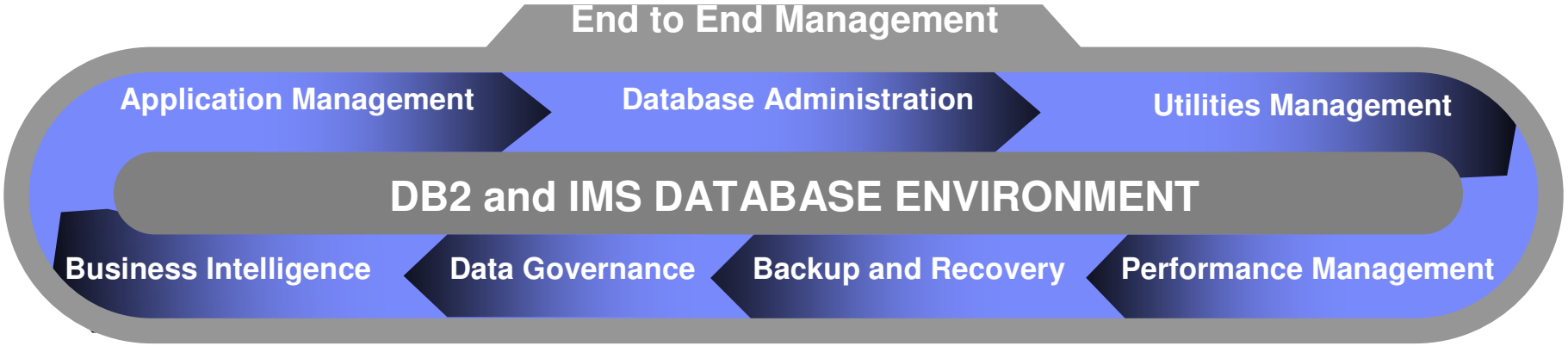
Improve Database ROI with IBM Optim Solutions

Bob Amble
rcamblej@us.ibm.com

Managing your business environment



Business Challenges

- Optimizing costs associated with maintaining existing applications
- Quickly responding to new business requirements and opportunities
- Ensuring that business and regulatory needs can be properly met
- Maximizing IT staff productivity to streamline business operations

DB2 for z/OS Tools

Making your life easier …

- ✓Provide autonomic features to add capability and simplify operations
- ✓Avoid tedious tasks and reduce errors
- ✓Preserve your investment in z/OS applications and databases
- ✓Free up valuable DBA resources to focus on business differentiators

Today: Maximize business value of System z

2000: Reduce TCO

2000-2004: Initial portfolio
- **DB2 V8 support**

2005-2008: Portfolio Expansion
- **DB2 9 support**

2008-2010: Portfolio Expansion & ROI Focus
- **DB2 X Support**

DB2 for z/OS Tools Portfolio

Application Management
- DB2 Administration Tool
- DB2 Path Checker
- DB2 Bind Manager
- DB2 Query Monitor
- DB2 SQL Performance Analyzer
- DB2 High Performance Unload
- DB2 Table Editor
- Optim Development Studio
- Optim Data Growth
- Optim Query Tuner
- Optim Test Data Management
- InfoSphere Data Architect

Utilities Management
- DB2 Utilities Suite
- DB2 Automation Tool
- DB2 Automation Toolkit SAP Edition
- DB2 Utilities Enhancement Tool
- DB2 High Performance Unload

Database Administration
- DB2 Administration Tool
- DB2 Object Comparison Tool
- DB2 Administration Toolkit SAP Edition

Performance Management
- OMEGAMON XE DB2 Performance Expert
- OMEGAMON XE DB2 Performance Monitor
- DB2 Query Monitor
- DB2 SQL Performance Analyzer
- DB2 Buffer Pool Analyzer
- DB2 Performance Toolkit SAP Edition
- Optim Query Workload Tuner
- Optim Development Studio
- Optim pureQuery Runtime

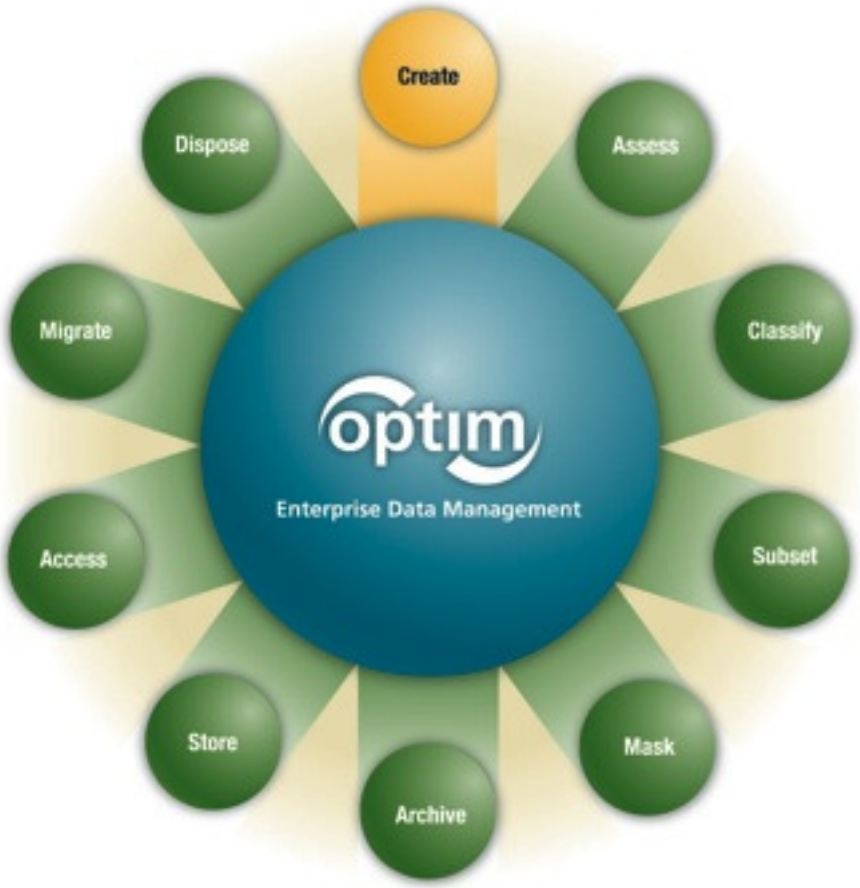
Information Integration
- InfoSphere Information Server
- InfoSphere Classic Data Event Publisher
- InfoSphere Classic Federation Server
- InfoSphere Classic Replication Server
- InfoSphere DataStage
- InfoSphere Replication Server
- InfoSphere Change Data Capture

Backup and Recovery
- DB2 Recovery Expert
- DB2 Log Analysis Tool
- DB2 Cloning Tool
- DB2 Change Accumulation Tool
- DB2 Object Restore Tool
- Application Recovery Tool for IMS and DB2 Databases

Data Governance
- Optim Data Growth
- Optim Data Privacy
- Optim Test Data Management
- DB2 Audit Management Expert
- Data Encryption for DB2 and IMS

Business Intelligence
- Cognos for Linux on System z
- DataQuant
- QMF

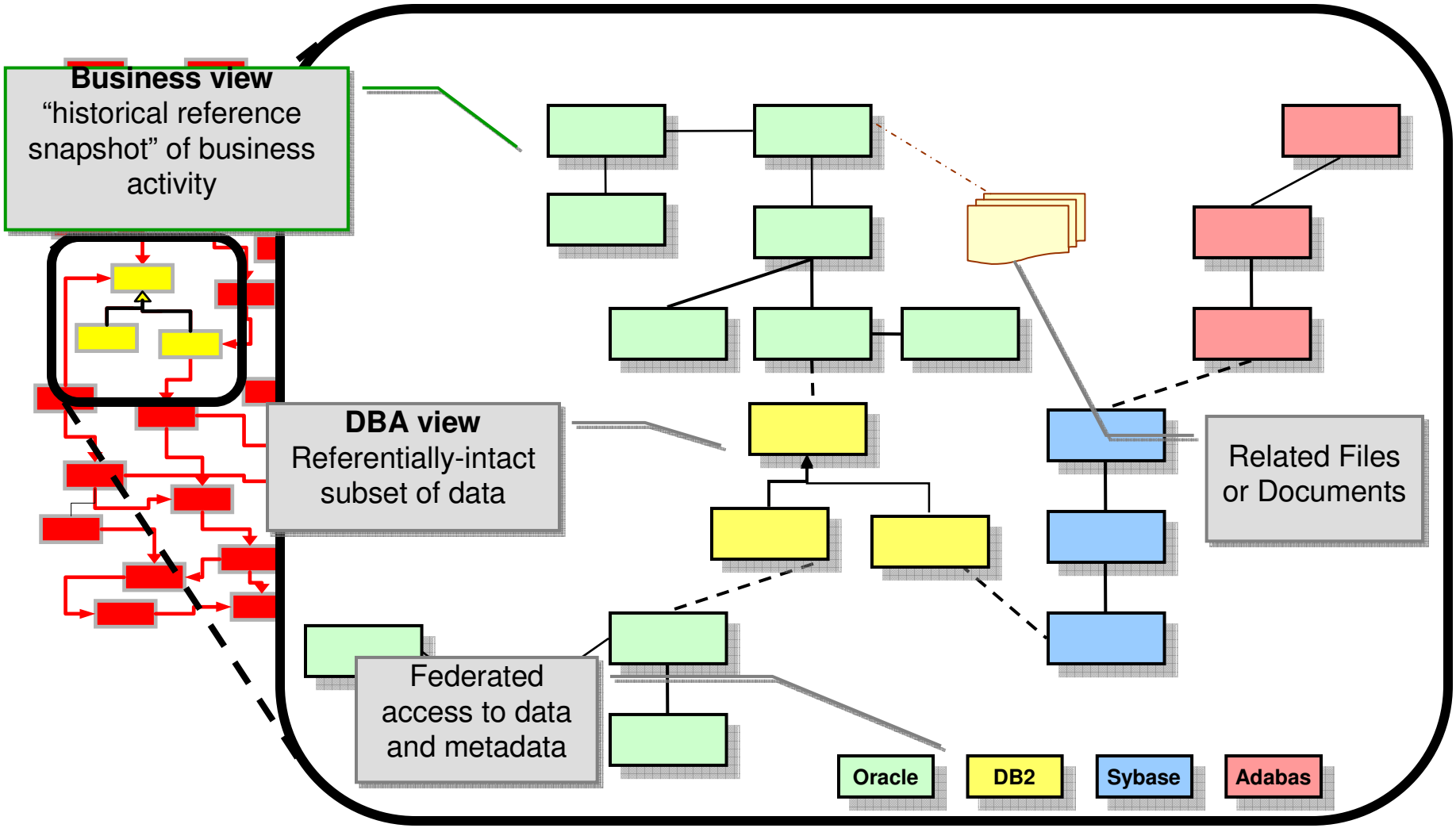
IBM Optim™ Solutions



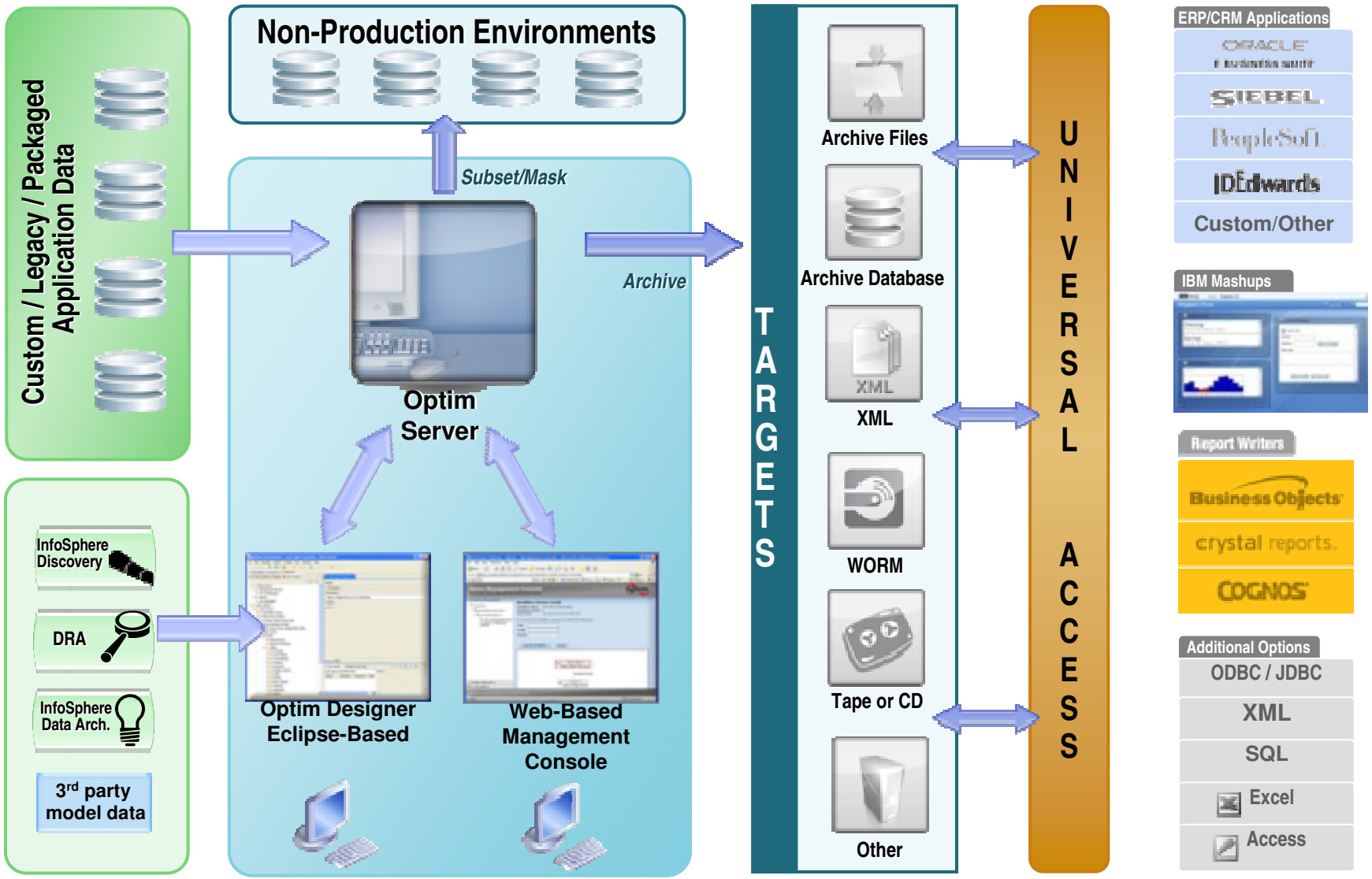
Recognized by Gartner, IDC, META as industry leader.
Most recent Gartner Archive report - 46% market share.

