Compliments of THALES

Key Management For Dummies®

Thales e-Security Special Edition

Protect the integrity and privacy of your critical information!

Making Everything Easier!™

FREE eTips at dummies.com®

Richard Moulds

About Thales e-Security

Thales e-Security is a leading global provider of information security providing data encryption solutions to the financial services, manufacturing, government, and technology sectors. With a 40-year unrivalled track record of protecting corporate and government information, Thales solutions protect our customers' most sensitive data and is used by four of the five largest energy and aerospace companies, government networks around the world, and secures more than 80 percent of the worldwide payment transactions. Thales e-Security has offices in Japan, Hong Kong, United Kingdom and the United States and is represented in over 90 countries around the world. For more information, visit `www.thales-esecurity.com`.

Key Management FOR DUMMIES®

THALES E-SECURITY SPECIAL EDITION

by Richard Moulds

WILEY

Wiley Publishing, Inc.

Key Management For Dummies®, Thales e-Security Special Edition

Published by
Wiley Publishing, Inc.
111 River Street
Hoboken, NJ 07030-5774
www.wiley.com

Published by Wiley Publishing, Inc., Indianapolis, Indiana

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

For general information on our other products and services, please contact our Business Development Department in the U.S. at 317-572-3205. For details on how to create a custom *For Dummies* book for your business or organization, contact info@dummies.biz. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

ISBN: 978-1-118-09190-6

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

WILEY

Table of Contents

Introduction

Encryption isn't just for the information security specialists anymore. The growing recognition that the loss of personal information really does have a personal value and a personal impact has brought accountability for its security into the foreground. The result is that almost no organization can avoid the need to establish a data protection strategy — protecting sensitive information within business applications, when in storage, and while moving over networks, particularly the Internet and onward to the cloud.

Encryption has emerged as a best practice and in some cases a mandated method for protecting sensitive data. To facilitate this demand, encryption is now making its way into many of the devices and applications used every day. IT security is entering the age of ubiquitous encryption — cryptography is emerging as the foundation for enterprise data security and compliance.

There's no doubt that encryption can provide powerful security, but without proper management, encryption can quickly become complex and costly. Getting it wrong, either from a technology or management perspective, can at best create a false sense of security and at worst leave critical data scrambled forever — the equivalent of a corporate document shredder.

This book helps you to plan how your company can manage the widespread use of encryption to secure your most valuable data. It examines the key management challenges that emerge as encryption is deployed more widely and provides you with a key management pathway that reduces risk, lowers deployment costs, and preserves your compliance with the growing raft of data protection regulation.

Although this book focuses on the use of cryptography in the context of encryption for confidentiality, it isn't our intention to overlook other forms of cryptographic security, such as the use of digital certificates for authentication and digital signing for integrity. Although these usage scenarios tackle different threat models and address different compliance needs other than privacy, many of the key management challenges remain the same. Readers with an interest in cryptography not directly driven by encryption will still find this book useful.

How This Book Is Organized

Key Management For Dummies, Thales Special Edition, is set up so you don't have to read it cover to cover, front to back. You can skip around and read just the sections that are of interest to you.

Chapter 1 concentrates on the basics of encryption, where it's used, and why the use of encryption is growing. Chapter 2 looks at the importance of keys and the components of the key management life cycle. Chapter 3 examines the pros and cons of different key management approaches, contrasting best in class point solutions and enterprise-wide systems.

Icons Used in This Book

This book uses the following icons to indicate special content:

TIP

These paragraphs point you in the right direction to get things done the fast and easy way.

REMEMBER

You want to pause and take note of these paragraphs.

WARNING!

These paragraphs offer practical advice to help you avoid making mistakes.

TECHNICAL STUFF

Paragraphs marked with the Technical Stuff icon contain information that's relevant and of interest, but you don't need to dwell on it if you're in a hurry.

Chapter 1

Getting Started with Encryption

In This Chapter

- Getting to know the ins and outs of encryption
- Protecting data at the network and storage level
- Understanding the factors that drive change

Because corporate information is incredibly valuable, the pressure to protect it is enormous. You may be tempted to lock it away in a safe place and throw away the key, but the information is valuable precisely because people and business systems require access to it. The trick is protecting data from harm while providing appropriate levels of access to only those that are suitably authorized. In this chapter, we discuss the implications of choosing to deploy encryption for the first time or to expand an existing encryption infrastructure.

Understanding the Where and Why of Encryption

Encryption transforms information into an unreadable format so it remains private from anyone or anything not approved to read it. An individual or application that needs access to encrypted information must possess the correct secret code, called a *key,* to convert the data to its original form. In this way, encryption provides a fail-safe mechanism where if all other security measures fail and data is stolen, or even if data is simply lost, the information contained is still protected.

The prevalence of encryption has risen in step with the following:

- Government regulations
- High-profile data breaches
- Proliferation of malware and advanced persistent threats
- Increasing mobility of sensitive information
- Industry regulations and best practices
- Insider attacks
- Softening consumer confidence

Encryption technology has been around for centuries, but compliance and best practices for data protection are fueling a significant increase in its adoption. Encryption is fast becoming a commodity — embedded as a native feature in applications, databases, devices such as laptops, and storage systems such as disk and tape drives.

Knowing your attacker

Your data (and by *your data* we mean *you!*) faces many threats. Here's a sample of the primary sources:

- **Lost or stolen media:** The loss of backup tapes or disk drives from data centers or on the way to off-site storage locations occurs frequently. Data selected for long-term archival or emergency recovery situations are often commercial records, intellectual property, and sensitive personal information — all attractive targets. Even data that is years out of date to a corporation may still be extremely valuable to an attacker.

  A recent case in the UK involved the loss of unencrypted data during the transfer of information from one government department to another. This loss of two CDs potentially exposed the personal details of more than 25 million benefit claimants.

- **Theft by privileged users:** The access rights granted to privileged users, such as database or system administrators, can lead to super-user accounts and capabilities being exploited by malicious employees or staff that are susceptible to social engineering.

One of the leading financial service companies in the United States, for example, discovered that a senior database administrator was trying to sell the personal information of more than two million customers contained on its database.

- **Identity theft and impersonation:** Even relatively unprivileged internal users can have limited (but legitimate) access to valuable information. Identity theft can be an internal issue and isn't limited to consumers. With the increase in outsourcing and use of contract staff, this risk is even more pronounced.
- **Data sharing and cloud computing:** In an interconnected business environment, data is at risk when shared with partners and service providers outside the traditional corporate perimeter. Validating identity, enforcing access rights, and proving that you are keeping control is fraught with challenges regardless of whether the data is sent by snail mail, e-mail, or a web service.
- **Advanced persistent threats:** A *hacker* (often an insider) can access or modify data within critical business applications by modifying application software or corrupting configuration settings, resulting in an attack that could go undetected for long periods.
- **Theft of mobile devices:** With the use of portable devices growing, including laptops, smart phones, USB memory sticks, and even iPods, so too does the risk of theft or loss. Compromise of such devices risks exposing the sensitive customer data and corporate intellectual property stored on them.

Protecting data in motion

Banks and online retailers value *network encryption,* which protects data moving over communications networks. The SSL standard (the technology behind the padlock symbol in the browser and more properly referred to as TLS) is the default form of protection for Internet communications that provides customers with peace of mind through its familiar icon. These days, many security-conscious companies go one stage further and protect not only their Internet traffic but also their internal networks, corporate backbone networks, and virtual private networks (VPNs) with network level encryption.

