

ENCRYPTION STRATEGIES

The Key to Controlling Data

White Paper

October 2007

Table of Contents

Executive overview.....	3
Key management.....	3
The key is the key.....	3
Redundancy works	4
Evolution of Encryption	4
The security landscape	5
The Sun strategy.....	6
Three main encryption methods	6
At creation: host-based encryption	6
In-band: appliance-based encryption	7
At rest: device-based encryption	7
Conclusions	8

Executive overview

The reality is this: digital data cannot be controlled. Anything that's created can and will touch countless media surfaces during its life on a LAN or SAN—servers, magnetic disk, optical disk, tape, even memory sticks. It can be sent around the world with a 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 data can be controlled with encryption keys, which allow the key holder to limit access to the data regardless of where it resides.

Sun's Storage team has been involved in commercial and government-funded projects involving data protection for many years. These projects 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, discusses the pros and cons of three common encryption methods used today, and discusses the challenge of key management in an encrypted world.

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 encryption.

The key is the key

Even in the era of “guns, 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 in thousands of hands, all under HIPAA guidelines for confidentiality (United States regulation). 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, not more keys. It does not 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 — so 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, or 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 industry-wide, 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 will not have to worry about where the data is located or how many copies have been made.

Evolution of Encryption

Moving to a key-based management system for encryption will take time, but Sun has a multiple phased plan. 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 now available.**
Our current solution includes strong security measures to safeguard the data environment, but has limited encryption key management capabilities. Encryption is handled by the T10000 tape drive; it does not intersect with applications and infrastructure, so implementation is accomplished quickly and easily.
- **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.

The security landscape

Not very long ago, hard-copy data was effectively safeguarded with “guns, guards and gates,” which refers to the strategy of keeping sensitive files under physical lock and key. This model does not work well in the digital world. A locked door will do little to protect against hackers, disgruntled employees, or simple human error in a world where files can be accessed or shared with a click. Consider:

- **January 2007** – Wellpoint’s Anthem Blue Cross Blue Shield reports cassette tapes containing customer information were stolen from a lock box held by one of its vendors. Data included names and Social Security numbers of 196,000 customers.
- **April 2007** – The Georgia Department of Health reports a computer disk containing personal information (including addresses, birthdates, dates of eligibility, full names, Medicaid or children’s health care recipient identification numbers, and Social Security numbers) was reported missing from a private vendor. Affiliated Computer Services (ACS) was the vendor contracted to handle health care claims for the state; 2,900,000 personal records were involved.
- **May 2007** – An IBM vendor lost computer tapes containing information on IBM employees — mostly former employees — including Social Security numbers, dates of birth, and addresses. Tapes were reported as missing in transit from a contractor’s vehicle.
- **June 2007** – The State of Ohio reports that a backup computer storage device with the names and Social Security numbers of every state worker was stolen out of a state intern’s car. The tape contained personally identifiable information of nearly 84,000 current and former Ohio state employees. In addition, the names and Social Security numbers of more than 225,000 taxpayers were also on the tape.
- **July 2007** – Transportation Security Administration (TSA) authorities realized in May 2007 that a storage device was missing from TSA headquarters. The drive contained historical payroll data, Social Security numbers, dates of birth, addresses, bank account numbers and routing information, and details about financial allotments and deductions.

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.

This type of device-based security is effective today, and will continue to be 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, evolving HIPAA regulations, and BASEL II in Europe, each of which places ever-higher emphasis on data encryption and an ever-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 falling 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 is building upon its 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 main encryption methods

There are three primary encryption methods available today. 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. Here's a brief overview.

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%, requiring additional processing power to preserve performance, and resulting in additional expense in the datacenter. Encryption-specific accelerator cards and grid computing platforms can help address these performance issues.

- To decrypt data when it is retrieved from a storage archive, host 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. For large scale data encryption purposes, it can be cumbersome and impact performance.

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, requiring some changes to the existing data infrastructure. While this method is an easy way to encrypt data, several considerations need to be kept in mind.

- In-band 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.
- It’s a costly option, requiring a dedicated appliance for every one to two current generation storage devices.
- In-band encryption is the least scalable of the three methods. It works well as an immediate fix, particularly for legacy storage devices, but it grows more expensive and is more difficult to manage as data volume increases.

Bottom line: In-band appliance-based encryption can be implemented across a variety of previous generation devices, and it is well suited as a quick method for localized encryption solutions. For extensive data storage encryption needs, the cost and management complexity of encrypting in-band can increase significantly.

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. It’s 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.

Device-based encryption is accelerating quickly due to its cost-effective and high performance nature, but several considerations need to be kept in mind.

- Data is transmitted unencrypted until it reaches the storage device.
- Previous generation storage devices need to be refreshed to support the technology.

Bottom line: Device-based encryption is easy to implement and cost-effective, best suited to static and archived data or encrypting large quantities of data for transport. As the technology matures, increasingly large numbers of devices will be manageable from a single key management platform.

Conclusions

Copies of digital data can be sent worldwide with a 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 or switch-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 it 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 can be widely implemented and puts little 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 has developed technologies to allow encryption at the storage device level and is aggressively expanding its portfolio of device-based encryption offerings.

In addition to supporting a variety of encryption methods, Sun is continuing to enhance technologies for key-based data management. The result is that data encryption and data management are becoming increasingly integrated, eventually delivering automated, policy-driven data encryption and key allocation, making it easier not only to secure sensitive data but also to manage it more effectively.

