving across an enterprise network. In this case, we can imagine data moving outside of highly secured segments to areas of the network designed for performance and ease of use. There are many ways to use and misuse an enterprise network. Acceptable uses can range: mobile users connecting to the network using virtual private networks (VPNs), contractors and business partners accessing business systems related to their work, developers creating and testing new applications, and systems administrators installing new software and experimenting with different configurations. All of these activities can create risks that do not exist in a highly controlled environment. In addition, there may be activities that violate policy but manage to "fly under the radar." For example:

- Web application developers may deploy a Web server on an extra workstation in the office without following IT procedures
- An analyst may decide it would take IT too long to develop reports for her, so she creates a database and replicates data as needed from production databases
- A team of consultants set up shop in a conference room for a short-term project and install a wireless access point for their convenience

Security professionals might cringe at these examples while business professionals might be more willing to the weigh the pros and cons of bypassing the "IT bureaucracy." Let us just assume that there are times when reasonable professionals will disagree about the merit of such actions. How should we protect information flowing through parts of the network that could harbor vulnerable systems that could be used for data breaches?

Ideally, we could eliminate all unofficial applications, databases, and make-shift servers; but even if we could eliminate all such systems, the same conditions that prompted their introduction in the first place will likely remain. Another tactic, and one that fits in a defense-in-depth strategy, is to encrypt communications on the enterprise network, at least when dealing with confidential and private data. By using SSL-encrypted communications for the most valuable data, we make it much more difficult for unauthorized persons or programs to capture that data in transit.

Movement Outside of the Enterprise Network

Once data leaves the controlled boundaries of the enterprise network, we cannot safely make any assumptions about the security of such external systems or the applications or servers for which the data is destined. In this situation, SSL technologies provide two types of protection: confidential communication and reliable authentication.

Suppose you would like to send confidential information to a business partner over the Internet. If it is a small amount of data, you might use email; for larger amounts, FTP may be the tool of choice. In either case, there is no way to ensure that the message or transfer could not be intercepted and read unless the message is encrypted. SSL is the standard method for doing so. Another concern is ensuring that the message actually reaches the intended party.

Authentication is the process of verifying a party's identity. Usernames and passwords are frequently used when someone wants to employ an application or service, but these authentication mechanisms are of little use when trying to negotiate a transfer between two servers. A better option is to use digital certificates. These are electronic forms of identification that are designed to be virtually tamper-proof. If you receive a digital certificate electronically signed by a trusted third party, you have sound evidence that the sender is who it claims to be. Figure 3.2 shows an example of a certificate with information about the domain of the server for which it was issued, the issuer, that is the trusted SSL certificate vendor, valid dates for the certificate, and cryptographic attributes that are used to detect tampering.

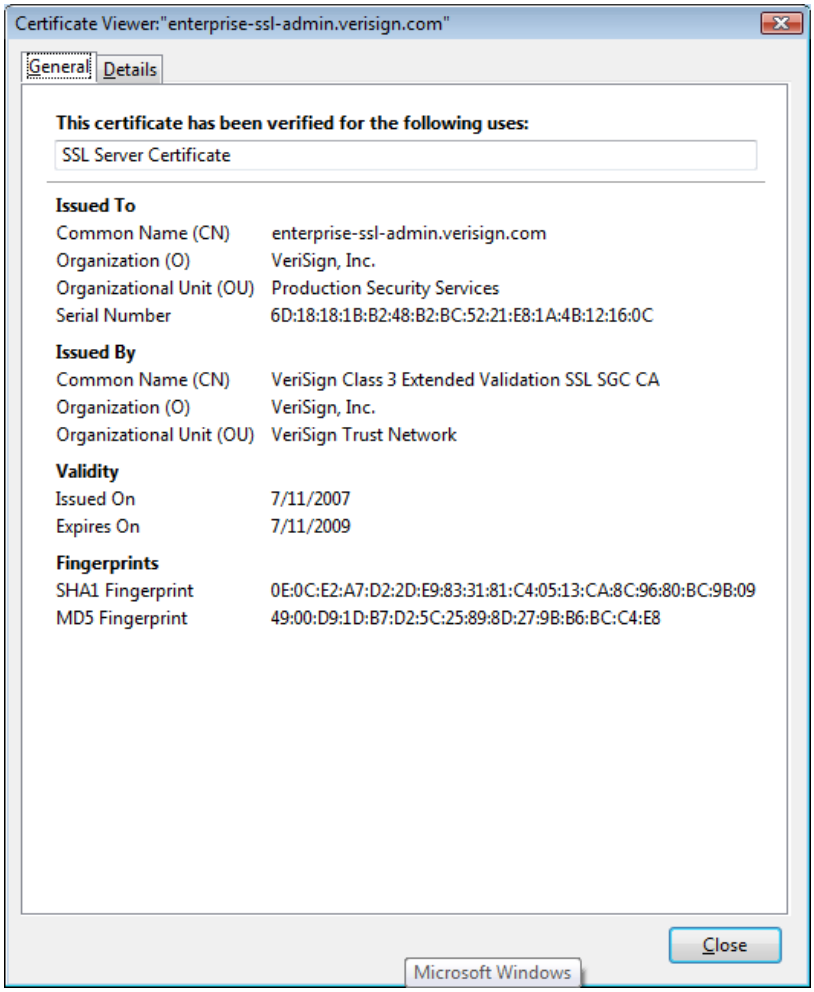

Figure 3.2: An SSL digital certificate is like a digital id card; it is evidence from a trusted third party that the server holding this certificate is actually part of the business it claims to be.

When data moves within highly secured segments of a network and performance considerations outweigh the marginal benefit of another security control, SSL encryption might not be used. However, when data moves outside the enterprise—and for confidential and private data, even within the enterprise network—SSL technologies can provide encryption for confidentiality and digital certificates for authentication purposes.

Data at Rest: Identify Servers Hosting Critical Applications

Another element of a high-impact security strategy is the proper management of servers hosting critical applications. Part of that management process addresses data and part addresses systems issues; here we will focus on data.

> **Cross-Reference**
>
> See the section, Server and Workstation Security Measures for more information about systems security.

Business processes and workflows copy, move, and delete data from many parts of the network. During the business process review, it is important to identify servers hosting critical applications and protected data. In the case of highly regulated data, it is important to be able to demonstrate that one knows where private and confidential data is located, how it is stored, and how it is protected. Part of that protection will often include encryption of data when stored persistently and ensuring that servers receiving protected data are properly authenticated, as discussed earlier.

Access to Information: Managing Identities and Authorizations

In addition to reviewing the flow of information and the servers that hold persistent copies of protected data, a high-impact security strategy begins with a review of identities and authorizations. Security technologies, such as SSL encrypted communications and digital certificates, depend on sound business practices that ensure authenticated users are legitimately authorized to view and manipulate information.