As with any low-level security technique however, network-level encryption is a fairly blunt instrument. The network is almost completely blind to the value of the data flowing over it and lacking this context is usually configured to protect either everything or nothing.

Protecting data as it moves over a network is only part of a comprehensive data protection strategy. You must consider risks to information at its origin — before it moves — and at its final destination. Stealing a car in a parking lot or private garage is much easier than on the freeway while traveling at 60 miles per hour!

Resting easy about data at rest

When valuable data collects in one place, hackers may target it for attack. For a hacker, *data at rest* — data in your databases, file systems, and storage infrastructure — is probably much more attractive than the individual data packets crossing the network. Data at rest in these environments tends to have a logical structure, meaningful file names, or other clues that betray that this location is where the "money" is — that is, credit cards, Social Security numbers, intellectual property, financial information, and so on.

Of course, even data "at rest" actually moves around. For a host of operational reasons, data is replicated and manipulated in virtualized storage environments and frequently "rests" on portable media. Backup tapes are transferred to off-site storage facilities and laptops are taken home or on business trips all of which increases risk. Regardless of whether the information has actually been compromised, organizations can take no chances and must act on the potential breach, which often results in significant cost and, in some cases, mandated public disclosure, corporate embarrassment, and customer dissatisfaction.

Storage encryption involves encrypting data while it passes to storage devices, such as individual hard disks, tape drives, or the libraries and arrays that contain them. Using this type of encryption along with database and file encryption goes a long way toward offsetting the risk of losing your data. Like network encryption, storage encryption is a relatively blunt instrument, typically protecting all the data on each tape or disk regardless of the type or sensitivity of the data.

Remembering your keys

How do you ensure that the correct keys are available when and wherever they are needed? Most data at rest encryption projects must implement key escrow (where a copy of the key is held) and key recovery techniques. Armed with a sound key management system, you can not only recover keys but also consciously destroy keys. This provides additional benefit in that destroying a key provides a way to destroy data by rendering it permanently unreadable, which is particularly useful when disk drives are re-deployed, repaired, or sold on eBay.

Although using storage encryption is a good way to ensure your data is safe by default in case it is lost, adopting a more granular approach and encrypting at the level of individual files, volumes, or columns in a database may be necessary, particularly if data is shared with other users or is subject to specific audit requirements.

Securing data in use at the application level

As privacy policies become more specific and tightly enforced, “all-or-nothing” encryption at the storage or network level may not provide sufficient defense against more sophisticated insider attacks. Protecting data within transactional systems and business applications while the data is in use attracts the attention of CIOs and auditors, not because of the deeper security, but because it can reduce the scope and cost of certain compliance obligations. As a result, the processes of building encryption into custom applications and taking advantage of native encryption capabilities of commercial applications are increasingly common.

Protecting data at the application or database level is appealing because you have the benefit of context — knowledge of what individual data elements actually represent, what they are worth to an attacker, and what policies govern their use.

Application-level access controls are already widely used to protect access to sensitive information, such as salary records and financial data, but encryption provides an additional tool for enforcing those access rules. This can be combined with other forms of cryptography, such as digital certificates and signatures, to strongly authenticate users wishing to access or publish documents, software or messages.

Securing remote users and mobile devices

Workers are increasingly working away from the office and at home, boosting efficiency for the organization and flexibility for the worker. But workers may still need access to sensitive information and will “do what needs to be done to get the job done” — copying data to portable storage, e-mailing attachments to home accounts, and accumulating many years of unnecessary data on laptops — just in case they need it.

As a result, the number of locations or devices that fall under the auditor’s scrutiny is growing exponentially. Not surprisingly, the answer for many organizations is to use encryption as a base level of protection for laptop hard drives or memory sticks, and in special cases institute a full-blown digital rights management (DRM) system for more granular control. In extreme cases, organizations turn to more visceral methods by literally sealing USB sockets with glue. All three are usually unpopular moves.

WARNING!

The challenge for the corporate IT department is to ensure the users safeguard this stored information and at the same time, ensure that data is accessible when needed, even when the user and the device are disconnected from the corporate network. Basic encryption functionality is often built in to these devices but all too often the responsibility for managing the encryption keys rests solely with the user and the process of recovering a lost key can be painful and drive up help-desk costs.

Ironically, although cloud computing quite rightly raises numerous security concerns, services such as Salesforce.com may actually help alleviate this particular issue because data is stored in the cloud, not by the user, actually reducing the amount of data in circulation.

"End-to-end" data protection

Recently, there has been interest in the even loftier goal of end-to-end encryption, where data is protected by default wherever it goes over its entire lifecycle. Sensitive data is encrypted the moment it is captured, in a point-of-sale (POS) device at a retail store, for example, and stays encrypted or is re-encrypted while it moves between systems and security domains. This notion of encryption as a data "bodyguard" that always accompanies data objects (files, documents, records, and so on) is appealing but raises questions about establishing trust relationships between regions and interoperability when it comes to key management, but more about that later.

Matching Protection to Threats

Deciding where to encrypt follows the consideration of different threat profiles and the cost of deployment in each scenario. The obvious first step is to encrypt data whenever it passes over public networks; to avoid doing so is reckless. For many organizations, the next step is to consider encrypting private networks or to encrypt backup media, particularly whenever it goes off-site. Encryption at this level has the advantage of being invisible to the applications that run above them, easing deployment and reducing integration costs.

REMEMBER

Storage-level encryption is valuable but only addresses the physical theft or loss of the media and network encryption only guards against eavesdropping. Neither protect against other higher-level threats, particularly malicious insiders.

WARNING!

Addressing the risks associated with user-level mobile storage devices, including laptops, is a logical next step from protecting backup storage. Although both offer transparency to applications, the operational costs can be considerably higher for mobile device encryption. Unlike backup encryption, which is typically administered by relatively few individuals subject to tightly controlled corporate policies, mobile devices are typically in the hands of non-expert and notoriously hard to control end users.

The threat of internal super-users, in the form of privileged administrators or application developers, increases the need to establish more selective, finer-grained controls and enforce strict separation of duties. This process typically requires greater contextual awareness of individual users and their entitlements over specific classes of data. Encrypting data at the application layer gives organizations the ability to apply and enforce finer-grained controls on a per-user or per-request basis. This approach to selective protection focuses on the threats and data that really matter, which usually minimizes the impact to less sensitive systems or data.

Unfortunately, this added protection comes at a cost. Unlike storage-level, network-level, and (to some degree) database-level encryption deployments, which are typically transparent to existing systems, application-level encryption is by definition application specific and can involve the modification of application code and the definition and enforcement of more complex security policies.

WARNING!

The potential for a significant performance hit also exists. Encryption is widely considered, sometimes unfairly, to be a performance drain. Performing encryption on all data streamed to a tape drive may have a minor performance impact but it's easily quantified and predictable. Adding additional, per transaction, encryption steps in a high-capacity data processing application can have a much more significant impact particularly if large numbers of user or data specific keys are in use and key management and key delivery processes haven't or can't be automated.

Although tool kits and management systems have simplified the integration task, retrofitting encryption can still be costly and time consuming. Applications that process sensitive or highly regulated data, such as point-of-sale systems and other transactional networks, can typically justify the increased disruption. However, for less sensitive systems that handle less sensitive data, you must decide whether you want to go this route based on a comprehensive cost-benefit analysis.

