Essentials for Test Data Management for Siebel

Steve Tallant
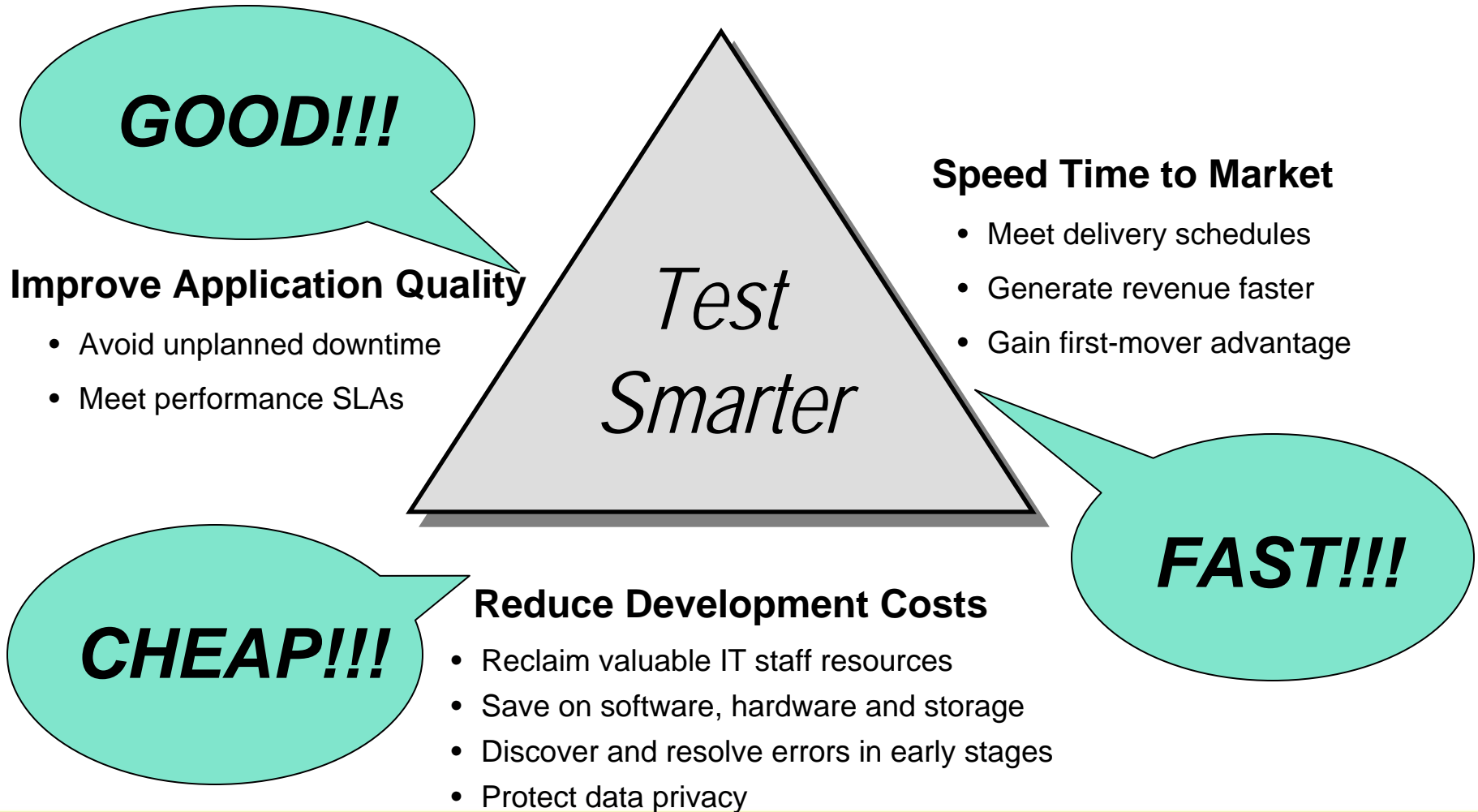
Product Manager

IBM Software Group

What is Test Data Management (TDM)?

TDM refers to the need to manage data used in testing and other non-production environments

- Extract related subsets of production data that are targeted to functionality under test

- Edit data to create error and boundary conditions

- De-identify (mask) test data to protect privacy

- Compare "before" and "after" images of test data

Benefit: Improving application quality & customer satisfaction

A Picture to Explain

GOOD!!!

Speed Time to Market

- Meet delivery schedules
- Generate revenue faster
- Gain first-mover advantage

Improve Application Quality

- Avoid unplanned downtime
- Meet performance SLAs

Test Smarter

FAST!!!

CHEAP!!!

