White Paper

Encryption and Tokenization

Which is best for your business?

ENCRYPTION AND TOKENIZATION ARE THE TWO LEADING TECHNOLOGIES used to protect sensitive data from loss and subsequent breach notification and legal liability. Organizations who try to meet compliance regulations struggle to understand when to use strong encryption, and when to use tokenization, to protect information. This white paper attempts to provide some guidance on how and when each of these technologies should be used. Many organizations will find both technologies helpful in different places in their IT infrastructure.

www.townsendsecurity.com

724 Columbia Street NW, Suite 400 • Olympia, WA 98501 • 360.359.4400 • 800.357.1019 • fax 360.357.9047 • www.townsendsecurity.com

Encryption - Protecting Sensitive Data Where It Lives

Encryption protects data by obscuring it with the use of an approved encryption algorithm such as AES, and a secret key. Once encrypted, the original value can only be recovered if you have the secret key. The use of strong encryption keys makes it impossible, from a practical point of view, to guess the key and recover the data. Almost all compliance regulations provide a safe harbor from breach notification if sensitive data is encrypted.

Encryption is a mature technology with a recognized body of standards, independent certification of vendor technologies, and it undergoes continual scrutiny by the professional cryptographic community. Organizations that deploy professional encryption solutions that have been independently certified enjoy a high level of confidence in the protection of their data assets.

Tokenization - Protecting Sensitive Data With Substitution

Tokenization works by substituting a surrogate value for the original sensitive data. This surrogate value is called a “token”. The token value does not contain sensitive information, it replaces it ,maintaining the original value. There is one and only one token value for any given original value. For example, a credit card number 4111-1111-1111-1111 might be assigned the token value of 1823-5877-9043-1002. Once this token is assigned it will always be used when the original value would have been used.

Tokens may contain some of the original value as a “hint” to help users. Following the above example, the token might be assigned the value 1823-5877-9043-1111 (the last 4 digits are the same as the original value). This is a “masked” token as it hides the first 12 digits of the original value, but exposes the last 4 digits.

Tokens are not just credit card numbers. Tokens can be created for social security numbers, driver’s licenses, passport numbers, zip codes, and many other types of sensitive data. Good tokens help minimize the impact on the applications and databases by providing a similar data type and length.

Recoverable and Non-Recoverable Tokens

There are two types of tokens – recoverable and non-recoverable. A recoverable token provides the ability to retrieve the original value of the sensitive data using the token. A non-recoverable token is one in which the original value is not stored in the token database, and which cannot be used to recover the original value. Both types of tokens maintain a relationship between the sensitive data and the token. But only recoverable tokens can be used to recover the original value.

Non-recoverable tokens are especially helpful in situations where you do not need to recover the original sensitive data. This might include developer environments, business intelligence databases, and so forth. Because non-recoverable tokens do not store the original sensitive data, the tokenization solution is itself not in scope for compliance.

Tokenization Solutions and Scope of Compliance

It is important to note that a tokenization solution that stores the original sensitive data is itself in scope for compliance. For example, you may tokenize data in several developer environments and they are no longer in scope for compliance, but the tokenization solution stores the original data so it IS in scope for compliance controls. This means that a tokenization solution must take the same care to protect sensitive data that you would take in your production databases. The tokenization solution should use industry standard encryption algorithms such as AES, and must use proper key management solutions.

Because tokenization solutions represent a higher value target, you should be sure the solution uses independently certified encryption and key management.

Tokenization to Reduce Database Impacts

One common use of tokenization is to minimize the impact on a database application that would be caused by encryption. Because encryption creates a different type of binary data representation, it can require that applications and databases be changed to handle this new data type.

Tokens usually have exactly the same data type and representation as the original sensitive data. Thus the use of tokens may greatly reduce the work needed to protect sensitive data.

Note that there are reasons why tokenization may not be optimal even when it reduces the work required to protect data. See the discussion below about performance.

Tokenization to Reduce Scope of Compliance

Another common use of tokenization is to remove one or more servers from the scope of compliance. A good example is a Business Intelligence database. Companies often provide their users with a copy of the production data so that they can perform business analytics on the data. Usually it is not necessary to have the real credit card number or social security number to do this type of analysis. By substituting sensitive data with a token, you can completely remove these applications and users from the scope of compliance. It you do not have access to sensitive data, you can't lose it!

In addition to business intelligence databases, it is common to tokenize data used by developers. Developers often make copies of production data in order to enhance applications, to extend business functionality, or to fix bugs. They need to have realistic data to properly test their work, but this can lead to exposure of sensitive data. By tokenizing the data the risk of exposure can be reduced or eliminated. It is worth noting that one of the largest historical losses of sensitive data (Card Systems) was due to data leakage from developer environments.