REMEMBER

What's right for one organization isn't necessarily right for another. The exact mix depends on the nature of the data itself, the security mandates that apply in your industry, the existing infrastructure that processes that data, the threats, and the IT resources and expertise available.

Recognizing the Drivers for Change

The use of encryption and other forms of cryptography is growing rapidly. At the same time, the degree to which internal and external auditors scrutinize these deployments is also changing. Both factors affect how organizations should think about key management. The sections that follow describe four drivers for change in the key management landscape.

When embarking on a deployment that involves new and potentially disruptive technology, such as encryption, always have a clear view about why you're doing it in the first place, and don't lose sight of this as you debate the numerous trade-offs and different scenarios. Are you seeking to genuinely manage risk in your business or are you just checking *yes* in the compliance box? Be honest!

Strategic versus tactical deployments

Information is the lifeblood of the modern organization. Indeed, for many companies their data assets are far more valuable than their physical assets. Loss of such information has potentially devastating implications. Banking, government, and other security-conscious sectors have led the way with a methodical approach to managing risk, deploying new security technologies in a proactive way and establishing best practices while they go. Nevertheless, even they have often deployed cryptography tactically to address specific regulatory requirements (for example, to protect ATM and POS networks). In most cases, these isolated deployments relied on manual key management processes.

However, examples of a more strategic approach to encryption and key management are emerging. With a more widespread and diverse use of encryption, many organizations are recognizing the downside of a piecemeal approach to encryption and have adopted a more top-down, strategic approach employing automation and centralized control when architecting their encryption infrastructure. These companies quickly conclude that encryption is actually quite easy and instead focus most of their attention on establishing a comprehensive key management infrastructure as a way to manage risk, help ensure business continuity, and gain control over escalating operational costs.

Regulation shines a light on best practices

Growing regulatory pressures — from the Payment Card Industry Data Security Standard (PCI DSS) to the HITECH Act to the European Union Privacy Directive — are forcing businesses to protect the integrity, privacy, and security of data irrespective of the actual or perceived risks. Other forms of legislation, such as the rapidly expanding set of regional and national data breach disclosure requirements, create a further incentive to invest in encryption. While not necessarily requiring the use of encryption, many regulations provide a safe harbor for organizations that lose information that had been encrypted and, therefore, was considered safe.

However, the days of auditors simply accepting the use of encryption at face value are numbered. While security mandates and privacy obligations evolve, there is a tendency to turn the spotlight from the use of basic encryption to the specifics of good or bad key management, and the language is getting far more prescriptive. The PCI DSS standard is a good example of how detailed key management recommendations have expanded in recent versions of the standard itself and in the guidance notes provided to auditors. As a result, best practices are quickly emerging in the area of key management.

REMEMBER

Even if the use of encryption is only grudgingly accepted as a cost of doing business (staying compliant), encryption still needs to be done right or valuable data easily can be lost forever. This concept is particularly important for companies that find themselves subject to the compliance requirements of data protection legislation but are new to the encryption game and have to deal with the associated management issues for the first time.

Keeping control in cloud computing

The potential cost benefits of utilizing cloud-based IT services, is attracting a great deal of interest and raising significant security concerns, too. There are obvious privacy issues associated with sharing sensitive data with external service providers over which your organization has relatively little control and there are similar concerns about account hijacking and data manipulation. The Cloud Security Alliance (`www.cloudsecurityalliance.org`) has boiled down these concerns to what they call the "Seven Deadly Sins."

While organizations consider the use of these cloud services, they must also recognize that the accountability for data protection still rests with the organization and is not the responsibility of the service provider.

Fortunately, organizations can take a number of steps to mask, encrypt, or otherwise protect sensitive data before passing it to the cloud. Leveraging encryption services in the cloud itself may also be possible, as part of a security aware platform as a service (PaaS) offering, for example. In all cases, keeping the keys associated with these processes in the control of the organization is vital, and the governance of these keys must be demonstrated to auditors.

As already noted, the use of cloud computing can provide significant security benefits by reducing the need for end users to retain personal copies of information because they can access the data online whenever needed. Reducing the volume of sensitive data stored on laptops and other mobile devices is definitely a good thing.

The myth of maturity

Organizations have plenty of reason to think that encryption has suddenly become easy. More and more applications and devices are delivered with built-in encryption functionality as a standard feature or an optional upgrade. You see it everywhere from file encryption on a laptop to data encryption in a tape drive, column encryption in a database, and SSL/TLS support in virtually everything. PKI-enabled operations underpin the security of a host of commercial software applications, and numerous standards groups and regulatory bodies are publishing recommendations about the use of cryptography and keys. Thinking encryption has become a risk-free commodity is easy, but don't be fooled.

The adage that security is part technology, but mostly process is still true. The basic number-crunching technology of encryption may indeed be a commodity, but poorly deployed encryption can easily lead to data-protection compromises and a false sense of security unless encryption keys are managed in line with corporate policy and best practices. If end users, business managers, or operations managers indiscriminately turn on encryption in an attempt to "do the right thing," then operational costs may spiral out of control when keys are lost and calls to the help desk skyrocket. Worse, if something goes wrong, important information may be lost forever. Literally, the key to successful encryption is how well you manage it, and few organizations have a good track record.

Chapter 2

Unlocking the Secrets of Key Management

In This Chapter

- Understanding the role that keys play
- Getting to know the key management life cycle
- Appreciating the differences between encryption and key management

When you use encryption to protect your data, you use cryptographic keys. You can't use one without the other. *Keys* are equivalent to the series of numbers that opens the combination lock of a safe. If a thief knows a safe's combination, even the strongest safe in the world provides no real security. Similarly, poor key management can easily compromise the strongest cryptographic algorithms. In this chapter, we explain how keys work. We tell you everything you need to know to get started on the right foot as you plan your approach to key management.

It Really Is All About the Keys

All encryption involves three aspects: the data to be protected, the algorithm or cipher that transforms the data so that it's unreadable, and the key, usually a random number. In most cases, the algorithm is not a secret, but the key certainly is. Just as in the physical world, it's the individual key in your pocket that makes your front door lock different from your neighbor's even if both locks operate in exactly the same way. Because the key is what makes the encryption process unique, the key becomes the secret to making the process reversible.

Anyone wanting to use encryption has a choice of cryptographic algorithms (DES, AES, and RSA are examples), all of which involve highly complex math that, thankfully, most of us don't have to worry about. Different algorithms suit different uses and have different security, performance, and power consumption properties. Limit your choice to algorithms that have passed the test of time and scrutiny of academics and other independent experts. Proving that an algorithm is truly secure is a difficult task, which is why they don't change very often and why there's a relatively small set of widely used algorithms.

No algorithm is unbreakable, especially given the huge leaps in computing power. The issue is how long it takes and what expenditure is required to crack the code. Organizations, such as the National Institute of Standards and Technology in the U.S., periodically make recommendations about the selection of algorithms. In late 2010, NIST published their latest guidance, NIST SP 800-131A. Given that all recommended algorithms are practically unbreakable using today's technology, it's reasonable to assume that the attacker will turn his or her attention elsewhere — to the key that locks (encrypts) or more importantly, unlocks (decrypts) the data you are trying to protect.

WARNING!

