

Optim[™]

Introduction to the IBM Optim Product Portfolio

Rickard Linck, Information Management, IBM SWG



Topics

- Fast Forward through
 - Test Data Management
 - Data Privacy
 - Edit, Browse and Compare
 - Data Archiving
- A very quick tour but will try highlight some of the most interesting featues of the products



Princeton Softech

IBM Completes Acquisition of Princeton Softech

 ARMONK, NY - 13 Sep 2007: IBM (NYSE: IBM) today announced it has completed its acquisition of Princeton Softech Inc., a privately held company based in Princeton, New Jersey.

Proven leader in Enterprise Data Management

- Solving complex data management issues since 1989
- In-depth functional knowledge of mission-critical applications and the business rules that govern them
- Over 2,400 customers worldwide
 - Including nearly half of the Fortune 500
- Only true enterprise solution: across applications, databases and platforms



"Har Ni ert Test Data Management (TDM) system up-to-date?"

- Tisdag den 3 juni kl.13-16.30 kommer IBM tillsamans med Mainsoft att anordna ett seminarie på IBM Nordic Forum, Kista Entre, Isafjordsgatan 1
 - Obs: mer fokus på fokus z/OS än distribuerat
 - Effektiv och lönsam TDM
 - Presentation och demonstration av vår lösning inom TDM
 - Exempel på en automatiserad Optim-lösning på z/OS
 - Hur Optim kan utnyttjas när man bygger upp ett TDM system, som betjänar både systemutveckling och förvaltning

Agenda:

12.30-13.00 Registrering

13.00-13.30 Optim i ett nötskal - en introduktion till IBM Optim / Leif Bäck, Mainsoft

13.30-14.30 Presentation och demo av Optim TDM / Jan Larsson, IBM

14.30-14.45 PAUS

14.45-15.15 Presentation av en automatiserad Optim lösning / Lauri Pietarinen, Relational Consulting

15.15-16.15 Hur beakta kraven från systemutveckling och förvaltning i en TDM-lösning / Hans Benjaminsson, IT-Arkitekterna

16.15-16.30 Frågor & Svar



Optim naming soup

older names newer names

Relational Tools for Servers?

Archive for DB2?

Data Growth Solution?

Relational Tools Server Option?

Test Data Management?

Move for Servers?

6/4/2008

otim Data Growth for z/OS

tim TDM Solution

tim TDM Solution

tim TDM Solution

otim TDM Solution

Option

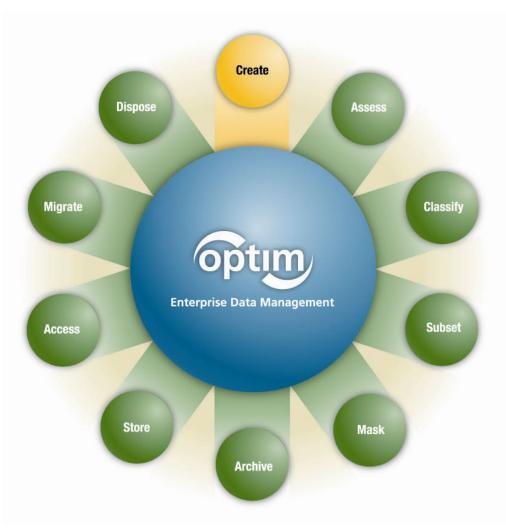
otim Data Growth

otim Data Growth

- Server Option



Optim[™] Solutions



Optim Test Data Management Solution

- Create targeted, right sized test environments
- Improve application quality
- Speed iterative testing processes

Optim Data Privacy Solution

- Mask confidential data
- Comply with privacy policies

Optim Data Growth Solution (Archiving)

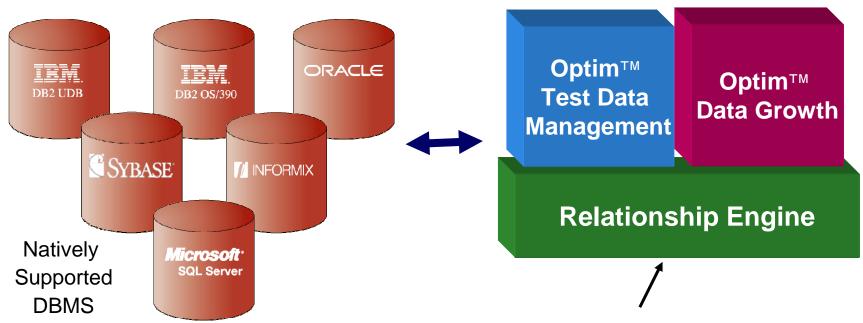
- Improve performance, Operational efficiencies
- Control data growth, save storage
- Support retention compliance