- **IBM Optim (formerly Princeton Softech)**
  - First product was for mainframe DB2 in 1989
  - Complements IBM's FileNet, Encryption, Cognos, Tivoli, Rational products
  - 2400 clients worldwide; 50% of Fortune 500
- **IBM Optim Solutions Address**
  - **Data Growth (Archiving)**
    - Control data growth
    - Performance
    - Storage Savings – Re-use
    - Version Upgrades/Migrations
    - Lower TCO
    - Discovery - Support retention compliance
  - **Decommissioning**
    - Enable application retirement, cost avoidance
  - **Optim Test Data Management (TDM)**
    - Create targeted, right sized test environments
    - Speed iterative testing processes
    - Improve application quality
  - **Optim Data Privacy (DP)**
    - Mask confidential data
    - Comply with privacy policies

Building Blocks of Optim - Complete Business Object



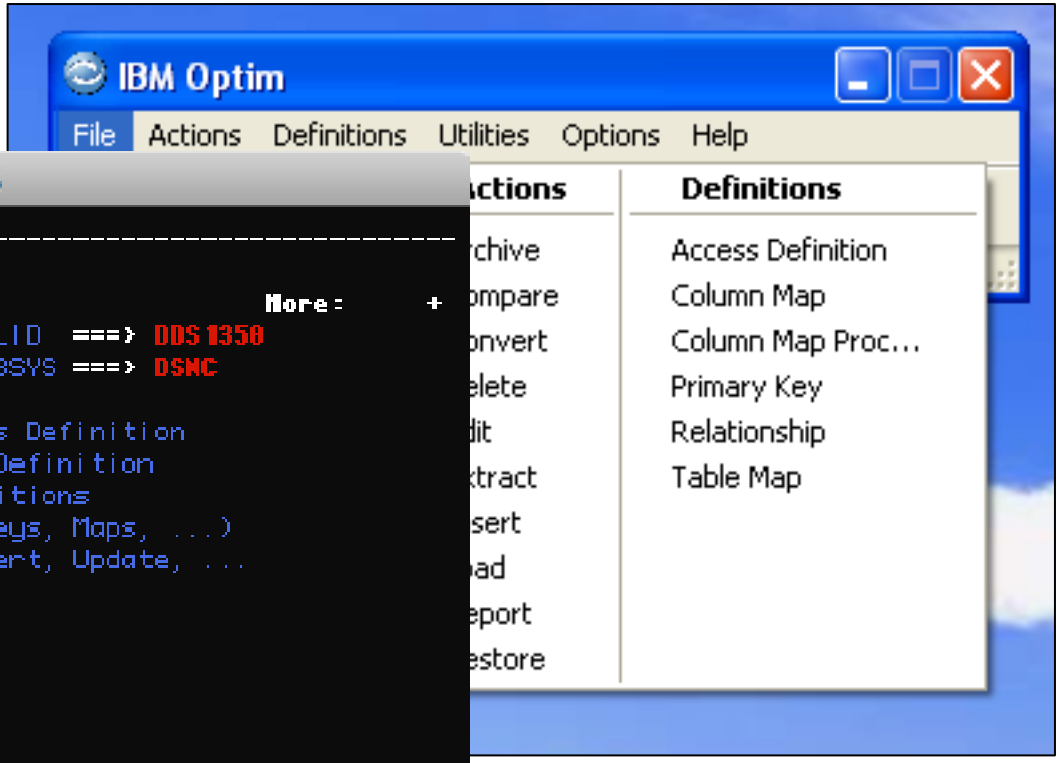
IBM Optim Solution Overview

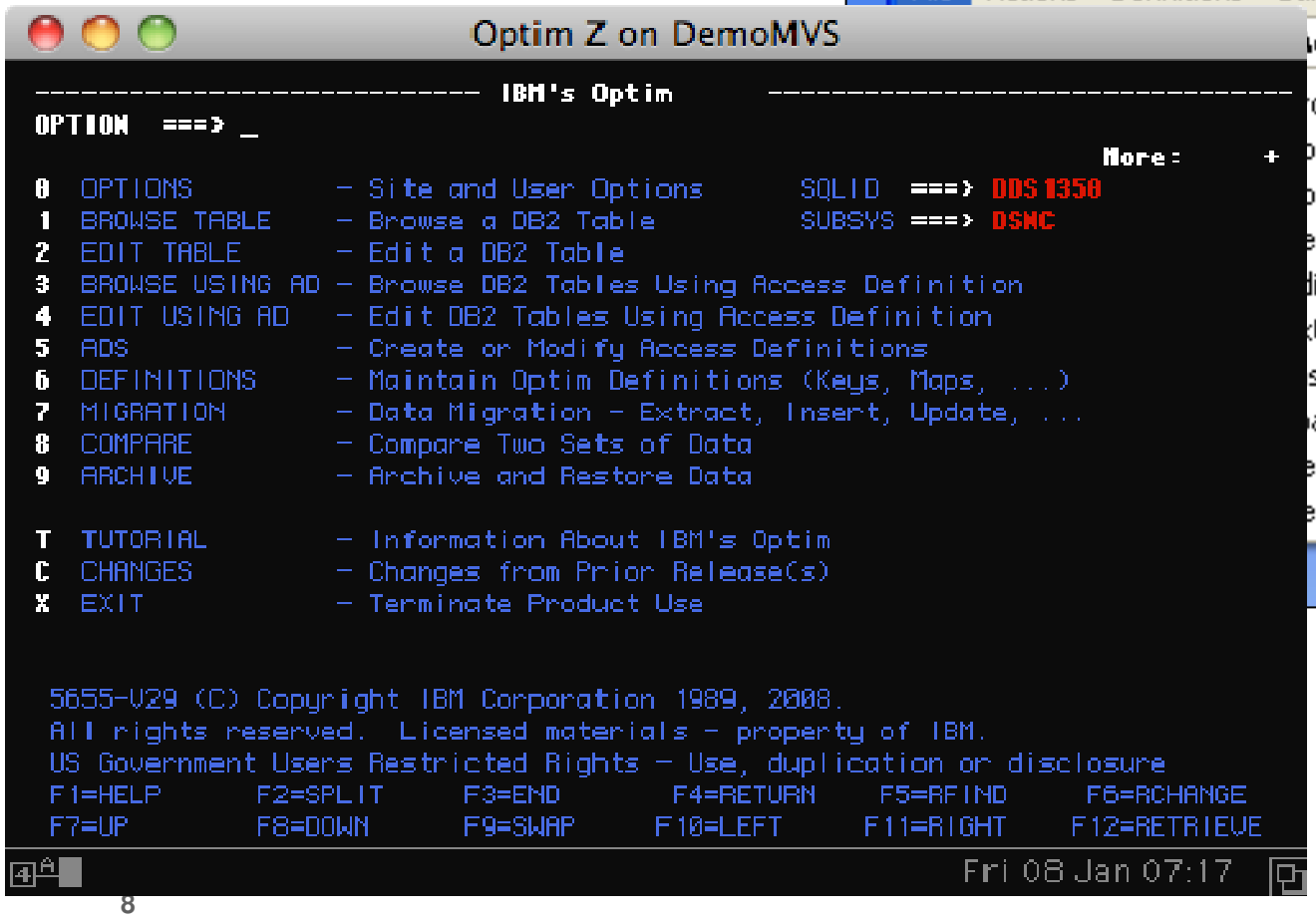


IBM Optim platforms

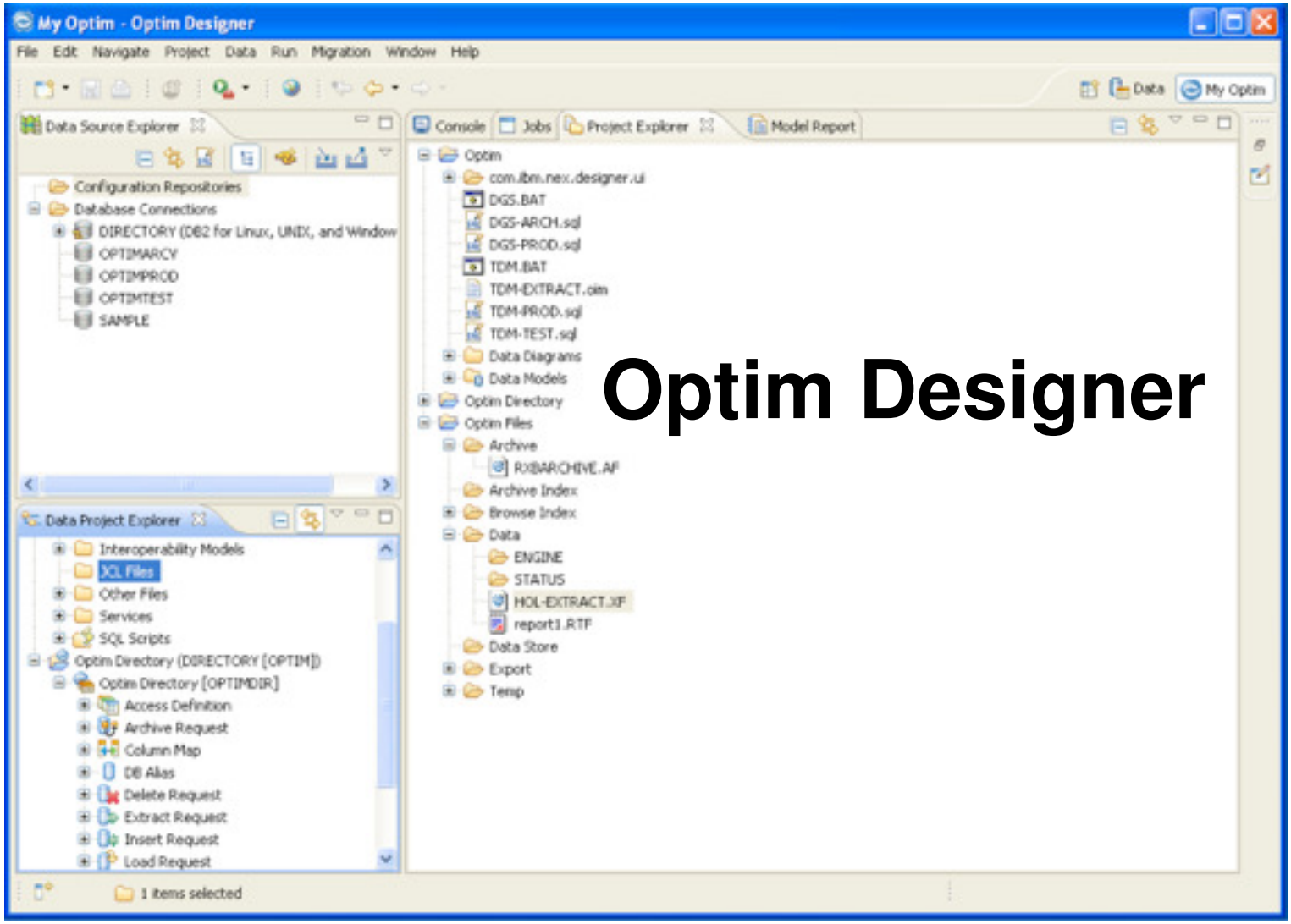
Optim for z/OS
(you can guess)

Optim Distributed
(Linux, Unix Windows)





Optim User Interface - Next Generation



IBM Optim Data Growth

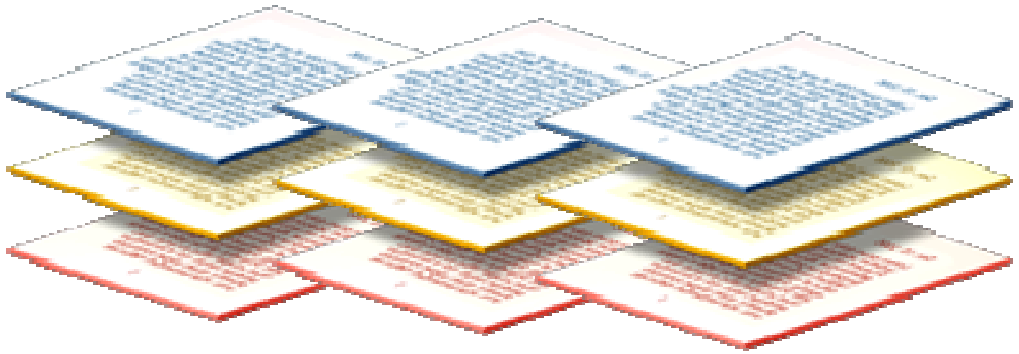
The Problem

- **Mergers & acquisitions**
- **Organic business growth**
  - eCommerce
  - ERP/CRM
- **The digital revolution**
- **Records retention**
  - Basel II
  - SOX
  - Euro-SOX
- **Data multiplier effect**
- ***Forrester estimates that 85%***
- ***of data stored in databases is inactive***