Reduce Development Costs

- Reclaim valuable IT staff resources
- Save on software, hardware and storage
- Discover and resolve errors in early stages
- Protect data privacy

Consequences of Choices

- Infrastructure Costs – higher HW storage costs
  - cloning databases requires more storage

- Development Labor - higher costs
  - greater data volume equates to longer testing cycles

- Defects – can be expensive
  - Costs to resolve defects in the production environment can be 10-100 times greater than those caught in the development environment

Test Data Management Projects

- Characteristics for Test Data Management Projects
  - Subset capabilities to create realistic and manageable test databases
  - Quickly refresh test environments
  - Edit data to create targeted test cases
  - Compare 'before' and 'after' images of the test data
  - Improve test coverage and quality
  - Speed application deployment
  - De-identify (mask) data to protect privacy

Siebel – Test Data Management Challenges

- **Siebel doesn't provide a solution or methodology for TDM**

- **Siebel has a very complex data model consisting of many tables with multiple relationships between tables**

- **Siebel Industry Applications share a common repository**
  - **Each application doesn't use all tables and relationships that are found in Siebel tools**

Benefits of Effective Siebel Application Testing

What if…

- Your business can deploy new/improved Siebel applications faster without sacrificing quality?
    - increase revenue generation

- Your business can benefit from using IT resources more effectively?
    - reduce costs

- Your company can implement a reliable Siebel upgrade?
    - ensure positive customer experience

TDM Benefits to Stakeholders

CIO

VP, Line of Business

Director, IT

- Speed-time-to-market without sacrificing quality.

- Ensure consistent testing methodologies and reduce costs.

- Minimize threat of data breach.

- Ensure a reliable, positive customer experience.

- Sustain or react to competitive situations quickly.
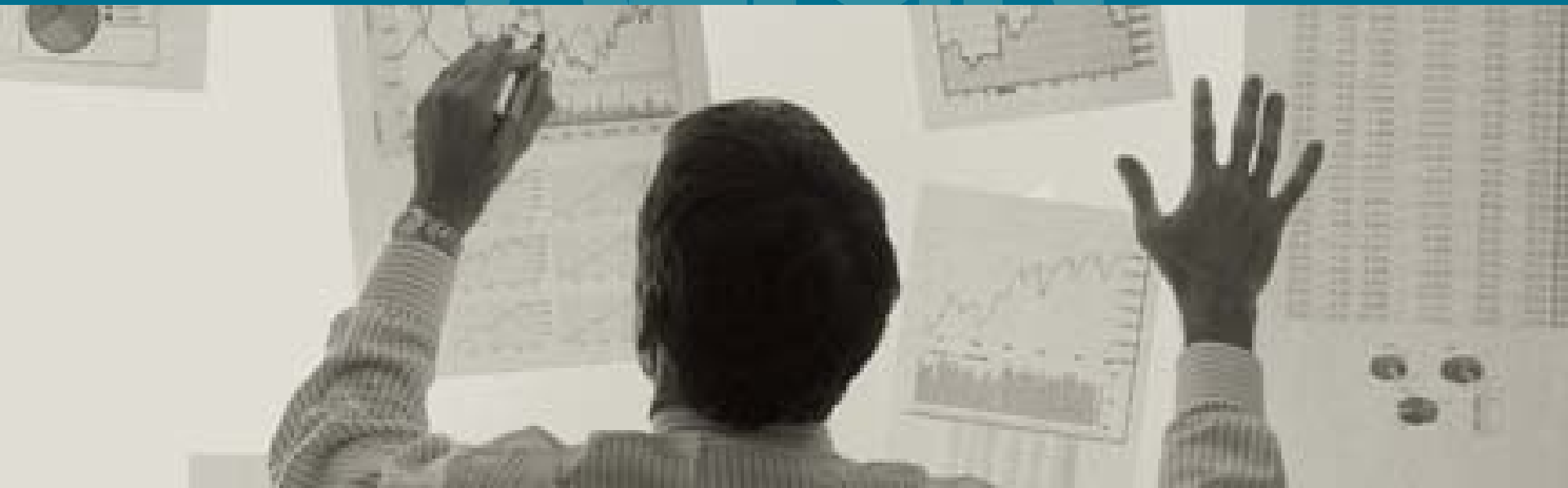
- Provide customers with sense of security.

- Populate realistic test data to improve testing and quality.

- Streamline testing processes for optimal environment.

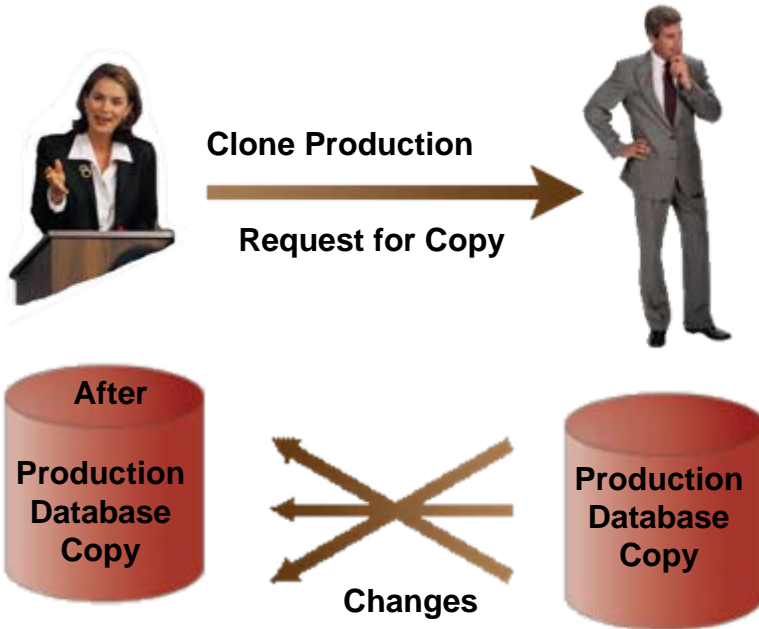- consistent methodology for privatization of data.

