New Approach to Enterprise Data Security: Tokenization

Jim Taylor
Liaison Technologies
What we will discuss

• Business drivers for data protection
• What is tokenization?
  – Approaching data security in the enterprise
  – Securing credit card numbers
  – Considerations for Personally Identifiable Information (PII)
  – Determining your tokenization platform
• Customer case study: insurance/financial services
• Question & answers
Business Drivers for Data Protection
If data can be monetized, it will be compromised

- **Citigroup:** Disclosure that names, account numbers and email addresses of 200,000 of its credit card customers were stolen

- On April 17, in what could be one of the largest known data breach ever, **Sony PlayStation Network (PSN)** was hacked. Sony acknowledged that PII data and possible credit card data of about 77 million users of PSN were impacted. This shattered TJX’s record of the largest customer data theft ever.
Business drivers for data protection

• Government
  – Sarbanes Oxley Act (SOX)
  – Gramm Leach Bliley Act (GLBA)
  – Healthcare Insurance Portability & Accountability Act (HIPAA)
  – Federal Information Security Management Act (FISMA)
  – State Breach Notification Laws (e.g. California State Bill 1386)

• Industry
  – Payment Card Industry Data Security Standard (PCI DSS)
  – Healthcare Insurance Portability & Accountability Act (HIPAA)
  – Health Information Technology for Economic and Clinical Health Act (HITECH)

• Company
  – Brand protection in general
  – High-wealth individuals, etc.
Tokenization business drivers

- **PCI DSS compliance**
  - Reducing scope as part of PCI DSS compliance
  - Reducing annual / ongoing PCI DSS audit costs

- **Reducing risk within and outside the enterprise**
  - Limiting exposure to sensitive data
  - Credit card numbers
  - Personally Identifiable Information (PII)
  - Protected Health Information (PHI)

- **Limiting application changes**
  - Encrypting data enlarges field sizes
  - Tokens, as we’ll see, do not require field sizes to be increased
What is Tokenization?
Approaching data security in the enterprise

- “If you don’t need the data, don’t store it.”
- “Compliance does not equal security”
- “Think Safe Harbors”
  - Render data useless to those who should not have access
  - If data is compromised that is rendered useless, you don’t have to let consumers (and the media 😞) know
- “If there is one thing you can do to protect your data, educate your organization about data security.”
- “If you need to maintain the data, protect it.”
Protecting data

- Hashing
- Masking
- Encryption
- Tokenization

Data Assets
- Databases
- Files
- Applications
- Archives
Encryption

- How granular? Field, column, table, database
- Where are the keys stored?
- What happens when data leaves the database?

Database Encryption

- Does the application have to be changed?
- What external components are used?
- Integration options – interfaces, languages

Application Encryption

- Field within a file?
- Entire file?
- Directory of files?

File Encryption
Key management

- Must address entire key management lifecycle
- Generate, distribute, expire, rotate, revoke, destroy...
- How do you activate keys and how many keys can be used?
- Where are the keys stored? (Active and archived)

Roles-Based Admin and Control
- Must have dual control of keys
- Users can never have direct access to keys
- User-level information must be available for key use

Event Logging
- Must log all administration functions pertaining to keys, users
- Must log all end user events (encrypt, decrypt)
- Integration with security information and event management (SIEM) systems
What kind of token are we talking about?

- It’s NOT the same as ‘token’ used for two-factor authentication
- It’s not the ‘token’ used for lexical analysis (creating a programming language)
- In data security, it’s a **surrogate value** which is substituted for the actual data (e.g. credit card) while the **actual data is encrypted** and stored elsewhere.
• Original data values cannot be mathematically derived from tokens
  – Tokens can be safely passed to databases, applications, mobile devices, etc.
• Solves the age-old problem of data for development and testing!
Centralized data vault and key management

- Protected data vault where sensitive data is encrypted and stored
  - Reduces the footprint where sensitive data is located
  - Eliminates points of risk
  - Simplifies security management

- Control over who accesses sensitive data
  - Rotate keys without having to decrypt and re-encrypt old data, and no system downtime
  - Keys are distributed to token server, not throughout enterprise
Tokenization still requires encryption / key management

- **Encryption**
  - Data must still be stored as ciphertext so it can be access when required, with the proper authority
  - Strong encryption, typically AES 256, but 128 and/or 3DES common

- **Key Management**
  - Keys are distributed to token server, not throughout enterprise
  - Full key lifecycle management still required
  - Rotate keys without having to:
    - decrypt and re-encrypt old data, and no system downtime
    - Re-tokenize - key rotation should not break the mapping between tokens and data.
Tokens can be formatted to:

- Preserve the format (length and data type)

<table>
<thead>
<tr>
<th>3752 5712250 3125</th>
<th>3752 433906 3125</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original data</td>
<td>head</td>
</tr>
</tbody>
</table>