Relationship Engine

Foundation of the Optim Solution

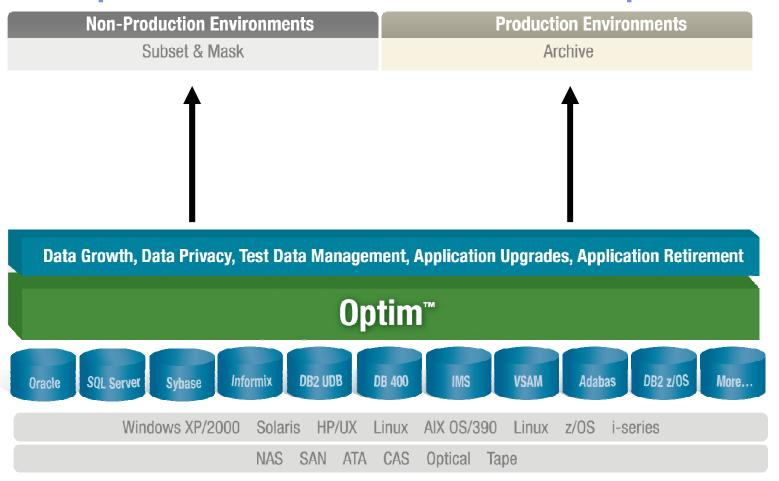
Technology to define, extract, access and selectively restore referentially intact subsets of related data from cross-platform relational database management systems



- Work with relationships that are defined in the database
- Work with logical relationships that are NOT defined in the database



Enterprise Architecture – across dbms/platforms



 Single, scalable, interoperable EDM solution provides a central framework to deploy policies to execute actions based on events within application data records from creation to deletion

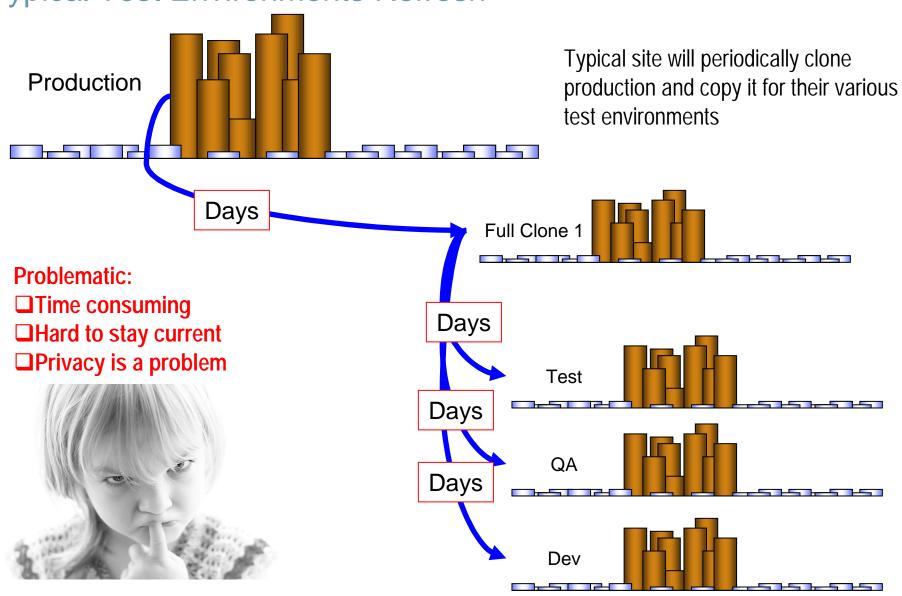


Optim[™]