The modern workforce is highly dynamic, in both good economic times and during downturns. Employees leave positions to join other firms or move internally, consultants and contractors augment staff during peak demand periods, and businesses form collaborative arrangements with business partners to more efficiently deliver goods and services to their customers. One of the tasks that cannot be fully automated is reviewing user accounts and the privileges they have. This task might sound relatively easy, at least once the responsibility is delegated, but it is often more complicated than it first appears.

There are several ways in which difficulties arise, including:

- Staff may change positions and require some, but not all, of their existing privileges as well as new privileges.
- Companies may use federated identity management, in which each company depends on the other to define the roles of their own employees. This may be difficult to monitor because each business depends on the other.
- Accounts may be shared, sometimes informally, within close working groups.
- Developers and systems administrators may establish common application and database accounts that are shared by pools of users. These accounts may not appear on a standard report of each employee's authorizations.

As these examples demonstrate, tracking identities and authorizations can lead to more complex arrangements than may be apparent at first.

The first step in developing a high-impact security strategy is to understand (1) how data moves through an organization and outside an organization; (2) how data is managed when it is stored; and (3) who has access to that data. As we can follow from this discussion, it sounds easier in theory than it is in practice. Once we have a handle on business process and information flows, it is time to tackle another substantial, but doable, challenge: review the technical infrastructure.

Review of Technical Infrastructure

With a solid understanding of how information flows, we can turn our attention to understanding how the infrastructure that supports those flows can be secured. In particular, we will examine three categories of infrastructure security:

- Network security measures
- Server and workstation security measures
- Application security measures

The goal in examining each of these areas is to identify particular security issues that should be addressed with respect to each of these segments of the IT infrastructure.

Network Security Measures

The overall goal of network security measures is to ensure that the flow of information over the network is authorized and limited to legitimate business purposes. This is a tall order. Some of the technologies that are required include gateways to control traffic in and out of the network, SSL encryption to protect the confidentiality of information flowing through the network, intrusion prevention and monitoring applications to detect unusual patterns in network activity, and vulnerability scanning tools to help identify weaknesses in infrastructure configurations and software. As we drill down deeper into more specific technologies, we can see how various technologies can help protect the network.

Perimeter Device Configuration

Gateways, or firewalls, have improved from relatively simple, stateless packet inspectors to devices that provide deeper and more complex analysis of data flowing over a network. Of course, gateways are still needed to control how data flows in and out of a network, and that starts with controlling which ports are open for use. The emergence of tunneling—the process of using one protocol to carry as its payload traffic in another protocol—is just one example of a data flow that is too complex for simple firewall rules to handle. A perimeter security strategy should consider ways such as these that basic security measures may be circumvented. More advanced gateway devices, such as application firewalls, include software that can analyze data en route to an application and determine whether it is appropriate traffic.

Network Monitoring

IPS may further improve overall network security by analyzing traffic patterns and detecting anomalous activity. When planning on the use of IPS, it helps to understand how they work. IPS can detect anomalous network activity through the use of rules, by comparison to baseline statistical patterns, or both. An advantage of rule-based approaches is that they can be shared across users of an IPS system. For example, an attack using a known vulnerability in an operating system (OS) component may require a particular sequence of actions to initiate, and an IPS could have a rule to detect that pattern. Statistical pattern methods are complementary and can help accommodate the unique activities on a network. For example, it may be perfectly normal for large data transfers to occur between servers during the middle of the night but not in the early morning. If the latter were to occur, it might be an indication of a data breach in progress.

Reporting and Alert Systems

It would be difficult to find a systems administrator or network manager complaining about not enough data or network activity. Security systems, applications, and OSs are profuse generators of logging data. The problem is not lack of data but extracting useful information from that data. Security information management (SIM) systems are tools for collecting, consolidating, and reporting from multiple devices. There are formidable challenges to building SIMs, and we should manage our expectations for these tools.

SIMs are useful today as consolidated reporting tools. Using protocols such as the Simple Network Management Protocol (SNMP), SIMs can collect data from multiple devices and help network administrators review data from across a variety of device types. As the technology advances, more complex analysis may be available, but in some cases, a good solid tool for reporting a diverse set of facts can be still be useful.

Network security at one level entails a combination of perimeter devices, network monitoring, and reporting systems. We have seen how security of data flowing over the network is enhanced with the use of SSL encryption. Next, we will examine the role of server and workstation security measures in strategies for protecting IT infrastructure.

Server and Workstation Security Measures

Servers and workstations are like factories in an industrial society: they are producers of specialized artifacts that depend on each other for inputs and use shared resources for distributing their outputs. Unlike the physical world where it would be difficult to masquerade as a factory, the digital world of servers and workstations do not have the same barriers to fraud. In terms of a high-impact security strategy, a key element is ensuring that servers and workstations can trust each other. For example, when a Web service receives a message requesting a service or piece of data, the server running that Web service needs to be able to trust the requestor if private or confidential information is being requested. SSL digital certificates are the standard means for establishing this trust. In addition to trusting that servers and workstations are what they appear to be, it is important to implement practices that protect the integrity of these devices.

Hardening OSs

A quick scan of a vulnerability database, such as the National Vulnerability Database (http://nvd.nist.gov/), will show many different types of vulnerabilities affecting a variety of components, including:

- Web servers
- FTP servers
- Media players
- Network management software
- Process monitoring applications

Some of the problems involve technical issues, such as buffer overflows, and the allowance of remote execution of code and privilege escalation. If systems administrators do not have enough to keep themselves awake at night, a visit to a vulnerability database will solve that problem. Modern OSs are all complex, multi-faceted applications and they have vulnerabilities. One of the best ways to mitigate the risks associated with these vulnerabilities is to harden the OS—that is, minimize the number of services running and the types of applications available on systems—and properly configure the OS.

A general rule of thumb is if a service is not needed, it should not be running. FTP servers, for example, have seen more than their share of vulnerabilities and exploits. If FTP is not required, do not run it. Similarly, production servers should not have compilers installed unless there is some compelling reason. Code should be developed and compiled on development servers and the binaries then ported to a production server. If an attacker were able to compromise a production server and had access to a compiler, the attacker could conceivably download code, compile it locally, and install it on the server. Of course, an attacker could also compile the code remotely and install it, but the attacker would need a compiler for every different type of system targeted; having access to a local compiler just makes an attacker's life easier.

Resource

For more information about hardening OSs, see the Bastille-Linux Project at http://bastille-linux.sourceforge.net/ and the benchmark tools at Center for Internet Security at http://cisecurity.org/bench.html.

