TOKENIZATION: WHAT'S NEXT AFTER PCI?

Rather than trying to protect cardholder data that is widely dispersed across the environment, a tokenization solution removes it altogether from any systems and applications that don't specifically require it. This is a major game changer.

Executive Summary

Until recently, the central theme for IT security has been: "Protect sensitive data wherever it resides." With the growing adoption of tokenization solutions, primarily in the payment card industry (PCI), a second principle is gaining equally wide acceptance: "Remove sensitive data wherever it's not required."

This paper examines the factors that have driven rapid adoption of tokenization among retailers and other merchants, and it offers lessons from the PCI experience that can be applied to other industries and use cases. Most notably, tokenization has helped reduce business risk and ease the compliance burden for securing credit card data. Looking beyond PCI, the paper explores where the next big wave of tokenization is likely to occur: in key vertical industries that need to safeguard personally identifiable information (PII) and protected health information (PHI).

The First Wave of Tokenization Was All About Payment Card Data

If necessity is the mother of invention, PCI compliance is the mother of tokenization.

First published in 2004, the Payment Card Industry Data Security Standard (PCI DSS) has imposed an enormous compliance burden on retailers, e-tailers, payment processors, and banks. It also affects any "merchant" that accepts credit cards as payment for goods and services including businesses, schools, educational and healthcare institutions and nonprofit organizations.

PCI DSS defines 12 major requirements and over 200 sub-requirements for protecting cardholder data. These must be applied across the entire Card Data Environment (CDE), meaning any system that accepts or stores payment card data plus any systems that access the data. Unfortunately, credit card numbers have long been used as a primary identifier for systems, applications and business processes that have no intrinsic need to access the number itself. (In many industries the same thing has happened with Social Security Numbers.) The staggering compliance burden this places on merchants becomes apparent in this description by Securosis of a typical retail environment:

> "As the standard reference key, credit card numbers are stored in billing, order management, shipping, customer care, business intelligence, and even fraud detection systems. They are used to cross-reference data from third parties to gather intelligence on consumer buying trends. Large retail organizations typically store credit card data in every critical business processing system.

"It is *incredibly* expensive to audit network, platform, application, user, and data security across all these systems — and then to document usage and security policies sufficiently to demonstrate compliance with PCI-DSS.[1]

The Greatest Risk Is In the Application Layer

According to a study conducted by the Verizon RISK team, 92% of all data breaches are the work of external agents, who target servers and applications most of the time. Drilling down further in the Verizon data, servers accounted for 80% of breaches and 95% of compromised records, with POS and web servers leading both metrics. Due to this, an organization interested in preventing data breaches or meeting compliance requirements must protect sensitive data in the application layer, where the majority of threats reside. To date, encryption, along with strong key management, has been the preferred method of enforcing data protection in applications.

Once an organization identifies which applications and business processes don't require use of the actual [credit] card number, tokenization can shrink the Card Data Environment significantly. In turn, this greatly reduces PCI compliance scope and costs.

However, tokenization has rapidly gained acceptance as an attractive alternative due to its compelling value proposition. The primary benefit of tokenization is that rather than trying to ***protect*** cardholder data that is widely dispersed across the environment, a tokenization solution ***removes it altogether*** from any systems and applications that don't specifically require it. This is a major game changer: Thieves can't steal what isn't there, and organizations don't need to protect what they no longer store. The result is a dramatic reduction in security and compliance requirements and costs.

Tokenization offers another significant advantage over encryption. Encrypting data often requires system software and business applications to be recoded so they can handle the added length of an encrypted value. Tokenization can be deployed with only minor application changes. This means data removal can proceed at a faster pace and far more cost-effectively than encrypting the same data would entail.

How Tokenization Works

In a typical tokenization scenario, card data is encrypted at the point of capture and transmitted to a secure, central repository, which may be operated by the merchant or a third-party service provider. (See Figure 2.) The system provides the merchant with a randomly generated substitute value, called a token, which cannot be traced back to the original. Because the token retains the same length and format as the original number, it can be seamlessly passed between applications, databases and business processes without risk.

