QM08

# Essentials of Test Data Management

Alan Johns

Optim Technical Presales Consultant

johnsala@uk.ibm.com


IBM Rational Software Development Conference 2008

WHERE TEAMS ARE R-HEROES

IBM

Rational. software

# Effective Test Data Management

Application testing for accuracy, reliability and quality is clearly important.

Today's enterprise applications drive business initiatives and require the utmost attention.

Application Development teams must test in an efficient, repeatable, accurate and cost effective manner.

Your Career and Sanity Depend on it!!

# The Development Lifecycle – What are you after?

**Reduce Business  Downtime**

**Get to Market Faster**

**Maximize Processes**

**Reduce Defects**

# Test target: What does the test seek to verify?

- Functionality
  - Features and capabilities
- Performance
  - Speed, availability, tolerance for load
- Usability
  - Ease with which the software can be employed
- Security
  - Vulnerability to unauthorized usage
- Compliance
  - Conformance to internal standards or external regulations

# Typical role/responsibility matrix

| | Developers | Testers on project teams | Testers in central group | Customers | IT operations |
|---|---|---|---|---|---|
| **Unit** | Sole | | | | |
| **Integration** | Sole | | | | |
| **Component** | Primary | Secondary | | | |
| **System** | Secondary | Primary | Secondary | | |
| **System integration** | | Secondary | Primary | | |
| **UAT** | | | | Primary | |
| **Implementation** | | | Secondary | | Primary |

# The Analysts know what you need

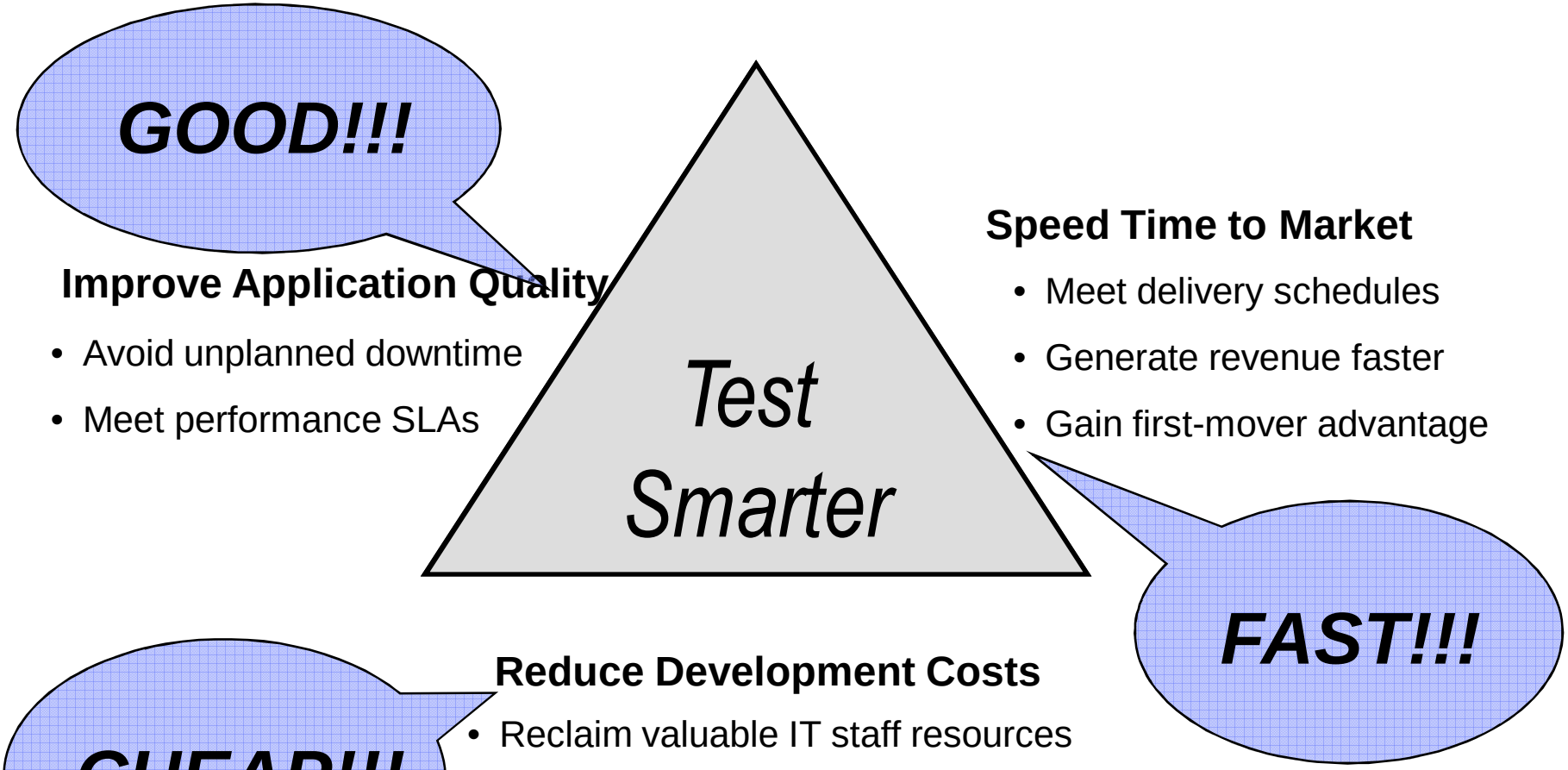
Forrester data shows that North American enterprises are planning to devote 10% more of their 2007 software budgets to new development.

But the success rate for app dev projects is famously low.

Nearly one-third of business stakeholders are dissatisfied with the state of custom app dev at their companies.

***To provide return on these additional investments, app dev organizations need to mend their ways — that is, their processes and methodologies.***

# A Picture to Explain




**Improve Application Quality**

- Avoid unplanned downtime
- Meet performance SLAs

**Speed Time to Market**

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

**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

# Current Practice?

## #1 - Clone Production

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

Clone Production

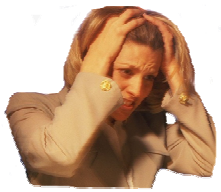
Request for Copy

Wait

After Production Database Copy

Changes

Production Database Copy



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

## #2 - Write SQL

Write SQL

Extract

- Complex
- Subject to Change

Extract

Changes

After

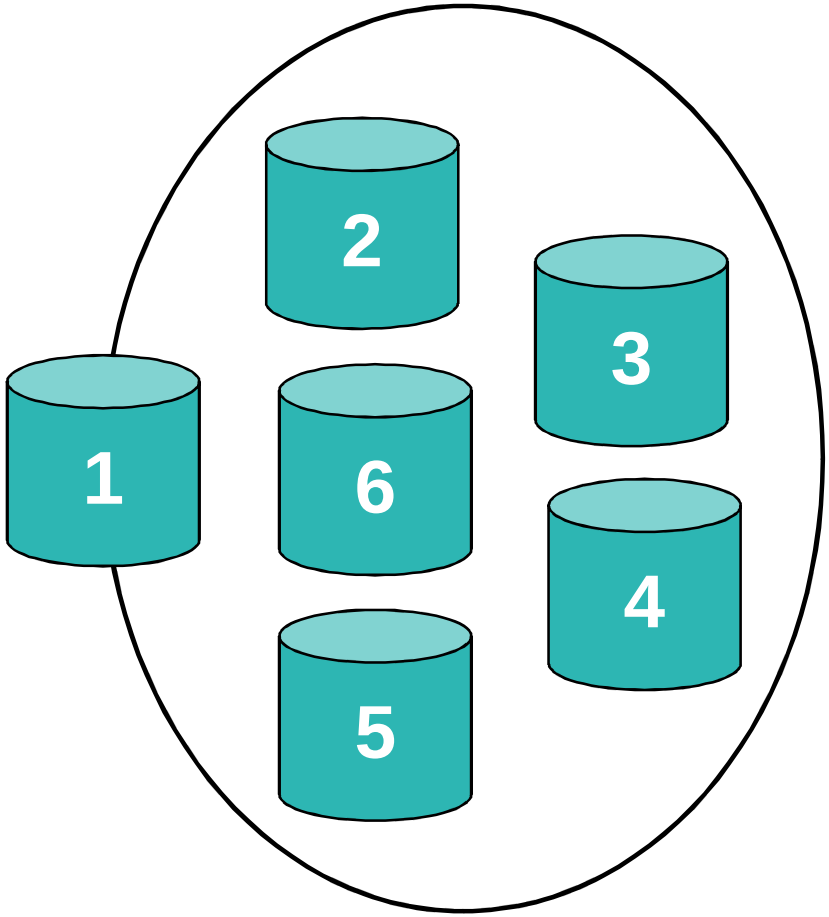
- RI Accuracy?
- Right Data?

Expensive, Dedicated Staff, Ongoing Responsibility.

Share test database with everyone else

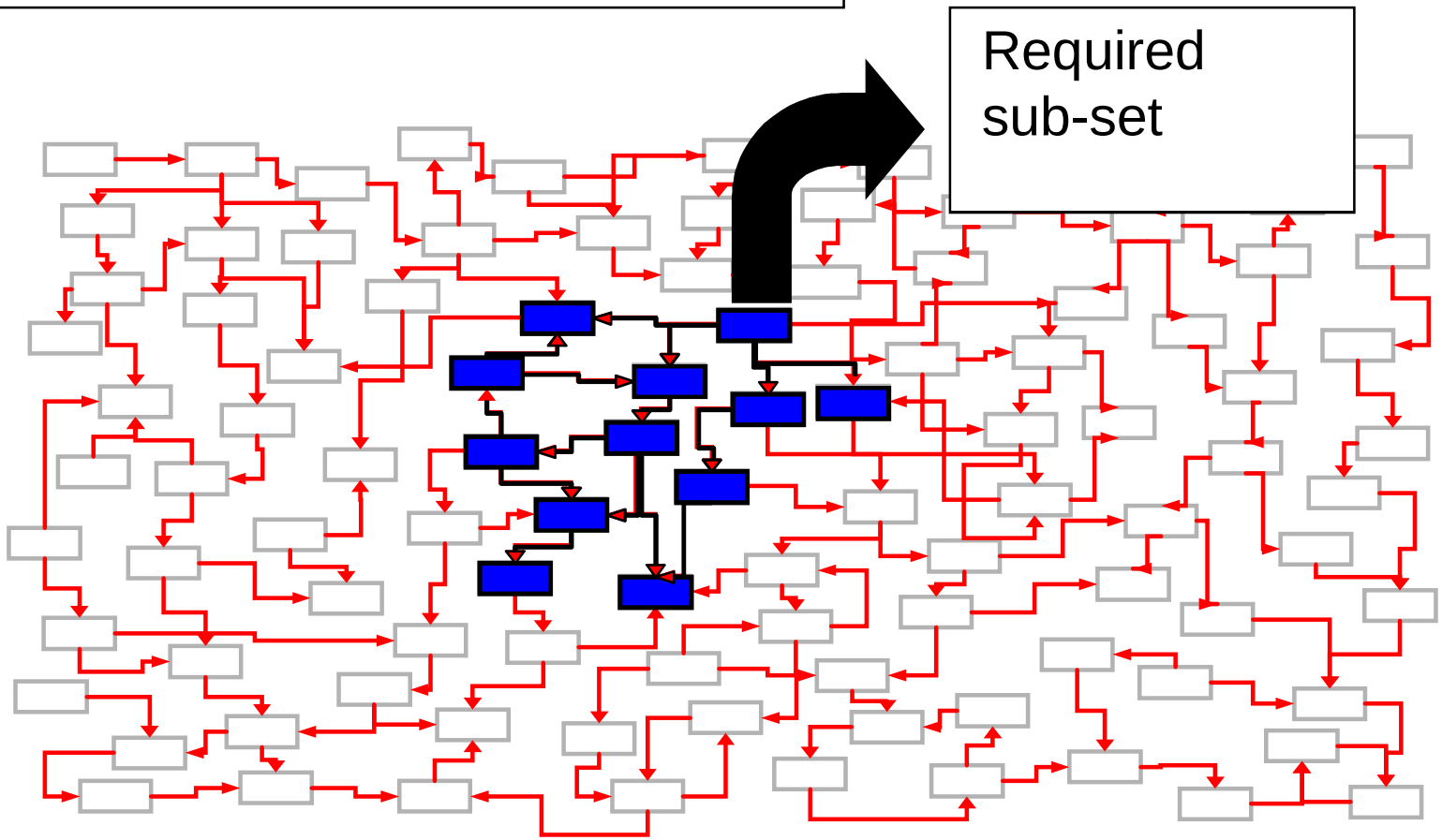
# Managing Application Data Growth

## The “Data Multiplier Effect”




| | |
|---|---|
| **Production** | **200 GB** |
| **Training** | **200** |
| **Unit** | **200** |
| System | 200 |
| UAT | 200 |
| Integration | 200 |
| Total Size | 1200 GB |

# What's in a Test Database



## Test Data Management

End cloning and duplicate copies of huge databases

Reduce storage requirements and associated expenses

Speed application deployment by streamlining the way you create and manage test environments

So what sort of process would give us this capability???

