Unit V : Basics of Tokenization

What is Tokenization

Tokenization replaces a sensitive data element, for example, a bank account number, with a non-sensitive substitute, known as a token. The token is a randomized data string that has no essential or exploitable value or meaning. It is a unique identifier which retains all the pertinent information about the data without compromising its security.

A tokenization system links the original data to a token but does not provide any way to decipher the token and reveal the original data. This is in contrast to encryption systems, which allow data to be deciphered using a secret key.

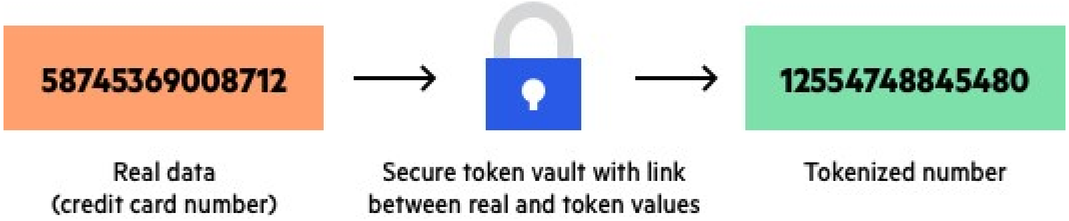
How Data Tokenization Works

Tokenization, in relation to payment processing, demands the substitution of a credit card or account number with a token. The token has no use and is not connected to an account or individual.

The 16 digits primary account number (PAN) of the customer is substituted with a randomly-created, custom alphanumeric ID. The tokenization process removes any connection between the transaction and the sensitive data, which limits exposure to breaches, making it useful in credit card processing.

Tokenization of data safeguards credit card numbers and bank account numbers in a virtual vault, so organizations can transmit data via wireless networks safely. For tokenization to be effective, organizations must use a payment gateway to safely store sensitive data.

A payment gateway is a merchant service offered by an e-commerce application service provider that permits direct payments or credit card processing. This gateway stores credit card numbers securely and generates the random token.



Tokenization in a nutshell

Payment Tokenization Example

When a merchant processes the credit card of a customer, the PAN is substituted with a token. *1234-4321-8765-5678* is replaced with, for example, *6f7%gf38hfUa*.

The merchant can apply the token ID to retain records of the customer, for example, *6f7%gf38hfUa* is connected to John Smith. The token is then transferred to the payment processor who de-tokenizes the ID and confirms the payment. *6f7%gf38hfUa* becomes *1234-4321-8765-5678*.

The payment processor is the only party who can read the token; it is meaningless to anyone else. Furthermore, the token is useful only with that single merchant.

Tokenization vs Encryption

The main difference between tokenization and encryption is that tokenization uses a 'token' whereas encryption uses a 'secret key' to safeguard the data.

Encryption

A core issue with data encryption is that it is reversible. Encrypted data is designed to be restored to its initial, unencrypted state. The safety of encryption is reliant on the algorithm used to protect the data. A more complex algorithm means safer encryption that is more challenging to decipher.

All encryption is, however, essentially breakable. The strength of your algorithm and the computational power available to the attacker will determine how easily an attacker can decipher the data. Encryption is thus better described as data obfuscation, rather than data protection. Encryption makes it more difficult to access the original information protected within the encrypted data, however not impossible.

The PCI Security Standards Council and similar compliance organizations treat encrypted data as sensitive data because it is reversible. Organizations are thus required to protect encrypted data.

Tokenization

Unlike encryption, tokenization of data cannot be reversed. Rather than using a breakable algorithm, a tokenization system substitutes sensitive data by mapping random data, thus the token cannot be decrypted. The token is a placeholder, with no essential value.

The true data is kept in a separate location, such as a secured offsite platform. The original data does not enter your IT environment. If an attacker penetrates your environment and accesses your tokens, they have gained nothing. Thus, tokens cannot be used for criminal undertakings.

The PCI and other security standards do not require organizations to safeguard tokenized data.

Benefits of Tokenization

Tokenization can provide several important benefits for securing sensitive customer data:

- **Enhanced customer assurance**—tokenization offers an additional layer of security for eCommerce websites, increasing consumer trust.
- **Increased security and protection from breaches**—by using tokenization, businesses do not have to capture sensitive information in their input terminals, keep it in internal databases, or transmit the data through their information systems. This safeguards businesses from security breaches.
- **Data tokenization improves patient security**—organizations can use tokenization solutions for scenarios covered under HIPAA. By substituting electronically protected health information (ePHI) and non-public personal information (NPPI) with a tokenized value, healthcare organizations can better comply with HIPAA regulations.
- **Tokenization makes credit card payments more secure**—the payment card industry needs to comply with extensive standards and regulations. Tokenization solutions provide a way to protect cardholder data, such as magnetic swipe data, primary account number, and cardholder information. Companies can comply with industry standards more easily, and better protect client information.

PCI Tokenization: Easing Compliance with Tokenization

The Payment Card Industry Data Security Standard (PCI DSS) ensures PAN data is protected by all organizations that accept, transmit, or store cardholder data. Failure to comply may result in fines and loss of brand authority.

Tokenization helps companies achieve PCI DSS compliance by reducing the amount of PAN data stored in-house. Instead of storing sensitive cardholder data, the organization only handles tokens, making for a smaller data footprint. Less sensitive data translates into fewer compliance requirements to comply with, which may lead to faster audits.

See how Imperva Data Masking can help you with data protection.

How Imperva Leverages Tokenization for Security and Compliance

Imperva's security solution uses data masking and encryption to obfuscates core data, so it would be worthless to a threat actor, even if somehow obtained.

We offer a holistic security solution that protects your data wherever it lives—on-premises, in the cloud, and in hybrid environments. We help security and IT teams by providing visibility into how data is accessed, used, and moved across the organization.

Our security approach relies on multiple layers of protection, including:

- **Database firewall**—prevents SQL injection and similar threats, while assessing for known vulnerabilities.
- **User rights management**—tracks the data movements and access of privileged users to identify excessive and unused privileges.
- **Data loss prevention (DLP)**—monitors and tracks data in motion, at rest, in cloud storage, or on endpoint devices.
- **User behavior analytics**—creates a baseline of data access behavior and uses machine learning to isolate and alert on abnormal and potentially dangerous activity.
- **Data discovery and classification**—discloses the volume, location, and context of data on-premises and in the cloud.
- **Database activity monitoring**—monitors relational databases, data warehouses, big data, and mainframes to produce real-time alerts on violations of policy.
- **Alert prioritization**—Imperva uses AI and machine learning technology to examine the stream of security events and prioritize the most important events.