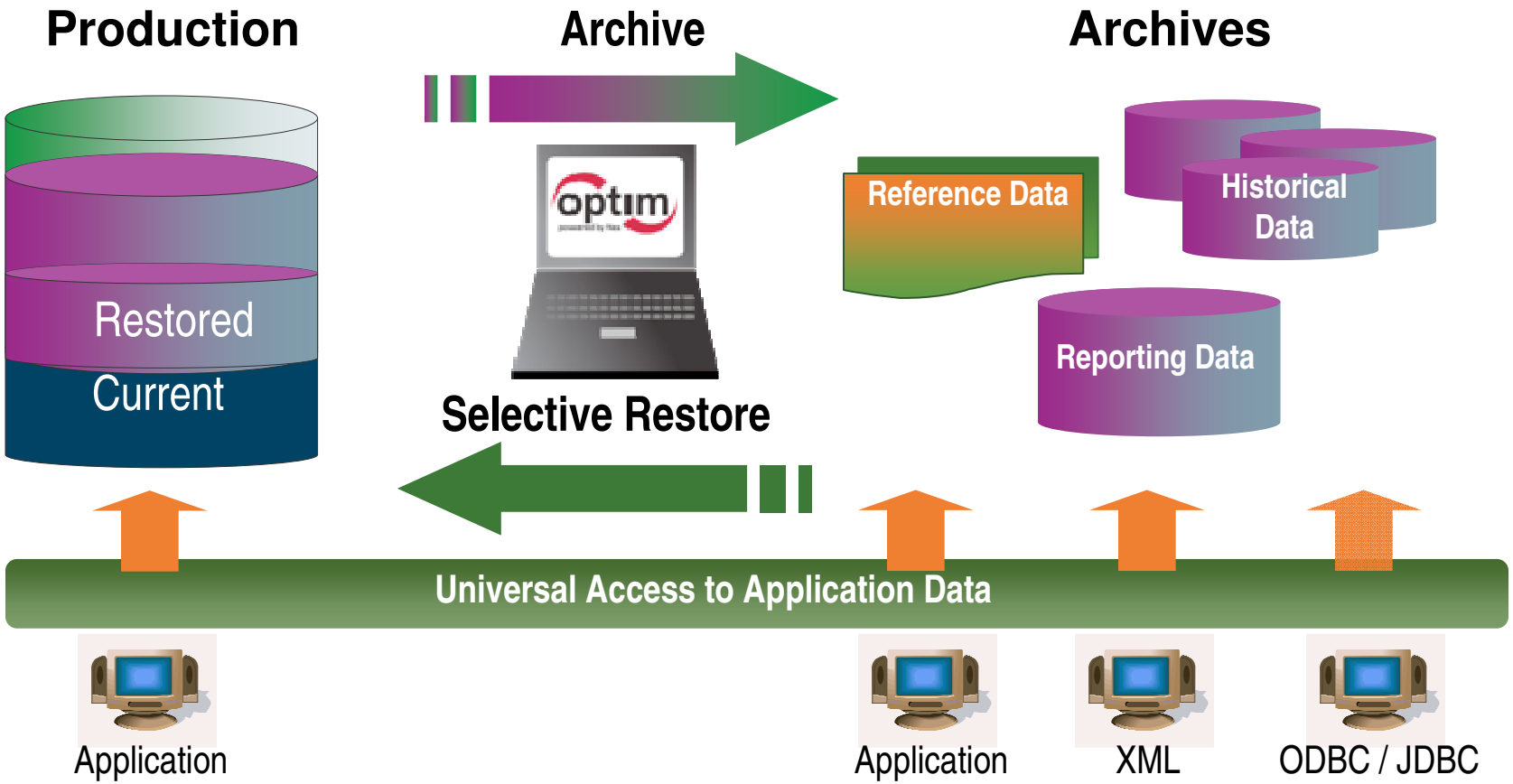
* Source: Noel Yuhanna, Forrester Research, Database Archiving Remains An Important Part Of DBMS Strategy, 8/13/07

A Definition of Archiving:

Archiving is an intelligent process for placing inactive or infrequently accessed data that still has ***business value*** on ***the right tier*** of storage, with the ***right class of service***, while maintaining ***search and retrieval*** capability during a specified ***retention period***.

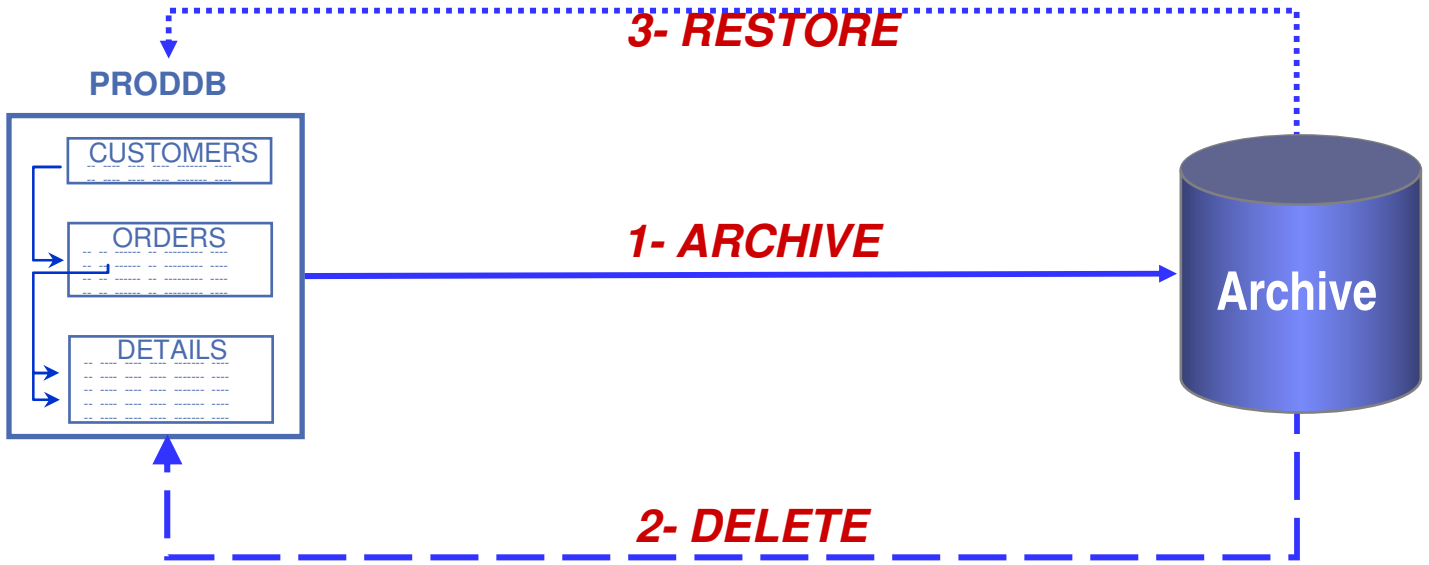


IBM Optim Data Growth: Archiving



- Complete Business Object provides historical reference snapshot of business activity
- Storage device independence enables Information Lifecycle Management
- Immutable file format enables data retention compliance

Process Steps



- ***Archiving Steps***

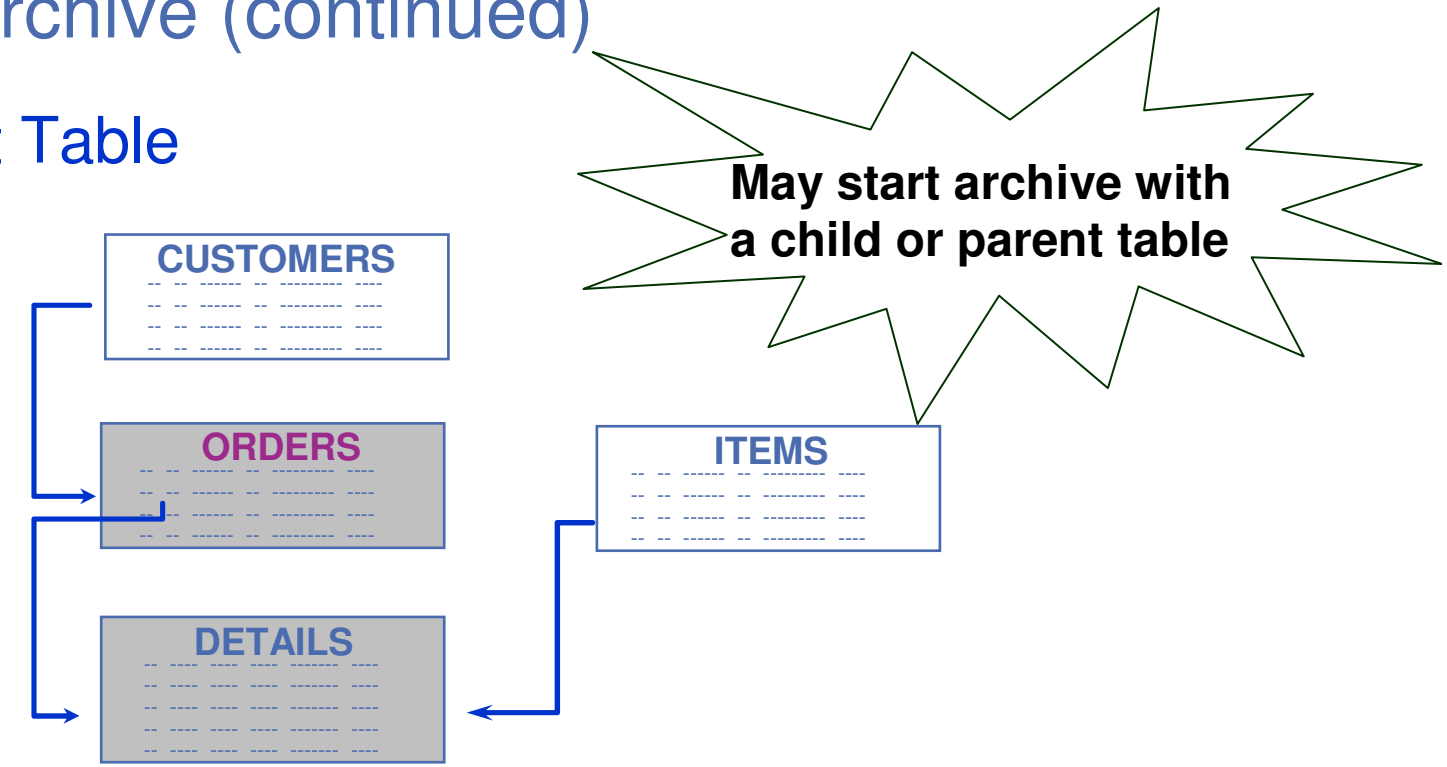
1. Identify the data to archive
2. Define the data to delete
3. Select Archive File storage
4. Choose a delete method
5. Run Archive Request
6. Create Delete Request if deferred
7. Run Delete Request if deferred

Restoring Steps

- Locate Archive File
- Create a Restore Request
- Run Restore Request

Defining the Archive (continued)

The Start Table

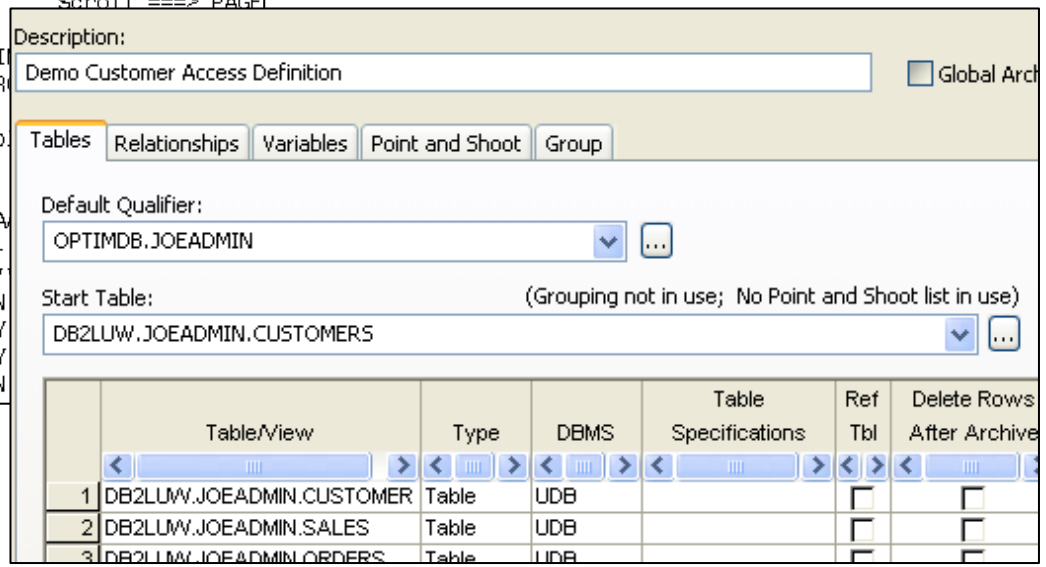


- Archive: All ORDERS older than four years and the related data in the other tables
- Delete: From ORDERS and DETAILS only

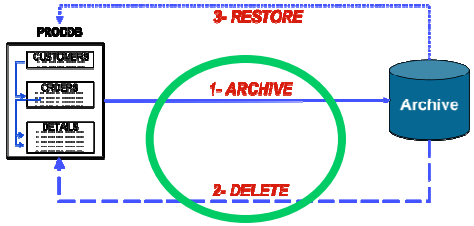
Defining the Archive (continued)

The Table List

```
Command ===>                                                  Scroll ===> PAGE

 Primary : COL, SEL, SQL, REL, POINT, GET TABLES RELATED, I
    Line : COL, SEL, SQL, ALL, GR(A), GP(A), GC(A), EXP, AR

Default Creator ID ===> PSTDEMO                          Tab
Start Table        ===> CUSTOMERS

                                                    Ref
Cmd   Status      (CreatorID.)Table/View Name       Tbl   DA
--- ----------- --------------------------------   ---   --
*** ********************************* TOP ***************
___             CUSTOMERS                                 N
___             ORDERS                             N      Y
___             DETAILS                            N      Y
___             ITEMS                              N      N
```



- Identify the **Start Table**
- Populate list with the **RELATED** functions
- Include selection criteria
- Indicate which tables will have rows deleted
- Specify Archive actions and indexes



Defining the Archive (continued)

Relationship Usage

```
Command ===>

Q1: If a Child Row is Included, Include its Parent Row to
Q2: If a Parent Row is Included to Satisfy any RI Rule, I

           Q Q Child
Cmd Status 1 2 Limit      Parent Table          Child Tabl
--- ------ - - ----- --------------------- -------------
*** ********************************* TOP ***************
___ SELECT Y N       CUSTOMERS             ORDERS
___ UNSEL  Y N       ITEMS                 DETAILS
___ SELECT Y N       ORDERS                DETAILS
___ SELECT Y N       ITEMS                 PARTS
*** ******************************** BOTTOM ***************************************
```

Tables | Relationships | Variables | Point and Shoot

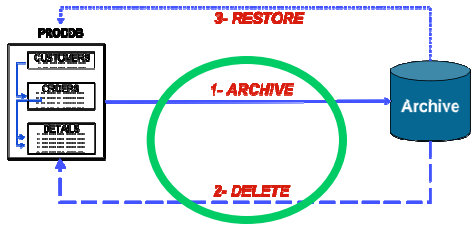
Option for each Relationship:

(1) If a child row is included, include its parent row to satisfy the RI rule

(2) If a parent row is included to satisfy any RI rule, include all child

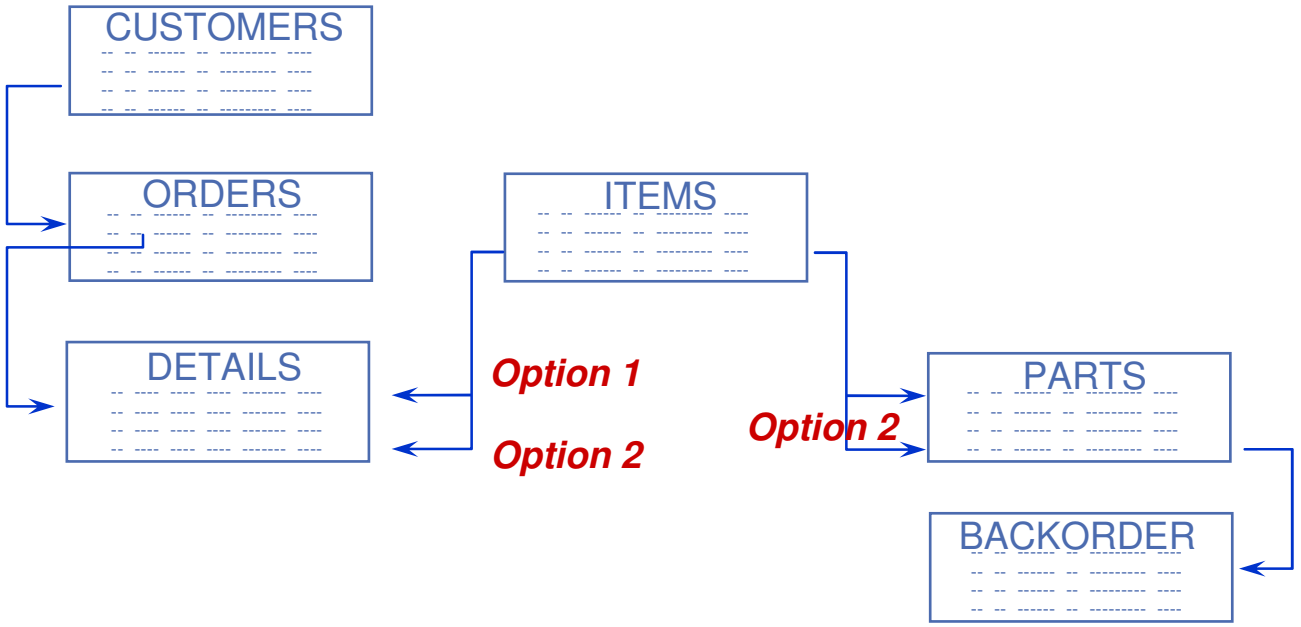
	Status	Select	Options (1)	Options (2)	Child Limit	Parent Table	Child Table	Constraint	Ty
1		☑	☑	☐		CUSTOMERS	ORDERS	RCO	UDB
2		☑	☑	☐		ITEMS	DETAILS	RID	UDB
3		☑	☑	☐		ORDERS	DETAILS	ROD	UDB
4		☑	☑	☐		SALES	CUSTOMERS	RSC	UDB

- **Select relationship paths**
  - Defined in the RDBMS catalog or Optim Directory
- **Designate relationship traversal**
- **Limit number of child rows archived**
- **Specify Access Method / Key Lookup Limit**



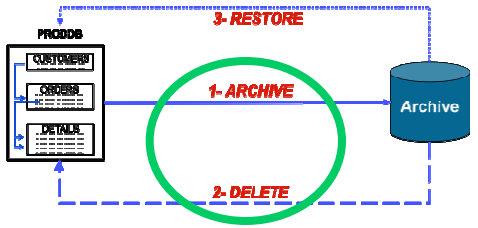
Defining the Archive (continued)

Relationship Traversal



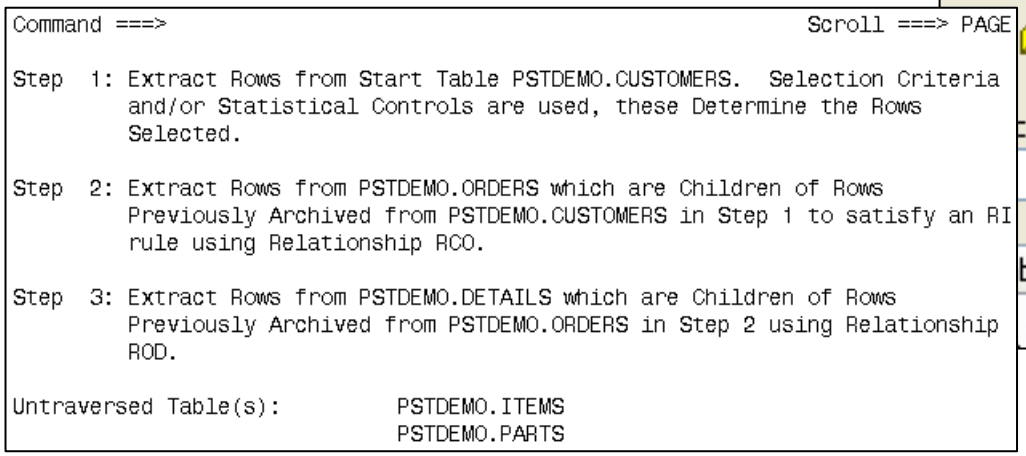
- **Option 1:** **Only ITEMS that are parents of DETAILS**

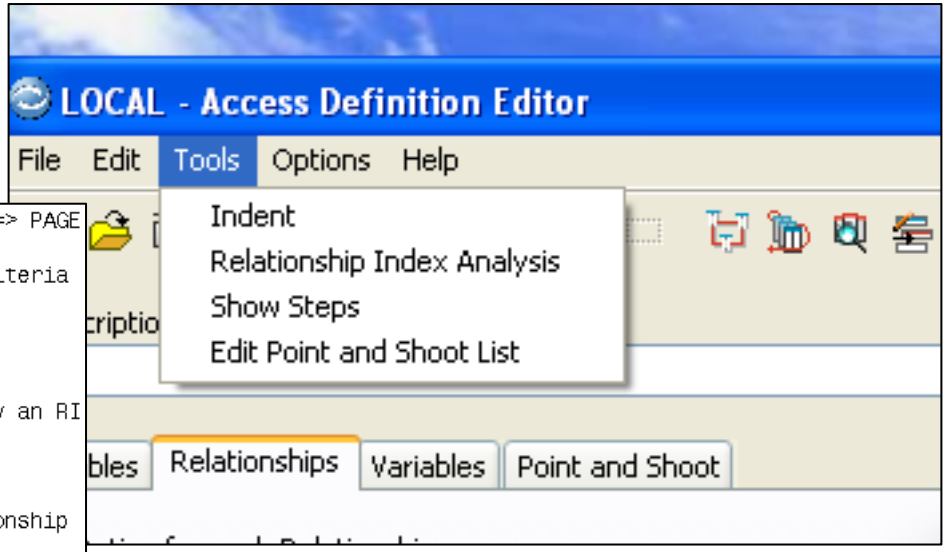
Option 2: All other DETAILS for those ITEMS ...
Each of the PARTS for those ITEMS



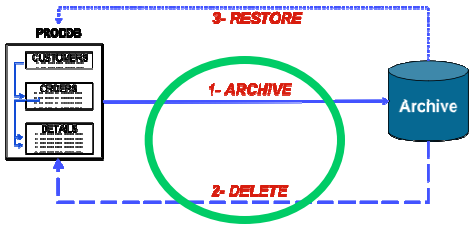
Defining the Archive (continued)

Show the Archive Steps



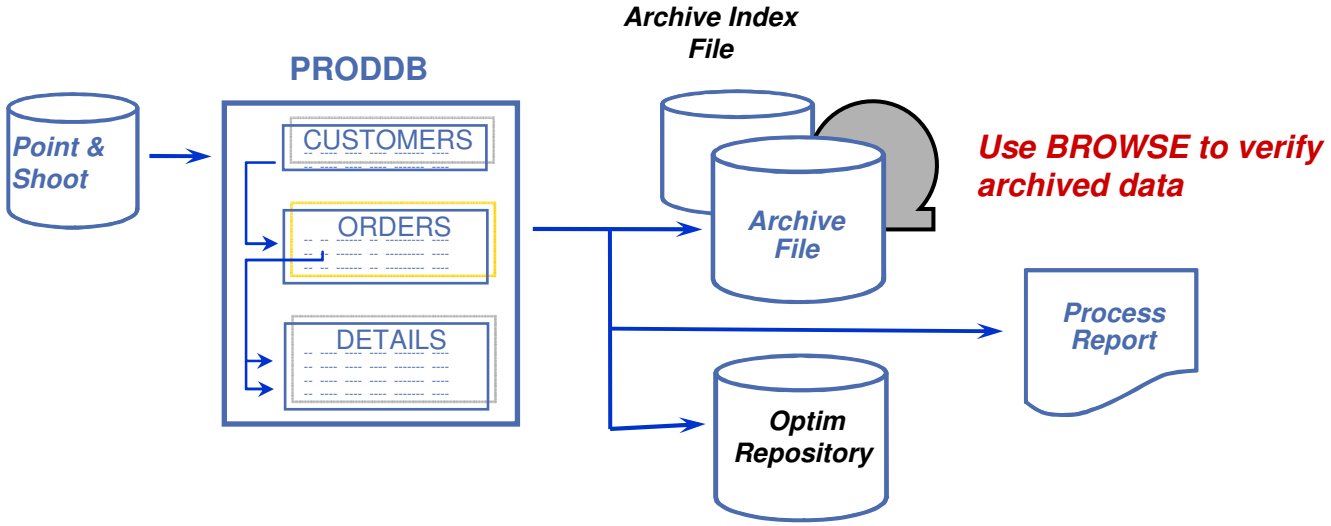


- Steps required to perform archive
- Cycles processed
- Untraversed tables

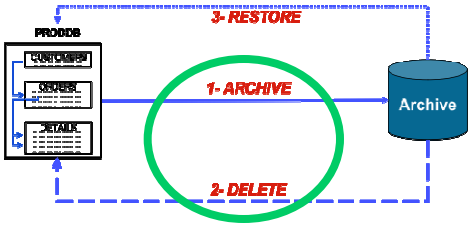


Defining the Archive (continued)

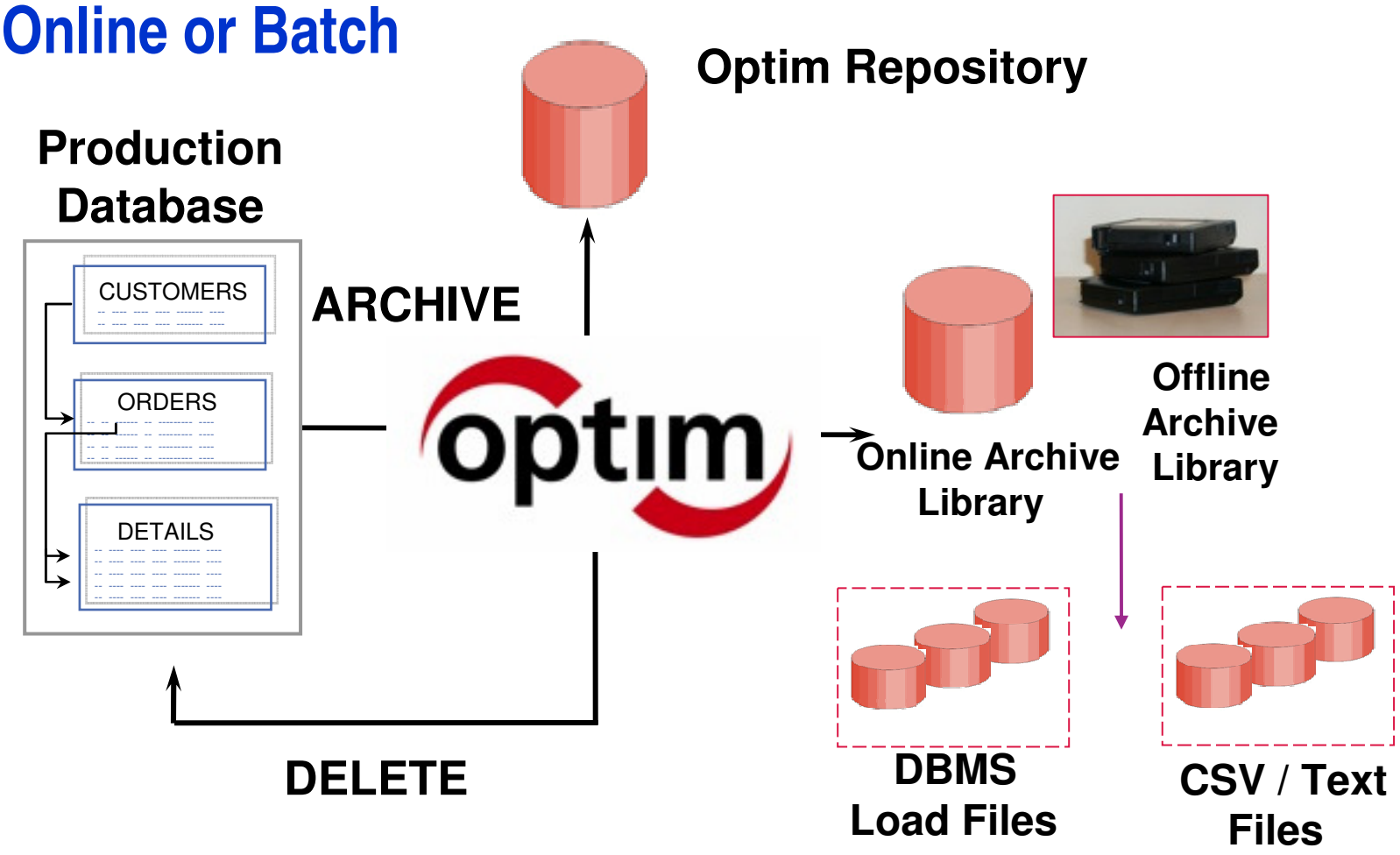
Archive Parameters



- Archive from source tables using DB2 High Performance Unload
- Archive both data and object definitions
- Execute Online or Batch