Hardening also requires proper configuration, which includes changing default passwords, not re-using passwords across administrator/root accounts, enforcing a strong password policy, and shutting down unnecessary services and daemons. Hardening an OS should be a standardized procedure. Consistency can help improve overall security and ease administrative overhead. However, there are times where some servers should have additional controls put in place. For example, access to database servers may warrant strong authentication, such as multi-factor controls or a challenge-response system.

Patching

A third element of a server and workstation security strategy is patching. We've already described the extent of vulnerabilities and one method for dealing with them (removing the vulnerable applications through hardening). Not all vulnerable applications can be removed, but many of them can be patched. Patching is a sufficiently complex process that it should be carefully considered and procedures formulated for patching in a high-impact security strategy. Some of the key elements of a sound patching strategy are:

- Procedures for monitoring the availability of patches
- Methods for assessing the importance of a patch and the speed with which it should be applied
- Rankings of different instances of systems that should be patched so that IT support staff can prioritize patching operations
- Procedures for testing and then rolling out patches
- Bypass procedures for fast-tracking emergency patches (this should be done judiciously due to the risk of disrupting production operations when sufficient testing is not undertaken)

Servers and workstations require support from several elements of a security strategy, including the use of digital certificates, OS hardening, and patching to reduce vulnerabilities.

Application Security Measures

The third leg of the infrastructure review triad is application security. For our purposes, the term "application" includes software that ranges from monolithic mainframe applications to individual Web services. As part of a security strategy, businesses should assess application-specific security measures, including:

- Access controls
- Security testing
- Hardening application components

As we shall see, these considerations parallel some of the issues in server and workstation security; however, these tend to address security more from a software engineering perspective than from a systems management point of view.

Access Controls

At the very minimum, application security entails specifying who can use an application and what can they do, or in security parlance, authentication and authorization.

Think of authentication and what comes to mind? Probably common scenarios such as a user logging into an email service or verifying the identity of a business running a Web site likely come to mind by checking an SSL certificate (See Figure 3.3, which is actually displaying an Extended Validation—EV—SSL certificate, a form of digital certificate that requires more extensive verification than conventional SSL certificates).

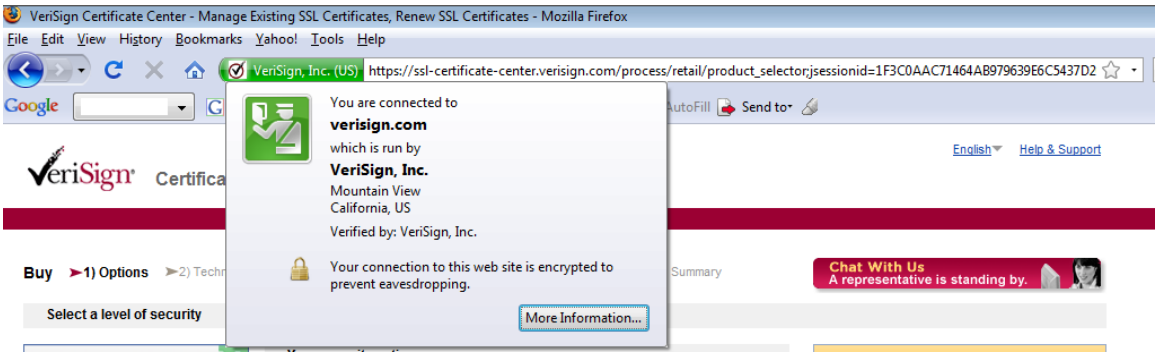

Figure 3.3: When we visit a site, we want to make sure we are dealing with the business we think we are dealing with. In other words, we want to trust the Web site. SSL also supports mutual authentication, which allows the Web site to trust its visitors.

Users trusting a Web site are only half of the authentication process. Business need to verify that business partners, customers, and others who are given access to their applications are who they claim to be. Just as customers want to be sure of the identity of a business behind a Web site before handing over a credit card number, businesses need to be sure of who they are dealing with before handing over data or granting access to services.

This can be done with mutual authentication. For example, a retailer might want suppliers to have access to inventory levels as part of a just-in-time delivery plan. Mutual authentication is in the interest of all parties. The retailer probably does not want competitors poking around its operational databases, and suppliers would not want to lose competitive advantage that access to detailed inventory information can provide.

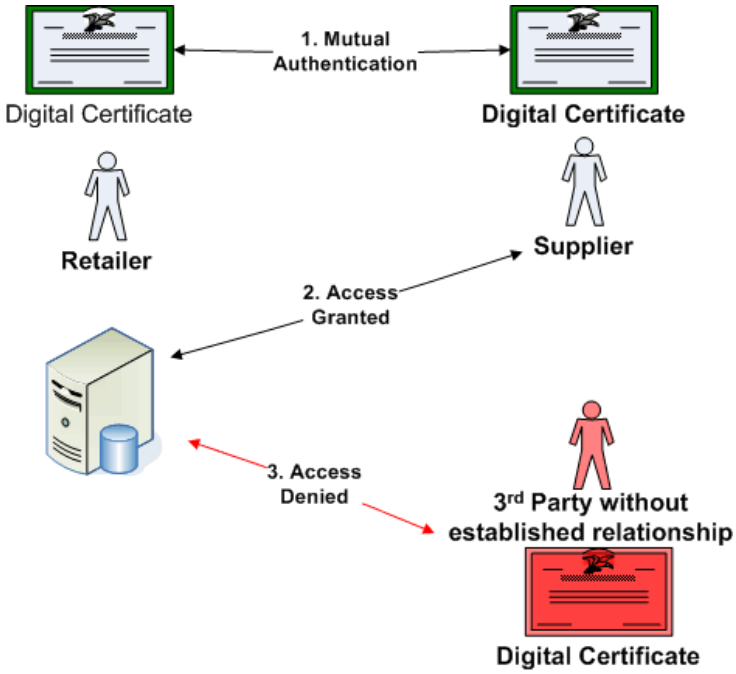

Figure 3.4: Mutual authentication via digital certificates can be used to control access to confidential information and services.

Access controls are based on some level of trust. We trust users not to share their passwords, to change them frequently, and to not reuse them. A business could conceivably just hand out passwords to business partners but that introduces new risks. For example, the business partner might deal with several retailers, each giving out passwords; to keep things manageable, the partner keeps the passwords written down on a sticky note, or worse, recorded in a wiki or other collaboration site. Digital certificates avoid this type of problem. Instead of trusting account users to keep passwords secret, we trust digital certificate providers to use reliable procedures to verify identities and to manage certificate operations, such as revoking certificates when needed.

