

ENCRYPTION STRATEGIES

The Key to Controlling Data

Sun Services
White Paper
August 2007

Table of Contents

Executive Overview	1
The Security Landscape	1
The Sun Strategy	2
Three Encryption Methods	3
At creation: host-based encryption	3
In band: appliance-based encryption	3
At rest: device-based encryption	4
Key Management	4
The key is the key	4
Redundancy works	5
Future directions	5
Conclusions	6

Executive Overview

When it comes to data management, today's reality is this: digital data cannot be controlled. Anything that's created can and will touch countless media surfaces during its life on a local area network (LAN) or storage area network (SAN)—servers, magnetic disk, optical disk, tape, even memory sticks. Data can be sent around the world with just a simple click, landing in places its creator never envisioned.

Thus the old method of controlling data by deleting it is an obsolete concept. In the digital age, data itself cannot be controlled, but access to it can be through the use of encryption keys, which allow the key holder to limit access to the data regardless of where it resides.

Sun has been involved in commercial and government-funded projects involving data protection that have yielded significant technology for encrypting data in transit and data at rest. They have also given rise to important strategies for using encryption key management not only to secure data, but also to manage it more effectively.

This paper outlines Sun's encryption strategy, presents the pros and cons of three common encryption methods used today, and discusses the challenge of key management in an encrypted world.

The Security Landscape

Not very long ago, hardcopy data was effectively safeguarded with “guards and gates,” which refers to the strategy of keeping sensitive files under physical lock and key. However, this model doesn't work well in the digital world. In a world where files can be accessed or shared with just the click of a mouse, a locked door will do little to protect against hackers, disgruntled employees, or simple human error. Consider the following¹:

- **January 2007** – Worldwide retailer TJX, the parent company of T.J. Maxx, Marshalls, and other stores, reveals an “unauthorized intrusion” into its computer systems where customer transactions are processed and stored. It's discovered that TJX had an outdated wireless security encryption system and had failed to install firewalls and data encryption systems, allowing perpetrators to access data for a period of 18 months. In one case, U.S. Secret Service agents found TJX customers' credit card numbers in the hands of Eastern European cyber thieves who had created high-quality counterfeit credit cards. It was ultimately revealed that more than 46 million account numbers were compromised, creating victims in the U.S., Europe, Asia, Canada, and elsewhere and resulting in at least one class-action lawsuit and an additional 21 U.S. and Canadian lawsuits against TJX. As of Q1 2007, TJX reported \$17 million in costs associated with this intrusion—and they are anticipating more.

¹ Source: Privacy Rights Clearinghouse, the nonprofit consumer education and advocacy project, 2007.

- **February 2007** – An employee of the U.S. Department of Veteran’s Affairs reports that a portable hard drive containing the personal information of 535,000 veterans, including social security numbers, is missing. Ten days later, the VA reports that billing information for 1.3 million VA doctors was also exposed, including names and Medicare billing codes. The breach potentially exposes the identities of nearly a million physicians and VA patients, and results in 46 separate investigations costing the VA more than \$20 million.
- **April 2007** – Using TurboTax online to access her previous returns, a Nebraska woman finds that she is able to view tax returns of other Turbo Tax e-filers across the country. The files contain complete personal tax returns that have been filed electronically with the Internal Revenue Service, including bank account numbers with routing digits and social security numbers.
- **May 2007** – Two laptop computers from Highland Hospital in Rochester, New York, are sold on eBay. One of the computers—later recovered—contained confidential records of 13,000 patients, including social security numbers.

Such incidents are just the tip of the iceberg. Sensitive information will continue to get into the wrong hands at an escalating pace. One defense is to “scramble” the data via encryption as it rests on a storage device (in a “data-at-rest” condition). This way, if a device is accidentally or intentionally breached, or if the storage device is lost, stolen, or misplaced, the data it contains will be unintelligible without a decryption key—rendering it useless to prying eyes.