# Test Data Management – required functions

We need to be able to establish the correct data at the correct volume, so a flexible extract/insertion facility is needed

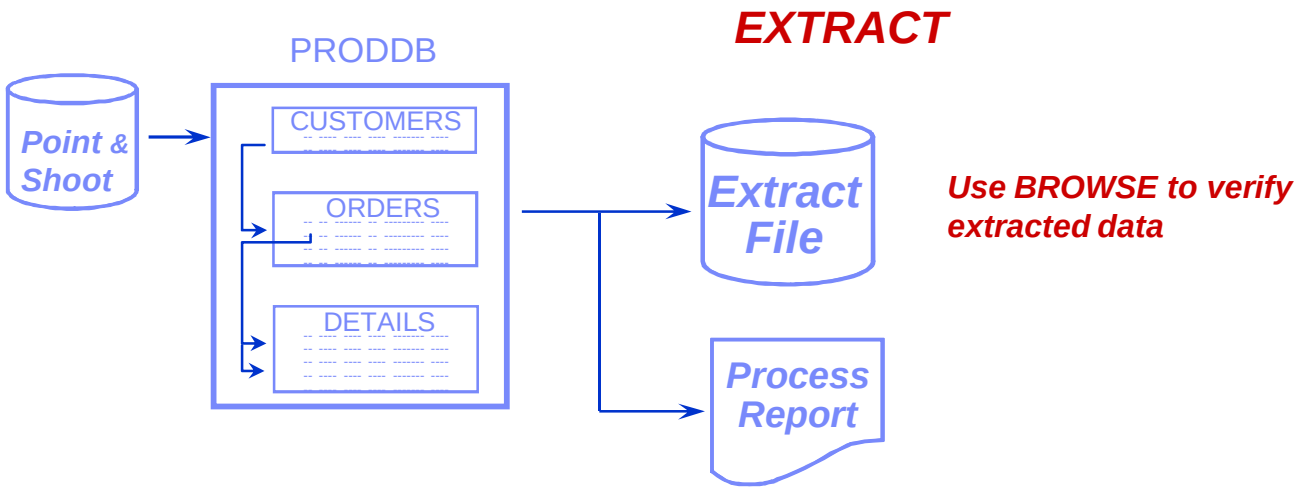
We need to be able to change data at the row level so that we can place error conditions etc into the data as required.

We need to be able to base line our test environments, once configured, so they can be re-established and managed quickly.

We need to see what our applications are doing to the data so we can manage and tailor our testing data efficiently

# Extract Process

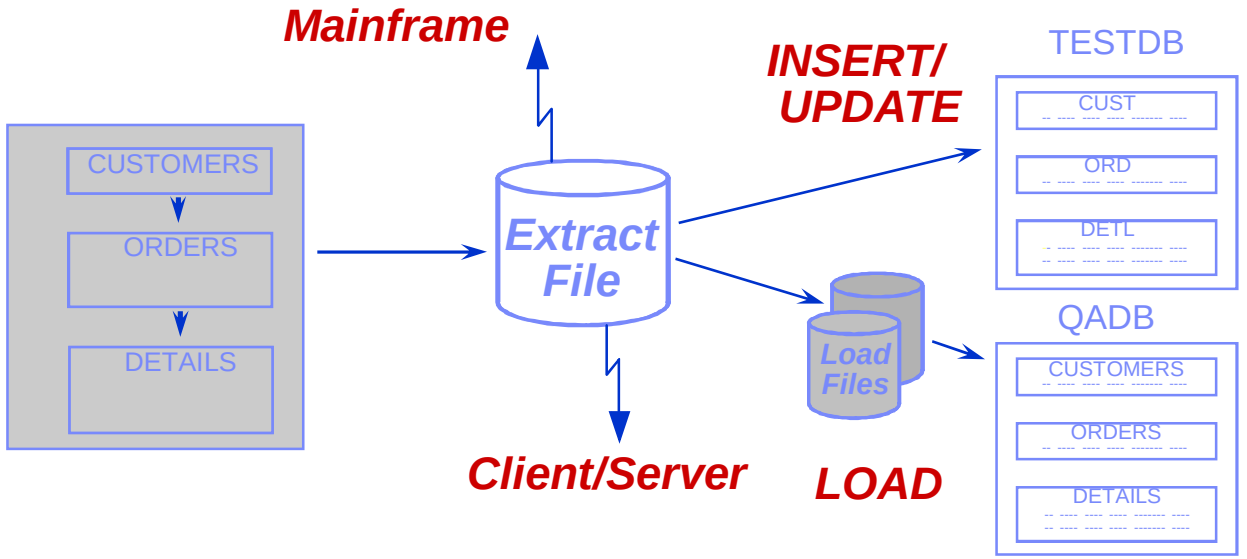
## Extract Parameters




Extract from source tables using dynamic SQL

Extract data and/or object definitions

# Populate Destination Tables




Dynamic SQL

Load utility for large volumes of data

Download to Client/Server RDBMS from MVS

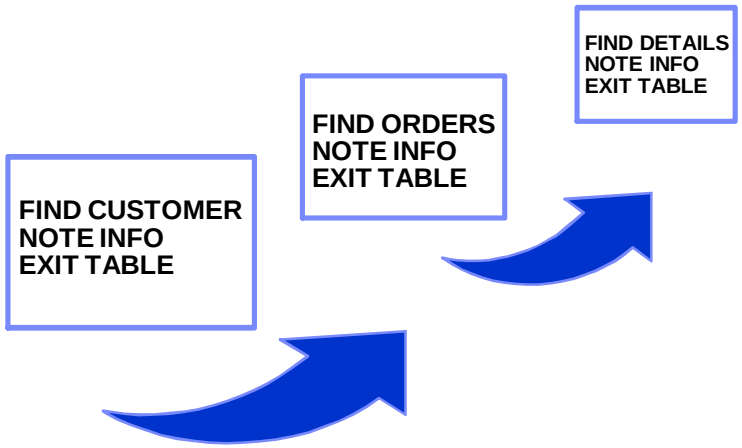
Upload to mainframe DB2 from Client/Server