Security Testing

Security testing is a complex subject but one that can and should be managed in the scope of a broad security strategy. Security testing should be done before a new application is released to production and throughout the life of the application. Initial testing should include tests to ensure:

- Processes run with the least privileges required to function
- Applications fail securely—for example, if an unexpected input is passed to an application, the application should gracefully fail and not suffer a buffer overflow or similar problem that leaves the application in a vulnerable state
- Expose only needed functionality; this reduces the number of ways an attacker can compromise the system and is known as "reducing the attack surface"
- Unusual or unexpected events are logged with sufficient detail to enable administrators and developers to diagnose the problem
- Applications function properly on hardened servers (see the earlier section on Hardening OSs)

Ongoing testing is required for several reasons. First, vulnerabilities in an application or constituent component may be discovered after the application is deployed. Second, during the course of routine patching, a new, unknown vulnerability could be introduced. Third, the configuration states of applications change over time and users may be granted elevated privileges that introduce additional vulnerabilities. Also, applications may be used in new ways, such as providing data to business partners outside the enterprise network, which should prompt thorough testing. Automated testing tools and vulnerability scanners should be used make this process more efficient than a completely manual operation.

Resource

See the Open Web Application Security Project (OWASP) for more information about best practices in application security testing at http://www.owasp.org/index.php/Main_Page.

Hardening Application Components

The last of the application security measures is hardening applications. As with OS hardening, the goal is to reduce vulnerabilities in an application. Security testing can reveal potential problems with software such as:

- Injection attack vulnerabilities—This can occur if inputs are not properly scrubbed before they are passed to modules or subsystems, such as database query processors; SQL injection attacks are perhaps the most well-known form of such attacks
- Insecure configurations in subsystems such as application servers and database listeners
- Hard-coded username and passwords for database accounts or other service accounts
- Unnecessarily elevated privileges

Many of these vulnerabilities can be corrected by changing code or configuration parameters. In other cases, additional measures, such as the use of application firewalls or database activity monitoring systems, may be warranted.

A review of technical infrastructure can help identify security measures for network security, server and workstation security, and application-specific measures. Not surprisingly, many fundamental security controls, such as the use of SSL for encryption, the use of digital certificates for authentication, and vulnerability scanning play prominent roles in protecting IT infrastructure.

IT environments are highly dynamic. Reviewing business processes, workflows, and IT infrastructure at one point in time is necessary but not sufficient for developing and maintaining adequate security. An ongoing governance process is required as well.

Security Policies and Governing Procedures

Security practices in an organization may begin with best practices established by the security community but will inevitably change to accommodate the particular needs of the organization. Costs and benefits are balanced. Compliance requirements are targeted. Business strategies are accommodated. Even given such dynamic constraints, it is important to formulate policies and governing procedures to avoid *ad hoc* responses to situations and to ensure that lessons learned over time are captured and incorporated into ongoing procedures.

In order to maintain a high-impact security strategy, well-defined policies and procedures should be established covering a number of topics:

- Use of encryption to protect the confidentiality of data at rest and data in motion
- Use of server authentication and mutual authentication for application services; these policies should describe when digital certificates should be used, limits of self-signed digital certificates (that is, digital certificates created by the user of the certificate, not a trusted third party), and the need for mutual authentication in Web services providing private or confidential data
- An overview of patch management procedures, including monitoring the release of patches, testing patches prior to use in production environments, and exceptions for emergency patching
- Processes for hardening OSs and applications to eliminate known vulnerabilities
- Use of vulnerability scanning and reporting tools
- Workstation security practices, including the use of antivirus, anti-spyware, personal firewalls, and disk encryption
- Secure use of mobile devices and limits on the types of data that may be copied to mobile devices
- Security awareness training for staff, contractors, and consultants as well as acceptable use policies clearly describing the types of activities that may be performed on the organization's IT infrastructure
- Auditing and monitoring requirements to maintain compliance with government and industry regulations

Policies addressing these areas and others related to security require maintenance. They have to be modified to accommodate changes in technology, business practices, and business strategy. Governing structures that include both IT and business executives familiar with the breadth of the business environment and current strategy are necessary to ensure that policies and procedures remain useful guides to security practices and not simply documents on a shelf shown to auditors once a year.

Summary

Creating and maintaining a high-impact security strategy begins with understanding business processes and workflow. This process is followed by an analysis of IT infrastructure, particularly networking services, servers and workstations, and applications. The final step is creating policies and governing procedures that shape and maintain a sufficiently secure environment. Throughout this chapter, we have seen recurring reference to fundamental security technologies such as SSL, encryption, and digital certificates as well as core security practices, such as application and OS hardening and vulnerability scanning. This should be no surprise. These technologies and practices are well-established elements of information security best practices that one will see over and over again.

Chapter 4: Best Practices for Implementing a Business-Centric Security Management Strategy

A business-centric security management strategy is multifaceted and takes into account both the technical and organizational aspects of information security. Throughout this guide, we have seen how security threats and vulnerabilities can undermine business operations and integrity, and we have discussed methods for developing a security strategy. In this, the final chapter of the guide, we turn our attention to examining best practices for implementing a business-centric security management strategy.

So how is business-centric any different from other approach to information security? The starting point is the business strategy. What are the goals and objectives of the business (or other organization) and how are they implemented? The answers to those questions start to frame the security discussion because we can assess risks to particular business processes and assets. Part of that assessment process is determining a relative value for an asset or process that is being protected. For example, we wouldn't invest more than the value of car in an anti-theft device for the vehicle. The same logic applies in information security. We mitigate risks to business information assets according to the value of those assets and the priority we assign to them.

Once we understand threats, vulnerabilities, and the risks and costs associated with them, we can then formulate a security strategy for protecting the business. This chapter examines specific methods for mitigating information security risks. As we shall see, one security control, or measure, can help reduce multiple risks, and every risk is ideally mitigated by more than one control. Of course, the reality of business is that we cannot always have our best case scenario, but we strive to get as close as possible.