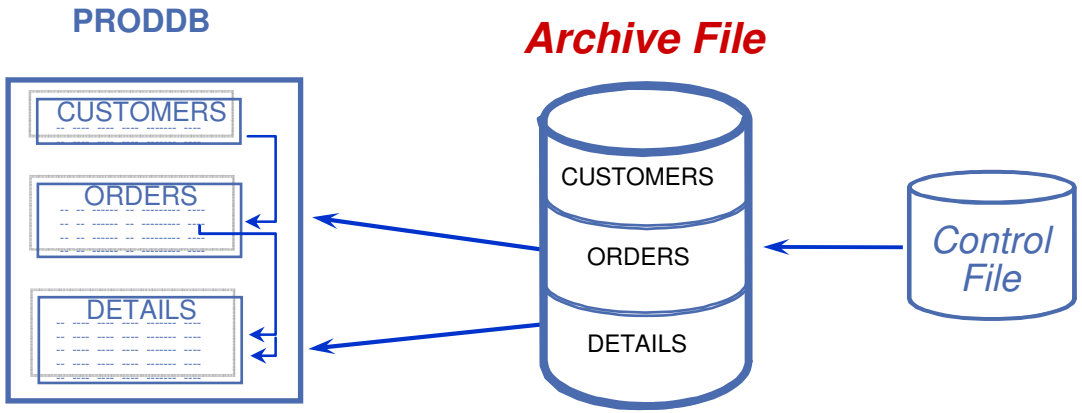


Run the Archive Request



Archive Process

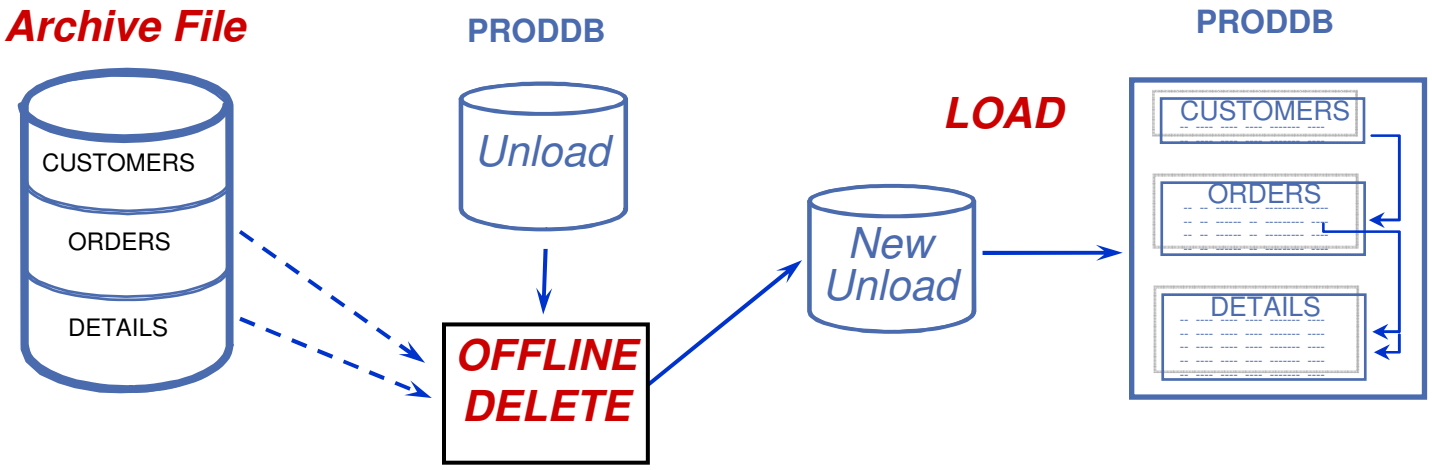
Delete the Archived Data



- Delete is automatic after successful archive OR can be deferred post archive verification
- Delete specifications define which data to delete
- Control File enables Retry/Restart of delete

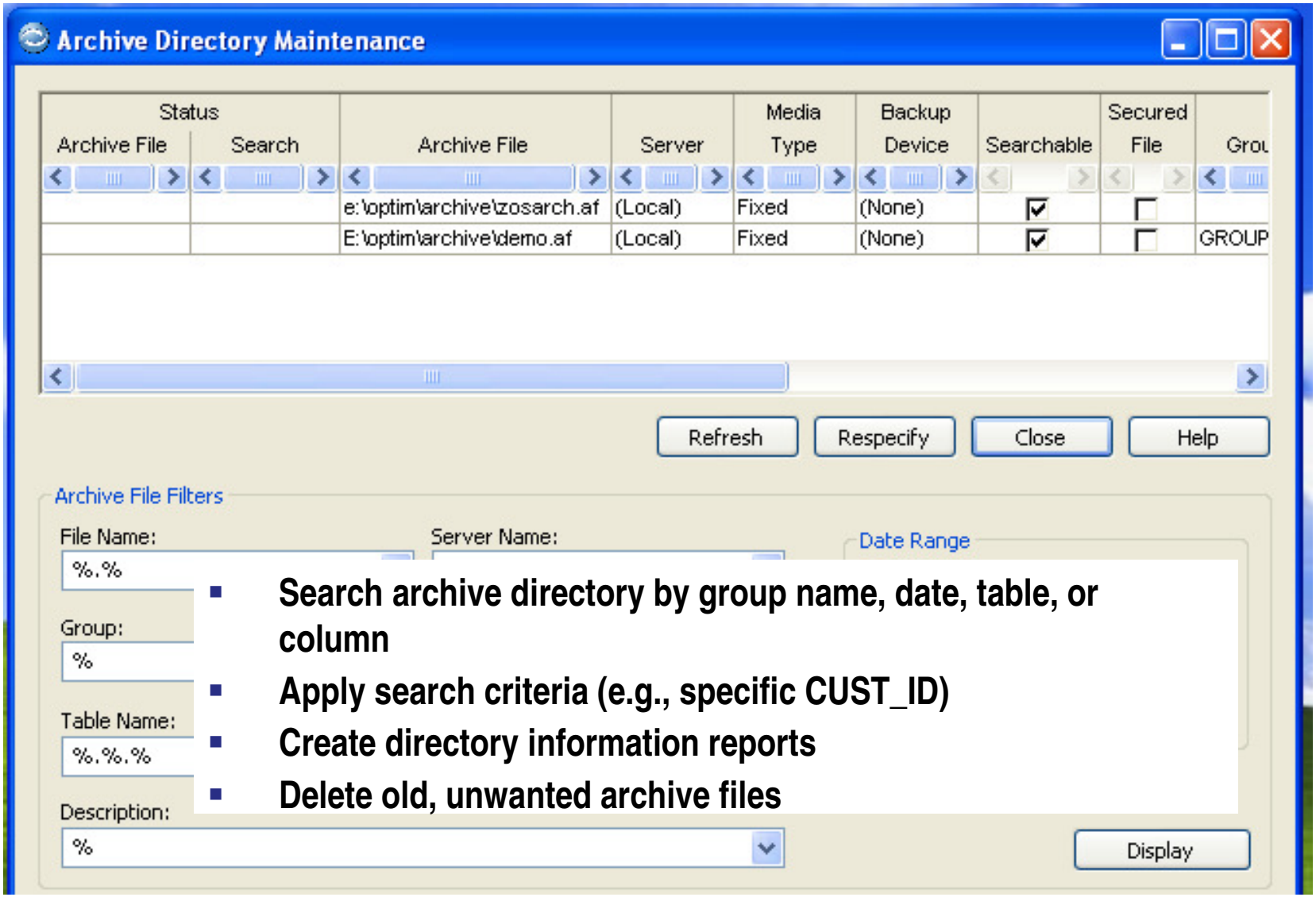
Archive Process (continued)

Delete the Archived Data - Offline Delete Method



- **Incorporate delete into normal database maintenance procedures**
- **Delete specifications define which data to delete**
- **Eliminates impact of logging during delete**

The Archive Directory - Managing Your Archived Data



The Archive Directory (continued)

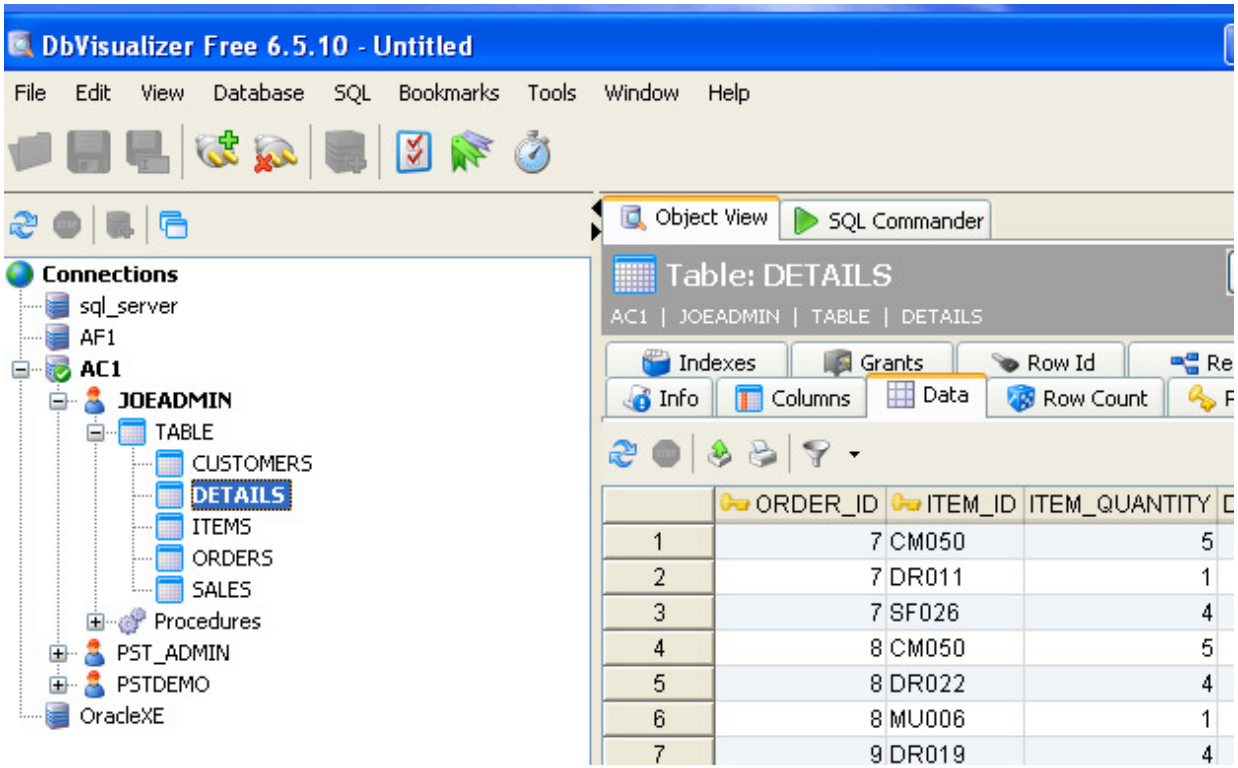
Archive Indexes

- **Enable rapid searches for archived data**
- **Defined in the Access Definition or can be added later**
- **Two index types:**
  - Sparse: only high/low column values are stored in the Archive Directory
    - Useful to locate candidate archive files in directory during search
  - Dense: all column values are stored in a file pointed to by the Archive Directory
    - Useful to speed searching archive for a particular record

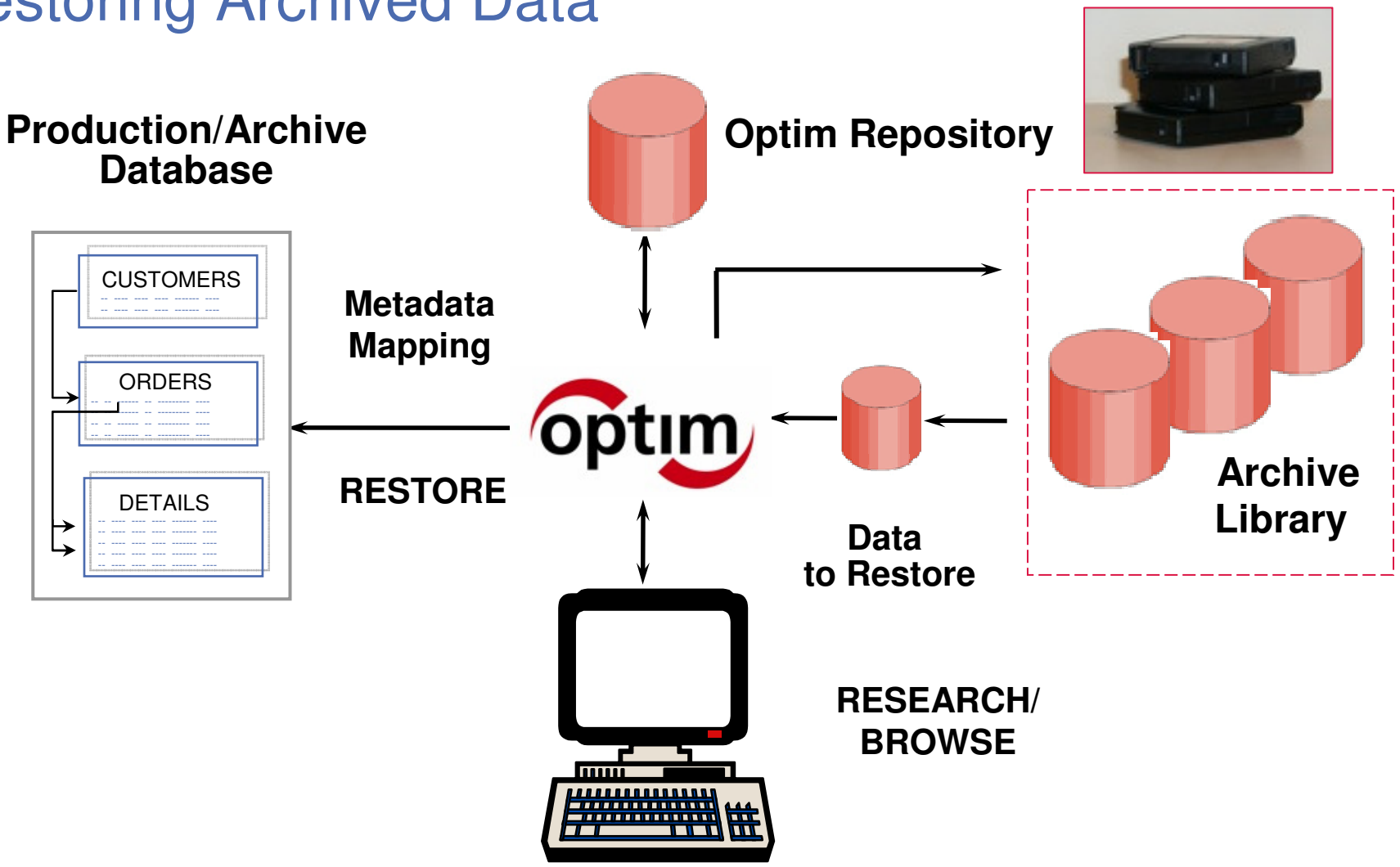
Browsing the Archive Files

Reporting Options

- **Convert Archive file to CSV file for input to other reporting programs or applications**
- **ODBC / JDBC access via Open Data Manager (ODM)**