- Preserve a number of leading and trailing characters

<table>
<thead>
<tr>
<th>3752 5712250 3125</th>
<th>3752 X4mbAdLQ 3125</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original data</td>
<td>head</td>
</tr>
</tbody>
</table>

- Mask a portion of the token when a full value is not needed or desirable

<table>
<thead>
<tr>
<th>3752 5712250 3125</th>
<th>3752 ***** 3125</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original data</td>
<td>head</td>
</tr>
</tbody>
</table>

Tokens that maintain the length and format of the original data don’t require applications to be modified.
• Formatted tokens can be used wherever masked credit card information is required

**USING CREDIT CARD NUMBER**

3752 5712250 3125

**USING TOKEN**

3752 4333906 3125

Determines card type – standard, private label, gift card

Last 4 digits retain confirmation info

• Therefore systems are removed from PCI DSS scope wherever tokenized data suffices
Tokens are surrogates for masked data - SSN

• Formatted tokens can be used wherever masked personally identifiable information is required

SOCIAL SECURITY NUMBER

375-57-2125

USING TOKEN

433-39-2125

“What are the last 4 digits of your Social Security Number?”

• Therefore wherever tokenized data suffices, risk is reduced
1:1 Token / data relationship

- Same token value is consistent for same data across entire enterprise; maintains **referential integrity** across applications
- Data analysis can be performed using token – e.g. data warehouse

**Before using credit card number**

Transaction: 1
- CC#: 3752 5712250 3125
- Item: Paper
- Item: Stapler
- Item: Staples

Transaction: 2
- CC#: 3752 5712250 3125
- Item: Paper
- Item: Notebook
- Item: Staples

**After using token**

Transaction: 1
- CC#: 3716 4136820 3125
- Item: Paper
- Item: Stapler
- Item: Staples

Transaction: 2
- CC#: 3716 4136820 3125
- Item: Paper
- Item: Notebook
- Item: Staples
The Tokenization Process
Credit card information is input for tokenization.
Ciphertext and tokens are stored in the data vault.
Tokens are distributed to applications for storage.

3752 5712250 3125

Token Engine

3752 4333906 3125

Loss Prevention

Data Vault

&1y'13JhM)7N56^$90
Tokenization considerations for PII
• Some situations may determine you do NOT want a 1:1 token relationship for obfuscation purposes

• E.G. salary
  – $65,000 is always seen as $65,000 today
  – If token = 18903, then anywhere 18903 is found, it equates to $65,000
  – In tokenized world, multiple tokens could be mapped to $65,000

• Business drivers and requirements will drive business use
• Separate credit card data from personally identifiable information
• PCI requires separation for transfer, but also good practice
Geographical data segmentation

- Data may not be able to be centralized across geographical borders
- Separate data vaults based on privacy and transborder laws
Data masking: Test systems use ‘production tokens’

Production HR System
Germany

Outsourced Development
India
Determining your tokenization platform
What platform should host tokenization?

1. Corporate Applications
2. Token Server
3. Data Vault
4. Key Manager
PCI Compliance: Financial Services Company

A Customer Case Study
Company profile – financial services / insurance

- Revenue: > $53 Billion
- Employees: >65k
- Products: Life and dental insurance, financial services
- Level 1 merchant

One of largest insurance / financial companies in the world
Objectives and concerns

- **Objective: PCI Compliance**
  - Utilize encryption, tokenization and/or key management to reduce risk and gain compliance
  - Balance security with business continuity
  - Minimize impact to the business (and systems)

- **Focus on the key applications in various environments**
  - Credit card capture and processing systems
  - Data warehousing and loss prevention applications

- **Concerns**
  - Legacy systems
  - Performance key
  - Limited programming resources to deploy
  - Limited operational resources to maintain ongoing PCI compliance
Credit Card and Token Flow

- Token Manager
- Data Vault
- Key Manager
- Loss Prevention and Other Corporate Applications

3752 5712250 3125
Key benefits of the solution

• Implemented globally
  – In over 15 countries
  – Standardized data protection globally

• Little and/or no additional costs
  – No infrastructure additions
  – No additional personnel
  – Risk reduction - isolated credit cards to a smaller foot print within the organization

• Limited program modifications
  – Tokens fit into applications where credit cards were previously
  – Non-invasive - No changes to database fields

• Minimal impact overhead and latency

• Easy to maintain-centralized key management
Current work

- PCI-DSS scope reduction
  - Reviewing options for entire credit card information supply chain

- Protection of Personally Identifiable Information (PII)

- Constant evaluation of ‘Best Practices’ as it relates to new technology
Enhanced Security
- Central, protected data vault
- Shrink footprint where data is stored
- Prevent unauthorized access to sensitive data
- Centralized key management

Flexibility & Control
- Maintain length and format of original data
- 1:1 and 1:Many token / data relationships
- Use different token formats for different data types

Ease of Management
- Fewer systems to keep track of
- Fewer changes to applications
- Must log all administration and user-accessible functions – integrate with SIEM