The encrypted credit card data is vaulted in a highly secure facility, with multiple layers of protection and appropriate redundancy for disaster recovery and business continuity purposes. Only applications that require the actual card number are authorized to access the vaulted data; this is the only point in the CDE where tokens and account numbers are correlated.

Like encryption, tokenization can be performed on the database layer, from the network or on the application layer. Tokenizing or encrypting data at the point of capture—in the application layer--provides the best protection as data exposure is minimized.

Tokenization Reduces PCI Compliance Costs and Business Risk

One of the major benefits of tokenization is risk consolidation, says Sam Curry Chief Technology Officer of RSA's Identity and Data Protection Division, "In essence, tokenization enables a merchant to consolidate sensitive data, and the related risk, from dozens or hundreds of systems, databases and networks to just a handful of points,"

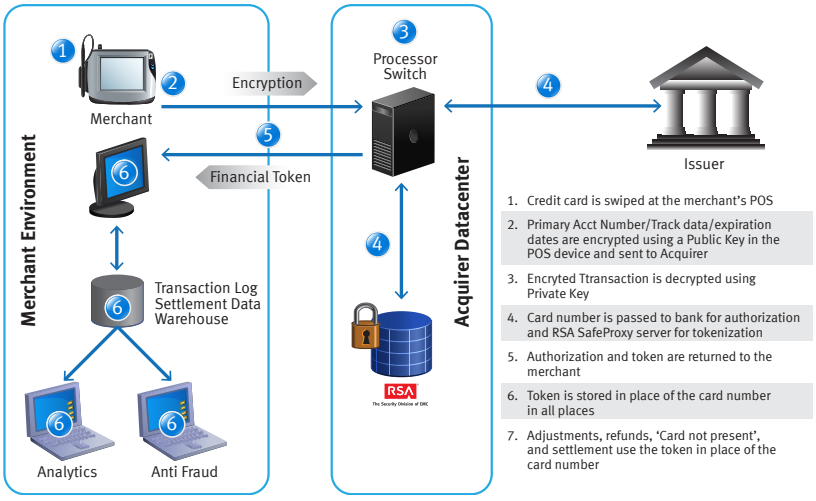
---

1 *Tokenization vs. Encryption: Options for Compliance*, Securosis, July 2011, page 3 [https://securosis.com/research/publication/tokenization-vs.-encryption-options-for-compliance]

2 *2011 Data Breach Investigations Report*, Verizon, June 2011 [http://www.verizonbusiness.com/resources/reports/rp_data-breach-investigations-report-2011_en_xg.pdf]

Figure 2: Tokenization as a Service

In the payment card world, tokenization can be implemented as an on-premise solution deployed by a merchant or as a third-party service offered by a payment processor as illustrated here. In this scenario, card data is encrypted at the POS using public/private key encryption to ensure safe transmission. The data is decrypted at the processor switch so the transaction can be authorized, and a token is returned to the merchant along with the authorization. The card number is then encrypted and centrally vaulted for maximum protection.



says Curry. "These points include the card processing infrastructure, primarily point-of-sale systems and the store network, and the secure vault. Companies can then focus security resources on safeguarding those high-risk points, making it easier to protect against intrusions."

Once an organization identifies which applications and business processes don't require use of the actual card number, tokenization can shrink the Card Data Environment significantly. In turn, this can greatly reduce PCI compliance scope and costs. For example, when a $5 billion global technology company outsourced its payment processing to a third party that tokenized cardholder data, the firm only had to comply with a few questions on the PCI Self-Assessment Questionnaire rather than the complete set of 200-plus questions. In turn, the company saved more than $3 million in PCI-related costs and months of internal development time. Similarly, an RSA customer in the government sector recently implemented tokenization with RSA Data Protection Manager and reduced its PCI scope and time spent on PCI compliance by 33%.

Although compliance has been the driving force behind tokenization, the result has been to enhance security and reduce business risk by greatly reducing the footprint of sensitive information across the enterprise. History shows that where only a small number of well-defended targets exist, criminals tend to move on to more vulnerable environments. Even if a data breach occurs and token values are stolen or exposed, the information is useless in perpetrating identity fraud and similar crimes.