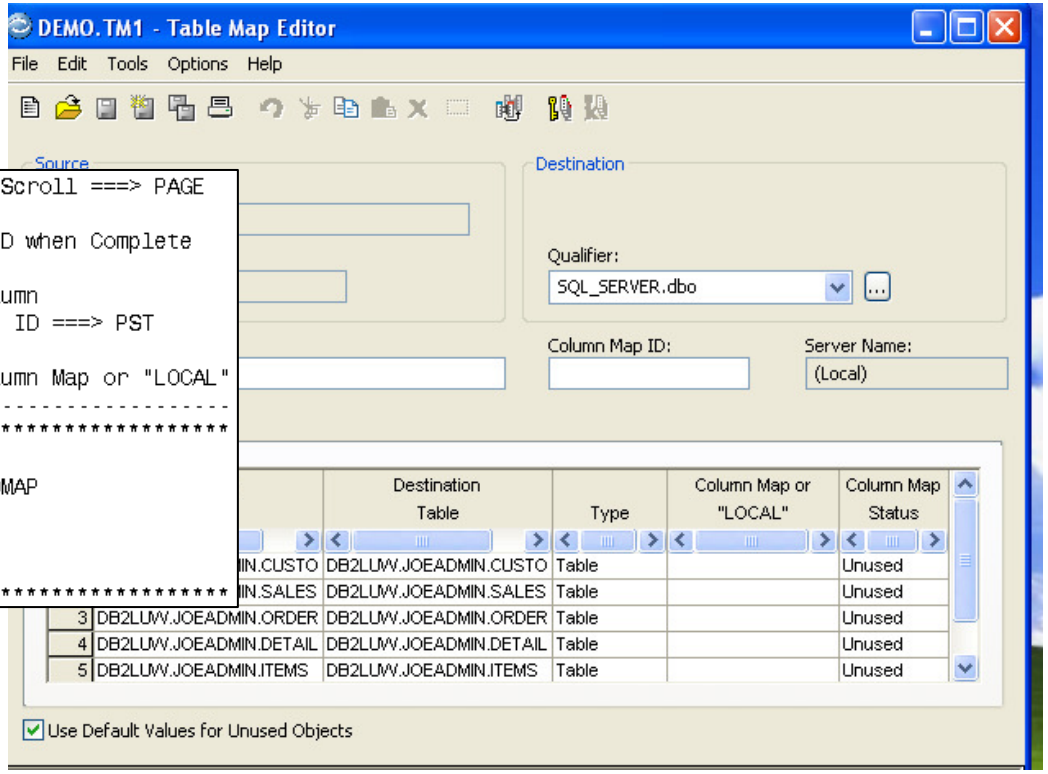
Restoring Archived Data



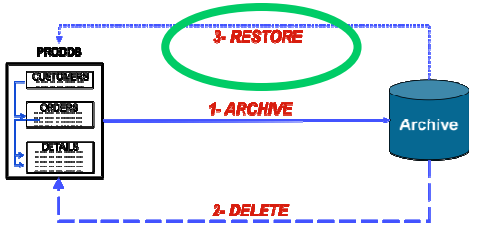
Restore Archived Data

Table Map

```
Command ===>                                                  Scroll ===> PAGE

Available Commands: APPLY, SAVE, LIST, MAP, POPULATE, REL, END when Complete
Destination May be any DB2 Tables or Views
                                                              Column
   Src CID: PSTDEMO     Dest CID ===> PSTDEMO2                Map ID ===> PST

   Archive Tables         Destination Table Name   Type    Column Map or "LOCAL"
------------------- ------------------------- ------- ----------------------
*********************************** TOP ***************************************
CUSTOMERS             CUSTOMERS                 TABLE
ORDERS                ORDERS                    TABLE    ORDMAP
DETAILS               DETAILS                   TABLE
ITEMS                 PSTTEMP.ITEMS             UNKNOWN
PARTS                                           UNUSED
******************************** BOTTOM ***************************************
```



- **Map unlike table names, qualifiers**
- **Exclude individual tables from restore**
- **Can be saved in Optim Directory**



Restore Archived Data (continued)

Column Map

Literals

Special Registers

Expressions

Default Values

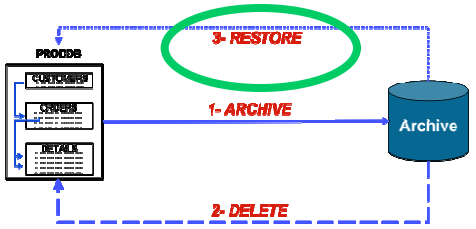
User exits

	Source Column	Source Data Type	Destination Column	Destination Data Type	
1	CUST_ID	CHAR(5)	CUST_ID	CHAR(5)	Equal
2	CUSTNAME	CHAR(20)	CUSTNAME	CHAR(20)	Equal
3	ADDRESS	VARCHAR(50)	ADDRESS	VARCHAR(50)	Equal
4	CITY	VARCHAR(15)	CITY	VARCHAR(15)	Equal
5	STATE	CHAR(2)	STATE	CHAR(2)	Equal
6	ZIP	CHAR(5)	ZIP	CHAR(5)	Equal
7	EMAIL	CHAR(70)	EMAIL	CHAR(70)	Equal
8	CCN	CHAR(19)	CCN	CHAR(19)	Equal
9	YTD_SALES * 1.2		YTD_SALES	NUMERIC(7,2)	Numeric Expression
10	EXIT PSTEXIT		SALESMAN_ID	CHAR(6)	Exit
11	PHONE_NUMBER	CHAR(10)	PHONE_NUMBER	CHAR(10)	Equal

```
Command ===>

     ----------PSTDEMO.ORDERS-----
 Cmd     Source Column      Data Ty
 --- -------------------- -------
 *** ********************************
 ___ ORDER_ID               DEC(5,0
 ___ CUST_ID                CH(5)
 ___ ORDER_DATE             DATE
 ___ CURRENT_TIMESTAMP
 ___ FREIGHT_CHARGES        DEC(4,2)
 ___ 'UNKNOWN'
 ___ ORDER_POSTED_DATE      TIMESTAMP
 ___
 *** ********************************
```

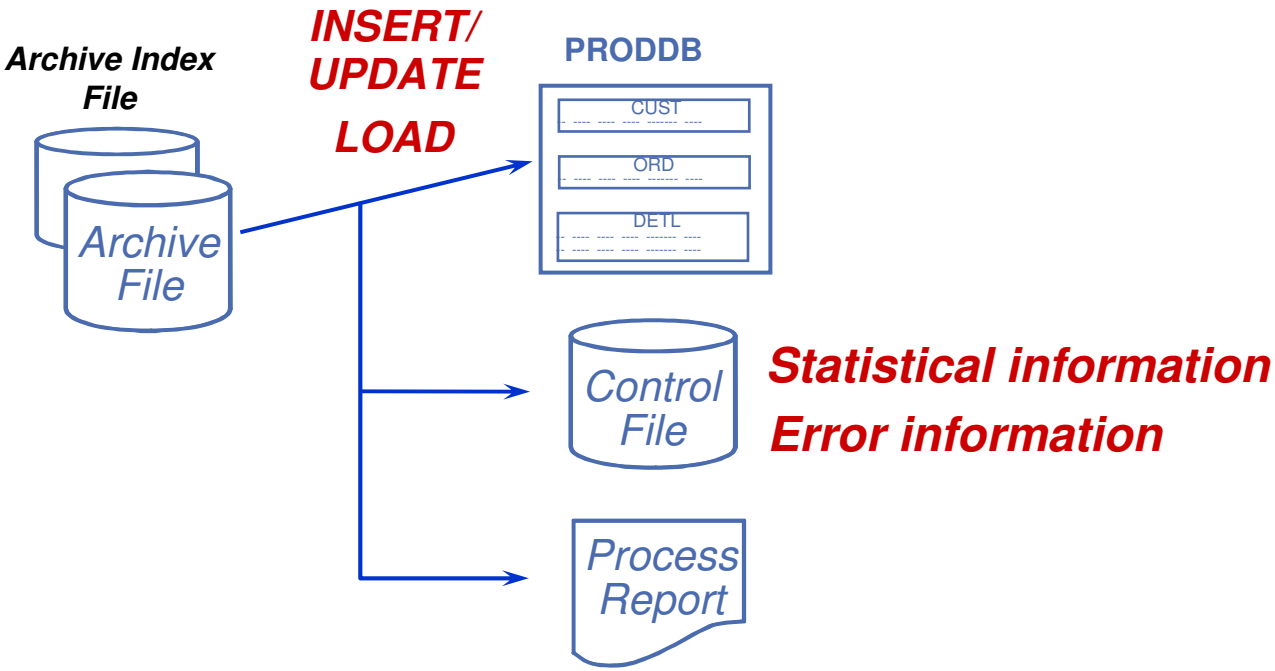
```
4 ORDER_TIME          TIME       SPC REG
5 FREIGHT_CHARGES     DEC(6,2)   MAPPED
6 ORDER_SALESMAN      CH(6)      LITERAL
7 ORDER_POSTED_DATE   TIMESTAMP  EQUAL
8 ORDER_SHIP_DATE     CH(8)      NOTUSED
*** BOTTOM ***********************************
```

- Ø **Map unlike column names**
- Ø **Datatype conversions**
- Ø **Populate new destination columns**



Restore Archived Data (continued)

Control File

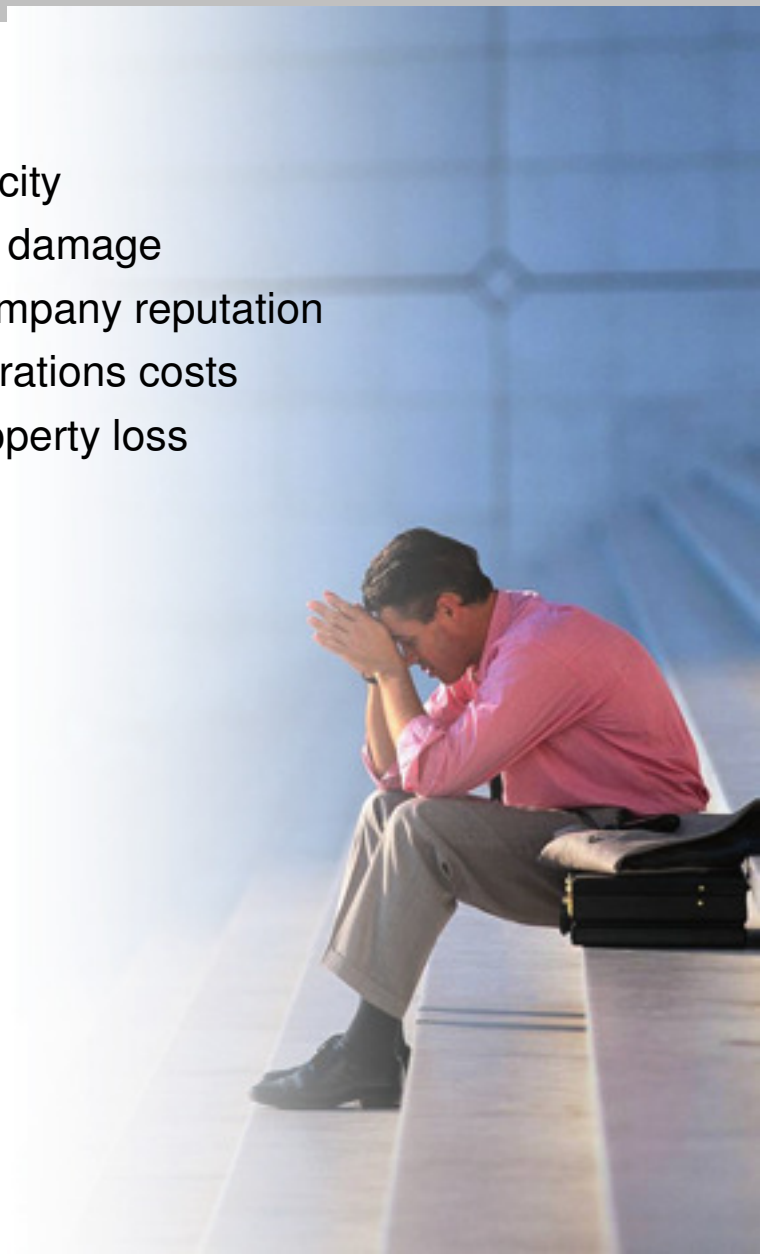


- **If errors occur during RESTORE:**
  - BROWSE the control file for error information
  - RETRY/RESTART the RESTORE process

IBM Optim Test Data Management / Data Privacy

What's at Stake?