Basic Principles for Testing Siebel Data

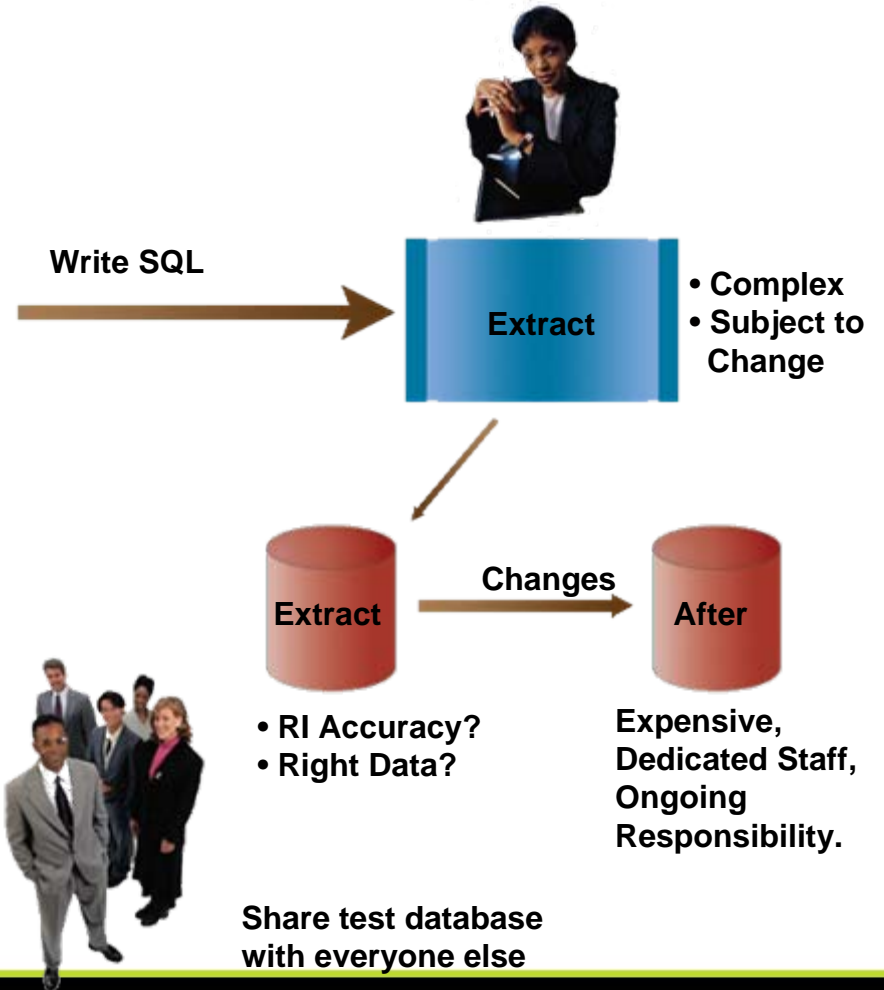# Current Practice

#1 - Clone Production

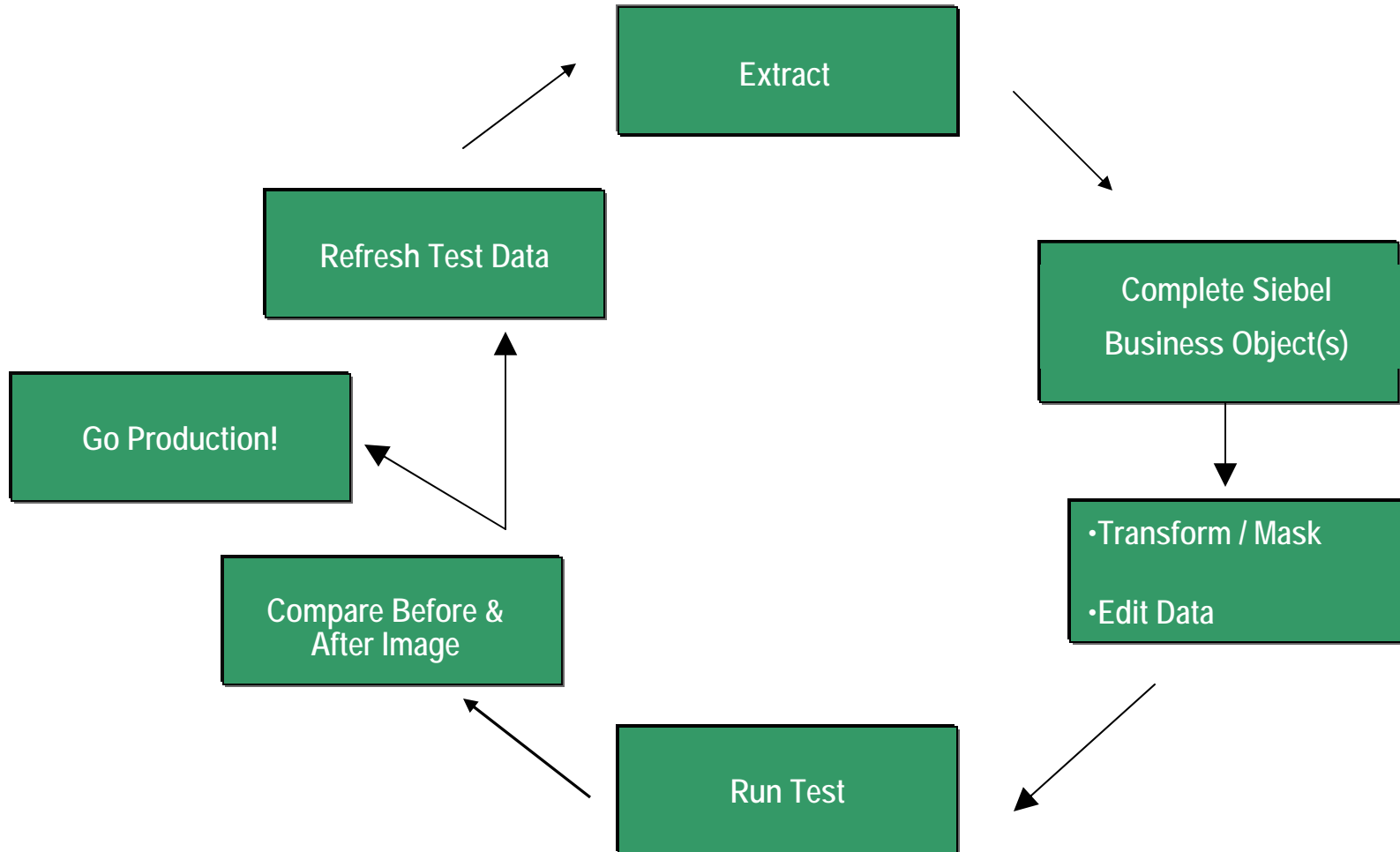***Repeat ?*%$!***

Clone Production

Request for Copy

After

Production Database Copy

Production Database Copy

Changes

Manual examination:
Right data?
What Changed?
Correct results?
Unintended Result?
Someone else modify?