This type of device-based security is effective today, but it will become less secure over time as more and more information is created digitally and as technologies emerge to gain access to that data. That’s the reason for legislation such as the Sarbanes-Oxley Act of 2002, California’s Senate Bill 1386, and the evolving Health Insurance Portability and Accountability Act (HIPAA) regulations, each of which places even greater emphasis on data encryption and an even greater burden on those who must protect data.

The Sun Strategy

After studying the data landscape, Sun has concluded that encryption can not only solve the current problem—that is, the problem of data getting into the wrong hands—it can also form the basis for an effective data management system through rules-based key management. The trick is in making key management simple and affordable, using levels of automation that reduce or eliminate the need to decide what data should be encrypted. Over time, encryption could become a more effective way to manage data than trying to track and eventually delete the data itself.

Sun has an emerging strategy to make this scenario come true using a variety of encryption methods that put control in the hands of the datacenter. The Sun strategy can be implemented at any one of three points in the life of data: at creation, at the time of transport, or at a time when it is at rest on a storage device. The technologies available from Sun allow companies to decide how, when, and where to encrypt their data.

Three Encryption Methods

Data can be encrypted when it's being created (host-based encryption), when it's being transported across the LAN (appliance-based or in-band encryption), or when it's at rest on a storage device (device-based encryption). Each of these methods has advantages and disadvantages. Following is a brief overview of each type.

At creation: host-based encryption

With host-based — or server-based — encryption, data is encrypted the moment it's created, providing the highest possible level of data security. Since data is encrypted at creation, there's no chance of unencrypted data being intercepted, either accidentally or maliciously. If data is intercepted, encryption renders it unreadable and worthless. Host-based encryption is a good fit for active databases where data changes constantly.

While host-based encryption is a highly secure approach to data encryption, several considerations need to be kept in mind.

- Current operating infrastructures need to be changed to implement this method. This is not the case with appliance-based or device-based encryption, both of which place the encryption burden on devices rather than on system infrastructure.
- Once data is encrypted, it can't be compressed, so the encryption infrastructure will expand over time as data volumes continue to expand.
- Encryption can increase data processing overhead by as much as 40 percent, requiring additional processing power to preserve performance and resulting in additional expense in the datacenter. Encryption-specific accelerator cards and emerging grid computing platforms will help address these performance issues.
- To decrypt data when it is retrieved from a storage archive, host-based encryption software must be maintained with the data. This is a challenge due to the constantly changing nature of software — one that affects both cost and maintenance. Bottom line: host-based — or server-based — encryption is highly secure and well-suited to active data files.

In band: appliance-based encryption

With appliance-based encryption, data is encrypted “in band” as it is being transported from the point of its creation to its destination. This method protects data at the network level, implementing security features on LAN-connected or SAN-connected encryption appliances or switches. Data leaves the host unencrypted, then goes into a dedicated appliance where it is encrypted. After encryption, it enters the LAN or a storage device. The technology for this method exists today. It is simple to install and requires no changes to the existing data infrastructure.

While this method offer an easy way to encrypt data, several considerations need to be kept in mind.

- Appliance-based encryption is not as secure as host-based or device-based encryption. It's relatively easy to bypass by changing the LAN infrastructure to intercept unencrypted data.
- Appliance-based is a costly option, requiring one dedicated appliance for every two to six storage devices.
- Appliance-based encryption is the least scalable of the three methods. It works well as an immediate fix, but it grows more expensive and is more difficult to manage as data volume increases. Bottom line: in-band appliance-based encryption is easy to implement and is well suited as a quick method for localized encryption solutions.

At rest: device-based encryption

Data at rest can be encrypted on a disk controller or dedicated storage server, making it easy to validate and eliminating the performance penalty on the server. This method is easy to implement and provides a good fit for mixed environments with a variety of operating systems.

Device-based encryption supports data compression. With this method, it's impossible to bypass encryption without detection. Since the storage devices handle the encryption task, no changes are required to the existing data infrastructure. Decryption code is built into the data storage container, so there's no need to maintain decryption software specifically for archived data.