# Traditional vs. Relational Tools

## *Single Table Editors*

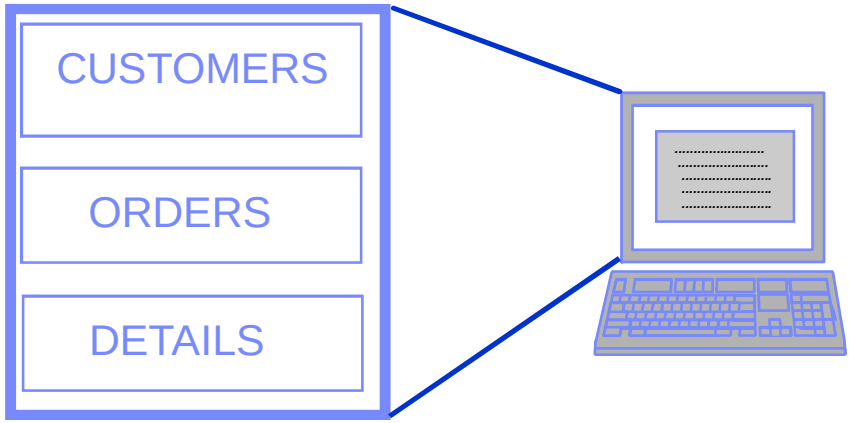
One table/view at a time

No edit of related data
from multiple tables

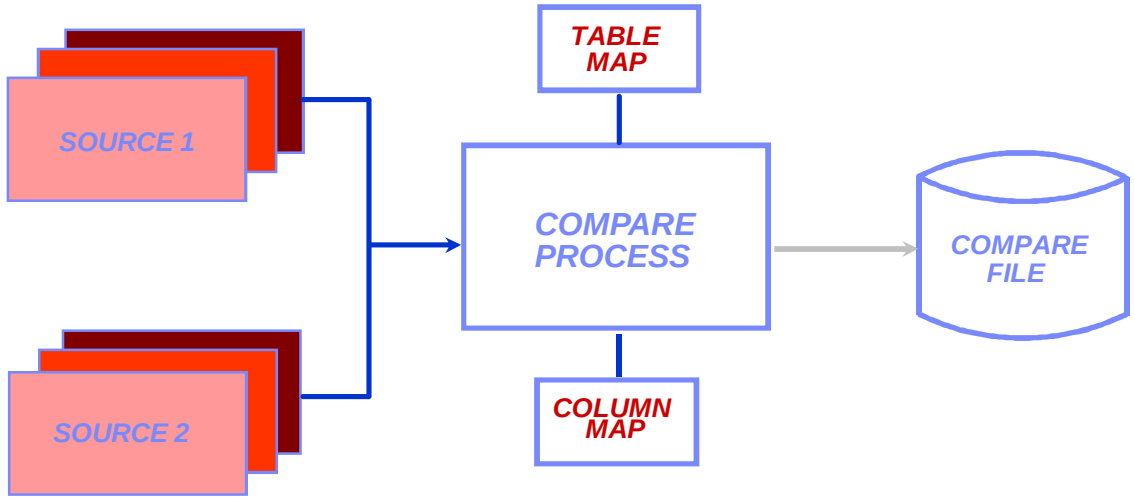



## *The Relational Editor*

- ***Simultaneous* browse/edit of related data from multiple tables**



# Relational Compare Facility




Single-table or multi-table compare

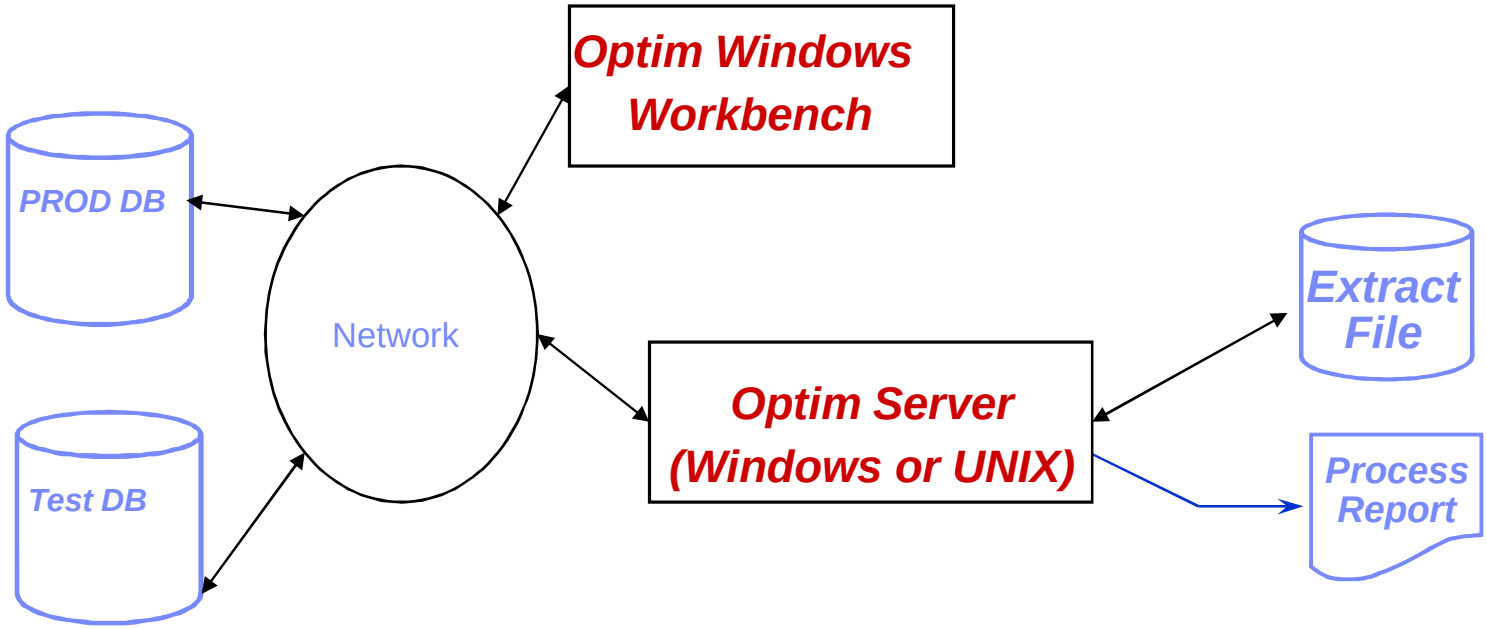
Creates compare file of results

Displays results on screen
For application testing, QA, and to verify database contents

Enhances productivity by finding unexpected changes in the data

# The Optim Server

## Optim Processes run on Named Servers



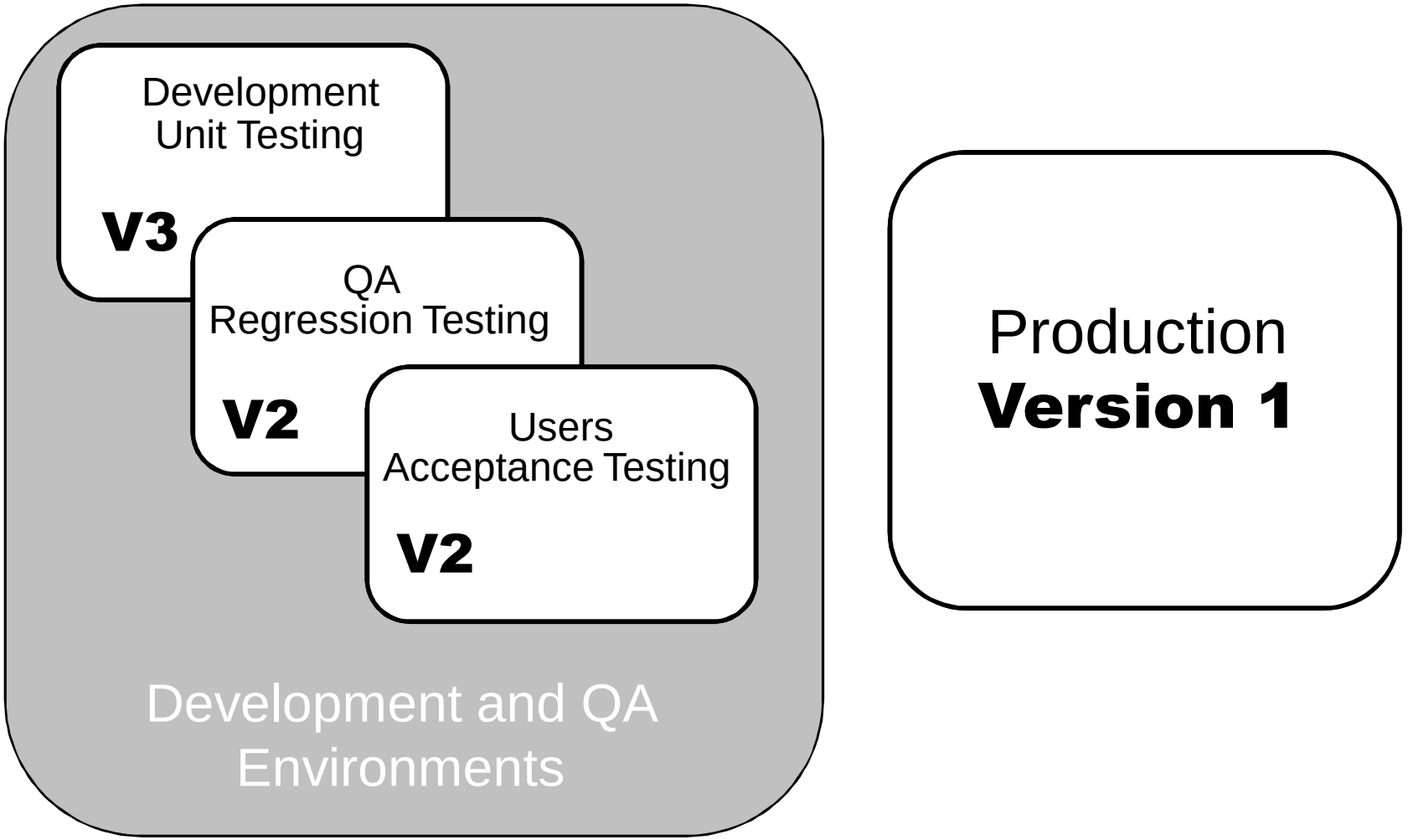

Extract from source tables using dynamic SQL

Extract data and/or object definitions

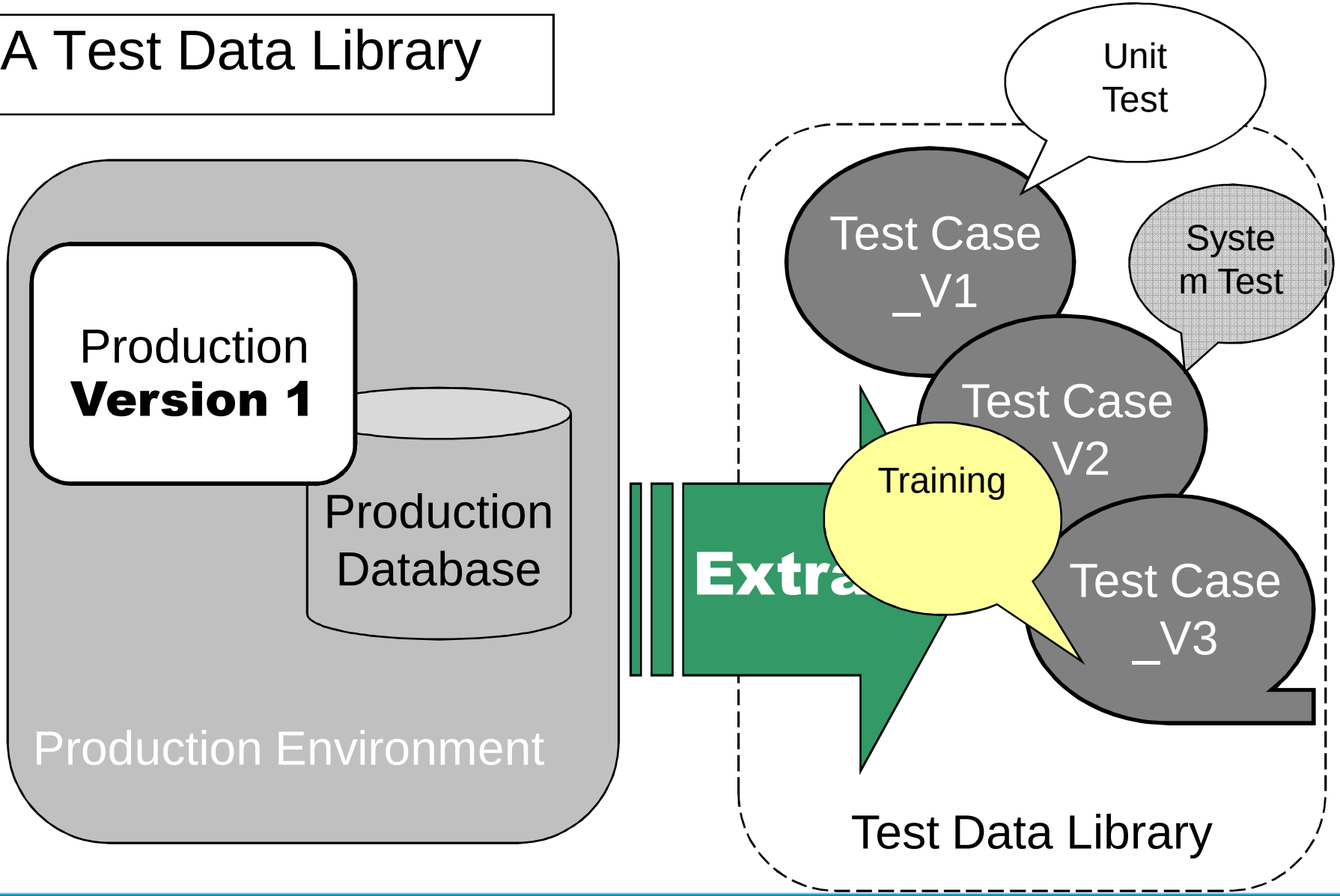
Run processes either from the Workbench or from a named server