#2 - Write SQL

Write SQL

Extract

- **Complex**
- **Subject to Change**

Changes

Extract

After

- **RI Accuracy?**
- **Right Data?**

Expensive, Dedicated Staff, Ongoing Responsibility.
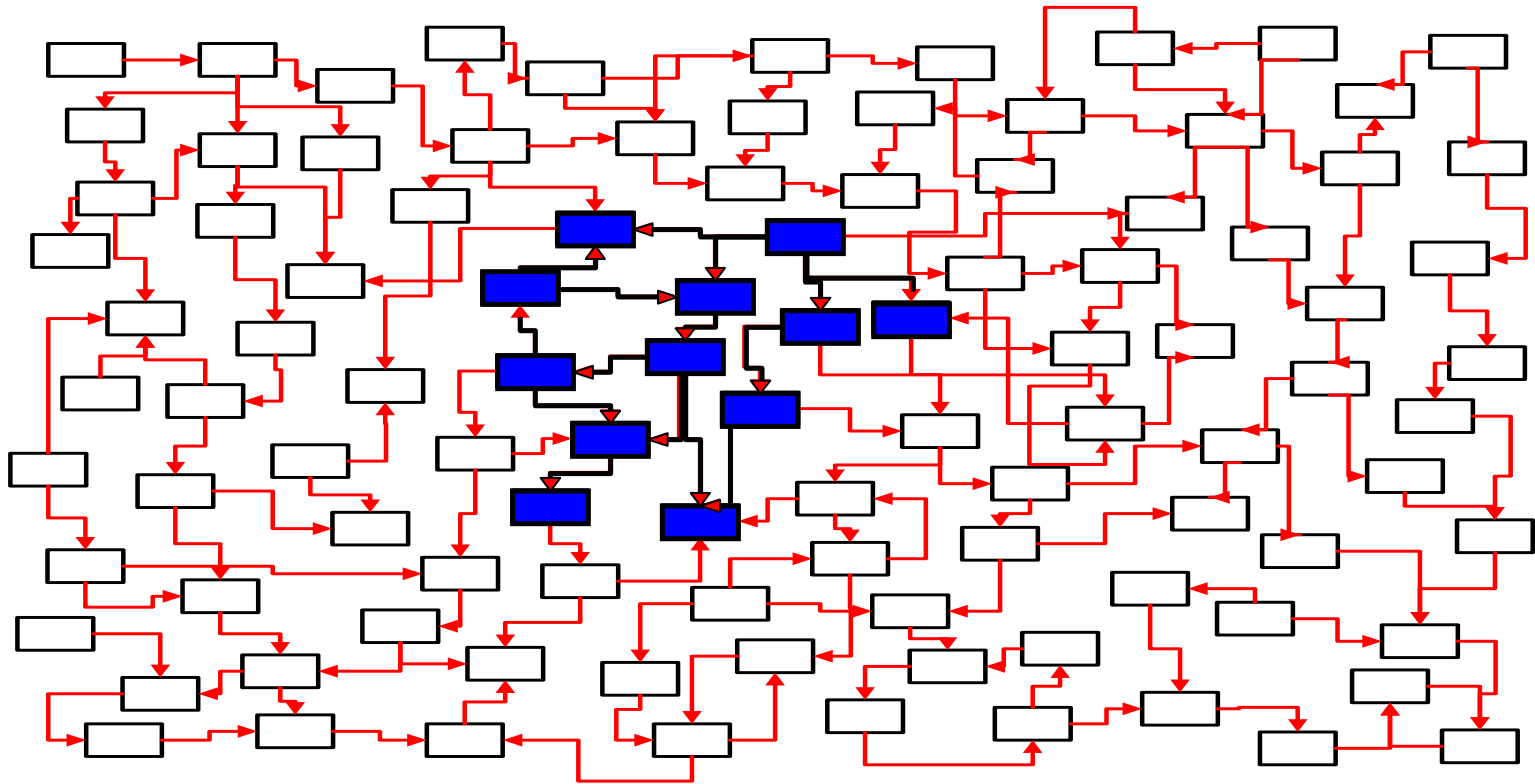
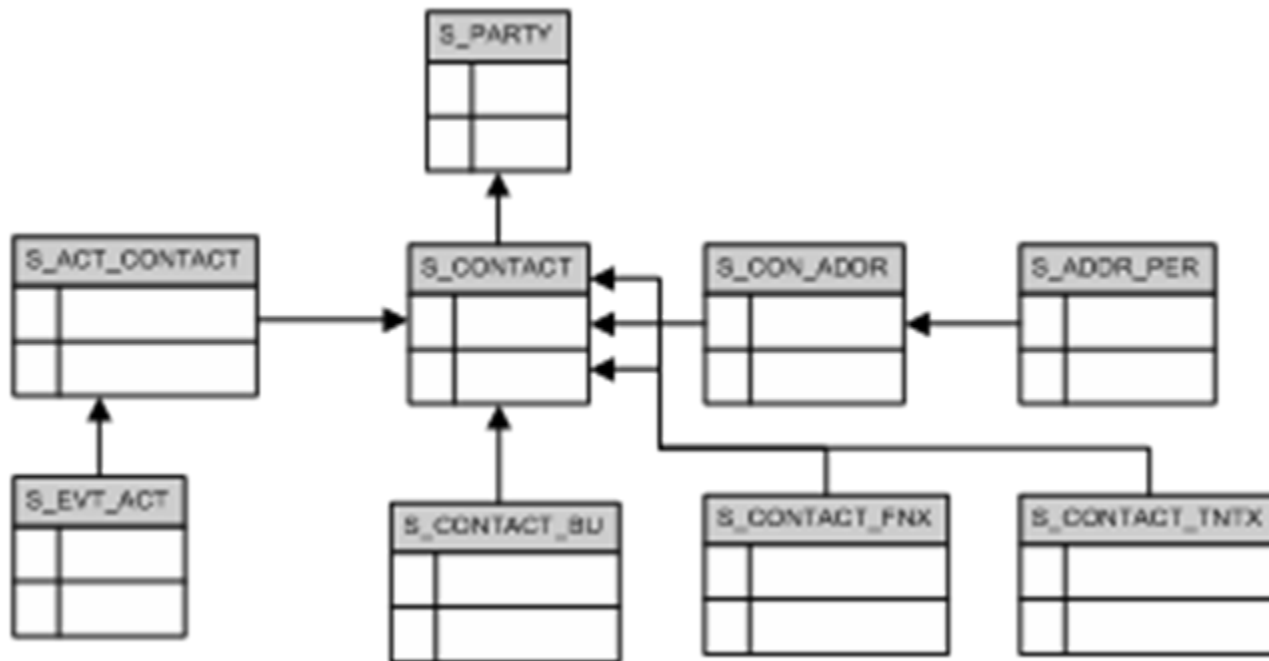**Share test database with everyone else**