Test Data Management



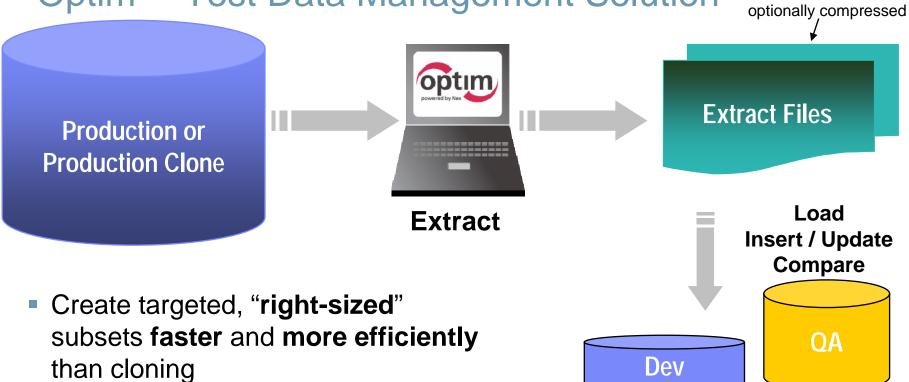
Typical Test Environments Refresh





Test

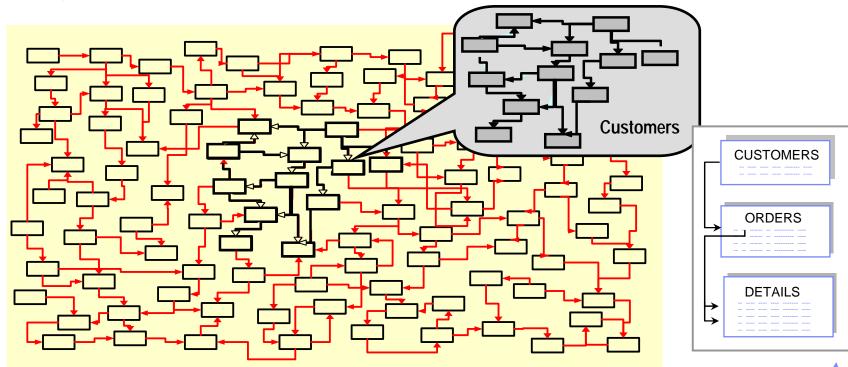
Optim[™] Test Data Management Solution



- Easily refresh, reset and maintain test environments
- Accelerate release schedule



How to get intact and useful data - Complete Business Objects

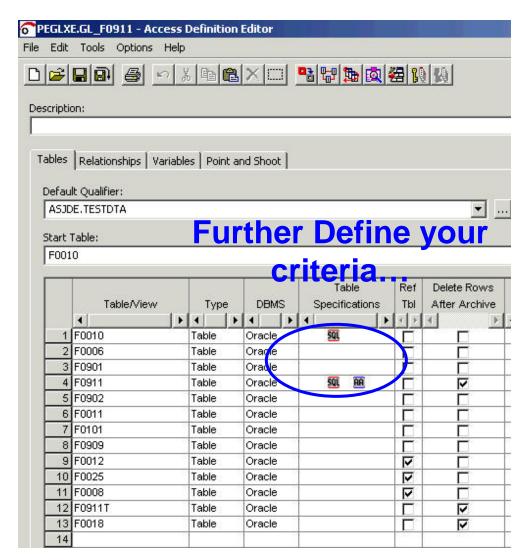


- Represents application data record orders, invoices, customer
 - Referentially(relationally)-intact subset of data across related tables and applications; includes metadata
- Provides "historical reference snapshot" of business activity
- Federated object support across enterprise data stores