# Enterprise Application

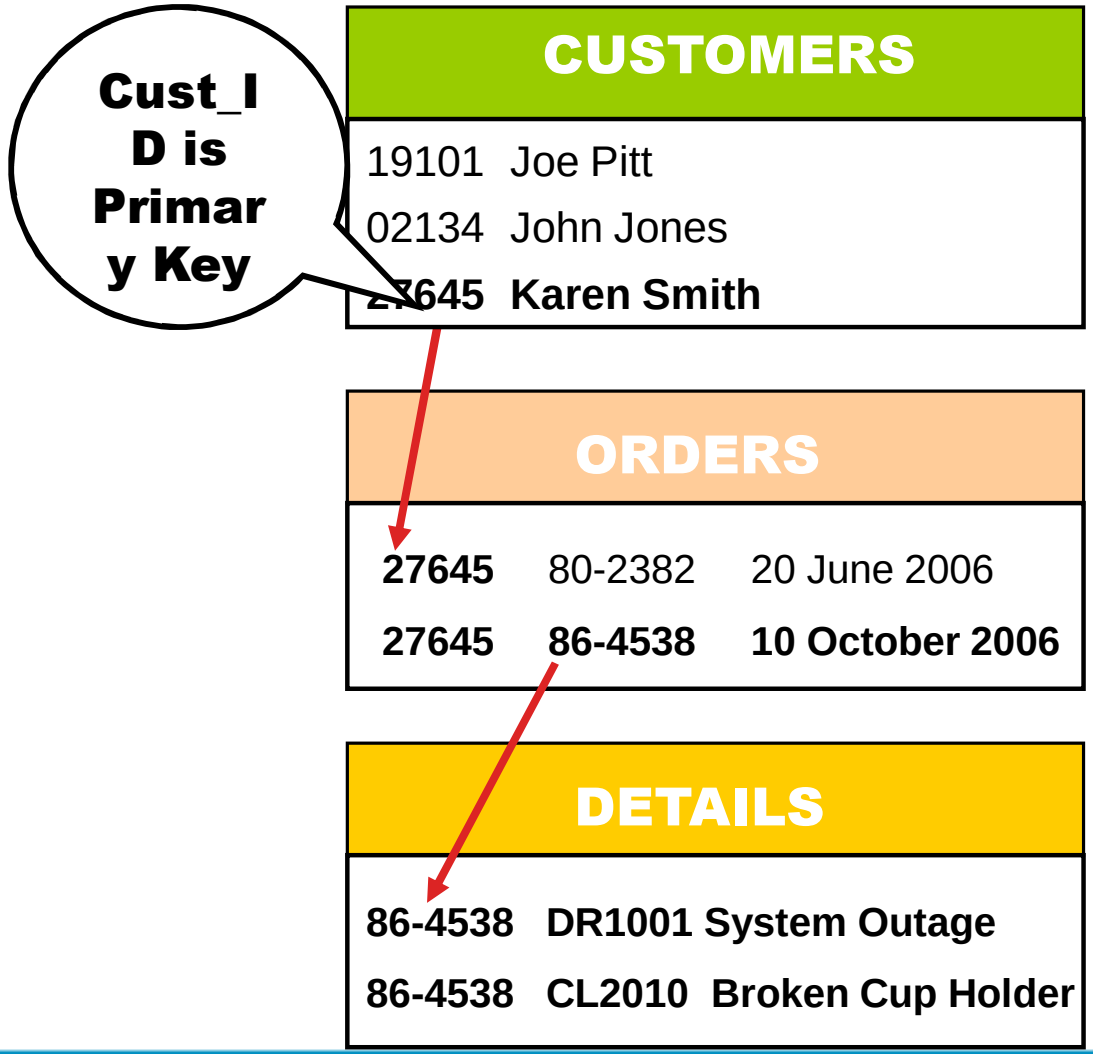
Snapshot in Time



# A Test Data Library



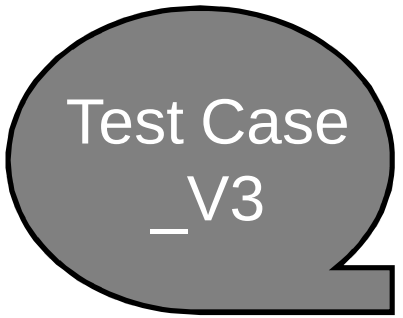
# What's in a Test Case?



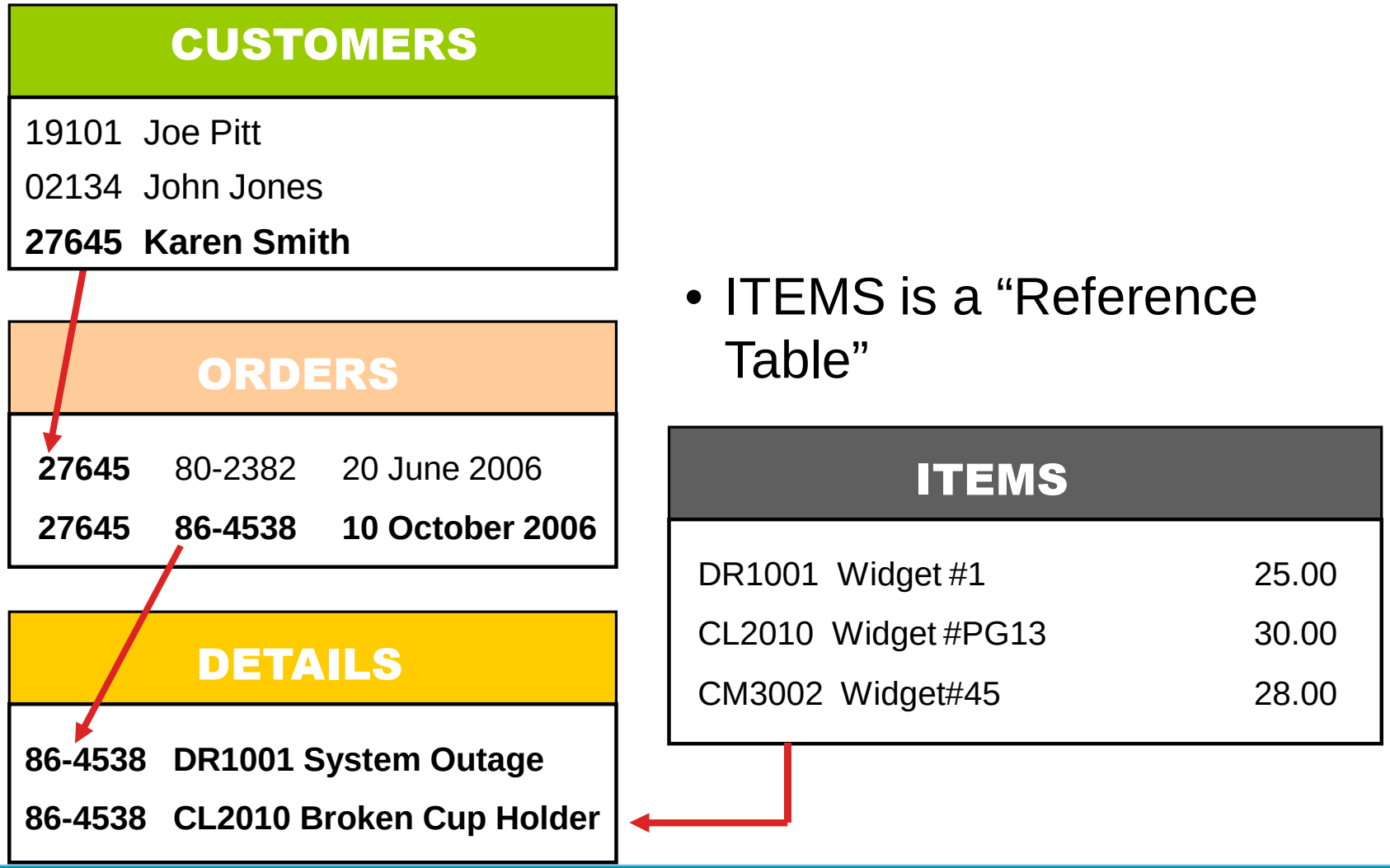

Referentially-intact subset of data

Example:

All Open –DN Call Back related to Cust_ID **27645** (Karen Smith)