Test Data Management – Building Blocks

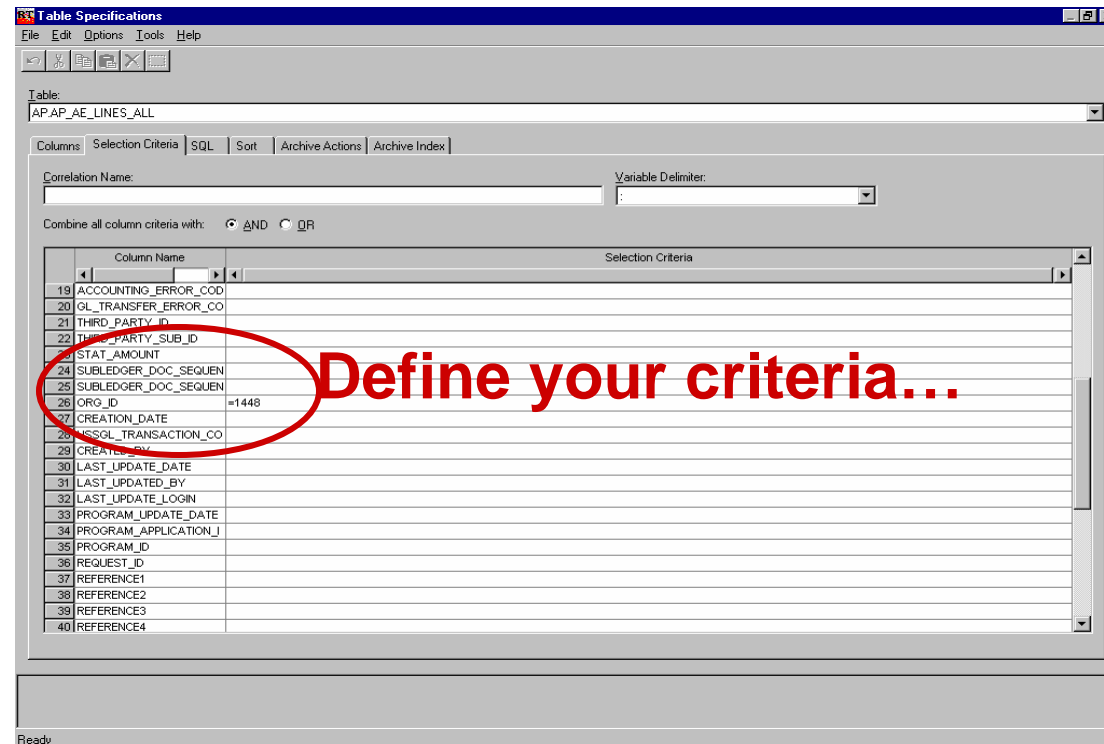# Ensure Referential Integrity in Subset

Ensure Referential Integrity in Subset

Creating a Subset with Provided Template

- Criteria can be based on one or more modules
- All Date Values
  - Create Date
  - Transaction Date
  - Effective Date
- Organizations
- Status
- Order number(s)
- "And/Or" combinations
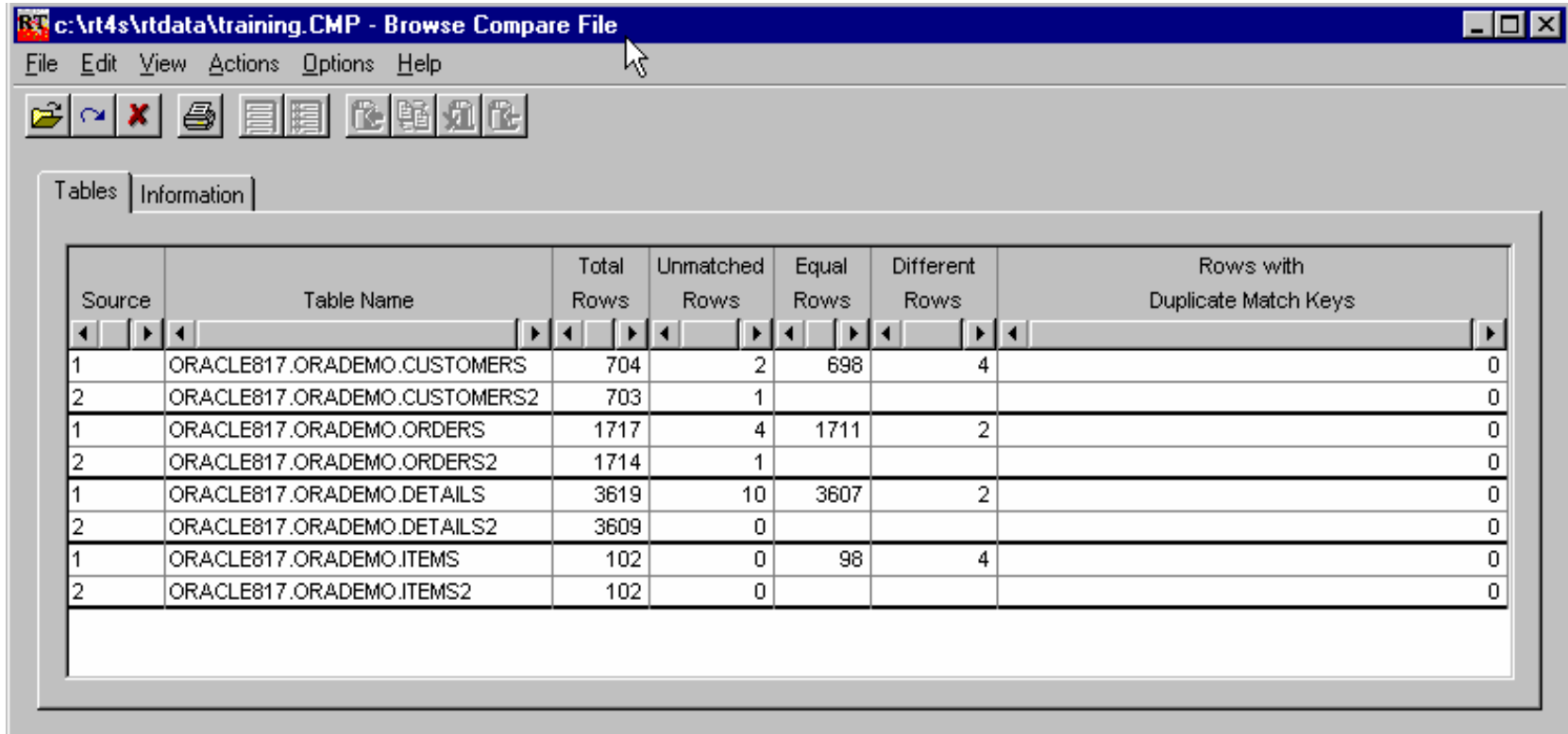