Aberdeen Group, which has been tracking PCI DSS compliance efforts for several years, reported in 2010 that "the current use of tokenization is strongly correlated with Best-In-Class results" in protecting cardholder data via PCI DSS implementation. "The top performers were 2-times more likely than all others to indicate current use of a tokenization solution....The average performance of tokenization users was even higher than that of the average Best-In-Class company in Aberdeen's study as measured by the number of known incidents of data loss, data exposure or audit deficiencies within the last 2 years." [3]

Industry Validation Speeds Adoption

Beyond these well-established benefits of tokenization, two other factors have contributed to widening adoption: industry validation and the emergence of third-party services. The principle of tokenizing credit card data was first demonstrated in 2005, but the technology did not gain traction for several years. In October 2009 tokenization got a big boost when Visa—a perennial leader in defining and enforcing best practices for card data security—published guidelines for encrypting card data and recommended the use of tokens to replace the primary account number (PAN) in payment-related business functions. Visa followed up in 2010 by publishing best practices for tokenization, stating that:

3 *Avoiding a Kick in the Head: The Value of Tokenization for Protecting Cardholder Data*, Aberdeen Group, February 2010

> "Entities that properly implement and execute a tokenization process to support their payment functions may be able to reduce the scope, risks and costs associated with ongoing compliance with the Payment Card Industry Data Security Standards (PCI DSS)."

Confirming that guidance, the PCI Security Standards Council in August 2011 issued its own guidelines for developing, evaluating or implementing a tokenization solution, offering this advice:

> "Storing tokens instead of PANs is one alternative that can help to reduce the amount of cardholder data in the environment, potentially reducing the merchant's effort to implement PCI DSS requirements..."

> "Tokenization solutions do not eliminate the need to maintain and validate PCI DSS compliance, but they may simplify a merchant's validation efforts by reducing the number of system components for which PCI DSS requirements apply."

"In essence, tokenization enables a merchant to consolidate sensitive data, and the related risk, from dozens or hundreds of systems, databases and networks to just a handful of points.... Companies can then focus security resources on safeguarding those high-risk points, making it easier to protect against intrusions."

SAM CURRY, CHIEF TECHNOLOGY OFFICER, RSA IDENTITY AND DATA PROTECTION DIVISION

Secure Payment Services Shift the Risk to Providers

In the earliest days of tokenization, there were two basic deployment models for tokenization solutions: Merchants could build a homegrown system, or they could buy, deploy and operate a vendor solution such as RSA Data Protection Manager. A third option emerged when RSA partnered with First Data, the largest payment processor in the payment card industry, to create a secure payment solution that offered both encryption and tokenization of cardholder data as a hosted service.

This new model offered two compelling benefits: First, it freed merchants from the significant complexity and cost of building and maintaining an on-premise payment processing infrastructure. Second, by removing cardholder data from the enterprise environment and vaulting it in a vendor's secure repository, an outsourced solution shifted much of the risks and burdens of PCI compliance to trusted third parties with proven capabilities for securing card data. Merchants' security obligations didn't vanish completely. For instance, they're still responsible for securing the in-house payment processing environment. However, the bulk of their PCI DSS scope is transferred to service providers.

Interest was immediate and enthusiastic, especially among Level 3 and Level 4 merchants who, due to their smaller size, typically lack the resources to implement and maintain a payment processing infrastructure on their own. Because First Data's embrace of tokenization conferred instant legitimacy, these merchants jumped at the chance to gain the cost, compliance and ease-of-deployment benefits of a hosted offering from a leading provider.

The First Data model of tokenization is not for everyone. Many Tier 1 and Tier 2 retailers tend to work with multiple payment processors and thus need an on-premise solution that is vendor agnostic. Even so, a number of larger merchants opted for a hosted solution. Within three months of availability, First Data had more than 100,000 merchants using the service, a number that has since more than doubled.