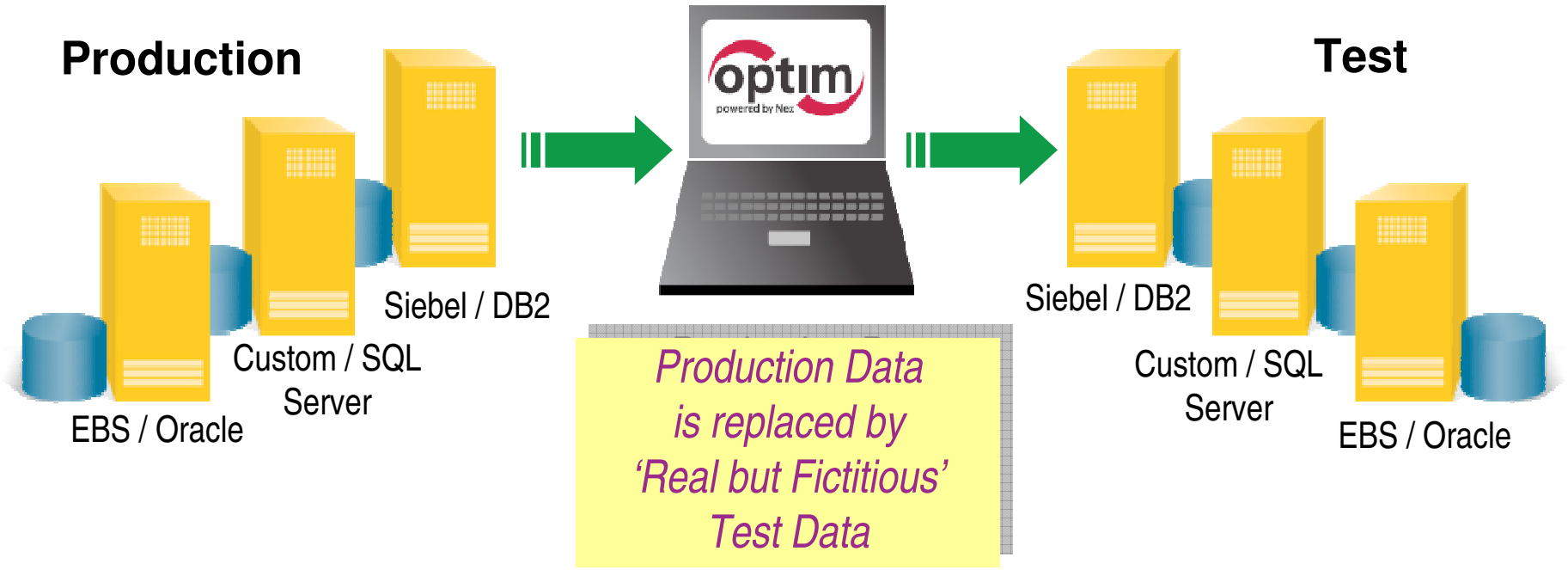
- **Fines and penalties**
- **Lawsuits**
- **Loss of customer loyalty**
- **Loss of revenue**
- **Share price erosion**
- Negative publicity
- "Brand equity" damage
- Damage to company reputation
- Increased operations costs
- Intellectual property loss



Where Data Theft Happens

- Data mistakenly left behind
- Laptops
- Hard drives
- Thumb drives
- Data exposed in testing and training
- Outsourcers
- Internal employees
- Application breaches

IBM Optim Data Privacy Option

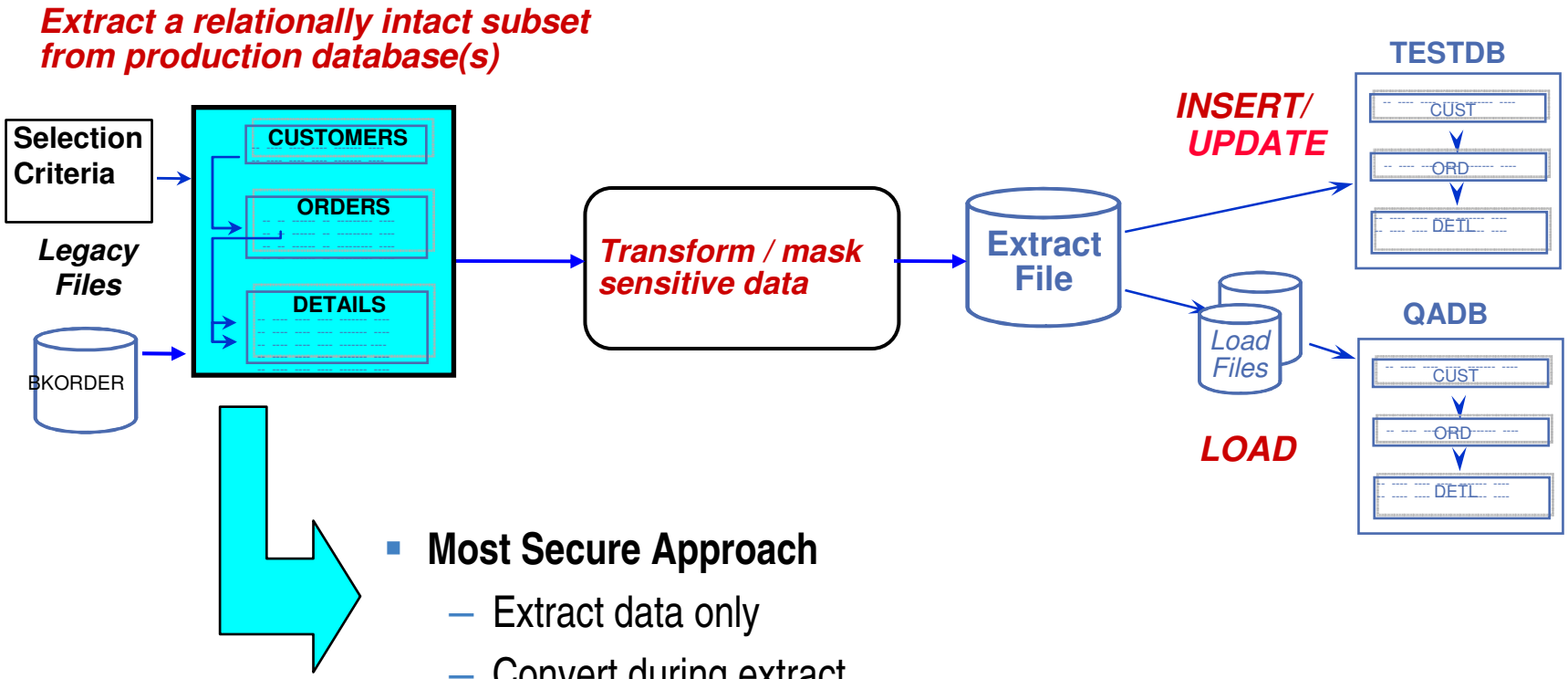


- Substitute confidential information with fictionalized data
- Deploy multiple masking algorithms
- Provide consistency across environments and iterations
- Enable outsourced development / testing
- Protect private data in non-production environments

De-Identifying Test Data

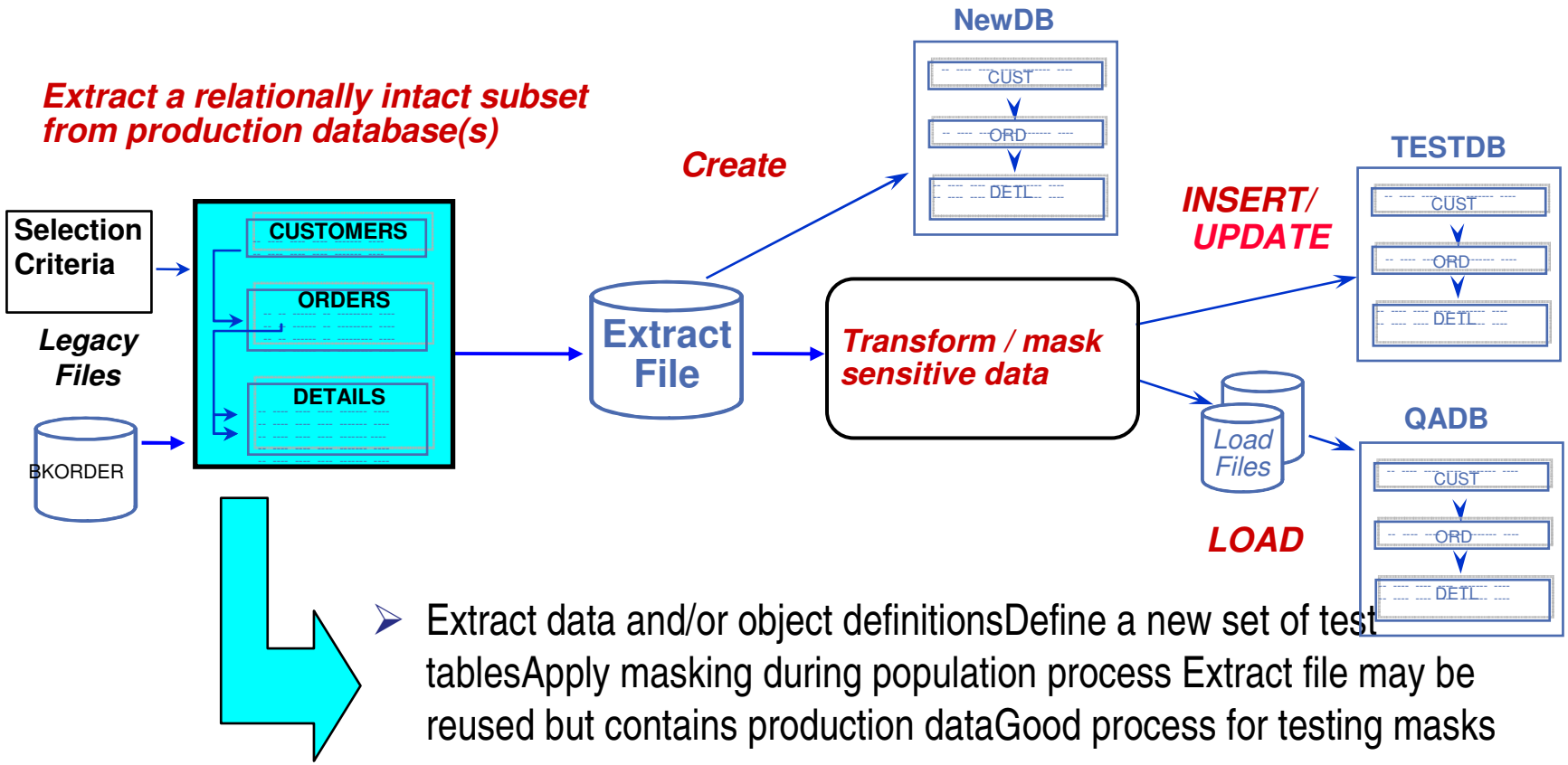
- **Removing, masking or transforming elements that could be used to identify an individual**
  - Name, address, telephone, SSN etc.
- **Masked Data is longer confidential; therefore acceptable to use in open test environments**
- **Masked or transformed data must be appropriate to the context**
  - Consistent formatting (alpha to alpha)
  - Within permissible range of values

Data Privacy in Application Testing



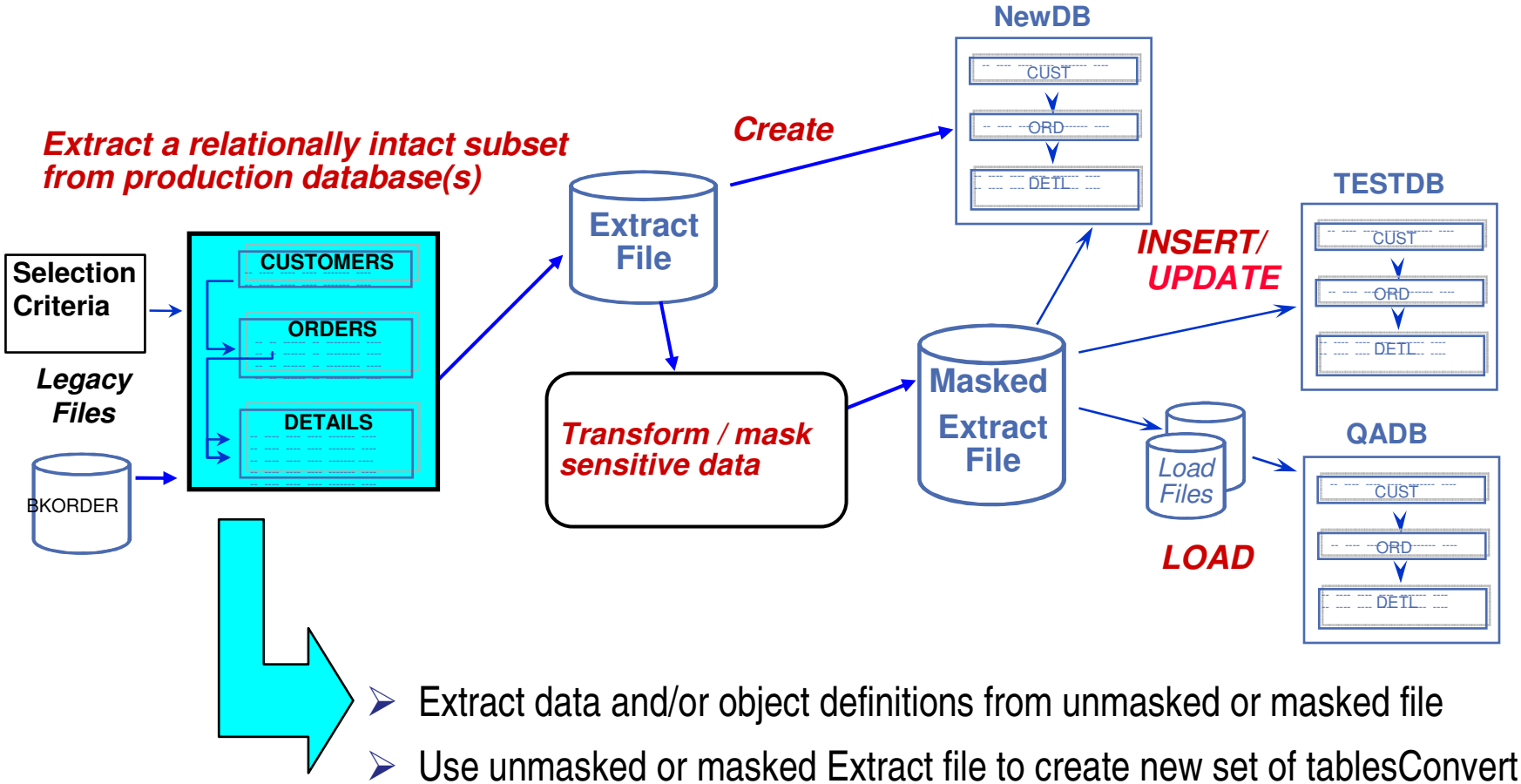
- **Most Secure Approach**
  - Extract data only
  - Convert during extract
- **Extract file already contains masked data**
  - Can be shared with testers to reuse