Define your criteria…

Comparing Data

- Compare the "before" and "after" data from an application test
- Compare results after running modified application during regression testing
- Identify differences between separate databases
- Audit changes to a database
- <u>Compare analyzes complete sets data</u> – finding changes in rows in tables
    – Single-table or multi-table compare
    – Creates compare file of results
    – Displays results on screen

Browsing the Compare File

Source	Table Name	Total Rows	Unmatched Rows	Equal Rows	Different Rows	Rows with Duplicate Match Keys
1	ORACLE817.ORADEMO.CUSTOMERS	704	2	698	4	0
2	ORACLE817.ORADEMO.CUSTOMERS2	703	1			0
1	ORACLE817.ORADEMO.ORDERS	1717	4	1711	2	0
2	ORACLE817.ORADEMO.ORDERS2	1714	1			0
1	ORACLE817.ORADEMO.DETAILS	3619	10	3607	2	0
2	ORACLE817.ORADEMO.DETAILS2	3609	0			0
1	ORACLE817.ORADEMO.ITEMS	102	0	98	4	0
2	ORACLE817.ORADEMO.ITEMS2	102	0			0

- Generated for each pair of tables
- Identifies tables containing unmatched rows
- Identifies tables containing duplicate match keys

Challenge - Revising Test Data

- You must be sure that all logic paths are tested

  BUT…

- Your production data may not contain all the needed test cases

  – Errors

  – Boundary conditions
  – Unusual combinations
  of data

De-Identifying Test Data

- Removing, masking or transforming elements that could be used to identify an individual
  - Name, address, telephone, SSN / National Identity number

- No longer confidential; therefore acceptable to use in open test environments

- Masked or transformed data must be appropriate to the context
  - Consistent formatting (alpha to alpha)
  - Within permissible range of values

How should companies Protect Privacy?

- Organizations need the ability to **de-identify, mask and transform** sensitive data

- Companies can apply a range of transformation techniques to substitute customer data with **contextually-accurate but fictionalized data** to produce **accurate test results**

- By masking personally-identifying information, you protect the **privacy and security** of confidential customer data, and **support compliance** with local, state, national, international and industry-based privacy regulations

Modular Subsetting Deployment Approach

Required Siebel Subsetting common ADs

Reference Tables Subset

Siebel Party Model Subset

Account, Contact, etc...