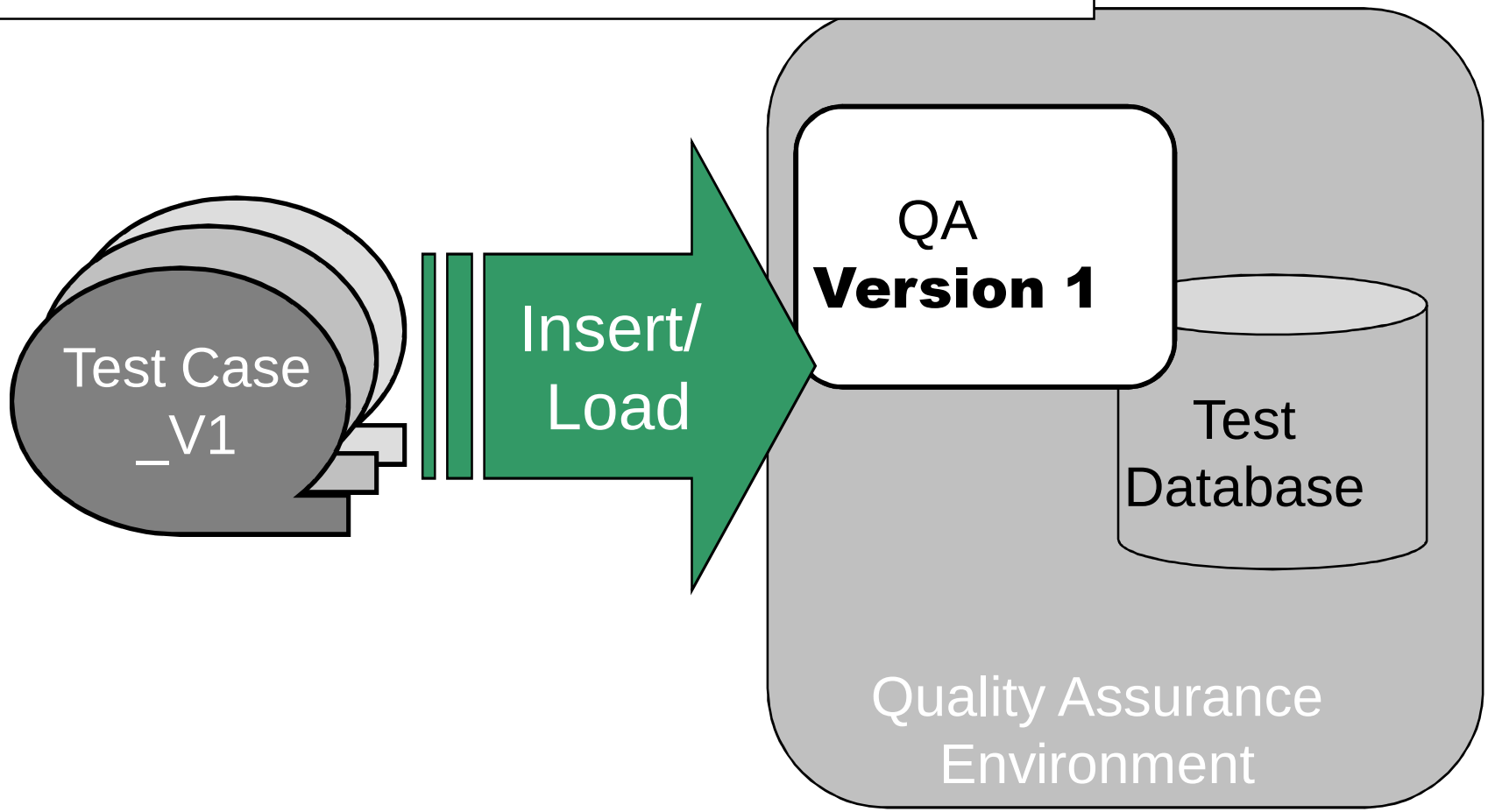


# What's in a Test Case?



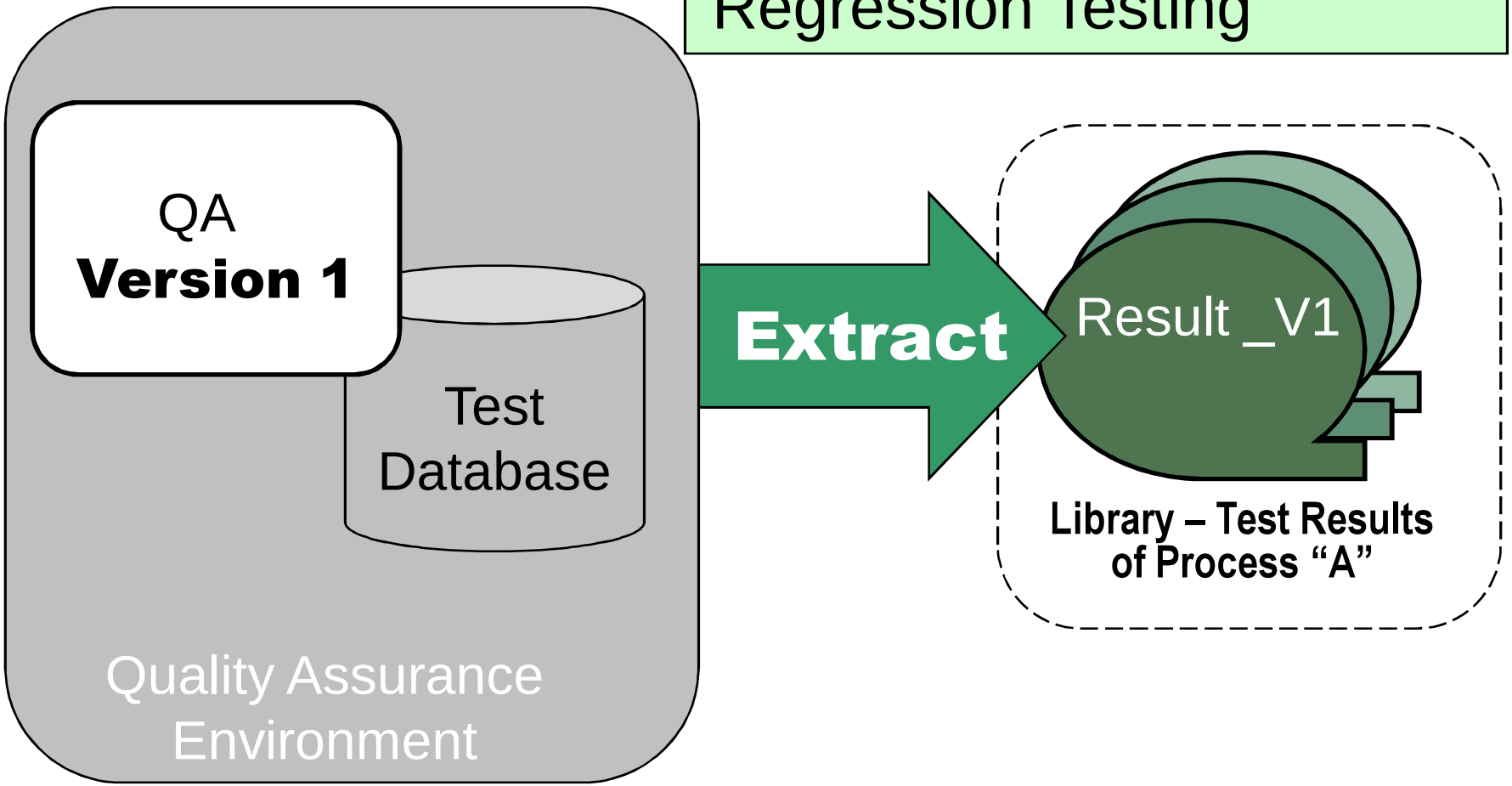

- ITEMS is a "Reference Table"

# Insert into Development Environment

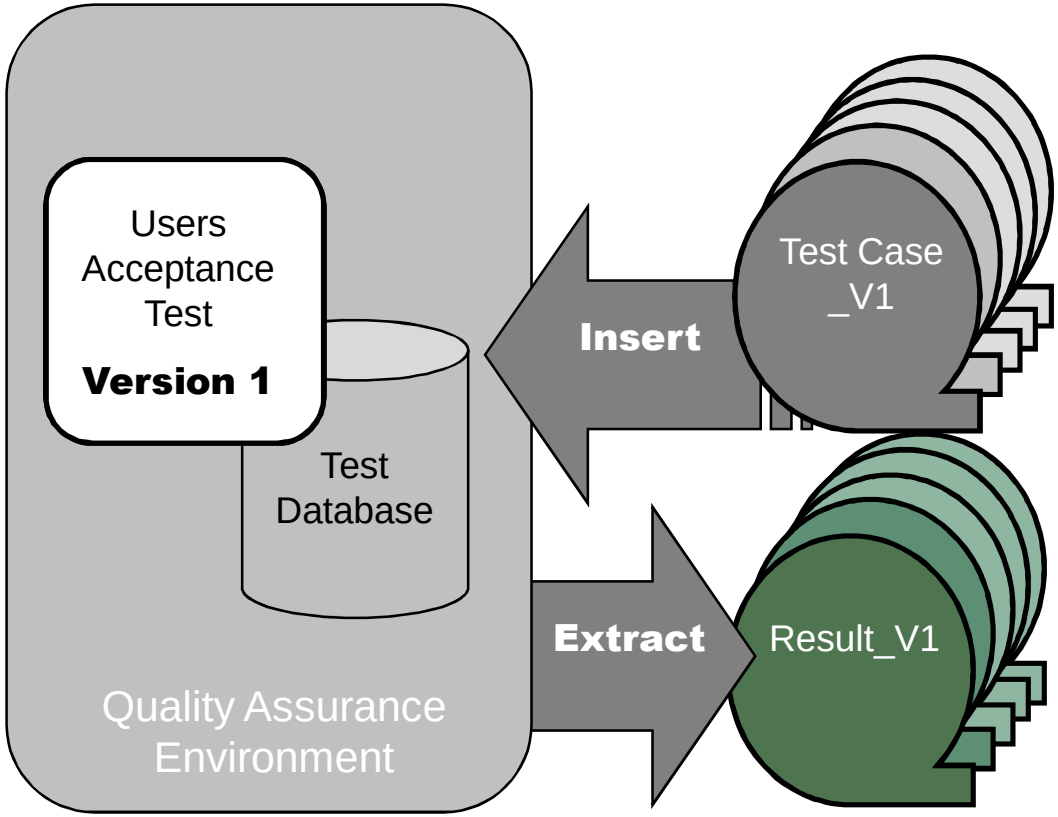


# Tracking the Results

## Regression Testing



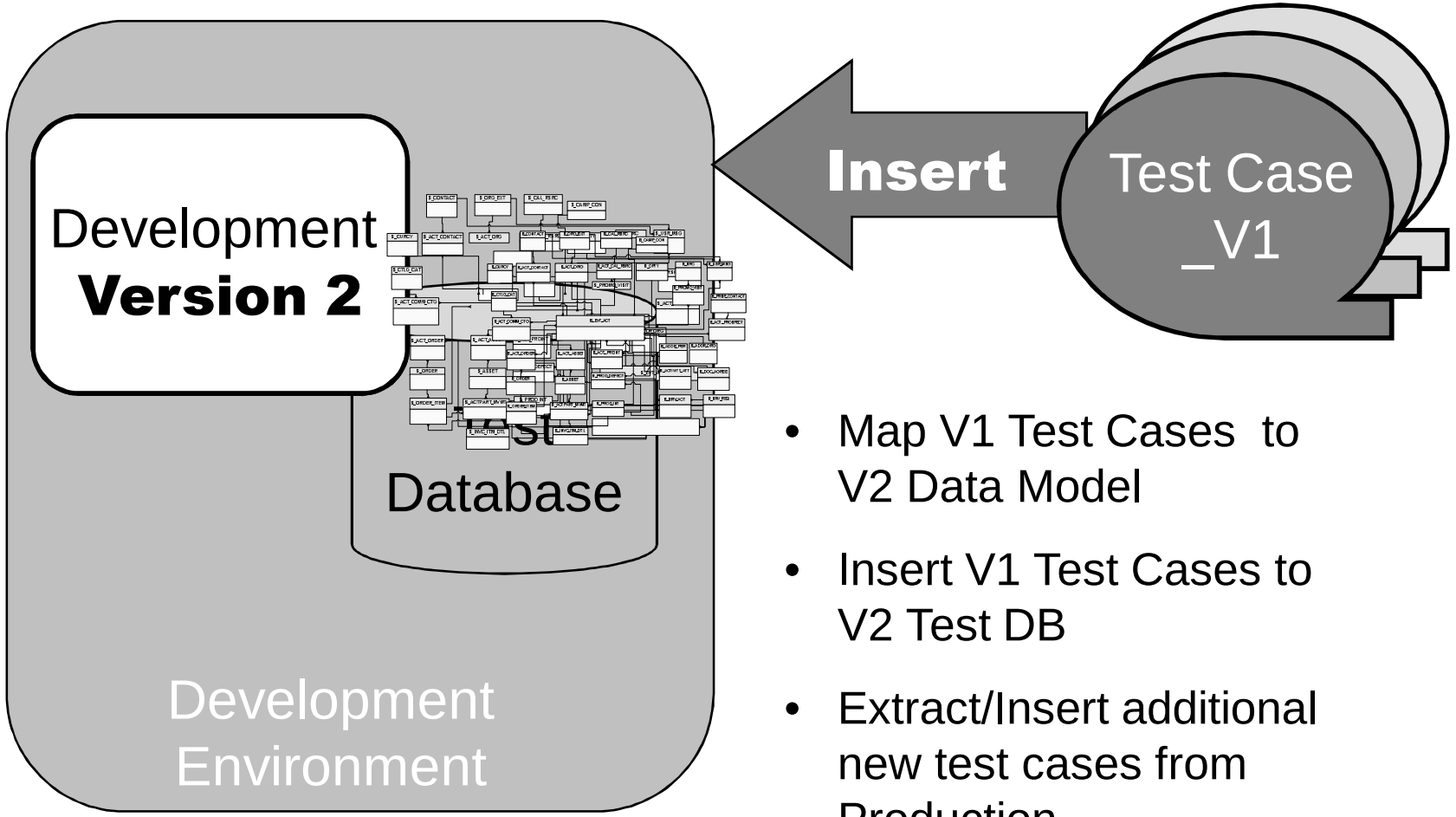
# Tracking the Results




- Reuse Test Cases from “Test Case Library”
- Take snapshots of process results for later comparison

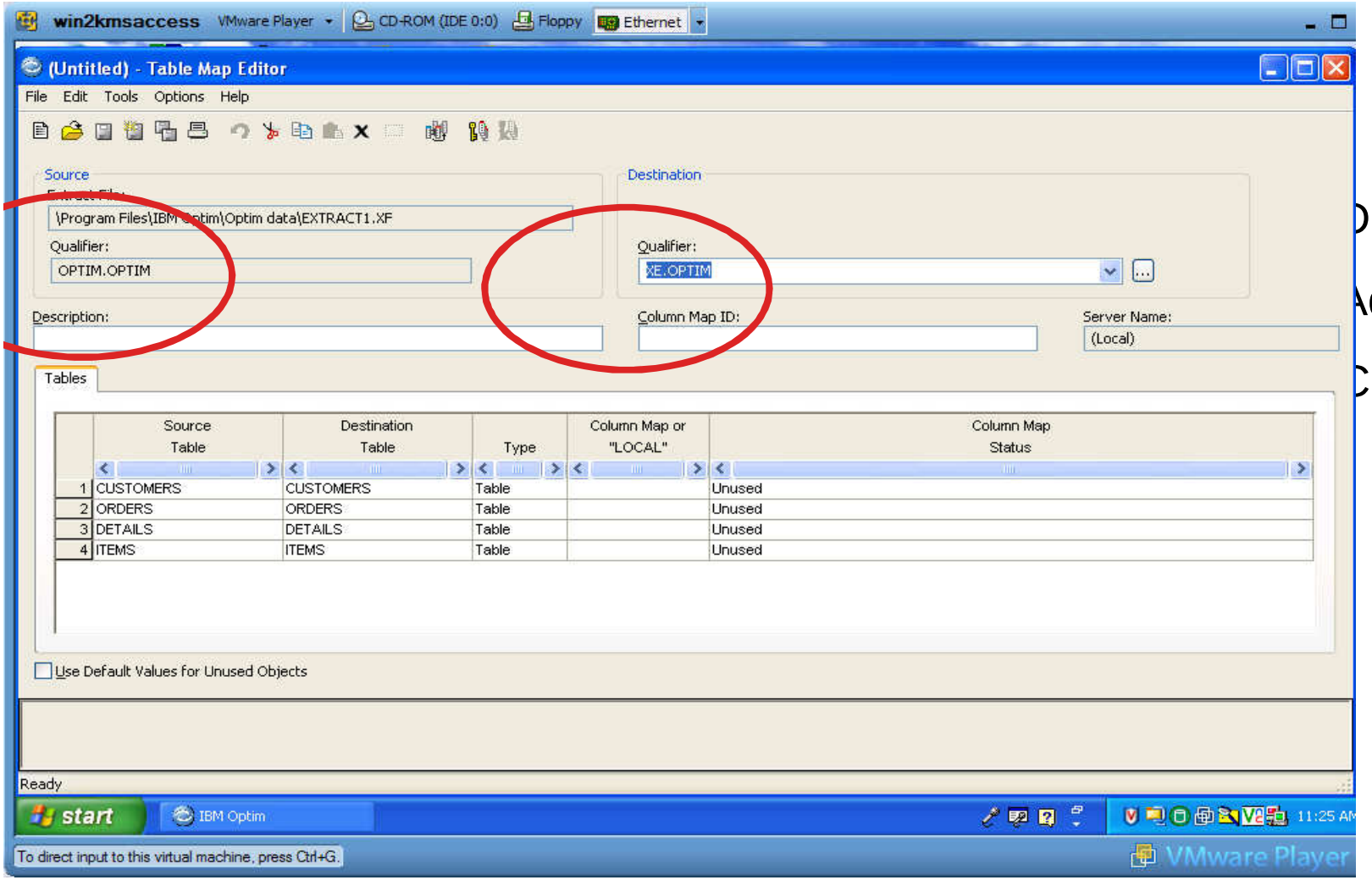
# Version 2: Unit Test – Reusability

**Big Application Changes!!!!**




- Map V1 Test Cases to V2 Data Model
- Insert V1 Test Cases to V2 Test DB
- Extract/Insert additional new test cases from Production

