Protecting Your Data in the Cloud

Ulf Mattsson
Chief Technology Officer
ulf.mattsson [at] protegrity.com
Ulf Mattsson

- 20 years with IBM Development & Global Services
- Inventor of 22 patents – Encryption and Tokenization
- Co-founder of Protegrity (Data Security)
- Research member of the International Federation for Information Processing (IFIP) WG 11.3 Data and Application Security
- Member of
  - Cloud Security Alliance (CSA)
  - PCI Security Standards Council (PCI SSC)
  - American National Standards Institute (ANSI) X9
  - Information Systems Security Association (ISSA)
  - Information Systems Audit and Control Association (ISACA)
Guidance from Cloud Security Alliance

Security Guidance for Critical Areas of Focus in Cloud Computing

Top Threats to Cloud Computing V1.0
Cloud Security Debate

Debate >> Encryption is better equipped than tokenization to secure data in the cloud.

October 01 2010

AGAINST

Ulf Mattsson
CTO, Protegrity

One of the biggest advantages of tokenization is a better security risk. Exposure of data is reduced, but tokenization makes it easier to replace sensitive data with anything with the same or similar properties. That said, analysts have taken this step by replacing sensitive data with anything with the same or similar properties.

Next Generation Tokenization for Compliance and Cloud Data Protection

By Ulf Mattsson – ISSA member, New York Metro, USA Chapter

This article will discuss how next-generation tokenization protects data as it flows across systems while minimizing PCI compliance costs.

Cloud Computing – Assessing Data Security Risks and Solutions

Ulf Mattsson
CTO
Protegrity

Ulf Mattsson, CTO, Protegrity: Making Sense of the Sony Breach

April 2011 by Ulf Mattsson, CTO, Protegrity Corporation
About Protegrity

- Proven enterprise data security software and innovation leader
  - Sole focus on the protection of data
  - Patented Technology, Continuing to Drive Innovation

- Growth driven by compliance and risk management
  - PCI (Payment Card Industry)
  - PII (Personally Identifiable Information)
  - PHI (Protected Health Information) – HIPAA
  - State and Foreign Privacy Laws, Breach Notification Laws
  - High Cost of Information Breach ($4.8m average cost), immeasurable costs of brand damage, loss of customers
  - Requirements to eliminate the threat of data breach and non-compliance

- Cross-industry applicability
  - Retail, Hospitality, Travel and Transportation
  - Financial Services, Insurance, Banking
  - Healthcare
  - Telecommunications, Media and Entertainment
  - Manufacturing and Government
Risks Associated with Cloud Computing

- Handing over sensitive data to a third party
- Threat of data breach or loss
- Weakening of corporate network security
- Uptime/business continuity
- Financial strength of the cloud computing provider
- Inability to customize applications

Source: The evolving role of IT managers and CIOs Findings from the 2010 IBM Global IT Risk Study
“It is fascinating that the top threat events in both 2010 and 2011 are the same and involve external agents hacking and installing malware to compromise the confidentiality and integrity of servers.”


Data Breaches – Mainly Online Data Records

- 900+ breaches
- 900+ million compromised records:

Source: 2010 Data Breach Investigations Report, Verizon Business RISK team and USSS
Compromised Data Types - # Records

- Payment card data
- Personal information
- Usernames, passwords
- Intellectual property
- Bank account data
- Medical records
- Classified information
- System information
- Sensitive organizational data

Source: Data Breach Investigations Report, Verizon Business RISK team and USSS
Industry Groups Represented - # Breaches

- Hospitality
- Retail
- Financial Services
- Government
- Tech Services
- Manufacturing
- Transportation
- Media
- Healthcare
- Business Services

Source: Data Breach Investigations Report, Verizon Business RISK team and USSS
Breach Discovery Methods - # Breaches

- Third party fraud detection
- Notified by law enforcement
- Reported by customer/partner
- Unusual system behavior
- Reported by employee
- Internal security audit or scan
- Internal fraud detection
- Brag or blackmail by perpetrator
- Third party monitoring service

Source: Data Breach Investigations Report, Verizon Business RISK team and USSS
Example of How the Problem is Occurring – PCI DSS

Encrypt Data on Public Networks (PCI DSS)

Public Network

Attacker

Private Network

Clear Text Data

Encrypt Data

At Rest (PCI DSS)

Application

Database

OS File System

Storage System

Clear Text Data

Source: PCI Security Standards Council, 2011
How can the problem be solved?
- Tokenization and other options can reduce the risk

Source: PCI Security Standards Council, 2011
Amazon Cloud & PCI DSS

- Just because AWS is certified doesn't mean you are
  - You still need to deploy a PCI compliant application/service and anything on AWS is still within your assessment scope

- PCI-DSS 2.0 doesn't address multi-tenancy concerns

- You can store PAN data on S3, but it still needs to be encrypted in accordance with PCI-DSS requirements
  - Amazon doesn't do this for you
  - You need to implement key management, rotation, logging, etc.