The fundamental areas of a business-centric security management strategy span a number of areas and include:

- Protecting critical servers
- Protecting mobile devices and communications
- Deploying sufficient network defenses
- Providing end user training

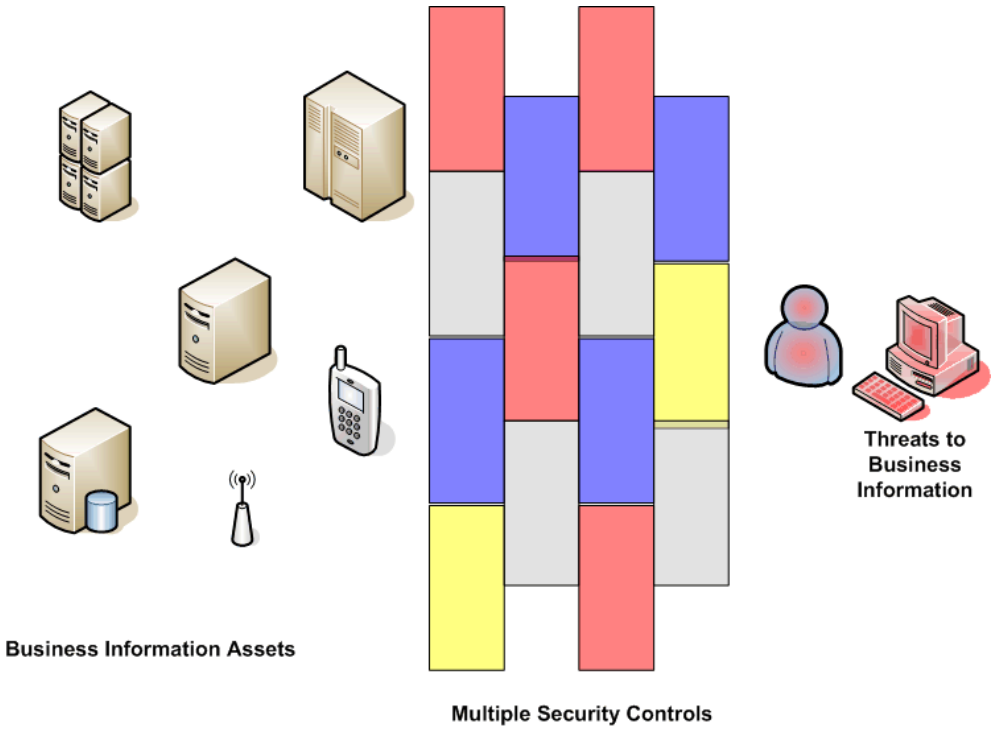

Figure 4.1: A guiding principle of best practices in business-centric security strategy is to apply multiple security controls in an overlapping manner to create a defense-in-depth approach to mitigating risks.

The drivers behind the best practices in each of these areas are the need to maintain the confidentiality of business information, integrity of that data, and the availability of information systems and assets. What follows is a non-exhaustive set of best practices that serve those drivers. Servers that support critical applications, maintain enterprise databases and perform other essential functions are a good place to begin our discussion.

Protecting Critical Servers

A critical server is one that, if it were to go down or otherwise have degraded performance, would have an adverse impact on business operations. Examples include email servers, database servers, and application servers used in production environments. It is important to classify servers in terms of their criticality because, as with data classification, some servers are more important than others; when it comes time to allocate information security resources, it is imperative to know how to prioritize server security spending.

What Constitutes a Critical Server?

How do we distinguish critical servers from non-critical servers? We need to start with business strategy and the business processes put in place to support them. Note that we do not start by asking opinions of users of those servers. Developers, for example, may consider their servers critical because in effect they are "production" servers from their perspective. If developers' database server goes down, they will not be writing much database code; that does not, however, make it a production server and therefore possibly a critical server. Of all the production servers, some of these are critical because business processes depend on them, and if they were to fail, the business process could not be executed or could only be executed at a significantly slower pace.

Clearly this is a gray area where reasonable people can disagree. For example, many of us might consider a Web server hosting a site on corporate charitable giving as a production system but not critical; if it were down for the day, it would be an inconvenience but work could be made up when the system is restored. In general, we can think of server categorization as a subset of all enterprise servers that are most important for business operations. Many, but perhaps not all, production servers may be categorized as critical.

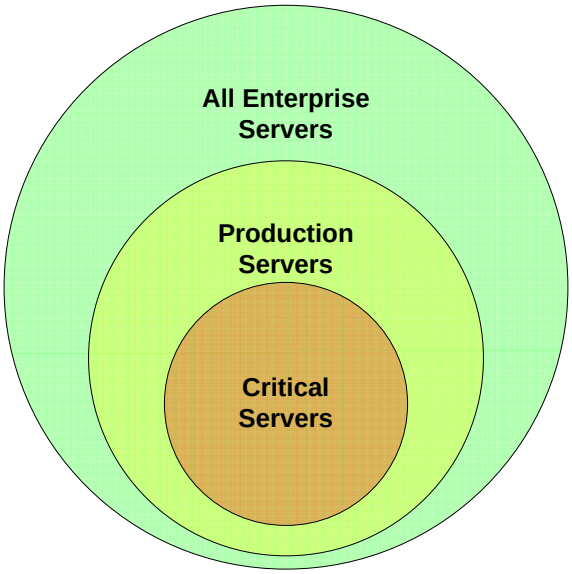

Figure 4.2: Servers can be categorized in terms of criticality to business operations. Critical servers have the highest priority for security measures because their disruption can have significant adverse impact on business operations.

Once critical servers have been identified, we can apply a defense-in-depth strategy to protect them. We should apply these principles to all servers if possible, but we should start with critical servers, then other production servers, and then to all other enterprise servers if there are sufficient resources.

Some of the multiple, overlapping security measures we can use include:

- Encrypted communications
- Hardened operating systems (OSs)
- Locked-down databases running on those servers

These three measures represent the types of controls that can be applied to protect information exchange between servers, reduce the attack surface of the OS, and reduce vulnerabilities in core applications running on these critical servers.

Using Encrypted Communications

Servers can house data from the various data classification categories, such as public, sensitive, private, and confidential.

- Public data can be freely disclosed; sensitive data should not be disclosed but would not cause significant harm to the company if it were disclosed.
- Sensitive data should not be disclosed, but if it were, that would not cause significant harm to the organization. Examples of sensitive data include project schedules and approved vendor lists.
- Private data is data about a third party, such as a customer or patient, that must not be disclosed outside of established procedures.
- Confidential data is company proprietary data, such as trade secrets, that need to be keep tightly controlled to prevent adverse affects on the organization. In theory, we might only be concerned about protecting communications when private and confidential data is involved; however, as servers may share different categories of data, we should apply security to protect the category of data warranting the most control.

Consider a Web application with a product and order database. The product catalog—including product lists, descriptions, and current pricing—is public information. Customer order data, including shipping addresses, billing addresses, and credit information, is private. Rather than risk disclosing private customer data, all communications between the application and the customer should be protected with encrypted communications.

SSL/TLS communication is the industry standard method for secure communications (TLS is also known as SSL version 3). It provides authentication so that we can verify the identity of the server we are working with as well as encryption of data communications between servers or between servers and clients. (Actually, the SSL/TLS standards do not require encryption of data to be compliant with the standards, but SSL/TLS is often used for encryption).