# Table Map

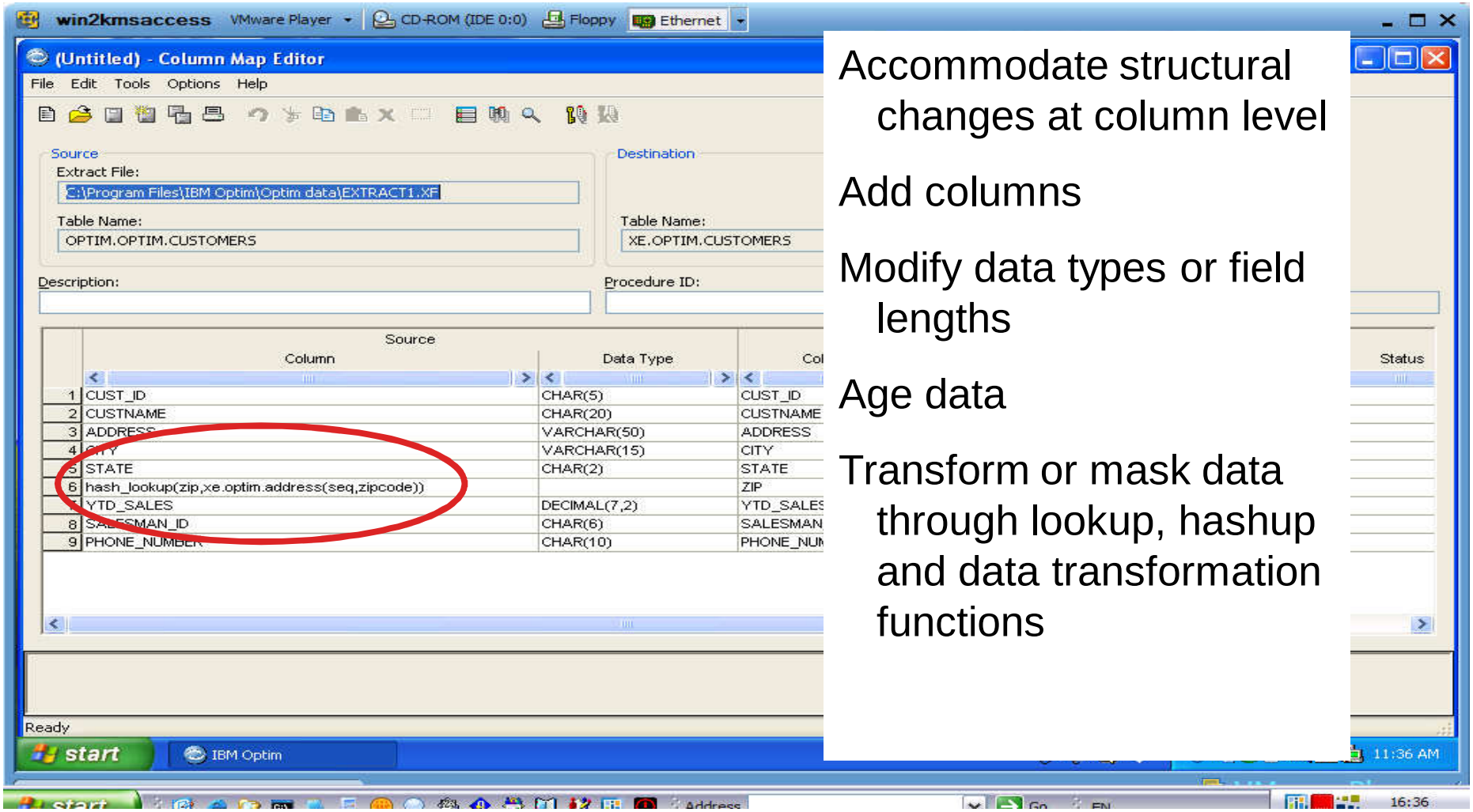



Define Target

Add tables

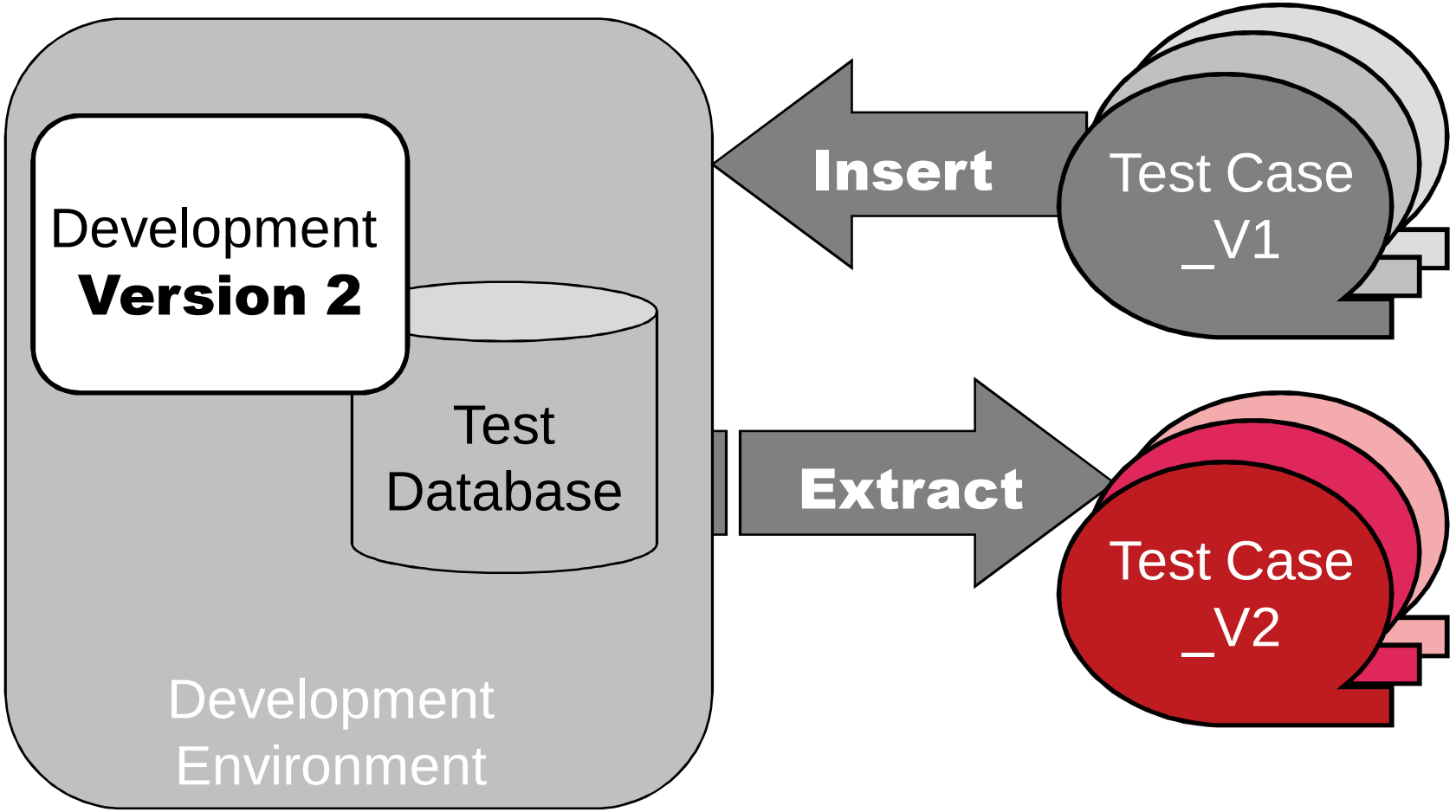
Change table names

# Column Map

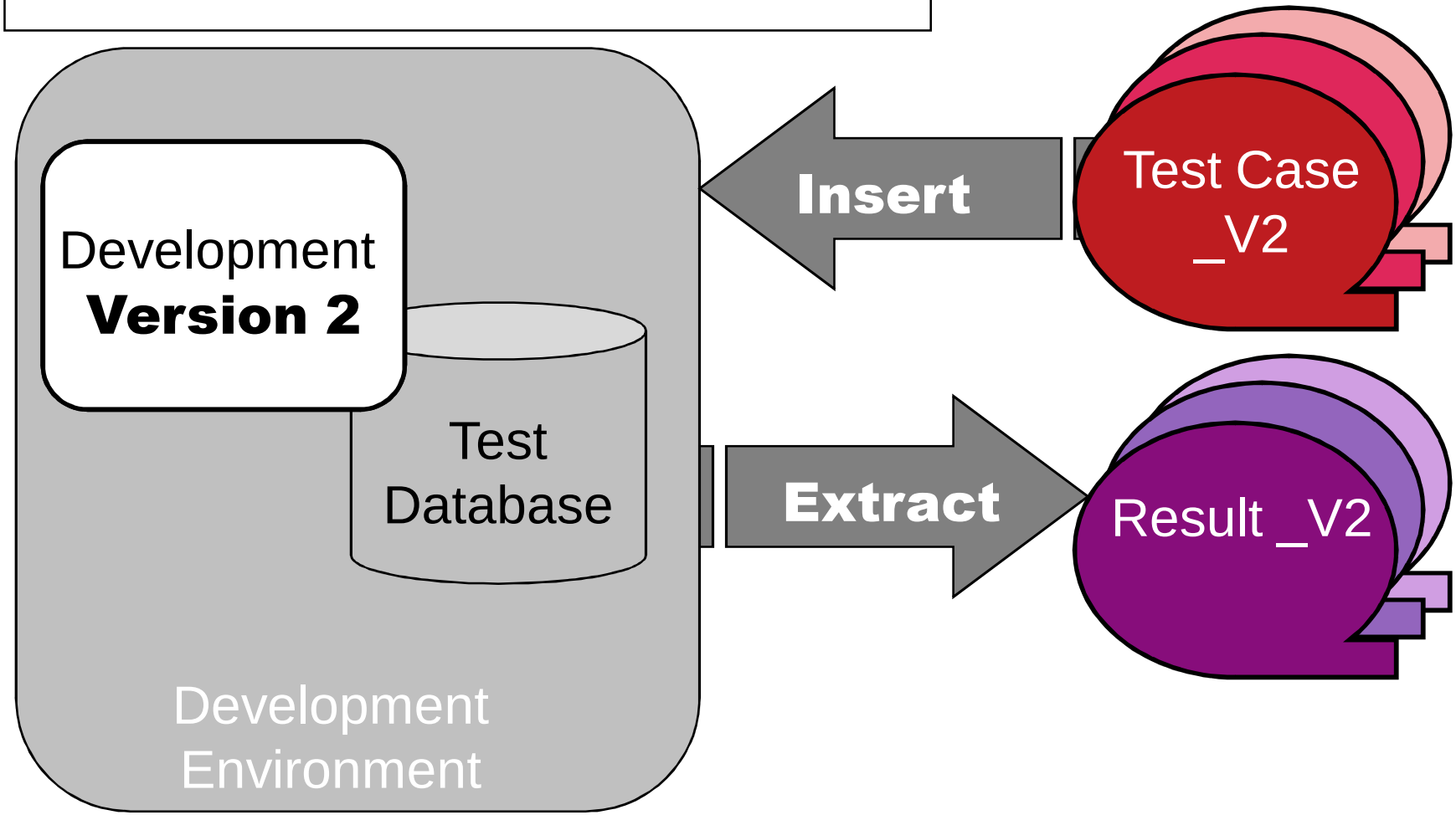



- Accommodate structural changes at column level
- Add columns
- Modify data types or field lengths
- Age data
- Transform or mask data through lookup, hashup and data transformation functions

