Oracle Key Vault
Data Subsetting and Masking

Paul White
Database Security Specialist
Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Oracle Database Security Solutions

Defense-in-Depth for Maximum Security

PREVENTIVE
- Encryption & Redaction
- Data Masking
- Privileged User Controls

DETECTIVE
- Activity Monitoring
- Database Firewall
- Auditing and Reporting

ADMINISTRATIVE
- Privilege Analysis
- Sensitive Data Discovery
- Encryption Keys and Certificates

Copyright © 2015 Oracle and/or its affiliates. All rights reserved.
Managing Master Keys in Oracle Wallet

- **CRITICAL**: Remember wallet password
- **CRITICAL**: Do not delete wallet. Retain copy of password-based wallet even if using auto-login
- **CRITICAL**: Do not have multiple databases share same wallet
- Set strong wallet password using numbers, capitalization, length >= 12 characters...
- Rotate master encryption key and wallet password approximately every six months
- Backup wallet before and after each rotation operation
- Keep wallet backup separate from encrypted data backup
- Restrict wallet directory and file permissions
- Keep wallet read-only for daily use, set immutable bit where available
- For RAC, consider storing wallet in ACFS (DB 11gR2) or ASM (DB 12cR1)
- For DB 12cR1, separate duties using SYSKM
Management Challenges: Proliferation

Primary Data Center

Failover Data Center
The Challenges of Key Management

**Management**
- Proliferation of encryption wallets and keys
- Authorized sharing of keys
- Key availability, retention, and recovery
- Custody of keys and key storage files

**Regulations**
- Physical separation of keys from encrypted data
- Periodic key rotations
- Monitoring and auditing of keys
- Long-term retention of keys and encrypted data
Key Management with Oracle Key Vault

- Centrally manage and share keys, secrets, Oracle wallets, Java keystores, and more
- Optimized for Oracle stack (Database, Middleware, Systems) and Advanced Security TDE
- Robust, secure, and standards compliant (OASIS KMIP) key manager
Oracle Key Vault Software Appliance Platform

• Full-stack solution based on hardened configuration
  – Easy to install, configure, deploy, and patch
  – Open x86-64 hardware to choose from

• Includes Oracle Database security options
  – Transparent Data Encryption, Database Vault, Virtual Private Database

• Separation of duties for administrative users

• Full auditing and alerts

• Preconfigured reports
Oracle Key Vault High-Level Architecture

Databases

Middleware

Servers

Standby

Administration Console, Alerts, Reports

Secure Backups

Oracle Wallet
Certificate
Java Keystore
Server Password
Credential File
Oracle Advanced Security Transparent Data Encryption (TDE)

Online Master Key Scenarios

- Single Instance
- Multiple DBs Same Machine
- RAC
- Data Guard
- GoldenGate
Provisioning Endpoints

1. One-time enrollment token
2. Endpoint package
3. Endpoint installation and configuration
4. Results: Endpoint certificate, binaries and configuration file
5. Sharing
Management Reports – Endpoint Activity
## User Activity

### All User Activity

<table>
<thead>
<tr>
<th>Time</th>
<th>Subject</th>
<th>Operation</th>
<th>Object</th>
<th>Object Type</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>23-OCT-15 14:48:14</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>APP_SERVER_2</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 14:39:56</td>
<td>OKVADMIN</td>
<td>Modify Wallet</td>
<td>FinanceWallet</td>
<td>WALLET</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:56:04</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>DC2_FINANCE_RAC_NODE_3</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:57:53</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>DC2_FINANCE_RAC_NODE_2</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:57:43</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>DC2_FINANCE_RAC_NODE_1</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:57:27</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>DC1_FINANCE_RAC_NODE_3</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:57:14</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>DC1_FINANCE_RAC_NODE_1</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:57:01</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>APP_SERVER_2</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:55:51</td>
<td>OKVADMIN</td>
<td>Modify Endpoint Attributes</td>
<td>APP_SERVER_1</td>
<td>ENDPOINT</td>
<td>✓</td>
</tr>
<tr>
<td>23-OCT-15 13:56:10</td>
<td>OKVADMIN</td>
<td>Add Endpoint Group</td>
<td>HR_DB</td>
<td>ENDPOINT GROUP</td>
<td>✓</td>
</tr>
</tbody>
</table>