With the right keys, any attacker can access the underlying data. Mismanagement and human error can result in keys falling into the wrong hands. Conversely, you need to ensure that keys are always available to an authorized user when required. For data retrieval or full-blown disaster recovery, you may find yourself relying completely on rapidly accessing keys in order to unlock previously encrypted information.

The Key Management Life Cycle

The task of *key management* is the complete set of operations necessary to create, maintain, protect, and control the use of cryptographic keys. The important thing to note is that keys have a *life cycle;* they're "born," live useful lives, and are retired. Of course, it's never quite that simple. Figure 2-1 shows the typical life cycle of a key, and the following sections cover each part of the key management life cycle in detail.

REMEMBER

Actual key lifecycles vary among different situations, the ordering of phases altered slightly, and certain phases skipped completely. This example is not prescriptive but rather a planning guide.

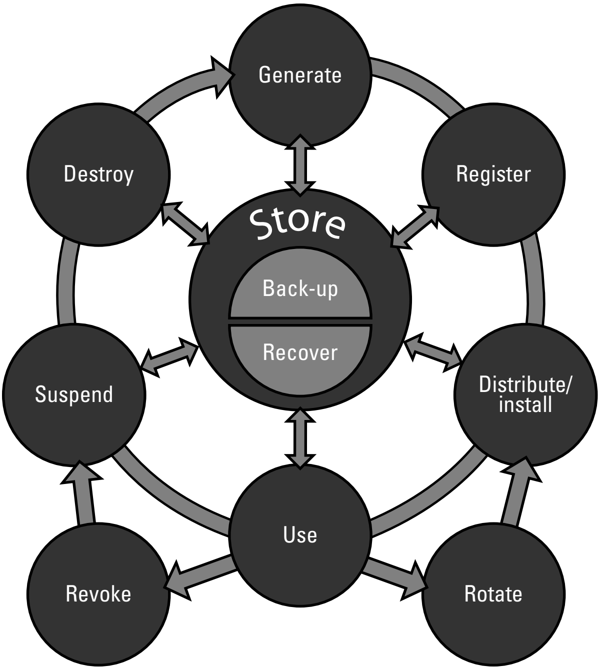

Figure 2-1: The key management life cycle.

Key generation

A new key life cycle always starts with key generation. The main challenge is to ensure that key generation produces unpredictable keys. If the key generation process isn't truly random, a new key can be predicted from knowledge of previous keys or the key generation process itself. The ability to make predictions dramatically reduces the number of combinations to test and makes cracking a key more likely.

Key registration

A key on its own is just a useless number. Before a key becomes useful, it must be registered or associated with a particular user, system, application, or policy. Typically, the strength and nature of this association determines the value of the key. For example, a key tied to an identity and used for the purposes of authentication or signing documents is often associated with a definition of that identity in the form of a digital certificate. The mathematical link between the secret (*private*) key and the public key contained with the digital certificate combined with the trustworthiness of the authority that issued the certificate defines the strength of that association.

Key storage

Whether keys are in the pre-operational or operational state, they need to be stored securely and, ideally, nowhere near the data that they protect. To do so would be like slapping a sticky note with your Amazon.com password on your monitor and labelling it “My Amazon.com Password”. You just don’t do that.

Pre-operational keys can be stored in a physical safe, electronically on a memory device, on optical media, or even a paper printout — all safe and all offline. Operational keys on the other hand need to be accessible in real time, online, and are often stored within live systems. To store these keys securely, they’re often encrypted by, you guessed it, more keys. This raises the obvious question of how well protected are the key encryption keys (KEK) or *wrapping* keys. This cycle can go on for many levels of protection creating a hierarchy of keys. No matter how many layers of encryption are applied to protect operational keys, the KEK must be stored somewhere safe and, ideally, that means in dedicated hardware.

REMEMBER

You may be tempted to think it’s safe to hide a key, buried in the massive amount of data stored on a disk drive, memory chip, or backup tape, but you’d be wrong. Unlike the rest of the data being stored on these devices, a key has one special property: It’s a perfectly random number. This very randomness can make it stand out like a beacon. If you use the right tool, you can easily locate a key on a hard drive or in the contents of memory chip. You can attempt to obscure the key or break it into fragments, but you’re really only delaying the inevitable. To be truly secure, you should store the key in a way that is inaccessible to malicious software and other scanning tools.

To resolve this issue, you may wish to establish a safe cryptographic zone, independent of the host computer or server and outside the scope of the operating system. For ways to approach this issue, see Chapter 3.

Key distribution and installation

Any key management system must address the secure transmission of keys from their place of storage or generation to the application or device requiring them. A secure link or process for sending the key is essential. The challenge is validating the identity of the person or application requesting the key and approving the request. There's little point in safely storing a key if you are going to give it out to anyone that asks for it. Similarly, the risk of accepting keys from untrusted sources is equally apparent.

The *enrollment* problem — the task of establishing trust between entities that share keys — can become one of the costliest aspects of key management, particularly in large-scale systems. Situations deserving serious consideration include those where entities are dispersed geographically or where multi-factor authentication and/or frequent renewal of authentication credentials is required.

Key use

When keys are in an operational state, in principle they can be used for many things. Good security practice is to restrict the use of a single key to only one purpose, such as encryption, authentication, key wrapping, or creating digital signatures. The use of the same key for two different cryptographic processes may weaken the security provided by one or both of the processes. Key usage policies must consider whether certain uses of a key conflict with one another. For the same reason that keys should be *stored* in an isolated or dedicated environment, it logically follows that they should be *used* in a safe environment; otherwise, keys are exposed every time they're used.

Generally, access to a key should be on a "need to know" basis or perhaps more appropriately, a "need to use" basis, in order to avoid the unnecessary proliferation of keys.

Key rotation

If the same key is used over a long period, the risk of that key being compromised increases. Correspondingly, the longer a key is used, the amount of data it's protecting increases and therefore, so does its value to an attacker. The risks multiply as time goes by and the prize gets richer — the available window to crack any one key increases along with the potential for leakage of critical information that helps the attackers with insider knowledge.

Companies must minimize these risks by periodically refreshing, or rotating, keys as often as their risk-management strategy and operational constraints allow them. It's a classic trade-off of risk mitigation versus cost.

Key backup

As with any backup system, backing up keys protects both the user and the company in the event that a key is lost. To safeguard against the loss or unintentional destruction of a key, certain keys must be replicated and stored offline or on an offsite backup. Mechanisms for regular synchronization of any backups with the primary key store are important. Rigorous controls must be in place to prevent misuse of the backup, but having a backup is vital.

In some cases, key backup may not be appropriate. For example, if keys are used to create digital signatures intended to carry legal weight, it's important that only the authorized signer have access to the signing key. The mere existence of a copy of that key casts doubt on who really signed the document. Similarly, in the context of encryption, if the destruction of a key is intended to serve as proof of the destruction of the original data, proving that all backups of the key are destroyed becomes necessary; otherwise, asserting that the data really is unrecoverable becomes impossible.

Key recovery

Whenever keys are backed up or archived there is the challenge of defining the key recovery policy; who can request a recovery of the keying material and under what conditions the recovery process should be carried out. For example, recovery may be required to happen in a physically secure, offline facility under specific supervision procedures. Often, key recovery, like many recovery scenarios, is performed with some urgency because systems have gone down following an emergency or a forensic request requires timely delivery of information. Recovery approval procedures and escalation policies need to be predefined and followed at a moment's notice. Strong auditing functions need to be in place to ensure the keying material is recovered only by pre-approved entities and that the appropriate procedures have been followed.