SSL encrypted communications can help mitigate a number of threats:

- Man-in-the-middle attacks in which an attacker intercepts a message between parties and alters the message stream. SSL encryption scrambles the content of messages and related services, such as digital signatures, and provide authentication and non-repudiation functions.
- Eavesdropping on communications. Protecting against this threat is especially important if the communications travel over unencrypted or weakly encrypted wireless networks. An early wireless encryption standard, WEP, is fairly easily cracked and should not be depended on to protect the confidentiality of server-to-server or server-to-client communication. Fortunately, if the server encrypts data using SSL before it is sent over a weakly protected wireless network, attackers will not be able to decipher the message in any reasonable period of time.
- Insider attacks from persons with access to internal communications. An internal attacker who does not have access to an application or database may still be able to capture data from those systems if the data were transmitted in unencrypted form. Even in the case of communications between internal servers, there is often a need for encrypted communications.

Remember, data in motion is not protected by the application and database access controls that help protect that data when it is at rest.

Hardening Server OSs

Hardening an OS reduces the potential vulnerabilities by using several techniques:

- Changing default configurations
- Removing default accounts
- Shutting down services that are not required
- Removing applications not needed in a production environment, such as removing compilers on production servers that run applications developed and compiled on other servers
- Reducing privileges on all accounts to the minimum set needed to perform business operations
- Patching the OS to apply security updates

Resources

For more information about hardening OSs, see the Bastille Hardening program at http://bastille-linux.sourceforge.net/ and Center for Internet Security Benchmarks at http://cisecurity.org/bench.html.

We should also apply the same principles to enterprise applications running on these servers. We'll consider databases as an example.

Locking Down Databases

Databases are a prime target for attackers because databases often store valuable information. Even if the server uses SSL-encrypted communications and the OS is hardened, attackers may be able to steal data by attacking at the application layer.

Locking down a database includes several steps:

- Removing or disabling default accounts and schemas
- Changing default passwords
- Removing unnecessary database options
- Securing the database listener, the process that establishes connections to the database
- Applying access controls to database files and directories
- Implementing strong password policies or other authentication measures to reduce the risk of password-cracking attacks

In addition to securing the database server, developers should be aware of coding techniques for avoiding SQL injection attacks. All the measures previously listed will not block an apparently legitimate query that is sent to the database by an approved application. It is the developers' responsibility to implement application code that is not vulnerable to such attacks.

> **Resource**
>
> See Colin Angus Mackay's *SQL Injection Attacks and Some Tips on How to Prevent Them* for more on this topic.

In addition to protecting servers, businesses should adapt their security measures to protect information when it is stored or used on mobile devices.

Protect Mobile Devices and Communications

Mobile devices are now commonplace. Smartphones, netbooks, laptops, and other mobile devices are de facto parts of the IT infrastructure. We do not generally consider mobile devices as part of the IT asset base, which includes servers, network hardware, desktop devices, and so on. This must change. Employees, business partners, contractors, consultants, and customers are using mobile devices to conduct business. Businesses with large consumer customer bases, such as banks, are creating mobile versions of their online services, such as online banking. Mobile devices are an established and widely adapted platform that we need to consider in a business-centric security strategy.

There is a significant difference between many mobile devices used for business and other IT hardware: the mobile devices are often not owned by the business. Obviously, this means that a business is not in full control of the device, thus,

- There is no standard platform for all mobile devices used for the business
- IT probably does not have an inventory of mobile devices
- These devices are not managed within a business' asset management program

There are, however, ways businesses can control what business data and business operations are allowed on non-business-owned mobile devices. This is done through a series of security policies that define security controls that should be in place before business is conducted with an employee-owned mobile device. These policies assert a business need to protect information assets while recognizing that the mobile device is ultimately owned and controlled by someone else. To distinguish these different types of devices, we will refer to employee or other third-party owned devices as semi-managed devices.

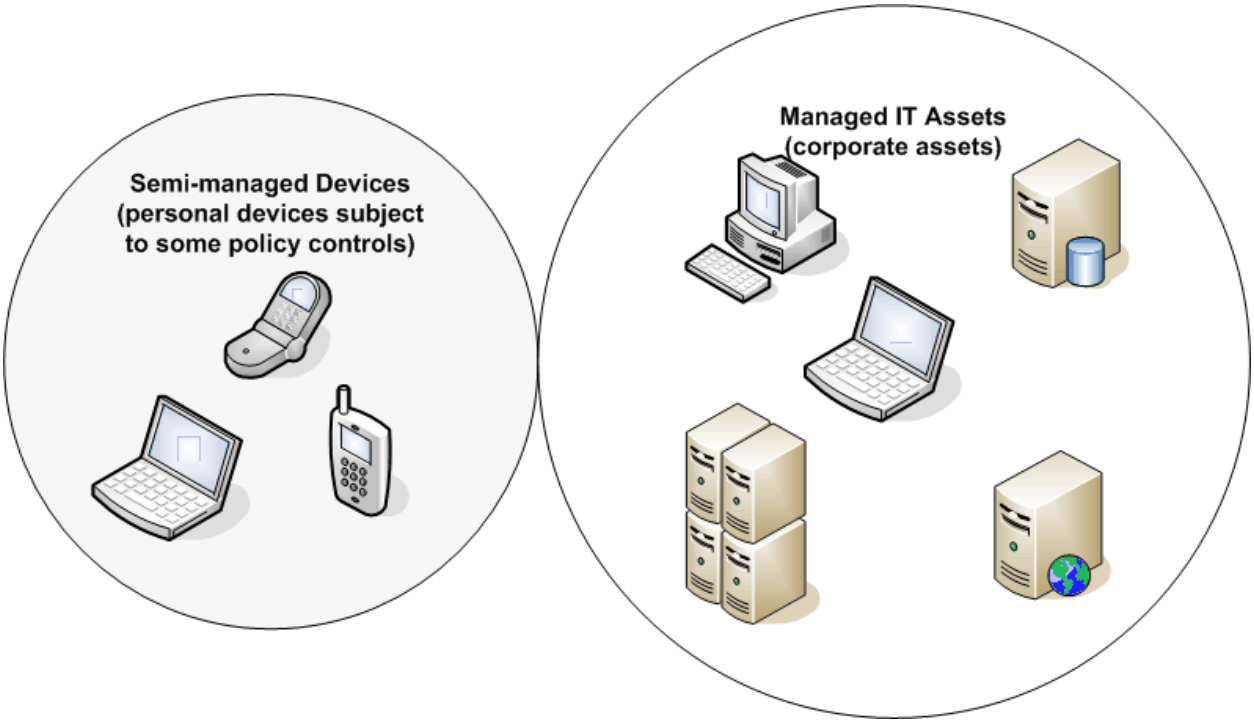

Figure 4.3: Mobile devices owned by employees are only semi-managed; however, they may be subject to policies governing conditions under which business data may be stored or transmitted to those devices.