### Specific User Group Activity

**Select User Group**

- No User Group Activity

### Specific User Activity

**Select User** OKVADMIN

<table>
<thead>
<tr>
<th>Operation</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add endpoint</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Modify Endpoint Attributes</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Add User</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Create Wallet</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Add Endpoint Group</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Modify User Attributes</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Modify Wallet</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Control Keys with **On-Prem Key Vault**

- Security and availability of keys
- On-premise key access control by customers
- Isolation of keys per tenant
- Maintain control and visibility

Customer Premise

Keys, Wallets, Credential Files

Customer DBCS
Database Security
Data Subsetting and Masking Pack Technical Overview
ORACLE DATABASE SECURITY
MAXIMUM SECURITY FOR CRITICAL DATA INFRASTRUCTURE

PREVENTION
- Encryption & Redaction
- Subsetting and Data Masking
- Privileged User Controls

DETECTION
- Activity Monitoring
- Database Firewall
- Auditing and Reporting

ADMINISTRATION
- Privilege Analysis
- Sensitive Data Discovery
- Encryption Keys and Certificates
# Need to Mask and Subset Data

## Use Cases

### Limit Sensitive Data Proliferation
- Test, Dev, QA, Training, Research, Outsourced, Cloud, and more

### Compliance
- PCI-DSS, HIPAA, European Data Protection, Canada PIPEDA, and more

### Save Storage Costs
- Non-Prod such as Test/Dev, Mega Data warehouses, and more

### Share What is Necessary
- With Subscribers, auditors, courts, partners, testers, developers, and more

### Right to be Forgotten/Erasure (New GDPR in Europe)

## Challenges

### How to Locate Sensitive Data?
- In the midst of numerous applications, databases, and environments

### How to Accurately Protect Sensitive Data?
- Data has different shapes and forms: VISA, AMEX, Discoverer, Master, SSN, and more

### Is the Protected Data Usable?
- To developers, testers, applications, and more

### We Do Not Have Resources?
- To develop and maintain such solution in this ever-changing IT landscape

### Will the Applications Continue to Work?
- and More .....
Oracle Data Masking and Subsetting Pack

Reduces Risk in Sharing by Obfuscating or Removing Sensitive Data

Discover Sensitive Data
Modeling Application Data
Mask Data using Format Library
Subset Based on Goal/Condition
Mask/Subset in Export or on Staging
Mask in Workload Captures & Clones
Pre-installed in Enterprise Manager
Data Masking and Subsetting Methodology

Create Data Model
• Discover sensitive data
• Discover data relationships

Select Formats & Criteria
• Masking formats and templates
• Goals & conditions for subsetting

Preview & Validate
• Preview masking algorithm results
• Preview subset reduction results

Execute Transformation
• In-Database or In-Export
• Rerun with same data model
Application Data Modeling

Sensitive Data Discovery

[Diagram showing data relationships, sensitive columns, automated discovery, and target database]
Extensive Masking Format Library

• Provides common masking formats
• Supports custom masking formats
  – Random numbers/strings/dates
  – Substitute
  – User defined PL/SQL function
  … and more
• Generates sample masked values
• Templates for specific versions of E-Business Suite and Fusion Applications
## Comprehensive Masking Transformations