Key revocation

Any compromised key — or even one suspected of compromise — must be revoked and replaced. The potential damage resulting from a compromise can be widespread, and restoring security can be both complex and expensive, particularly where keys are shared, where they affect many users, or where they're tied to physical tokens that may now need to be replaced. The challenge is communicating that a key has been revoked and providing an escalation path when a key becomes unavailable, potentially bringing business operations to a halt.

Key suspension

When a key reaches the end of its defined operational life, it should be removed from service, potentially leading to the destruction phase. However, in many cases, it may be necessary to suspend the key from use but not destroy it. For example, following a key rotation, an expired key that no longer performs encryption may be required to decrypt data that it previously encrypted unless that data has been entirely rekeyed.

Key destruction

A key should be destroyed when there's no need to preserve the use of that key. In all cases, proving that the key has been deleted, including backups, is important. This means you need to have had control of the keys since their creation and sufficient audit trails to know explicitly when and where copies and backups were made — impossible to determine in retrospect.

Knowing when to throw away the key

Many government regulations and industry standards mandate keeping certain types of data for specific lengths of time (seven years or more isn't uncommon). Storage beyond the mandated period becomes a liability and encryption provides a convenient way of destroying the data simply by destroying the key.

Policy enforcement

Defining and enforcing policies affects every stage of the key management life cycle. Each key or group of keys needs to be governed by an individual usage policy defining which device, group of devices, or types of application can request it, and what operations that device or application can perform — for example, encrypt, decrypt, or sign. In addition, policies may dictate additional requirements for higher levels of authorization to release a key after it has been requested or to recover the key in case of loss.

TIP

Adopt the same level of security to the enforcement of key management policies as you would to protect the keys themselves. In the end, if attackers can subvert the process that enforces the key management policy more easily than attack the keys directly, they will likely take the path of least resistance. As with any security system, attackers look for a weak link, and all phases of the key life cycle provide opportunities to expose a weakness. Enforcing a key management policy is as much about establishing consistency as it is about focusing on the threats at any one stage in the life cycle.

Chapter 3

Strategies for Successful Key Management

In This Chapter

- Choosing between localized and centralized key management
- Mapping out requirements for key management
- Achieving a real return on your investment

The topic of key management is new to many organizations and often treated as a specialist activity confined to the administration of a few isolated applications. That is changing as the use of cryptography, particularly encryption, emerges as a mainstream IT security tool.

This chapter discusses the various trade-offs and considerations that enable an organization to chose the right approach to key management to maximize the likelihood of a successful deployment and shorten the time to achieve a genuine return on investment.

Different Approaches to Key Management

You have choices in deciding how to manage keys. A host of variables drives this choice but three typical scenarios emerge:

- **Native key management tools:** At the simplest level, it is possible to utilize the basic key management capabilities that are native to the individual encryption product or products being deployed.

- **Localized key management platforms:** In order to better manage risk and to ensure control of the entire key life cycle, secure localized key management devices can be used to augment the deployment of individual cryptographic applications.
- **Centralized key management:** In larger scale deployments where automation is a primary concern or where unification of key management policies is desirable, it's possible to adopt a more strategic, top-down, approach to key management. This approach draws a clear distinction between the key management tasks that are performed centrally and the distributed locations and applications where those keys are actually used.

Looking at native key management tools

Native key management tools utilize the basic key management capabilities that are inherent in the individual encryption product or products being deployed.

Most organizations adopt cryptography incrementally to solve specific security requirements — addressing particular points of risk or external mandates. Common examples include

- The use of SSL/TLS on externally facing Web servers
- VPNs for remote access
- Data/file encryption on laptops or tape/disk drives
- Database encryption
- Specialist situations, such as POS or ATM networks

Many commercial products used in these situations require keys to be managed through at least some of the phases of the key life cycle, and they include software functionality to handle some or all of those tasks. The emphasis, policies, and general security properties of these key management utilities will differ enormously among products, and it's fair to say that some are not much more than an afterthought. A good analogy is the numerous remote controls in a home theatre set-up. They all perform similar functions that skew toward the specific device they're designed to control. The result is a high level of inconsistency, which, in a home theatre setup with multiple controllers, is nothing more than frustrating, but in a corporate IT setting drives up cost and complexity and, because we're talking about security, introduces risk.

Enhanced security through localized key management platforms

Concerns over potential attacks or specific industry mandates drive organizations to consider independent, or best-in-class key management tools. These tools typically manage and protect keys in dedicated hardware on behalf of software applications instead of using the native key management tools found within those applications. Typically, the functionality of these devices goes beyond just managing the keys and includes performing cryptographic operations using those keys.

TIP

Ideally, all management aspects associated with the key life cycle should be performed on a secure platform that provides some level of physical and logical protection for the keys and the processes that utilize and manage those keys. The ability to create an isolated or "trusted" zone for cryptographic functions is best achieved with dedicated cryptographic hardware.

Unlike native key management tools that are delivered as part of the software application itself, the use of a separate hardware-based key management platform overcomes the inherent weaknesses of software-based cryptography. Using an independent security platform for key management also creates a powerful separation between the tasks of managing the keys and managing the applications that use those keys. The notion of a "separation of duties" to mitigate the threat of a single super-user is a common best practice in defining key management policies and is strongly recommended in security standards such PCI DSS.

When deciding to utilize hardware key management devices, you have different options available:

- **Hardware Security Module (HSM):** The use of HSMs is a well-established approach for data center security and for protecting server and mainframe hosted applications.

Mandated in government and certain financial/payment markets for decades, HSMs provide the following capabilities:

- Protect cryptographic keys and perform various cryptographic functions in a secure tamper-resistant hardware environment.
- Overcome the threat of software-based attacks and provide robust tools to enforce key management policies across the life cycle.
- Provide a simple mechanism for introducing strong authentication for key management administrators.
- Establish and enforce powerful dual controls and separation of duty schemes where multiple administrators can mutually supervise key management activities.
- Incorporate high-speed cryptographic processors to improve performance, remove processing bottlenecks, and increase system capacity.

HSMs are marketed in a variety of form factors to suit different deployment needs. The most common examples are

- **Network attached appliances:** These devices essentially provide a network-based service that a number of applications can access in order to have cryptographic operations performed on their behalf (for example, a shared signing or encryption service). There are a number of advantages in taking this approach: keys can be managed in one location rather than on disparate application servers, yielding many of the benefits of centralized key management systems; HSM capacity can be shared across multiple application instances, improving efficiency and resilience; and that this service-oriented or abstracted model for deploying HSMs fits very well with the trend towards application virtualization.
- **Embedded plugin cards:** Typically, these devices follow the PCIe interface and form factor definition, plugging directly into a peripherals slot on the server motherboard. They are ideally suited to situations where HSM resources need to be dedicated to a single server platform rather than shared over a network. Typically, usage is driven by a need to maximize performance or to simplify deployment,

where the HSM can be delivered to a remote location embedded within a pre-configured system.

- **Portable security devices:** These devices are physically smaller and connect via a USB interface to the application host, which may be a server, desktop, or laptop computer. They are ideally suited to offline situations where the HSM is stored in a physical safe when not in use or where the HSM is required to be portable, to be used by an individual performing occasional tasks, such as code signing, application development, or remote system management and authorization.