For both managed and semi-managed mobile devices, businesses can use several policy and technical measures to reduce risks to data related to mobile devices:

- Encrypting communications with mobile devices
- Authenticating mobile devices with digital certificates
- Maintaining OS patches
- Keeping antivirus software up to date

As with server security, we use multiple measures in to implement our defense-in-depth strategy. That strategy enables us to mitigate multiple risks with a single security control and to apply multiple controls to individual risks.

Encrypt Communications with Mobile Devices

Data that is transmitted to and received from mobile devices may be sent over wireless communication providers' private cell phone networks as well as the Internet. This may not be a concern for many types of communications, but when dealing with private and confidential data, especially when there is a regulatory responsibility to protect this data, encrypting communications to mobile devices may be necessary.

Unlike some of the other security controls we can dictate for mobile devices, this one is well within the control of the business. Private and confidential data is sent only over an SSL-encrypted communication channel. Part of the SSL protocol defines a handshaking procedure between the server and client, so we can be sure that the client will receive the data only after establishing a secure connection. Of course, without sufficiently strong authentication, we run the risk of transmitting data to a spoofed device.

Authenticate Mobile Devices with Digital Certificates

Digital certificates are a key part of ensuring we are communicating with the mobile device we believe we are communicating with. This allows parties in a communication session to authenticate the identity of the devices with which they are dealing. Let's take a look at digital certificate capabilities on a couple of mobile device platforms.

When using the Windows Mobile 6 OS, it is relatively easy to install digital certificates. Windows Mobile is preconfigured to manage three types of certificates:

- Personal certificates maintained in the MY store
- Intermediate Certification Authority (CA) certificates, which are stored in the CA store
- Root CAs, which are stored in the ROOT store

These certificates are used by applications communicating with the device. For example, Microsoft Exchange ActiveSync verifies the trustworthiness of a device by examining its digital certificate. Windows Mobile also provides a cryptography application programming interface (API) for working with digital signatures and digital certificates.

The popular Blackberry smartphone also supports digital certificate-based authentication. These mobile devices support the Extensible Authentication Protocol – Transport Layer Security (EAP-TLS) for mutual authentication and the use of client digital certificates.

Mobile device vendors and software developers are providing some of the tools needed to secure mobile device communication and provide for digital certificate-based authentication. It is business' role to create policies defining when this functionality should be used.

Maintain OS Patches

Mobile devices are like other devices on the corporate network: they run with complex OSs that occasionally need to be patched. Not all patches are security patches, but when a patch is released to correct a vulnerability, the patch should be installed in order to reduce risks.

Many smartphones are personally owned, so businesses cannot force device owners to patch. There may be legitimate reasons for not patching. For example, if a patch to correct one problem introduces another, the user may conclude from their perspective that they would rather live with the security vulnerability. This may not be in the best interest of the business.

In general, there are two methods for ensuring devices are properly patched and configured. Network access control servers can query a device trying to connect to the network and determine whether the device's configuration meets minimum requirements. This works well when a conventional laptop running Windows connects to the network, but may not be sufficient for all the smartphone platforms that could try to establish a connection. An alternative method, and one that should be in place even if network access controls are in place, is a policy that dictates the configuration and security expectations for using a corporate network. This obviously does not have the same enforceability that technical controls have, but it at least puts users on notice that there are minimum security requirements for using a smartphone within the business network environment.

Keep Antivirus Up to Date

In a similar manner, mobile devices should maintain up-to-date antivirus and other anti-malware applications. Applications running on smartphones can open infected documents and inadvertently download malicious content as easily as they can on laptops, so comparable protections should be in place on both platforms.

Use Encryption on Mobile Devices

It would be unfortunate if the data communicated over an SSL-encrypted channel were leaked once it arrived at the mobile device because of unencrypted storage. SSL encryption only protects data in motion; once it lands on the device and is decrypted during the last steps of the SSL communications protocol, it is up the platform OS and applications to protect the data. Device encryption is one part of the solution to the problem. When selecting an encryption program for a mobile device, be sure to consider the need to encrypt data on permanent and removable media, such as SD cards.

Protecting mobile devices and communications requires multiple layers in accordance with a defense-in-depth strategy. These layers include encrypting data during communication with SSL technologies, authenticating devices with digital certificates, maintaining OS patches, keeping anti-malware up to data, and mitigating the risk of a data leak by using device encryption for both permanent and removable storage devices.

Network Defenses

Network security was at one time practically synonymous with information security. We've moved well beyond those days as we see from the demands for mobile device and communication security. Network security is still an essential element in information security, of course, and no description of business-centric security management, no matter how brief, would be complete without it. The following discussion is not exhaustive but does highlight the controls that can be used to mitigate threats to the network and to devices on the network. These include:

- Deploying and configuring network perimeter devices
- Filtering content on the network
- Monitoring and auditing network activity

Deploying and Configuring Network Perimeter Devices

The purpose of network perimeter devices is to keep malicious attackers, content, and software off the network while preventing valuable data from leaking. This goal is easily stated but the implementation is somewhat more complex. For starters, the types of material that should be blocked range from malicious software to content that is offensive or inappropriate for the business environment. Preventing data leaks is challenging because it requires policies and rules that define the type of content that should not be sent unencrypted outside the network as well as how to identify that type of content. (One of the additional benefits of using SSL encryption is there is significantly less risk of the content that legitimately leaves the network being compromised by data thieves).

Blocking malicious content and unauthorized access requires a number of security controls:

- Firewalls
- Intrusion prevention systems (IPSs)
- Network access controls

These controls complement each other by addressing different types of threats.

Firewalls

Firewalls are still a staple of network security although architectural changes have made the perimeter more porous than it has been in the past. Firewalls have evolved from stateless devices that could block or not block a port or filter out a particular type of network traffic to systems that can inspect deep into the contents of the packet, use information about the state of a session, and apply application-specific rules to identify and block unwanted content.

Network firewalls can still act as gateways between network segments and should be deployed where clear lines of separation are needed. Application firewalls should also be used when there is a need to filter content to critical applications. For example, an application firewall may be used to scan input to a Web application in order to block user input designed to conduct a SQL injection attack. This type of application firewall would provide one line of defense against SQL injection attacks. Developers who write code that cleanses user input, use stored bind variables, and other techniques for avoiding SQL injection vulnerabilities constitute another line of defense. Both are needed when practicing defense in depth. As more programmatic services depend on HTTP to send and receive data, the traditional role of the port-blocking firewall is changing. Blocking most ports but allowing HTTP data still allows a great deal of traffic into the network. Application firewalls and other means of deep packet inspection are required to detect threats tunneling in on HTTP traffic.