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditional masking</td>
<td>Masks rows differently based on condition</td>
<td><em>E.g. Mask national identifiers based on country</em></td>
</tr>
<tr>
<td>Consistent masking</td>
<td>Ensures the masked values is same across multiple databases</td>
<td><em>E.g. Mask employee identifiers consistently across schemas and databases</em></td>
</tr>
<tr>
<td>Compound masking</td>
<td>Reduces masking time by grouping related columns</td>
<td><em>E.g. Mask related columns: address (street, state, country), or dates (birth, joining)</em></td>
</tr>
<tr>
<td>Format preserving</td>
<td>Generates random values that maintain the original format</td>
<td><em>E.g. Mask passport numbers or tax identifiers while keeping the format</em></td>
</tr>
<tr>
<td>Perturbation</td>
<td>Generates random values within a user defined range</td>
<td><em>E.g. Generate random dates within a specific range</em></td>
</tr>
<tr>
<td>Shuffling</td>
<td>Shuffles the values within a column</td>
<td><em>E.g. Shuffle clinical data or PII between electronic health records</em></td>
</tr>
<tr>
<td>Reversible masking</td>
<td>Encrypts and decrypts data based on a passphrase</td>
<td><em>E.g. Outsourced data processing team can unmask data</em></td>
</tr>
</tbody>
</table>
Masking Examples

- **Mask Based on Condition**
  - Country | Identifier
  - CA | 226-956-324
  - US | 610-02-9191
  - UK | JX 75 67 44 C

- **Shuffle Records**
  - Health Records

- **Generate Deterministic Output**
  - HR
  - Emp ID | First Name
  - 324 | Albert
  - 986 | Murali
  - FIN
  - Emp ID | First Name
  - 324 | Charlie
  - 986 | Murali

- **Generate Random Values Preserving Format**
  - Company | Closing Price
  - IBFG | $36.92
  - XKJU | ¥789.8

- **Mask Operating System Files stored as Blobs**
  - BLOB
  - 3178973456
  - 6509876745
  - Search : [0-9](10)
  - Replace : *
  - BLOB
  - **********
  - **********

-and more...
Data Subsetting Use Cases

• Share relevant subset of data with internal and external teams
• Reduce storage cost for test/dev
• Extract subscriber data from SaaS
• Perform research and analysis on a subset of data
• Extract subset of data as part of e-discovery requests
Goal or Condition Based Subsetting

**Relative Database Size**
- 100%: 1024 GB
- 25%: 256 GB
- 10%: 102 GB

**Relative Table Size**
- 100M Rows
- 20M Rows
- 2M Rows

**Condition Based**
- Extract ASIA Sales
- ASIA
### Preview and Validate Subset Results

Impact of subset rules on tables are displayed below. The values shown here are based on estimates and may not be accurate.

<table>
<thead>
<tr>
<th>Name</th>
<th>Table Rule</th>
<th>Source Size</th>
<th>Estimated Subset Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MB</td>
<td>Rows</td>
</tr>
<tr>
<td>Applications</td>
<td></td>
<td>0.066</td>
<td>648</td>
</tr>
<tr>
<td>OE(OE)</td>
<td></td>
<td>0.056</td>
<td>433</td>
</tr>
<tr>
<td>HR(HR)</td>
<td></td>
<td>0.01</td>
<td>215</td>
</tr>
<tr>
<td>EMPLOYEE</td>
<td>Some Rows</td>
<td>0.007</td>
<td>107</td>
</tr>
<tr>
<td>LOCATION</td>
<td>Some Rows</td>
<td>0.001</td>
<td>23</td>
</tr>
<tr>
<td>JOBS</td>
<td>Some Rows</td>
<td>0.0006</td>
<td>19</td>
</tr>
<tr>
<td>DEPARTME</td>
<td>Some Rows</td>
<td>0.0005</td>
<td>27</td>
</tr>
<tr>
<td>COUNTRY</td>
<td>Some Rows</td>
<td>0.0004</td>
<td>25</td>
</tr>
<tr>
<td>JOB_HISTC</td>
<td>Some Rows</td>
<td>0.0003</td>
<td>10</td>
</tr>
<tr>
<td>REGIONS</td>
<td>Some Rows</td>
<td>0.0001</td>
<td>4</td>
</tr>
</tbody>
</table>
Deployment Options

**In-Database**
- Production
- Staging
- Test/Dev

Minimal impact on the production environment

**In-Export**
- Production
- Export
- Test/Dev

Sensitive data remains within the production perimeter
Hardware and Software

Engineered to Work Together