TEST DATA MANAGEMENT
CONCEPTS & METHODS
Test Management
Key requirements

- Automation: Define once, execute regularly
- Create different environments that do not interfere
- Easily add more environments

- Goal: Better QA with reasonable costs
Limitations of Tools for Test Data Management

- Test cases
- Application logic
- What needs to be tested?
- How often a copy is required?
Who needs what?
Who needs what?

Amount of Data
Who needs what?

- Amount of Data
- Degree of Modification

PRODUCTION  ACCEPT  REGRESSION  INTEGRATION  COMPONENT  FUNCTION  DEVELOPMENT
Row Level Processing
Unit Tests

Table Copying
Component / Integration Tests

Database / Subsystem Cloning
Release / System Tests / Preproduction
UBS HAINER  
TEST DATA MANAGEMENT

SUBSYSTEM CLONING | TABLE COPYING | ROW LEVEL PROCESSING

Database / Subsystem Cloning  
Release / System Tests / Preproduction

Table Copying  
Component / Integration Tests

Row Level Processing  
Unit Tests
Subsystem/Database Cloning
• Create a pre-production environment
• Data source for further copy operations
• Strict separation between prod and test
• Acceptance tests
• RDBMS upgrades
Preproduction Clone for DB2, Oracle, SQL Server, etc.
How does a clone become consistent?
How does a clone become consistent?

FlashCopy
SnapShot
EMC TimeFinder
IBM PPRC
HDS ShadowImage
Softek TDMF
Fujitsu Equivalent Copy

Time
Transactions
Prepare & (Re)Start

Rebuild current status

Time
Cloning

- Well defined set of scripts / jobs that read production
- Executed periodically by scheduler
- Strict decoupling of production from preproduction
- Rework after cloning is required
Characteristic of Cloning

• Need access to physical hard drives
• No need to stop or start databases
• Hardware assisted copy tools are very beneficial
• DBMS tools can also be used alternatively
UBS HAINER
TEST DATA MANAGEMENT

SUBSYSTEM CLONING | TABLE COPYING | ROW LEVEL PROCESSING

- **Table Copying**
  - Component / Integration Tests

- **Database / Subsystem Cloning**
  - Release / System Tests / Preproduction

- **Row Level Processing**
  - Unit Tests
Table Level Copying
Feeding the other Environments

Production → Preproduction → Test Environments

Clone → Copy
Wish list for file system level copies:

- Specify one schema, copy 1000 tables

1 Database
100 Tables
1000 Indexes
Characteristic Table Copying

- Replace/copy/rename selected objects
- Need authorization to move TS / to use utilities
- Hardware assisted copy tools not applicable
- But optimized for copy of mass data
- Masking/Anonymization feasible
Database / Subsystem Cloning
Release / System Tests / Preproduction

Table Copying
Component / Integration Tests

Row Level Processing
Unit Tests
Row Level Processing
UBS HAINER
TEST DATA MANAGEMENT
SUBSYSTEM CLONING | TABLE COPYING | ROW LEVEL PROCESSING

Production
DB2, Oracle, etc.  Clone  Pre-Production
DB2, Oracle, etc.  Copies  Test Environments

Inserts

Unit Testing
Function Testing
Provide data for the next application version

Version 1.0:

<table>
<thead>
<tr>
<th>ID</th>
<th>INTEGER</th>
<th>NOT NULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRSTNAME</td>
<td>VARCHAR(256)</td>
<td>NOT NULL</td>
</tr>
<tr>
<td>LASTNAME</td>
<td>VARCHAR(256)</td>
<td>NOT NULL</td>
</tr>
</tbody>
</table>

Version 2.0:

<table>
<thead>
<tr>
<th>ID</th>
<th>INTEGER</th>
<th>NOT NULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRSTNAME</td>
<td>VARCHAR(256)</td>
<td>NOT NULL</td>
</tr>
<tr>
<td>LASTNAME</td>
<td>VARCHAR(256)</td>
<td>NOT NULL</td>
</tr>
<tr>
<td>DATE_OF_BIRTH</td>
<td>DATE</td>
<td>NOT NULL</td>
</tr>
</tbody>
</table>
Automatic RI check

- All RI dependencies should be brought to test/development
Insert Options

- Completely replace target
- Append to target tables
- Merge based on primary key
Characteristic of Row Level Processing

- RI identification and transfer of relevant data to test system
- Flexible data customization
- Automates refresh of test case data
- Not suitable for mass data
- Developer tool
Manage Test Data Versions
Make Tests Reproducible
The Dilemma

- Tests should run under same preconditions
- Tests modify the data
VERSIONING OF TEST DATA

Backup before test execution:

Test Env.1 ➔ FROZEN TEST DATA
VERSIONING OF TEST DATA

- **Backup Before Test Execution**
  - Test Env.1
  - Frozen Test Data
  - Test Env.2

- **On Demand**
Characteristic of Versioning Management

- Availability of data at the push of a button
- Separate test/development from backup procedures
- Reduce error rates when doing acceptance tests
- Fast provisioning of test data instead of finding modified rows
Anonymization/Masking
What is the benefit of Anonymization?

• Compliance with legislation
• Preservation of reputation
• Avoid to provide competition with critical data
Judicial directives

European Privacy Directive 95/46/EC
European Union

PIPEDA, Privacy ACT
Canada

HIPAA, PCI DSS
USA

IFAI
Mexico
What must be anonymized and where?
Anonymization/Masking in the data landscape

- Strict separation
- Enforce anonymization
Anonymization/Masking in the data landscape

- Strict separation
- Enforce anonymization
Existing methods
Unload – Modification - Load

While copying
- Fast
- Automated
- Higher safety standard
Rewarding of Anonymization/Masking?

- Keep the danger potential low
- Match compliance rules: yes, we anonymize!
- Improvement of process with every new release
Characteristic of Anonymization/Masking

- Anonymization/Masking cannot be right or wrong
- Every application is different, no silver bullet existing
- Risk / Effort / Costs / Effect must be considered realistically
- No definite rules what exactly has to be anonymized exist
Thank you for your attention!

Questions?

Solutions

- TDM & Fast Data Refresh
  - CLONE
  - BCV5
  - Icebox
  - XDM

- DB2® Performance
  - BPA
  - TUC
  - ULT
  - XM

- Cost Savings, Performance
  - LPAR
  - MSU
  - zDynaCap

Contact us to learn more at info@ESAIGroup.com or 866-464-3724
www.ESAIGroup.com  www.ubs-hainer.com