Data Privacy in Application Testing



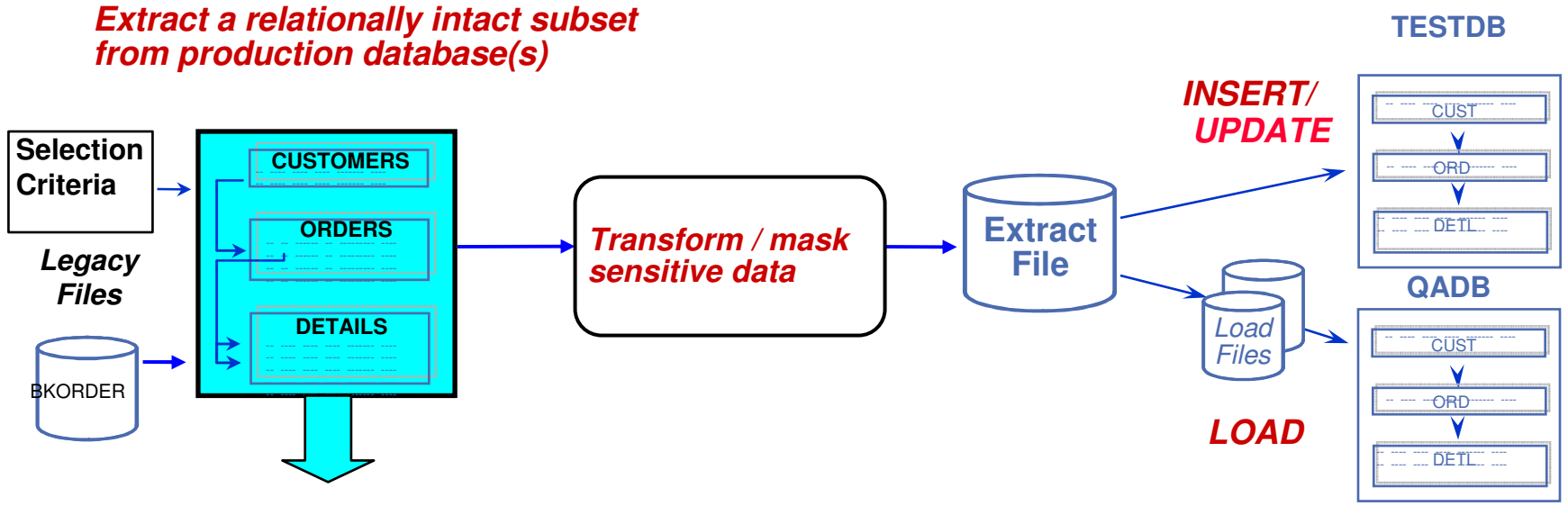
- Extract data and/or object definitionsDefine a new set of test tablesApply masking during population process Extract file may be reused but contains production dataGood process for testing masks

Data Privacy in Application Testing



- Extract data and/or object definitions from unmasked or masked file
- Use unmasked or masked Extract file to create new set of tablesConvert pre-masl extract file data into second masked extract fileMasked extract file to be reused fo population stepGood practice for testing masks using COMPARE

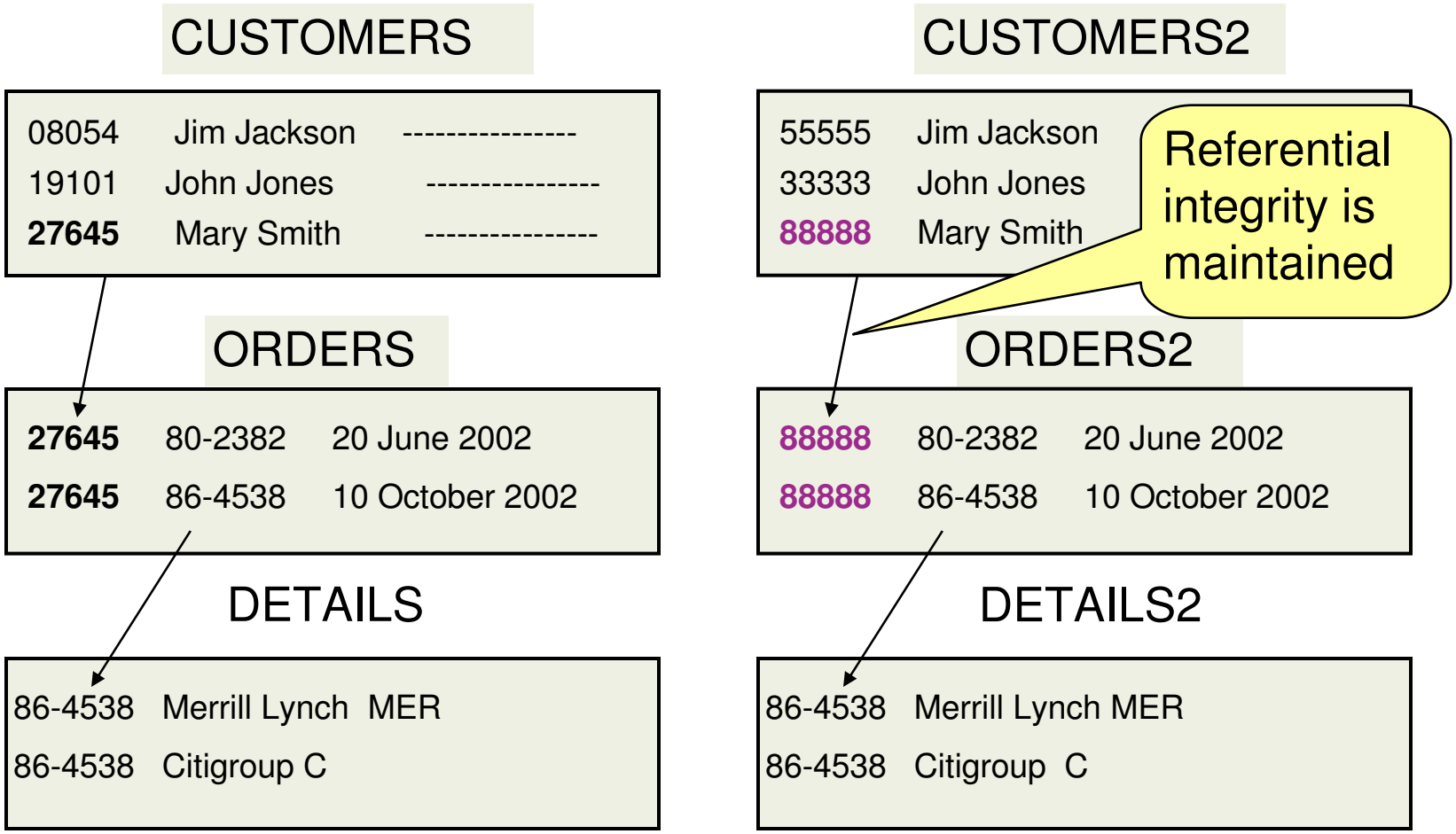
Data Privacy in Application Testing



Data transformation functions:

1. Hard-coded literals
2. Substring and/or concatenation of values
3. Special registers such as date, time
4. Arithmetic calculations
5. Sequential number generation
6. Random number generation
7. Lookup Table Functions (Random, Specific or HASHed)
8. Personally Identifying Information Lookup Tables Provided
9. Consistent Masking for non-related tables
10. Transformation Library (SSN, CCN, Email)
11. Propagation of masked primary keys to foreign keys
12. Client-defined exit routines for complex algorithms

Propagating Keys



Intelligent Masking Capability

Production Database

F. Name	L. Name	Credit Card	SSN
John	Denver	5298774132478860	254-77-6644
Vanessa	Jones	4324115574123650	154-74-7788

Data before Masking

Test Database

Valid

Valid

F. Name	L. Name	L. Name	SSN
John	Denver	5326458711224960	854-77-7234
Vanessa	Jones	4972584612457740	154-74-4186

Data after Masking… Masked with Valid CCN and SSN

How are these numbers valid?

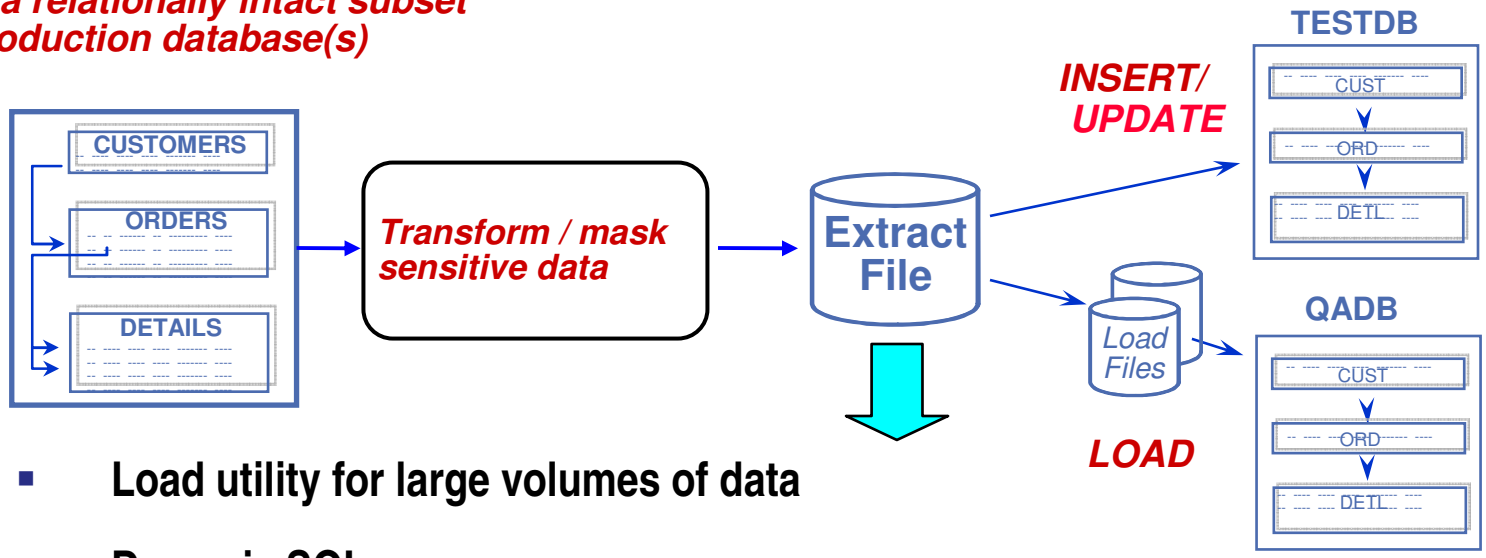
For Social Security Numbers	For Credit Card Numbers
A Social Security Number (SSN) consists of nine digits. The first three digits is called the "area number'. The central, two-digit field is called the "group Number". The final four-digit field is called the "serial Number". All numbers must fit the latest available criteria for each section.	Most credit card numbers are encoded with a "Check Digit". A check digit is a digit added to a number (either at the end or the beginning) that validates the authenticity of the number. A simple algorithm is applied to the other digits of the number which yields the check digit.

Delivered Lookup Tables

- **Name Lookup Tables**
  - First Name tables containing a list of more than 5,000 male names and 5,000 female first names
  - Last Name table containing a list of more than 80,000 last names
    - Clients can randomly mask any first name or last name with these provided lookup tables without having to supply their own.
- **Street Address/City/State/Zip Code Lookup**
  - Table containing corresponding Street address/City/State/Zip Codes for over 100,000 locations in the US.
    - Clients can randomly mask any street address, city, state or zip code with this provided lookup table.
    - Clients can mask an entire address row with a valid address row from the lookup table. (i.e. street address/city/state zip code)

Data Privacy in Application Testing

Extract a relationally intact subset from production database(s)



- **Load utility for large volumes of data**
- **Dynamic SQL**
- **Insert new rows**
- **Update existing rows; insert others**
- **Refresh from the Extract File**
- **Extract File maintains consistent baseline**

Success Story – A Large Worldwide Financial Services Company

Experiences Significant Cost Savings by Archiving Historical Data

Challenges

- Client had mainframe insurance claims application that contained data in the production database that was infrequently accessed

Solutions

- IBM Optim Data Growth

Benefits

- $799,166 NPV
- ROI = 188%
- IRR = 60%
- Payback = 14 Months

Assumptions

-Tier one costs $14 per GB per month

-Tier two costs $.75 per GB per month

Success Story – Midwestern Based Insurance Provider Experiences

Significant Cost Savings by Optimizing Non-production Databases

Challenges

- Client had a testing environment was cloned from rapidly growing production data base

Solutions

- IBM Optim Test Data Management
- IBM Optim Data Privacy Option

Benefits

- Cumulative Overall Savings = $986,610
- Test Data Management Cumulative Savings = $716,360
- Data Privacy Custom Build Cost Avoidance = $270,250

Assumptions
- Currently maintaining 4.8 TB in test, of which 50% is eligible for sub setting
- 20% Right-Sized Database (subset to 20% of former size)
- Annual (carrying) Cost per GB: $51.00

Resources

- **IBM Optim for z/OS Proof of Technology Sessions**
  - IBM Optim for z/OS Data Growth
  - IBM Optim for z/OS Test Data Management / Data Privacy
- **Details**
  - One day (per topic)
  - Up to 15 students per session
  - Hands-on labs with Optim
  - Held at a local IBM facility

Summary

- **Optim is a recognized market leader and used successfully by customers in almost all industries**
- **IBM Optim enables effective ILM (Information Lifecycle Management)**
  - The IBM Data Growth solution maintains application performance in the face of explosive data growth
  - Once archived, Optim supports prompt, accurate responses to audit and discovery requests
- **Pre-built modules for many popular applications are supported by IBM Optim (e.g. Lawson)**
- **Test data management can speed delivery of developed applications**
- **IBM Optim's data masking capabilities protect privacy by de-identifying sensitive data**