Organizations stand to save money by removing systems from scope of compliance. As the number of systems contracts, the cost of software remediation and the cost of audits can fall dramatically.

Encryption for Performance

Encryption of data in place can be indicated where there is a need for optimal performance. There is a certain amount of overhead in the process of retrieving a token and this can have unacceptable impacts on performance. When the encryption solution performs well, encrypting data in place can be the best solution.

Most tokenization solutions use SSL/TLS communications sessions in order to retrieve token information. The overhead of a round-trip secure communications session can vary from a few milliseconds to 100 milliseconds or more. Assuming that the round-trip request requires 10 milliseconds, this can add several hours of elapsed time to process 1 million records in a database table. For large databases, encrypting data in place is preferable to tokenizing the data.

Tokenization for Business Intelligence

As mentioned before, tokenization can be a very powerful way to protect data in business intelligence databases. It can remove the BI solution from the scope of compliance, reduce the number of users with access to sensitive information, and reduce the cost and time of compliance audits.

Because tokenization retains the integrity of database table relationships, BI users can still create complex business scenarios and get answers to complex analytical questions. This is true even when several fields in a BI database are tokenized. For example, a retail customer may want to analyze product sales information by credit card number, product concept, and region. Tokenizing the credit card number and zip code can retain the needed connection between tables to allow for this type of analysis.

Tokenization for Outside Service Provider Data

Many organizations send information outside of the company to external service providers. In the retail industry information is frequently sent outside for retail analytics. In the medical industry, sensitive information may be reported to governmental or insurance providers. And everyone is providing information to their outside accounting firm for standard audit activity. In many cases we can send tokenized data outside the company to prevent its loss by these service firms.

It is interesting to note that a data loss by an outside service provider almost always leads to breach notification by the originating company. The originating company may be the only source of information needed for breach notification.

Encryption for Compliance Clarity

For many years compliance regulations have included encryption as the main technology providing a safe harbor

from breach notification and legal liability. Many of these regulations specify encryption and key management based on the standards published by the National Institute of Standards and Technology (NIST), or similar widely accepted standards. Organizations can know that they are meeting these standards and can require that their vendors provide proof of compliance. This leads to a great deal of clarity to the organization trying to protect sensitive data.

No such regulatory clarity exists for tokenization. Tokenization has not been the subject of any recognized standards body, and has not had the many years of use that typically results in an understanding of best practices. Therefore there is not the same clarity about tokenization as a solution for data protection. A good case in point is the HIPAA / HITECH data security standard. This standard points to the use of NIST approved encryption and key management solutions as the only mechanism that provides a safe harbor from breach notification. Would tokenization of data also provide the same safe harbor? This is not at all clear.

For some organizations that need very clear guidance on data protection, encryption will be the preferred method for achieving compliance.

Compliance of the Tokenization Solution

A tokenization solution is implemented on a separate server with a database to store the tokens, the encrypted original data, and it provides the SSL/TLS services for user authentication and secure transfer of data. Because a tokenization solution consolidates sensitive data to one location, the tokenization solution will be in scope for compliance, and should provide the highest level of security for the sensitive data. In essence, a tokenization solution becomes a high-value target for the bad guys. You should be sure that your tokenization solution meets all compliance guidelines and requirements for encryption and key management. Only encryption that meets NIST standards should be used to protect data, and you should only use key management that meets NIST best practices.

Many encryption and tokenization vendors claim they provide you with NIST compliant AES and Key Management solutions. When evaluating encryption and tokenization vendors ask them to show you their proof of certification. The certifications give you the proof you need for compliance, protect you from weak security, and protect you from spurious patent or intellectual property claims.

Combining Encryption and Tokenization

For most organizations there will be appropriate uses for both encryption and tokenization. Encryption will be the right solution for one set of applications and databases, and tokenization will be the right solution for others. The appropriate technology will be decided by each organization's technical, compliance, and end-user staff working together.

In order to ease the development and compliance burden, organizations may wish to source encryption and tokenization solutions from the same vendor. There are many overlapping technologies in both encryption and tokenization, and you will probably want a common approach to both.

Townsend Security

Despite an orgnization's best efforts, their data will get out. The best way to secure critical information is with strong encryption and key management.

Townsend Security provides NIST-certified encryption and logging solutions for the Enterprise. Our encryption, key management, tokenization, and logging solutions protect sensitive data from loss, whether it rests within, or is transmitted outside, of your organization.

Our certified database encryption and key management solutions are guaranteed to meet the encryption and key management meet or exceeds the standards in PCI, HIPAA, HITECH and more. Organizations worldwide rely on Townsend Security for their data privacy needs.

Web: www.townsendsecurity.com
Phone: (800) 357-1019 or (360) 359-4400
International: +1 360 359 4400
Email: Info@townsendsecurity.com