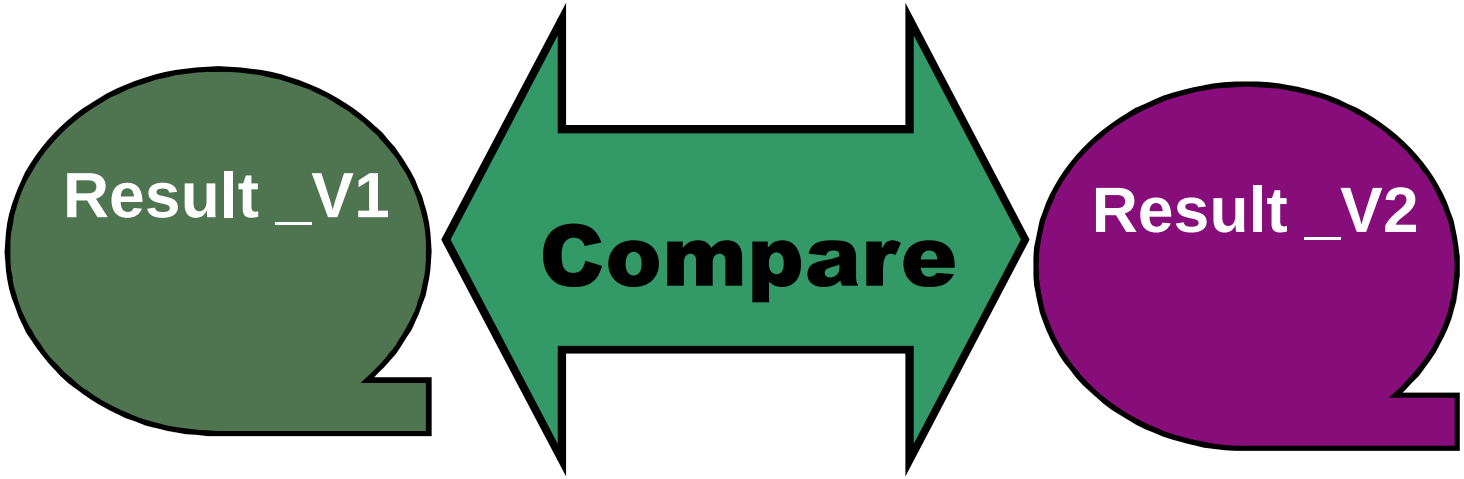
# Version 2: Unit Test



# Version 2: Regression Test



# Version 2: Regression Test of Process “A”



# Analyzing Test Data Results

## Version 1

| INVOICES | | | |
|---|---|---|---|
| 27645 | 86-4538 | Widget#1 | $80.00 |
| 27645 | 86-4538 | Widget#PG13 | $20.00 |
| | | Invoice Total | $100.00 |

## Version 2

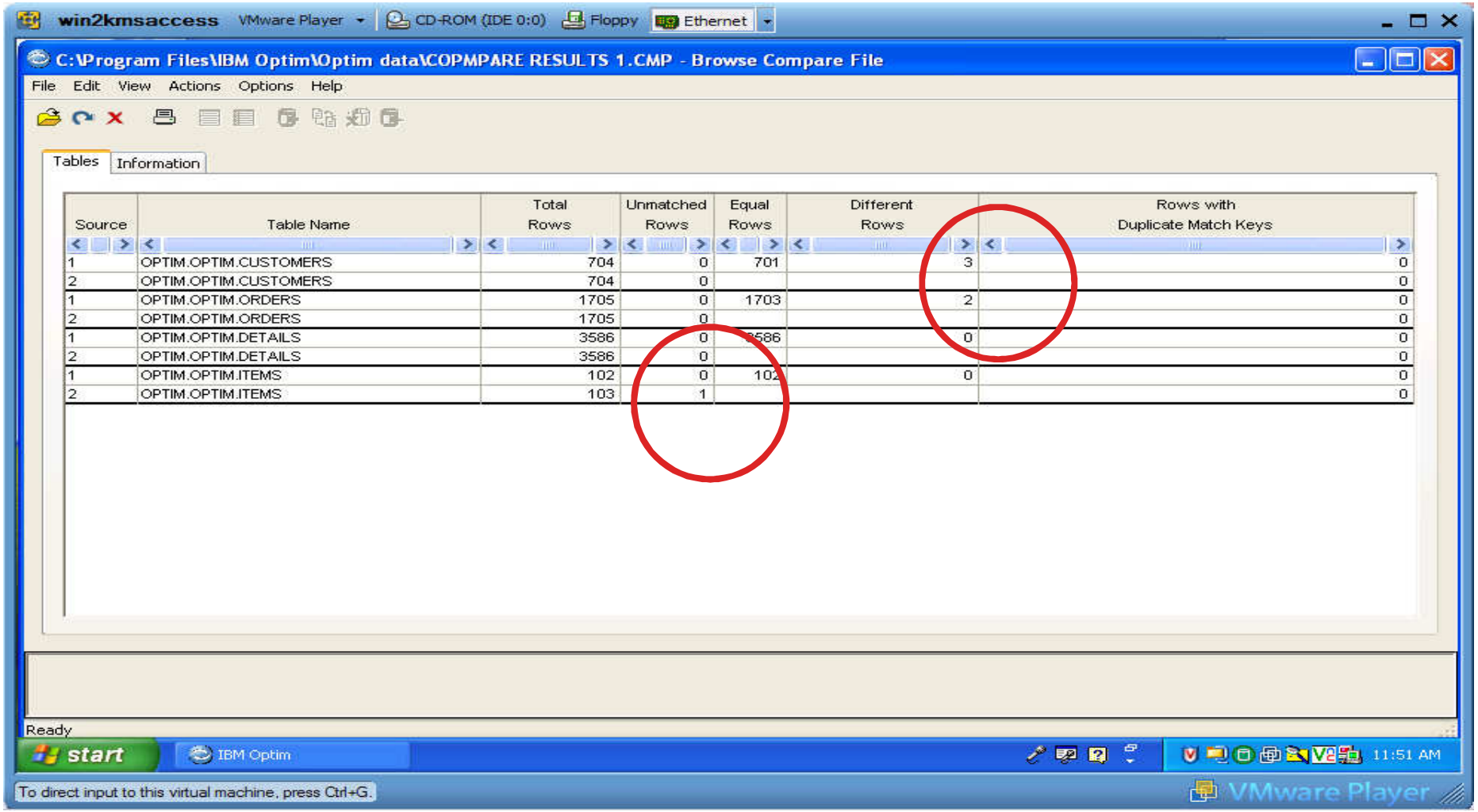
| INVOICES | | | |
|---|---|---|---|
| 27645 | 86-4538 | Widget#1 | $50.00 |
| 27645 | 86-4538 | Widget#PG13 | $50.00 |
| | | Invoice Total | $100.00 |

Both Invoices total $100

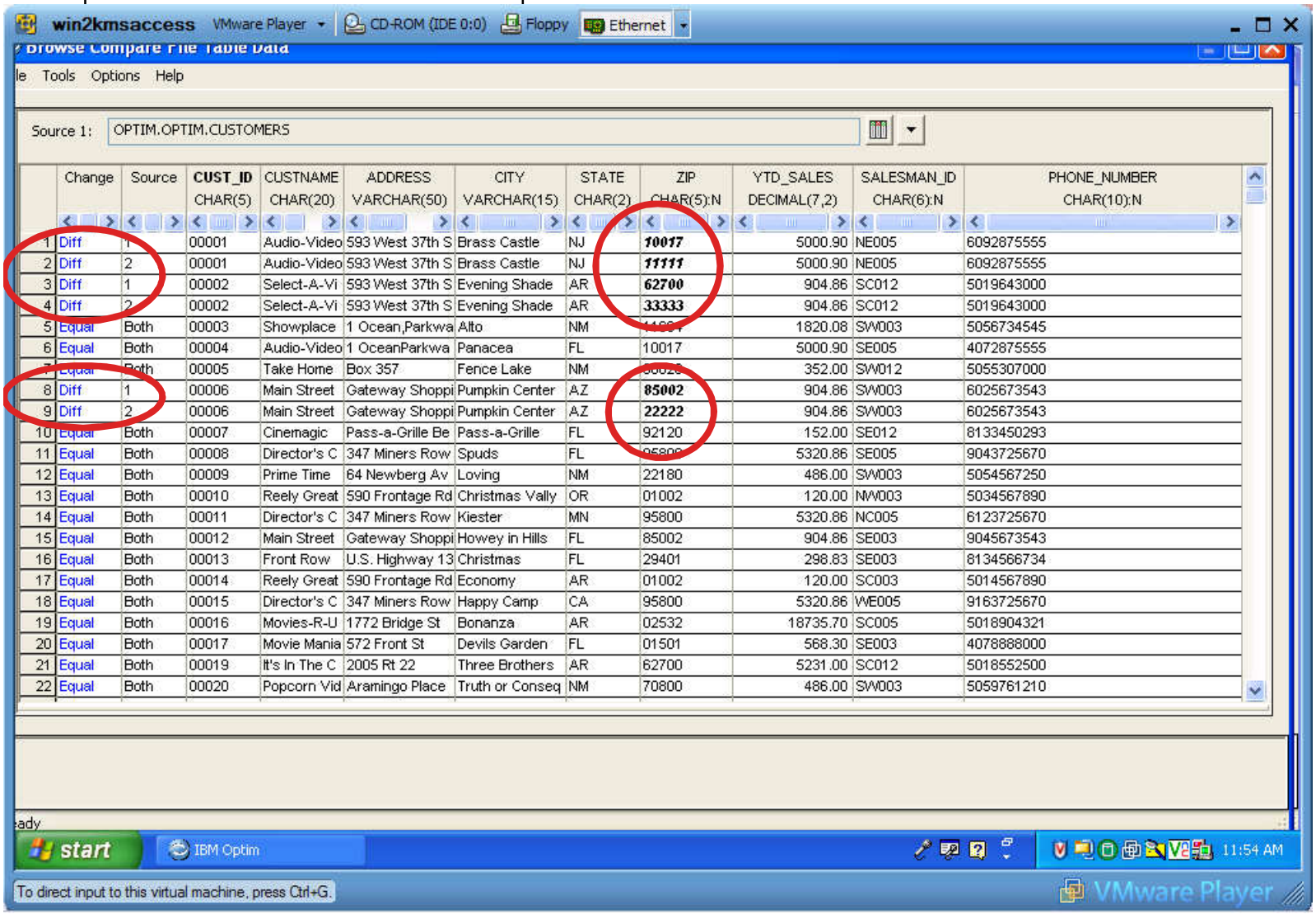
Composition is different

Could we have missed an error?