- If you deploy a server instance in EC2 it still needs to be assessed by your QSA (PCI auditor)
  - Organization's assessment scope isn't necessarily reduced

- Tokenization can reduce your handling of PAN data

Tokenization Use Case Example

❉ A leading retail chain
  • 1500 locations in the U.S. market

❉ Simplify PCI Compliance
  • 98% of Use Cases out of audit scope
  • Ease of install (had 18 PCI initiatives at one time)

❉ Tokenization solution was implemented in 2 weeks
  • Reduced PCI Audit from 7 months to 3 months
  • No 3rd Party code modifications
  • Proved to be the best performance option
  • 700,000 transactions per days
  • 50 million card holder data records
  • Conversion took 90 minutes (plan was 30 days)
  • Next step – tokenization server at 1500 locations
What is Tokenization and what is the Benefit?

Tokenization

- Tokenization is process that replaces sensitive data in systems with inert data called tokens which have no value to the thief.
- Tokens resemble the original data in data type and length

Benefit

- Greatly improved transparency to systems and processes that need to be protected

Result

- Reduced remediation
- Reduced need for key management
- Reduce the points of attacks
- Reduce the PCI DSS audit costs for retail scenarios
# Token Flexibility for Different Categories of Data

<table>
<thead>
<tr>
<th>Type of Data</th>
<th>Input</th>
<th>Token</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Token Properties</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credit Card</td>
<td>3872 3789 1620 3675</td>
<td>8278 2789 2990 2789</td>
<td>Numeric</td>
</tr>
<tr>
<td>Medical ID</td>
<td>29M2009ID</td>
<td>497HF390D</td>
<td>Alpha-Numeric</td>
</tr>
<tr>
<td>Date</td>
<td>10/30/1955</td>
<td>12/25/2034</td>
<td>Date</td>
</tr>
<tr>
<td>E-mail Address</td>
<td><a href="mailto:bob.hope@protegrity.com">bob.hope@protegrity.com</a></td>
<td><a href="mailto:emp0.snaugs@svtiensnii.snk">emp0.snaugs@svtiensnii.snk</a></td>
<td>Alpha Numeric, delimiters in input preserved</td>
</tr>
<tr>
<td>SSN delimiters</td>
<td>075-67-2278</td>
<td>287-38-2567</td>
<td>Numeric, delimiters in input</td>
</tr>
<tr>
<td>Credit Card</td>
<td>3872 3789 1620 3675</td>
<td>8278 2789 2990 <strong>3675</strong></td>
<td>Numeric, Last 4 digits exposed</td>
</tr>
<tr>
<td><strong>Policy Masking</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credit Card</td>
<td>3872 3789 1620 3675</td>
<td>clear, encrypted, tokenized at rest **3872 37## **** ******</td>
<td>Presentation Mask: Expose 1st 6 digits</td>
</tr>
</tbody>
</table>

*Note: The highlighted numbers (3872 and 3675) are example values for demonstration purposes.*
Data Tokenization – Reducing the Attack Surface

Unprotected sensitive information:

Protected sensitive information

Tokenization

Server

User

Tokenization

Server

Application

Databases

Data Token
PCI DSS - Ways to Render the PAN Unreadable

- Two-way cryptography with associated key management processes
- One-way cryptographic hash functions
- Index tokens and pads
- Truncation (or masking – xxxxxx xxxxxx 6781)
## Positioning Different Protection Options

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Strong Encryption</th>
<th>Formatted Encryption</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security &amp; Compliance</td>
<td>●</td>
<td>○</td>
<td>●</td>
</tr>
<tr>
<td>Total Cost of Ownership</td>
<td>○</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Use of Encoded Data</td>
<td>○</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

Best: ● ● ● ● ○  Worst: ○ ○ ○ ○ ○
Different Approaches for Tokenization

Traditional Tokenization
- Dynamic Model or Pre-Generated Model
- 5 tokens per second - 5000 tokenizations per second

Protegrity Next Generation Tokenization
- Memory-tokenization
- 200,000 - 9,000,000+ tokenizations per second
- “The tokenization scheme offers excellent security, since it is based on fully randomized tables.” *
- “This is a fully distributed tokenization approach with no need for synchronization and there is no risk for collisions.” *

*: Prof. Dr. Ir. Bart Preneel, Katholieke University Leuven, Belgium
## Tokenization Summary