- **Trusted Platform Module (TPM):** Whereas an HSM is a dedicated security module aimed at server-based applications, a TPM is typically a dedicated security chip that fits as a motherboard component inside a laptop or desktop computer creating a "root of trust" similar to the way that a SIM card works in a mobile phone. The TPM chip provides security services, such as key generation, encryption, decryption, and secure key storage that can be exploited by applications running on the host device. TPMs have the potential to act as a standardized way to securely protect keys on desktop and laptop machines in the event they are stolen or otherwise compromised. Although most commercial grade machines include this functionality, the actual adoption by end users and the organizations they work for has unfortunately been very limited.

- **Smart card:** A smart card or chip card is a credit card–sized electronic device with a built-in microprocessor and memory that can act as a dedicated system for storing keys and performing low-speed cryptographic processes. Smart cards can be accessed by applications when inserted into a card reader, typically a USB connected peripheral. The small size of a smart card is ideally suited for situations where keys need to be protected within a portable "wallet" (for example, keys used for user identification and encryption keys for data stored on PCs).

In addition to providing tangible security improvements when compared to purely software-based systems, hardware key management devices, and in particular HSMs, provide the opportunity to define common key management practices that can be applied across diverse applications — unifying policies as well as enforcing them.

Wrapping your brain around FIPS 140

The value proposition of any hardware-based solution, which in the end justifies the increased cost, is the tangible increase in security compared to a pure software solution. This raises the obvious question of why should anyone believe the security claims made by the vendors that build these products?

Fortunately, there's FIPS (Federal Information Processing Standard) 140-2. It's the benchmark for validating the effectiveness of cryptographic hardware. If a product has a FIPS 140-2 certificate you know that it has been tested and formally validated by the US and Canadian Governments. Although this is a US/Canadian Federal standard, it's been widely adopted around the world in both governmental and non-governmental sectors as a practical security benchmark and realistic best practice.

Organizations use the FIPS 140-2 standard to ensure that the hardware they select meets specific security requirements. The FIPS standard defines four increasing, qualitative levels of security:

- Level 1: Requires production-grade equipment and externally tested algorithms
- Level 2: Adds requirements for physical tamper-evidence and role-based authentication. Software implementations must run on an Operating System approved to Common Criteria at EAL2.
- Level 3: Adds requirements for physical tamper-resistance and identity-based authentication. There must also be physical or logical separation between the interfaces by which "critical security parameters" enter and leave the module. Private keys can only enter or leave in encrypted form.
- Level 4: This level makes the physical security requirements more stringent, requiring the ability to be tamper-active, erasing the contents of the device if it detects various forms of environmental attack.

The FIPS 140-2 standard technically allows for software-only implementations at level 3 or 4, but applies such stringent requirements that none have been validated.

For many organizations, requiring certification at FIPS 140 level 3 is a good compromise between effective security, operational convenience, and choice in the marketplace. For more information, visit `http://csrc.nist.gov/groups/STM/index.html`.

The unification of key management policies has obvious operational benefits, such as facilitating common training schemes and greater flexibility with regard to staff allocation. With the prevalence of more stringent privacy mandates, the value of unification is a motivator for the use of independent key management systems.

Stepping up to a centralized model

As organizations deploy ever-increasing numbers of encryption solutions, they find themselves managing inconsistent policies, different levels of protection, and experience escalating costs. When the pain gets sufficiently high, the best way through the maze is often to transition into a centralized key management model. In this case, and in contrast to the use of HSMs described earlier, the key management system performs only key management tasks, acting on behalf of other systems that perform cryptographic operations using those keys. With this in mind, a hybrid or hierarchical approach where HSMs perform localized cryptographic processing but are managed centrally by a common key manager is conceivable.

The approaches to centralized key management

In general, two subtly different approaches exist to centralized key management — the difference hinging on where authority really lies. In many cases, this comes down to where the key life cycle was initiated:

- **The end point driven approach:** The simplest approach is when individual applications or end points retain responsibility for many aspects of the key life cycle. Keys are generated locally by the application or end point that intends to use them. Contextual knowledge about what the key is used for and what policies apply to the key tends to remain at the end point. In this case, the key management system acts as a trusted repository or vault for keys, storing keys (and metadata) on behalf of the end points and releasing the key to the end point on request.
- **The manager driven approach:** In this more sophisticated case, the key management system assumes responsibility for the entire life cycle and literally becomes the "key authority." Keys and their associated policies are centrally generated and stored. Keys are distributed to suitably

authenticated and authorized applications or end points on request where keys are used but not retained. The management system is responsible for key recovery, revocation, and destruction.

The benefits of centralized key management

A number of significant benefits exist in using a centralized key management system. These include

- **Unified key management and encryption policies:** Consistent policies can be established and uniformly updated without the need to train local administrators or worse still fly people around the world.
- **System-wide key revocation:** A central key management system is well placed to affect system-wide changes. For example, if a key has been compromised — even suspected of compromise — the key must be revoked. A centralized key management system, in theory, can enable system-wide revocation of all the instances where a given key has been in use. It's never quite that simple but it does mitigate the risk and cost of making changes region by region or even device by device.
- **Single point to protect:** In a centralized approach, you only have to protect one key repository. Although keys are still used at the end points, they're only retained on a temporary basis, alleviating concerns over large numbers of historic keys accumulating at end points and potentially subject to misuse by local administrators.
- **Automation to lower costs:** As scale increases, the case for automation becomes overwhelming. Centralized key management reduces the number of expert personnel needed to run disparate systems, freeing up resources to perform other tasks.
- **Consolidated audit information:** By providing a system-wide auditing capability, a centralized approach to key management avoids the need to extract and consolidate audit logs across different end points, application silos, and key management utilities.
- **Single point of recovery:** Key recovery tends to occur hurriedly — either the organization is in general disaster recovery mode or an urgent need for forensic analysis was requested (such as a request from law enforcement to review encrypted information). In most cases, it's

easier to access and recover keys from a single repository than from geographically distributed end points.

- **Convenient separation of duty:** Compliance and security policies increasingly require that a clear separation of duties exists in critical applications — the need to ensure that no single administrator or privileged user can subvert the system. The process of abstracting key management from the application or end point that uses the keys creates a natural separation and, by centralizing the key management function, creates a "keeper of the keys." Sensitive information can be read only by users or applications granted access to the encrypted data by the application administrator *and* given access to the keys to unlock that data by the key manager.
- **Mobility of keys:** As more sensitive information is encrypted, keys will increasingly be moved around the organization in order to ensure that keys are available to systems or organizations that receive the information and need to decrypt it. A centralized key management system can be used to tackle this challenge by acting as a broker of keys, linking different systems and organizational structures.

The challenges of centralized key management

Although the benefits of centralized key management described in the preceding section sound compelling in high-scale deployments, centralized key management systems do present their own challenges. The most obvious comes down to security.

Unlike most other security management systems, such as those that manage software patches or network alarms, a key management system manages secrets. Therefore, any key management system must be secure and ideally tamper-resistant.

Security risks increase because of the sheer number of keys in the system. Furthermore, many of these keys are managed on behalf of applications or end points that are physically secured with hardware such as an HSM. This places considerable emphasis on the security properties of the central key vault and the enforcement of policies and administrative practices within the key management system as a whole.