Finding the Industry Sweet Spot

Because payment processors' core business places them directly in the stream of payment transactions—and because they already had the infrastructure in place to handle billions of transactions annually—they were the most logical place to implement tokenization on a large scale and as an add-on to their existing services. By recognizing that fact early on, First Data gained a first-mover advantage over competitors.

RSA believes that similar opportunities will emerge in other industries as thought leaders leverage existing infrastructure to profitably deliver tokenization services, both to targeted markets and to a general business audience.

The Next Wave of Tokenization Will Be About PII, PHI and the Cloud

The payment card industry has been a giant proving ground for tokenization. The external pressures of PCI compliance created an urgent market need and shaped how the technology has evolved and matured. Because PCI compliance initiatives have been pervasive in businesses, government and institutions—which all accept credit cards as payment for goods, services and fees—PCI projects have also created a small army of IT and security professionals across all sectors who know firsthand how effective tokenization is at removing high-value data from the environment.

An RSA customer in the government sector recently implemented tokenization with RSA Data Protection Manager and reduced its PCI scope and time spent on PCI compliance by 33%.

That group is starting to explore how they can leverage and extend their investments in tokenization to safeguard other types of structured data such as Social Security Numbers, motor vehicle license numbers, and banking and investment account numbers. The ultimate goal is to secure the transactions and communications that utilize the data including healthcare and employment records, financial transactions, and government records related to voting, taxes, and the criminal justice system.

This next wave of adoption is already under way and will touch multiple industries—led by the insurance, healthcare and hospitality industries—and focus on a broader range of personally identifiable information (PII) and protected health information (PHI).

Tokenizing SSNs in the Insurance Industry

RSA is currently seeing significant activity in the insurance industry where, for decades, Social Security Numbers were used as the primary identifier for policyholders. SSNs became so deeply embedded in business processes and application infrastructures that IT systems could not function without them, even if there was no other reason to retain such sensitive data.

With the advent of data disclosure laws and other regulations protecting PII, insurers have strong reason to remove SSN data from the network environment. A number of insurers that RSA works with are doing just that by tokenizing SSNs that previously were stored as clear text in protected databases. One of the U.S.'s largest insurers has already deployed tokenization within their SOA to support PCI compliance. Now, with RSA's assistance, they are rolling it out to other parts of the business to protect PII, starting with a large-scale project to tokenize hundreds of millions of SSNs, which are accessed by hundreds of applications across the organization. A standard component of the insurer's business process evaluation has been to determine where SSNs are being used simply for lookup purposes and therefore can be replaced by tokens with no impact.

The insurer is leveraging a centralized, web services model for distributing tokens to requesting applications. The centralized infrastructure is more efficient to operate and easier to defend than a distributed model, and it provides a platform for tokenizing other types of PII, such as enrollees' policy numbers or health record numbers, if the company decides to do so.

Tokenizing PHI in the Health Care World

Health care is another area where tokenization is being explored as an alternative to encryption or inaction. As of March 2011, more than 8 million people in the U.S. had been impacted by breaches of protected health information (PHI)[4], with medical and healthcare groups accounting for 16 percent of all identity records exposed nationwide. Stolen healthcare records contain identity and financial data that can be used to commit financial fraud. Information on the patient's health status, insurer, and medical providers allows an imposter to pose as the patient in order to obtain "free" health care. Security

4 As reported by the U.S. Health and Human Services Office for Civil Rights, which is responsible for privacy and security enforcement under HIPAA and certain provisions of the Health Information Technology for Economic and Clinical Health (HITECH) Act.

executives are not only concerned about intrusions by external hackers; insider threats– from negligent or malicious employees, partners and contractors and from process breakdowns are also major causes of data breaches.

Addressing these kinds of concerns, the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 increased privacy and security requirements introduced under HIPAA[5]. For example breaches involving unencrypted PHI now require affected individuals to be notified within 60 days of discovery, and incidents involving more than 500 records must be reported to the Department of Health and Human Services and publicly posted. In contrast, if breached data is encrypted or tokenized, notification requirements do not apply because the data is considered unreadable. Eliminating the notification requirement reduces the financial costs and brand damage associated with notification.