<table>
<thead>
<tr>
<th></th>
<th>Traditional Tokenization</th>
<th>Protegrity Tokenization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Footprint</strong></td>
<td>Large, Expanding.</td>
<td>Small, Static.</td>
</tr>
<tr>
<td></td>
<td>The large and expanding footprint of Traditional Tokenization is it’s Achilles heel.</td>
<td>The small static footprint is the enabling factor that delivers extreme performance, scalability, and expanded use.</td>
</tr>
<tr>
<td></td>
<td>It is the source of poor performance, scalability, and limitations on its expanded use.</td>
<td></td>
</tr>
<tr>
<td><strong>High Availability,</strong></td>
<td><strong>Complex replication required.</strong></td>
<td><strong>No replication required.</strong></td>
</tr>
<tr>
<td></td>
<td>Deploying more than one token server for the purpose of high availability or scalability will require complex and expensive replication or synchronization between the servers.</td>
<td>Any number of token servers can be deployed without the need for replication or synchronization between the servers. This delivers a simple, elegant, yet powerful solution.</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td><strong>Prone to collisions.</strong></td>
<td><strong>No collisions.</strong></td>
</tr>
<tr>
<td></td>
<td>The synchronization and replication required to support many deployed token servers is prone to collisions, a characteristic that severely limits the usability of traditional tokenization.</td>
<td>Protegrity Tokenizations’ lack of need for replication or synchronization eliminates the potential for collisions.</td>
</tr>
<tr>
<td><strong>Performance,</strong></td>
<td><strong>Will adversely impact performance &amp; scalability.</strong></td>
<td><strong>Little or no latency. Fastest industry tokenization.</strong></td>
</tr>
<tr>
<td>Latency, and Scalability</td>
<td>The large footprint severely limits the ability to place the token server close to the data. The distance between the data and the token server creates latency that adversely effects performance and scalability to the extent that some use cases are not possible.</td>
<td>The small footprint enables the token server to be placed close to the data to reduce latency. When placed in-memory, it eliminates latency and delivers the fastest tokenization in the industry.</td>
</tr>
<tr>
<td><strong>Extendibility</strong></td>
<td><strong>Practically impossible.</strong></td>
<td><strong>Unlimited Tokenization Capability.</strong></td>
</tr>
<tr>
<td></td>
<td>Based on all the issues inherent in Traditional Tokenization of a single data category, tokenizing more data categories may be impractical.</td>
<td>Protegrity Tokenization can be used to tokenize many data categories with minimal or no impact on footprint or performance.</td>
</tr>
</tbody>
</table>
## Evaluating Encryption & Tokenization Approaches

<table>
<thead>
<tr>
<th>Area</th>
<th>Impact</th>
<th>Encryption</th>
<th>Tokenization</th>
<th>Memory Tokenization (Protegrity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalability</td>
<td>Availability</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td></td>
<td>Latency</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td></td>
<td>CPU Consumption</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td>Security</td>
<td>Data Flow Protection</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td></td>
<td>Compliance Scoping</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td></td>
<td>Key Management</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td></td>
<td>Data Collisions</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
<tr>
<td></td>
<td>Separation of Duties</td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
<td><img src="image" alt="Circle" /> <img src="image" alt="Circle" /> <img src="image" alt="Circle" /></td>
</tr>
</tbody>
</table>

**Best** 🍊 🍊 🍊 🍊 🍊 **Worst** 🍊 🍊 🍊 🍊 🍊
Data Protection Challenges

- The actual protection of the data is not the challenge
- Centralized solutions are needed to manage complex security requirements
  - Based on Security Policies with Transparent Key management
  - Many methods to secure the data
  - Auditing, Monitoring and Reporting
- Solutions that minimize the impact on business operations
  - Highest level of performance and transparency
- Rapid Deployment
- Affordable with low TCO
- Enable & Maintaining compliance
Protegrity Data Security Management

- File System Protector
- Application Protector
- Tokenization Server
- Database Protector
- Secure Archive
- Enterprise Data Security Administrator
- Policy
- Audit Log

: Encryption service
Protegrity Data Protection Platform

**Coverage**
- Supports heterogeneous environments across operating systems, applications, file systems and databases

**Protection Methods**
- Encryption, Tokenization, DTP2 (data type/format preserving encryption), Masking and Monitoring

**Packaged Data Protectors for Databases, Files and Applications**

**Next Generation Tokenization**

**Separation of Duties/ Roles**
- IT Security resource responsible for defining policy is different than Administrator’s managing the sensitive data

**Policy and Key Management**
- Central and secure solution to manage the keys to the data across the enterprise

**Central Reporting**
- Reporting for security management compliance assessment
Please contact me for more information

ulf.mattsson [at] protegrity.com

Visit our table