Modular subsetting AD approach

Opportunities Subset (Activities)

Serv. Req. Subset (Activities)

Quotes Subset (Activities)

Orders Subset (Activities)

Example: TDM Approach

Table 4. Data Classification Example

Service Request Status	Opened 2006	Opened 2005	Opened 2004	Opened Pre-2003
Unscheduled	578	32		
Scheduled	2,356	211		
Pending	322	3		
Open	5,093,750	456,542	1,211	21
Open – DN Call Back	3,245	211	23	
In Progress	123,552	23,475	4,695	939
Closed	36,383,928	32,485,650	29,005,045	53,654,629
Cancelled	3,768,637	3,364,854	3,004,334	5,557,531

One Module, plus common
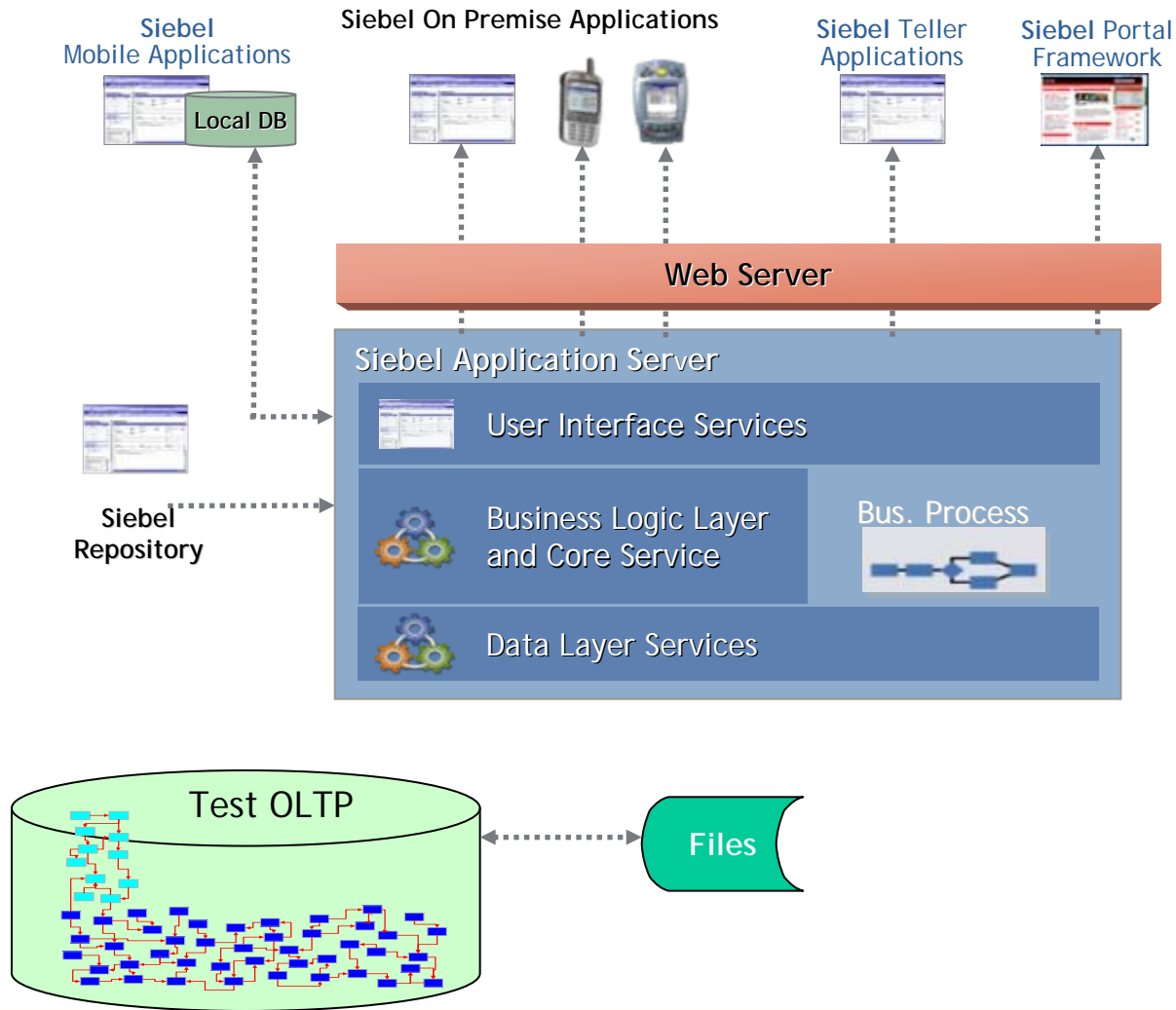
Effective Test Data Management for Improving Efficiency

- Save effort and expense
  - Compared to building new database at each stage
- Enable scalability and flexibility
  - Redeploy resources as needs evolve
- Promote consistency
  - Test against required conditions every time
  - Ensure reliable results
- Rapidly locate differences in data across successive product versions
  - Table to table
  - Multiple sets of related tables
- Identify, investigate, resolve errors
  - Avoid propagating to Production
  - Faster, easier, cheaper to fix in Test

TDM ROI Benefits

Projected ROI = 504% (3 years), Payback Period = 13 months

Best Practices with Proposed Solution	Projected Improvement or Benefit
Reduce storage costs for development / test databases	20.0%
Speed cycle times for testing and related activities	
Define test environments (tablespace, tables & relationships)	20.0%
Write scripts	20.0%
Load tables and columns / refresh test databases	20.0%
Edit test conditions	20.0%
Run test jobs	20.0%
Validate test results	20.0%

Countrywide ROI Analysis

Thank You!