In the same way that a centralized key management system can be viewed as a single point to attack, it could equally be considered a single point of failure. With a host of remote encryption systems and applications relying on the central manager for keys, essentially on demand, the concern that keys aren't available for a even a short period of time is very real. Resiliency across the key management system is vital. The ability for end points to have multiple paths to request keys provides a fail-over mechanism in the event that a component of the key management infrastructure fails. In some cases the situation may be further complicated by the need for end points to go offline and yet continue to operate, requiring some form of local key caching.

Envisioning a Key Management Solution

The topic of key management is relatively immature once you step outside the industry sectors that have been using it for years. By considering the benefits and challenges of adopting one or both of the two main approaches — HSM-based and centralized key management — and recognizing that encryption technologies, standards, best practices, and external mandates are all evolving, an organization can find it difficult to predict its needs with any degree of certainty.

TIP

Therefore, as you consider how to approach the key management challenge, bear in mind the current best practices. Most importantly, build for the future — flexibility, scalability, and security are critical selection criteria. You can be certain of three things:

- The amount of encryption and cryptography inside your organization will increase.
- In the future, you will deploy and manage types of encryption and cryptography that today you haven't heard of and in places you don't expect.
- The security and audit requirements on the key management process or system will become more stringent over time.

The following sections outline the primary requirements of any key management solution (localized or centralized) in the context of these three themes. The intent here is to provide a generic checklist of capabilities through which to filter the various alternatives that exist.

Scalability is key

The issue of scalability covers a number of aspects arising from the sheer number of managed keys or number of end points or application instances in play. The following criteria are requirements to assess:

- **Limitations on number of keys:** The architect must consider how many keys can be managed by the system. Ideally, the key management system should have an architecture that enables deployments to grow significantly in terms of the number of keys in use and in archive. For centralized systems, look for scale that supports many millions of keys.
- **Limitations on the number of end points:** In a centralized system, the number of managed end points is perhaps the greatest determinant of system architecture. This issue is less applicable in localized key management systems using HSMs because these tend to be tied to individual instances of an application (effectively the end point) or shared by a well-bounded group of machines or virtual machines.
- **System latency and performance:** Encryption-based systems already have a reputation for creating performance bottlenecks (sometimes unfairly), and the key management system shouldn't worsen the situation. The issue of latency has different dimensions when considering localized HSMs and centralized systems.
  - **Localized key management:** When using an HSM, not only are the keys managed but also the actual encryption (or signing function) is performed by the HSM. In this case, performance tends to be measured in terms of signatures per second, megabytes per second for symmetric encryption, and transactions per second (tps) for asymmetric encryption.

  The need for high performance is one of the primary motivations for using localized HSMs and performance ratings above 5,000 tps are common.

  - **Centralized key management:** Since these systems don't actually perform encryption, the issue is how quickly individual requests for keys can be serviced. A performance requirement of more than ten keys per second isn't unreasonable.

- **High availability:** Key management usually requires a high level of availability. The inability to access keys essentially results in the inability to access information. At a minimum, with localized HSMs, there's the need to replicate key stores to establish fail-over, and this process should be as easy as possible; whereas centralized systems require that the key delivery infrastructure, in addition to the key store, is resilient with individual end points having multiple paths to request and access keys.
- **Disaster recovery:** Looking beyond operational resiliency, the need to be able to recover the key management system itself in the event of complete failure is critical. Beware of approaches that rely only on a cold standby or surrogate device where there is as much chance of that device failing as the primary key manager.

  Insist that the architecture of the key management system allows for rapid recovery, subject to the appropriate authorization; otherwise, it may present a roadblock to the recovery of other systems.
- **Ease of use:** As any system scales, ease of use becomes an increasingly important factor, particularly in the case of a centralized approach where scale is at its highest. The ability to group keys, group end points, and assign roles and policies to those groups are the only ways to manage what may amount to millions of keys. Similarly, the ease with which new end points can be enrolled into the system will be an important requirement. In some cases, the ability to import existing keys from legacy systems and export keys in bulk to other systems may also be important. In the case of localized HSMs, the ability to share keys between HSMs and to establish backups of application keys is critical. HSMs that employ expensive physical tokens for sharing keys and creating backups of application keys should be avoided.

- **Automation:** An important aspect of any high-scale system is the ability to automate operational tasks. In a centralized system, the ability to automate the delivery of keys to end points, on request, is a basic tenet of the system.

  Even with localized HSMs, it is possible to exploit automated existing backup processes to create backups of application keys after they have been protected by the HSM.

Flexibility in an uncertain world

The broad requirement for flexibility comes from the fact that key management is a relatively immature topic in most if not all organizations. The requirement for flexibility spans three important areas:

- **The ability to support new applications and different end-point technologies:** Organizations tend to employ encryption and other forms of cryptography incrementally (for example, beginning with storage level encryption and adding support for database or mobile device encryption over time). Whether keys are managed locally in HSMs or centrally, the introduction of new cryptographic applications could expand the scope of the key management system, for example to add support for symmetric or asymmetric keys, Elliptic Curve Cryptography (ECC), or to manage digital certificates in addition to managing keys. Additionally, there is the need to adopt new technical standards as they emerge, the most prominent being the OASIS Key Management Interoperability Protocol (KMIP) standard, initially released in 2010.
- **Supporting new use cases and organizational policies:** Looking beyond the ability to simply support new classes of end points or applications, different uses of cryptography (for example signing) represent different policy scenarios. As the key management system evolves it must accommodate and enforce the security policies relating to these new scenarios and use cases. The recovery policies, escalation procedures, and audit obligations could vary significantly and test the system's flexibility.

- **Integration with back-end systems:** As the use of cryptography becomes more prevalent, the key management system is likely to become a useful integration point to other back-end management systems. The simplest example of this is a public key infrastructure (PKI) and specifically the Certificate Authority (CA). CAs are complementary to key management systems because they issue certificates that attest to the integrity and use of the private keys under management. Direct and automated integration with the CA or CAs of choice can significantly reduce the cost of renewing certificates. Further, integration with third-party systems, such as identity management, digital rights management, and storage management systems as well as other examples of custom-built business logic will become important.

Raising the security bar

The key management devices and systems are among the most security critical components of a typical organization's IT infrastructure and can easily represent a single point of risk. The following sections attempt to cover the core capabilities that can collectively deliver adequate security.

Secure key storage

Quite clearly, secure key storage is a primary requirement of all but the most basic key management systems. The adoption of best practices requires the use of dedicated hardware devices to protect keys. HSMs are the preferred method of performing localized key management tasks. However, HSMs are also used to underpin the security of centralized key management systems, avoiding the risks associated with purely software-based key management. In some cases, the HSM is used not only to protect the keys within the key management system but also to protect the core operating procedures and auditing functions within the system.

REMEMBER

Dedicated hardware needs to be part of any robust key management system. A centralized system needs to be at least as secure as any end point that it manages. Consider formal security certifications when writing key policy statements. FIPS 140-2 level 3 and CC EAL4+ are good starting points.

Strong authentication of administrators

There's little point in protecting the keys if system administrators are weakly authenticated with low-grade credentials, such as passwords. The use of stronger authentication schemes, typically using two factors (usually something you have and something you know) such as smart cards or other hardware tokens is often a mandatory requirement for key management administrators.

Split knowledge and dual control

In many key management systems, going beyond strong authentication and instituting mechanisms to separate the roles of administrators and to enforce concepts such as dual control or split responsibility is necessary. The goal is to make sure that no single person has access to any information, device, or function that allows him or her to subvert system security. In practice, this means that at least two people have individual credentials that must be brought together to "reconstruct" the authority to administer the system.