IPSs

IPSs should be deployed to monitor the state of the network and hosts. IPSs can use signature patterns, behavioral analysis, or both to detect anomalies on the network, such as:

- Large volumes of traffic from a server that normally has low traffic activity at that time of the day
- Attempts at password cracking
- Known OS vulnerability attacks
- Denial of Service (DoS) attacks
- Web application exploits

Unlike firewalls, IPSs are not about just blocking content by packet type or port but analyzing content and its impact on devices. This functionality is important because not all malicious content can be blocked by gateway devices. Some malicious content is not apparent until it enters the network and begins to interact with devices on the network; that is when an IPS can provide additional measures to detect and block that kind of activity.

Network Access Controls

Network access controls are gatekeepers for allowing and blocking access to network resources. Whereas firewalls operate at the packet level to block content, network access controls determine who and what devices will be allowed to establish a connection to a corporate network. Ideally, a deployed network access control will enforce established policies, such as:

- Who is allowed to access the network based on their identity
- User roles to determine what resources users may access once they have established connections to the network
- Ensure devices connecting to the network meet minimum configuration requirements
- Vary access privileges based on the type of device; for example, allowing only limited access to network resources from unmanaged devices

Network access controls are recommended when remote users regularly connect to the network, especially when unmanaged devices are used to work with corporate assets.

Filtering Content on the Network

Content filtering is a network-based method for scanning content as it enters or leaves the network to prevent unwanted material such as:

- Viruses, worms, Trojans, and other malware
- Spam and phishing emails
- Spyware and adware
- Content that is offensive or inappropriate for a business environment

Most endpoint devices today, such as desktop workstations and laptops, run a full suite of anti-virus, anti-spam, and anti-spyware applications but network protection is also advised. The combination of endpoint-based security measures and network-based measures provide defense in depth against these threats.

Network content filters have an added benefit of keeping employees and others from downloading content from or surfing to inappropriate sites while on the job. For an additional layer of defense, businesses can use third-party Web content-filtering services, such as the free OpenDNS service (http://www.opendns.com/). This service provides domain name services but also allows users to block access to specific types of sites, such as adult, gambling, shopping, and other user selectable categories.

Monitoring and Auditing Network Activity

An unintended consequence of deploying various network security devices is that these devices can generate a great deal of log data. This presents a set of all-too-common problems:

- Each type of device generates log data specific to the device
- The data is distributed across different systems
- There is so much data that it is sometimes difficult to cull out useful information

One way to help improve the management efficiency of network monitoring is to use a log aggregation tool. These can collect data from multiple devices using common protocols, such as Simple Network Management Protocol (SNMP), and perform basic data transformations, such as normalizing timestamps. The advantage of these log aggregation tools is that a network manager can retrieve multiple types of log data from a single application, and basic integration has already been performed. The quality of integration and the ability to detect and highlight important events will likely improve in the future, but these tools can still reduce the burden on network management today.

Network security measures are like common goods, all parts of the infrastructure and business processes benefit from their use. If we start with a business-centric view of network security, we would want many of the standard network security controls, such as firewalls, IPSs, and network access controls. Also, monitoring network activity can become time consuming without tools that can help managers keep up with the volume of log data that these other security measures generate.

The collection of technical controls we have discussed, from measures to protect critical servers and securing communications to protecting mobile devices and network assets, are just one part of a business-centric security strategy. Another part is a focus on end user training on information security.

Security Awareness

The old adage says a chain is only as strong as its weakest link—the same goes for information security. Too often, it is the users, and not technical controls, that fail us. A business-centric security strategy needs to consider security awareness topics and training delivery methods to mitigate threats due to human error and poor judgments.

Security Awareness Topics

The range of security awareness topics that could be covered in training is as broad as the threats, vulnerabilities, and countermeasures that security professionals deal with on a day-to-day basis. We do not need to turn all users into security professionals, and it is sufficient to focus on several fundamental topics that together can help mitigate threats:

- Training on security policies within the organization
- Types of threats to the devices commonly used in business, including mobile devices
- The need to protect data in motion with SSL-encrypted communications
- Threats from spoofing and mistaken identity and how to prevent it with the use of digital certificates
- Threats of data breaches from lost or stolen mobile devices and the need for encrypting stored data
- Phishing and other forms of social engineering attacks
- Malware, infected documents, malicious Web sites, and drive-by downloads

Admittedly, some of these topics can be a bit dry (only some of us care to delve into the details of things like SSL/TLS handshake protocols). How we present security awareness training is as important as what we present.

Effective Security Awareness Training

Effective security awareness training is delivered in a business context, not a technical context. Business users do not need to know the intricacies of asymmetric encryption, but they do need to understand that their business data is threatened if they lose their laptops or someone intercepts their wireless communications while emailing from a coffee shop. Another important aspect of context is the security policies that a business establishes. Those policies are formulated for a reason that must be conveyed to the users. In general, security awareness training should focus on business fundamentals, such as protecting the confidentiality, integrity, and availability of systems. With those as framing principles, the training can then move on to examine high-level threats, including malware, phishing and social engineering attacks, and data breaches. Next, we can focus on solutions, such as SSL encryption, digital certificates, safe browsing practices, and clues to watch for in phishing scams. Not everyone finds information security an engaging topic, and they shouldn't have to in order to understand the impact of security risks to the business.

Checklist of Practices and Technologies

We have covered quite a few topics in this chapter; to recap, the following quick checklist of practices and technologies can be incorporated into your business-centric security strategy:

- Secure communications with SSL—Data in motion does not have the advantage of the access controls in place with data at rest; encryption provides added protection against a number of threats
- Use digital certificates to authenticate devices from servers to mobile devices—Rather than assume we can trust the device to which we are about to send confidential data, verify the device's identity first
- Protect against malicious content with anti-malware and content filtering on the network and on endpoint devices
- Use network security controls such as firewalls, IPSs, and network access controls
- Develop a patch management plan to ensure OSs and critical applications, such as databases, are patched against security vulnerabilities
- Monitor network and host activity—The volume of log data from devices can be substantial; data collection and reporting tools can help
- Train end users by focusing on delivering information from a business-centric, not a technical, perspective
- Think in terms of defense in depth and use multiple security controls to protect against a single threat—Fortunately, many security controls protect against multiple threats as well

To summarize, a business-centric security strategy starts with the requirements of business, assesses the threats and vulnerabilities to the business, and formulates a combination of organizational and technical controls to mitigate risks. Several technologies, such as SSL-based encryption, digital certificates, anti-malware, and network security controls, as well as organizational controls, including polices and end user training, can be used collectively in a defense-in-depth manner to improve the security of the enterprise.