# Browse Compare File

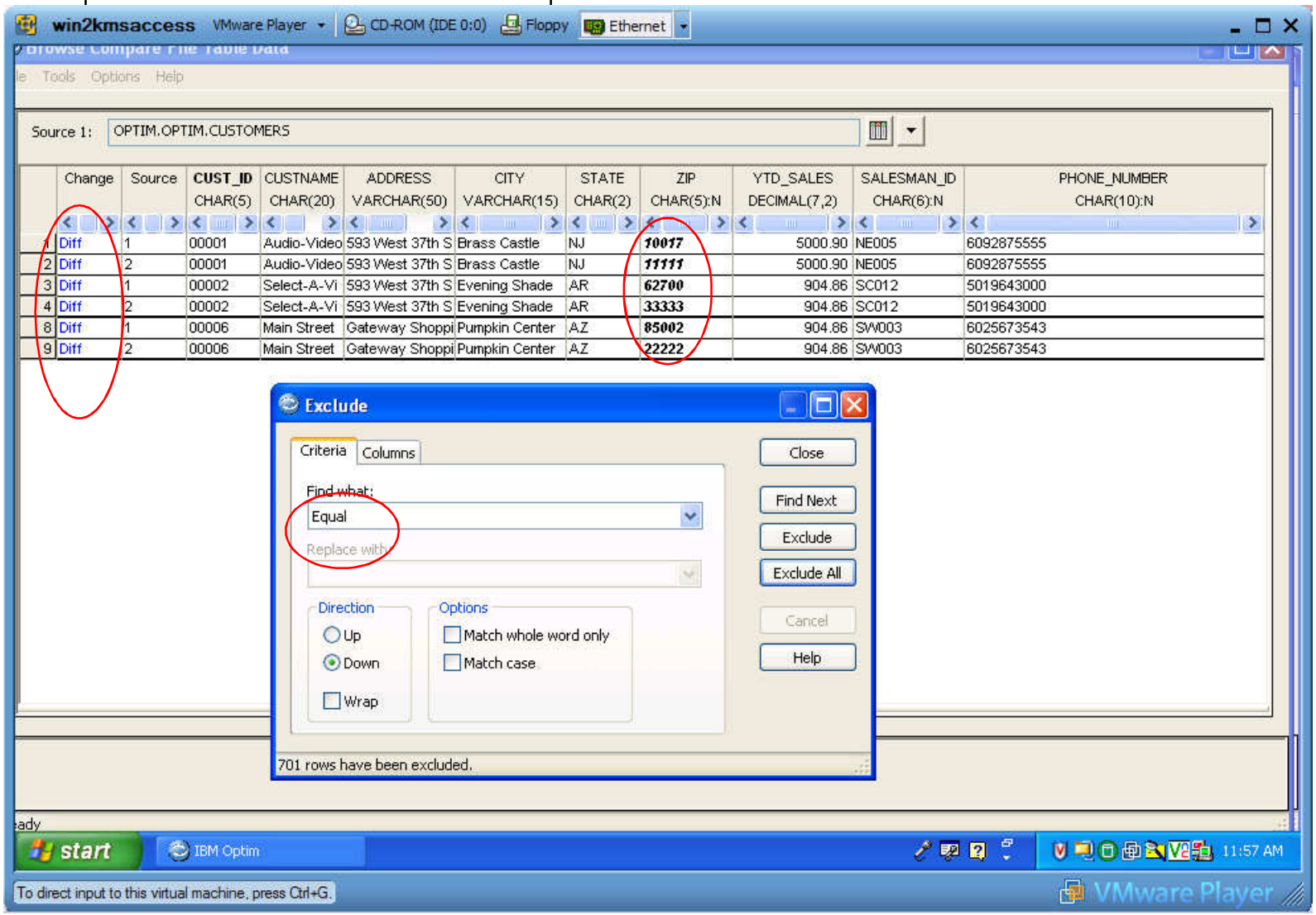


# View Details




| | Change | Source | CUST_ID CHAR(5) | CUSTNAME CHAR(20) | ADDRESS VARCHAR(50) | CITY VARCHAR(15) | STATE CHAR(2) | ZIP CHAR(5):N | YTD_SALES DECIMAL(7,2) | SALESMAN_ID CHAR(6):N | PHONE_NUMBER CHAR(10):N |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Diff | 1 | 00001 | Audio-Video | 593 West 37th S | Brass Castle | NJ | ***10017*** | 5000.90 | NE005 | 6092875555 |
| 2 | Diff | 2 | 00001 | Audio-Video | 593 West 37th S | Brass Castle | NJ | ***11111*** | 5000.90 | NE005 | 6092875555 |
| 3 | Diff | 1 | 00002 | Select-A-Vi | 593 West 37th S | Evening Shade | AR | ***62700*** | 904.86 | SC012 | 5019643000 |
| 4 | Diff | 2 | 00002 | Select-A-Vi | 593 West 37th S | Evening Shade | AR | ***33333*** | 904.86 | SC012 | 5019643000 |
| 5 | Equal | Both | 00003 | Showplace | 1 Ocean,Parkwa | Alto | NM | [illegible] | 1820.08 | SW003 | 5056734545 |
| 6 | Equal | Both | 00004 | Audio-Video | 1 OceanParkwa | Panacea | FL | 10017 | 5000.90 | SE005 | 4072875555 |
| 7 | Equal | Both | 00005 | Take Home | Box 357 | Fence Lake | NM | [illegible] | 352.00 | SW012 | 5055307000 |
| 8 | Diff | 1 | 00006 | Main Street | Gateway Shoppi | Pumpkin Center | AZ | ***85002*** | 904.86 | SW003 | 6025673543 |
| 9 | Diff | 2 | 00006 | Main Street | Gateway Shoppi | Pumpkin Center | AZ | ***22222*** | 904.86 | SW003 | 6025673543 |
| 10 | Equal | Both | 00007 | Cinemagic | Pass-a-Grille Be | Pass-a-Grille | FL | 92120 | 152.00 | SE012 | 8133450293 |
| 11 | Equal | Both | 00008 | Director's C | 347 Miners Row | Spuds | FL | [illegible] | 5320.86 | SE005 | 9043725670 |
| 12 | Equal | Both | 00009 | Prime Time | 64 Newberg Av | Loving | NM | 22180 | 486.00 | SW003 | 5054567250 |
| 13 | Equal | Both | 00010 | Reely Great | 590 Frontage Rd | Christmas Vally | OR | 01002 | 120.00 | NW003 | 5034567890 |
| 14 | Equal | Both | 00011 | Director's C | 347 Miners Row | Kiester | MN | 95800 | 5320.86 | NC005 | 6123725670 |
| 15 | Equal | Both | 00012 | Main Street | Gateway Shoppi | Howey in Hills | FL | 85002 | 904.86 | SE003 | 9045673543 |
| 16 | Equal | Both | 00013 | Front Row | U.S. Highway 13 | Christmas | FL | 29401 | 298.83 | SE003 | 8134566734 |
| 17 | Equal | Both | 00014 | Reely Great | 590 Frontage Rd | Economy | AR | 01002 | 120.00 | SC003 | 5014567890 |
| 18 | Equal | Both | 00015 | Director's C | 347 Miners Row | Happy Camp | CA | 95800 | 5320.86 | WE005 | 9163725670 |
| 19 | Equal | Both | 00016 | Movies-R-U | 1772 Bridge St | Bonanza | AR | 02532 | 18735.70 | SC005 | 5018904321 |
| 20 | Equal | Both | 00017 | Movie Mania | 572 Front St | Devils Garden | FL | 01501 | 568.30 | SE003 | 4078888000 |
| 21 | Equal | Both | 00019 | It's In The C | 2005 Rt 22 | Three Brothers | AR | 62700 | 5231.00 | SC012 | 5018552500 |
| 22 | Equal | Both | 00020 | Popcorn Vid | Aramingo Place | Truth or Conseq | NM | 70800 | 486.00 | SW003 | 5059761210 |

# Filter Details



# Summary

We capture REAL data from Production

We have built a library of test cases

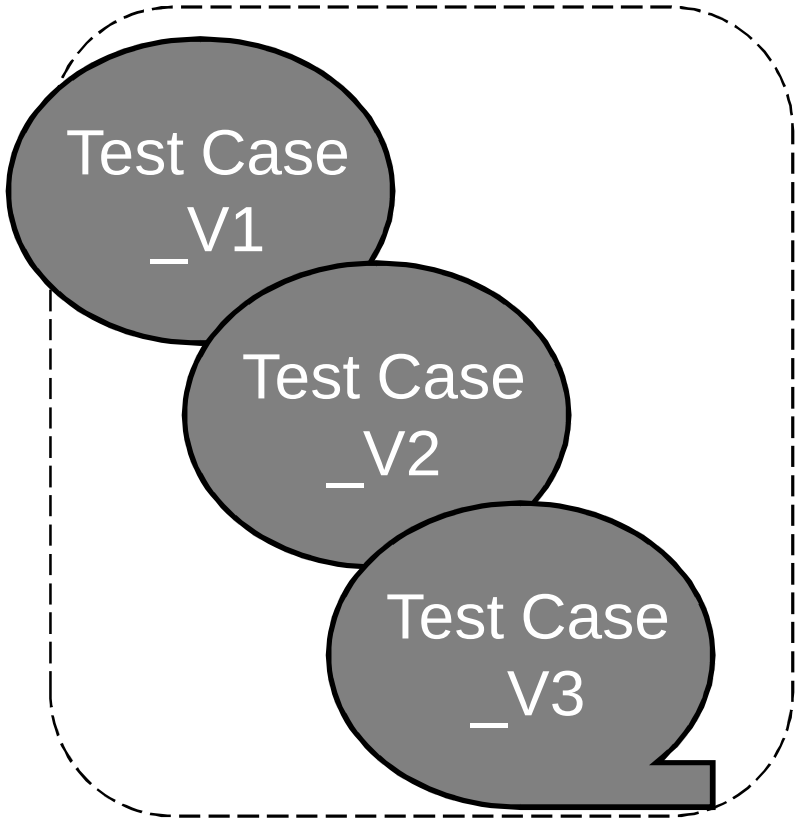
- Sets of related tables satisfying a particular condition

Test Data is re-used across test phases and iterative development versions

We can compare results across time

We have built a library of test data **results**

- Result of Process A;
- Run on Test Cases B, C, and D;
- In Version ***n***
- Test data results are compared across Versions 1 - ***n***




Test Data Library

# 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

# Effective Test Data Management for Improving Costs

Maximize allocated disk space

Reduce infrastructure costs

**Shift errors from production to test**

- An error in Production
  - Loss of business
  - Hard to find
  - Expensive to fix
- Find errors during:
  - Unit testing
  - Integration testing
  - Stress testing
  - Regression testing

# 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*

# Optim Core Solutions – Optim zOS/Optim Client Server

Optim zOS

Data Sources: DB2, IMS DB, VSAM, Seq

Environment: TSO/ISPF

Functional Elements

- MOVE
- ACCESS
- COMPARE
- MOVE for LEGACY

Optim client server

Data Sources (native): DB2 zos, U2, Informix, Oracle, Sybase, SQL Server

Data Sources (other): ADABAS, Ingres, flat, MySQL, DB2 400, Teradata, others to come

Environment: Windows (XP, 2000, 2003, Vista); Unix (Sun Solaris 8,9,10; HPUX 11,11i,12; AIX 4.2,4.3; Linux Redhat)

Functional Elements

- MOVE
- EDIT
- COMPARE
- ARCHIVE

# Supporting Enterprise Environments

**Test Smarter**

## IBM Optim

Any Questions???