REMEMBER

A robust key management system allows multiple smart cards or other hardware tokens to be used by multiple administrators to control use of a given key or set of keys. The application of such controls applies to local as well as centralized key management systems.

Mutual authentication of end points

In any scenario where keys are managed centrally on behalf of remote end points or applications, you should authenticate the identity of the end points requesting cryptographic services or keys. Again, there's little point in actively managing keys according to the corporate security policy if keys are then distributed without establishing the credentials of the requesting application or end point device. Without mutual authentication, an attacker can perform a man-in-the-middle attack.

Strongly authenticating the remote end points to the key management system can be achieved with remote hardware, such as an HSM in the case of remote servers, TPM chips within desktop or mobile devices, or smart cards for remote user authentication.

TIP

In centralized systems it may be possible to utilize HSMs at the remote end points to gain the additional benefit of localized cryptographic acceleration along with strong authentication of the end point. This is a good example of a hybrid approach encompassing both centralized and hardware secured local key management.

System integrity and dependencies

You need to understand what assumptions the key management system makes about the security of the environment in which it operates. Ideally, the key management system will not rely on the security of other systems to underpin its own security profile. System security is only as strong as its weakest link; systems not designed with strong security in mind can lead to significant system risk.

For example, the key management system shouldn't assume any level of inherent confidentiality exists in IP networks over which it communicates with applications or end points. Similarly, in the event that the key manager uses external file storage systems or databases for key and policy storage, it should place no requirements on these systems to provide any level of security. By not relying on the security of these other systems, it is the responsibility of the key management device or the system itself to perform its own protection and integrity checks prior to sending to and after receiving data from these external storage services.

Secure auditing

Any key management system needs to provide high-integrity audit logs and reporting tools. Internal system logging activities should span the entire life cycle of each key, where it was created, where it was used, and what administrative tasks were carried out on it. Such audit logs need to be accurate, comprehensive, and trustworthy. To achieve this goal, audit logs should be cryptographically secured, digitally signed, and time stamped to ensure integrity and to support long-term validation.

Tasting Success in Key Management

Encryption has moved from a niche security technology to a mainstream method for protecting sensitive data at multiple points across an organization. It represents best practice in information security and goes a long way toward meeting compliance requirements of data protection regulations.

Encryption functionality now ships as a standard feature in many applications and devices, making it practical for business users to adopt encryption and protect data. However, organizations face a new set of challenges as encryption deployments spread from a few niche applications to a broader range of distributed devices and fragmented business services.

Encryption is an incredibly powerful technology that protects information, but without proper management of the keys, the technology can raise new risks and not deliver the level of security that's required. This makes a strategic approach to key management a mission-critical priority.

In many cases manual key management processes can no longer handle the complexity and scale associated with the requirements of today's enterprise. Data needs to be available at all times, secured wherever it resides, and protected whenever it moves across the network.

The native key management capabilities packaged with encryption-enabled applications are sometimes inadequate and don't provide an overarching framework to unify and ensure the enforcement of security policies. This drives up costs and complexity particularly as recordkeeping for regulatory compliance becomes more burdensome.

The traditional silo-based approach to key management is not only costly but also can introduce security loopholes when data moves between applications. In a silo-based approach, each application or layer of infrastructure has its own encryption mechanism and key management approach.

Whether deployed locally or as a centralized system, best practice key management allows encryption and other forms of cryptography to be effectively and consistently controlled anywhere in the organization. This mitigates the risk of breaches due to human error and lost data due to lost keys.

A decrease in overall system risk is a major benefit of a joined-up approach to key management and a tangible return on investment accompanies improved security. Estimates put the typical cost of the loss of an individual data record between $10 and $200 and considerably more if the content is of particularly high value, such as corporate intellectual property. This estimate doesn't even consider the costs associated with the damage to the corporate brand and reputation.

Centralized key management and automated key distribution solutions can deliver greater scalability, lower operational costs, and lower compliance costs through consistent policy enforcement and auditing. Professional key management of any kind can protect the corporation from legal liabilities and penalties imposed by compliance regulation and data privacy statutes.

Some security professionals have avoided the widespread deployment of encryption because of its perceived complexity and concerns about data loss through losing control of keys. However, the data protection benefits that encryption delivers outweigh the risks. Data protection legislation such as the Payment Card Industry Data Security Standard (PCI DSS) specifically recognizes the power of encryption and more and more applications and devices are encryption enabled.

Enterprise-wide key management strategies help organizations to manage the perceived risks and realize the benefits of encryption. They provide robust, automated, and reliable systems that ensure organizations can enforce their own corporate policies consistently and save money in the process.

Appendix

Key Management Checklist

The Data:

- What data needs protection?
- What value does data have to your organization?
- What value does data have to an attacker?
- What would be the financial/reputational/legal cost of data exposure or modification?
- Who or what needs to access the data?
- Who or what actually has access to the data?

The Policy:

- Who sets the data access policy?
- Who's responsible for demonstrating that the policy is correctly followed?
- What controls are in place to ensure that the policy is followed?

The Cryptography:

- Which cryptographic algorithms are in use?
- Are all the algorithms published, peer reviewed, and widely tested?
- Are the algorithm implementations tested and validated independently?
- Does the cryptography provide confidentiality, integrity, authenticity, or a combination?
- Are the cryptographic processes appropriately protected?

The Keys:

- Where are the keys generated?
- Are the encryption keys backed up? If so, who releases copies of the encryption keys?
- Are signing keys backed up? If so, are signing keys expected to provide any degree of non-repudiation?
- Are keys manually loaded or imported? If so, how is this process protected?
- Are keys exported or displayed? If so, is it necessary and how is the process secured?

The Distribution Mechanism:

- How are keys transported between generation/storage/backup and usage locations?
- Are they encrypted in transit? If so, where does the transport encryption key come from?
- How is the source and destination of the key authenticated?

The Audit:

- What actions involving keys produce and do not produce audit entries? What governs this requirement?
- Is the audit trail cryptographically protected and time-stamped?
- How are the keys that protect the audit trail protected?

The Tools:

- What tools are used to manage policies, roles, and key usage end points?
- What tools are in place to help enforce policy, ensure secure key distribution, and aid key backup and recovery?
- What tools are used for auditing these policies and key management procedures?

The Integration Point:

- ✓ What crypto APIs (if any) do applications use to access keys and perform cryptographic functions?
- ✓ What scope for modification or abuse of in-house or commercial applications exists?

Certification:

- ✓ What standards and certifications are required (FIPS, PCI DSS, ISO, KMIP, and so on)?

The Scale of the Problem:

- ✓ How many end points or application instances need protecting?
- ✓ How many keys exist?
- ✓ How will these figures grow over time?
- ✓ What's the arrival rate of new end points?
- ✓ What's the rate at which end points are revoked and for what reasons?

About the Author

Richard Moulds is the Vice President of Product Strategy for Thales e-Security, a strategic business line of Thales Group. Richard joined Thales as the result of its acquisition of nCipher in October 2008 and is based in Boston, USA. Previously Richard has worked in the networking and video communications markets in various marketing, business development, and general management roles. Richard is a frequent contributor to the blog Key Management Insights, which can be found at `www.keymanagementinsights.com`.