When a $5 billion global technology company outsourced its payment processing to a third party that tokenized cardholder data, the firm only had to comply with a few questions on the PCI Self-Assessment Questionnaire rather than the complete set of 200-plus questions. In turn, the company saved more than $3 million in PCI-related costs and months of internal development time.

The data security challenges

In health care, the data security challenges are far more complex than in the payment card industry. Under current laws, medical providers, insurers, and other stakeholders are accountable for safeguarding PHI, defined as any information about health status, provision of health care, or payment that can be linked to a particular individual. HIPAA specifies 18 categories of identifiers, but an actual electronic health record may include hundreds of data points relating to medical conditions, medications, provider certificate or license numbers, medical device identifiers and serial numbers, and so forth.

Securosis points out a particularly thorny problem in securing PHI:

> "Many different groups need access to different subsets (or all) of [a patient's health record]: doctors, hospitals, insurance providers, drug companies, clinics, health maintenance organizations, state and federal governments, and so on. And each audience needs a different slice of the data — but must not see the rest of the data."[6]

Early experiences with tokenization

Some firms have applied tokenization to health records in a limited way, typically using a single token to represent an individual's name, address and SSN while other data in the health record is stored in the clear. In the long run, this is likely to prove insufficient since it has been demonstrated that a patient's identity may be deduced by correlating as few as two or three key identifiers that have been left unprotected.

Martin Sizemore, an enterprise architect with the IT management consulting firm Perficient, has argued that tokenization of health records, implemented within an SOA environment, is the right technology to address the exchange of PHI over public and private networks.

> "The big question is how to implement the tokenization of protected healthcare information? The short answer is make it a 'service' in a service-oriented architecture that talks to a tokenization server.... The tokenization server would contain the 18 or more key protected items and their corresponding tokens. The service would retrieve the protected information temporarily for healthcare applications and updates, but would prevent local storage of the information to maintain control."[7]

Whether this or another model emerges as the leading deployment scenario for tokenization of PHI, RSA expects to see a good deal of activity and progress in the next two to four years.

---

5 Health Insurance Portability and Accountability Act

6 *Tokenization vs. Encryption: Options for Compliance*, Securosis July 2011, page 7

7 *Is Tokenization the solution for Protected Healthcare Information (PHI)?*, February 2011, Perficient blog post, Martin Sizemore, February 2011 [http://blogs.perficient.com/healthcare/blog/2011/02/22/is-tokenization-the-solution-for-protected-healthcare-information-phi/]

Protecting the Hospitality Sector

RSA is also seeing increased interest in tokenization among customers in the hospitality industry. This may be in response to a recent rise in smaller attacks on companies in the hospitality and retail sectors. Verizon's 2011 data breach report notes that such organizations represent smaller, softer, and less reactive targets than, for instance, financial institutions and speculates that criminals may be deciding to "play it safe" in light of recent arrests and prosecutions following large-scale intrusions in financial services.[8]

PCI-compliant Cloud Computing

With several years of experience under its belt, the payment card industry will continue to be a bellwether for other industries, pushing the boundaries of what tokenization can do. One promising area is PCI-compliant cloud computing.

For the most part, cloud architectures have proven to be impractical for applications using payment card data. The high investment needed to build a PCI DSS-compliant environment in the cloud exceeds the cloud's efficient resource allocation benefits in most cases. Consequently, merchants have been limited in the types of business processes and applications they can move to the cloud, because many nonpayment-related systems use card numbers for look-up values.

Tokenization promises to open up new business models for PCI-compliant cloud computing, enabling merchants to shift IT-based services to the cloud by allowing applications that previously used PANs (and were thus governed by PCI DSS) to use tokens instead. The ability to take many business processes and IT systems out of PCI scope enables merchants to better leverage the cloud to achieve significant advantages in IT efficiency, cost and flexibility.