While device-based encryption is a promising technology, several considerations need to be kept in mind.

- The technology is still emerging and is not yet ready to be implemented.
- Data is transmitted unencrypted until it reaches the storage device.
- Existing storage devices need to be replaced to support the technology.

Bottom line: device-based encryption is easy to implement and cost-effective, and is best suited to static and archived data.

Key Management

In all three methods, data access is controlled with an encryption key. Therein lies the risk: lose a key and you lose your data. That fact alone makes key management one of the most important aspects of data security in an encrypted world.

The key is the key

Even in the era of "guards and gates," key management was an issue. When a key to a file cabinet was lost or misplaced, access to crucial information was delayed. This problem is greatly magnified in the digital age. Consider a scenario with medical information such as X-rays.

A decade ago, X-rays were kept in a locked file cabinet in a hospital. They were signed out by one person who was authorized to view the films, such as a doctor, therapist, or neurologist. One key and a backup provided access to all films. Each X-ray existed in just one location. Today, digital X-rays are shared with colleagues and specialists around the world who are asked to help with real-time diagnoses. Millions of X-rays need to be put into thousands of hands, all under HIPAA guidelines for confidentiality.

The challenge, suddenly, is not managing one or two keys for a locking file cabinet, but managing thousands or even millions of keys that allow access to data files. A successful data encryption strategy must address this challenge and make sure that the keys to unlock data are never lost. But how?

Redundancy works

The initial answer is fewer keys. It doesn't help to replicate a key five times, turning a thousand encryption keys into five thousand keys that must be managed and protected. Instead, a successful key management system involves redundancy of keys, so fewer keys are used to manage more data. That way, a select group of people manage a select group of keys. If a key is lost, it can be replaced without jeopardizing data.

A redundant key strategy allows key management to be device independent, so IT managers are free to choose where and how data will be encrypted—whether at the host, on a LAN-attached encryption appliance, or on a storage device. As an added bonus, this strategy will require no change to existing applications or processes.

As encryption becomes more prevalent industrywide, key management will become more automated, driven by a set of rules in operating systems and applications. Once these rules are in place, file-based encryption with automated key management will allow encryption keys to be used to manage data, so companies won't have to worry about where the data is located or how many copies have been made.

Future directions

Moving to a key-based management system for encryption will take time, but Sun has a plan that will roll out in multiple phases. Properly architected, a key management system from Sun will allow encryption technologies to be implemented without significant hardware or process changes.

- **Phase 1: Limited key management.** Early solutions will include strong security measures to safeguard the data environment, but they will have limited encryption key management capabilities. Encryption will be handled by devices; it will not intersect with applications and infrastructure, so implementation will be fast and uncomplicated.
- **Phase 2: File-level key management.** Encryption keys will be dynamically allocated within the Sun storage software architecture.
- **Phase 3: Rules-based automated key management.** Common toolbox commands will be used for device-level encryption, to be enabled at the file level or block level.

Conclusions

Copies of digital data can be sent worldwide with a single click, so the old method of controlling data by deleting it is an obsolete concept. In the digital age, the best way to control data is to encrypt it. Data can be encrypted when it's being created (host-based encryption), when it's being transported across the LAN (appliance-based encryption), or when it's at rest on a storage device (device-based encryption).

- **Host-based encryption** is highly secure, but it adds to data processing overhead and requires changes to system infrastructure. Still, when data security is paramount, host-based encryption is a good choice.
- **In-band appliance-based encryption** is costly and not scalable, but it's easy to implement and puts no burden on the processing pipeline. It's acceptable as a short-term fix, but it may have trouble meeting long-term encryption needs.
- **Device-based encryption** is easy to implement and places no burden on the processing pipeline. Sun is developing technologies to allow encryption at the storage device level.

In addition to new encryption methods, Sun is developing technologies for key-based data management, so data encryption and data management can be part of the same IT process. These emerging technologies will eventually automate the process of data encryption and key allocation, making it easier not only to secure sensitive data but also to manage it more effectively.