The RSA Approach to Tokenization

Given its proven effectiveness, tokenization should be an important component of any layered strategy for protecting high-value, structured data from end to end. With this principle in mind, RSA incorporated comprehensive tokenization functionality into the RSA Data Protection Manager platform, combining tokenization with industry-leading application encryption, data-at-rest encryption, and comprehensive lifecycle key management.

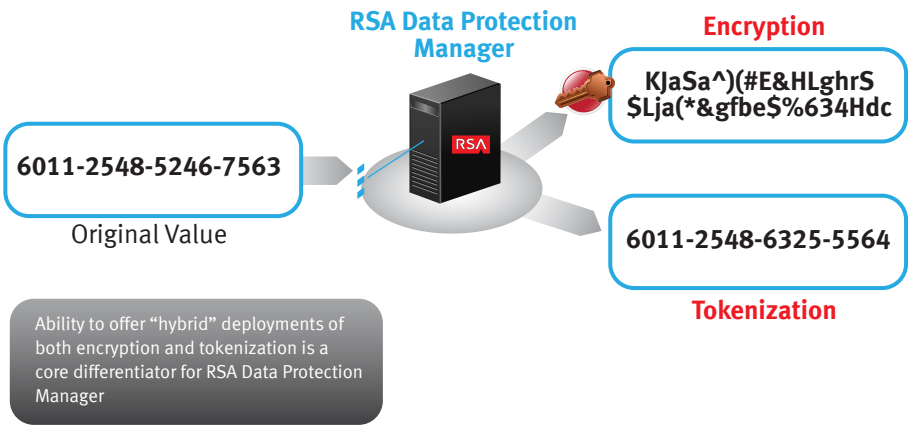


Figure 1: Protecting Credit Card Data

Encryption and tokenization are complementary technologies that use different mechanisms to protect structured data—in this case a credit card number—and are suitable for different scenarios. With tokenization, the original card number is replaced with a substitute value created by a random number generator that preserves the 16-digit data format. This allows applications to handle tokens without any special coding. A token can optionally include an element of the original value for identification purposes, for example the first 4 digits of the credit card number, as shown here.

8 *2011 Data Breach Investigations Report*, Verizon, June 2011 page 12

The RSA solution is unique in allowing flexible, hybrid deployments that leverage the distinctive and complementary characteristics of encryption and tokenization: tokenizing data at the point of capture when operational efficiency is the main concern, or immediately encrypting data when performance is paramount—for example, when the network connectivity required for generating tokens is interrupted—and then tokenizing the information later on. As a result, high-value information is secured throughout its lifecycle. Built-in enterprise key management simplifies provisioning, distribution and management of tokens and encryption keys and streamlines the deployment and administration of tokenization-enabled applications.

The RSA solution aligns with global best practices for PCI DSS compliance published by Visa in 2010 including guidelines for deploying the token server, card data vault, token generation and mapping, and cryptographic keys. Token values are created using non-algorithmic and non-reversible techniques, so tokens cannot be cryptographically linked back to the original data element. Flexible formatting allows for the use of pre-built token formats, customized formats, and partial disclosure of the original data, such as the last four digits of an SSN or credit card number.

Validating the power of this approach, the RSA solution has been adopted with great success by the largest retailers in the world and leaders in many other industries.

What's Next: Explore New Applications for Tokenization

Organizations that have already deployed a tokenization solution should look for opportunities to leverage their current investment to further reduce business risk, streamline compliance, enhance customer trust and enable new ways of doing business.

Those who have not yet implemented tokenization may want to identify a proof of concept project that promises a high level of business value and ROI.

To learn more about RSA tokenization solutions, contact your Account Representative today.

About RSA

RSA is the premier provider of security, risk and compliance solutions, helping the world's leading organizations succeed by solving their most complex and sensitive security challenges. These challenges include managing organizational risk, safeguarding mobile access and collaboration, proving compliance, and securing virtual and cloud environments.

Combining business-critical controls in identity assurance, data loss prevention, encryption and tokenization, fraud protection and SIEM with industry leading eGRC capabilities and consulting services, RSA brings trust and visibility to millions of user identities, the transactions that they perform and the data that is generated.

H11918 TOKEN WP 0312