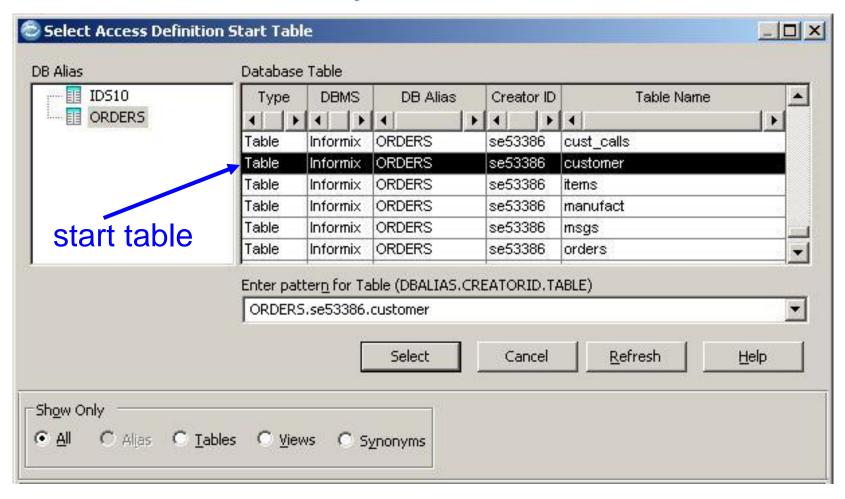
Optim Approach for Subsetting

- Add relationships
 - Same instance
 - across instances and database platforms
- Define Relationship Traversals
- Criteria can be based on one or more modules
- Date Values
 - Accounting Date
 - Fiscal Year
 - SO Trans Date
- Business Unit, Ledger Groups, Status
- "And/Or" combinations
- Grouping, Every Nth row



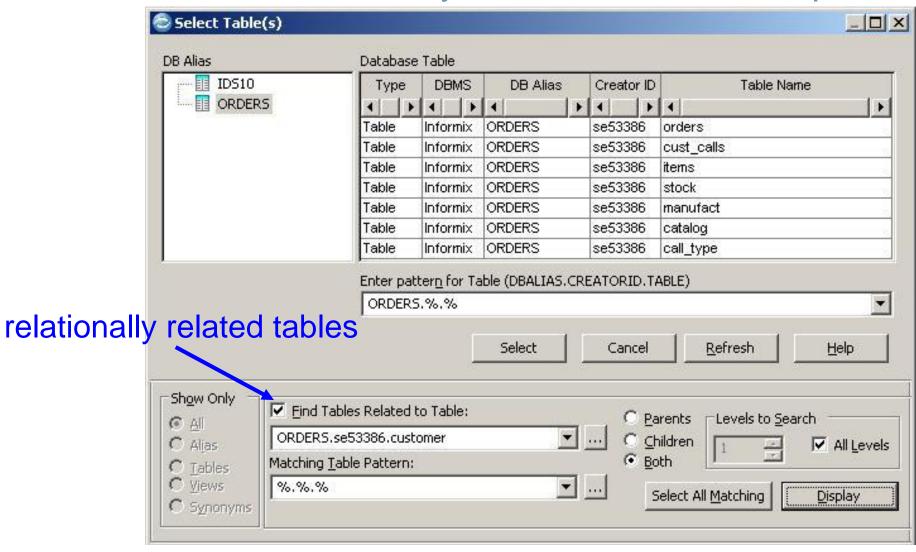


Access Definition – select your Start table



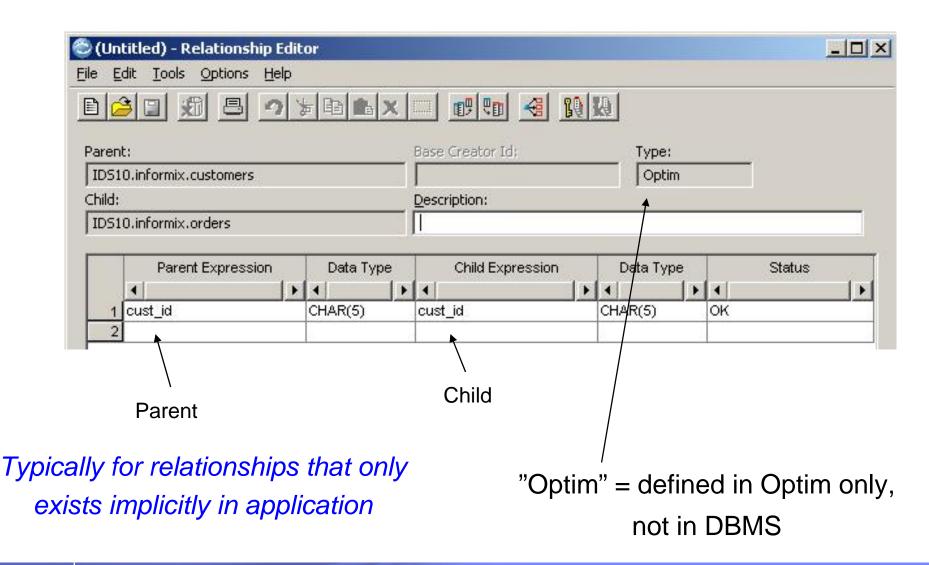


Access Definition – select your DBMS relationships



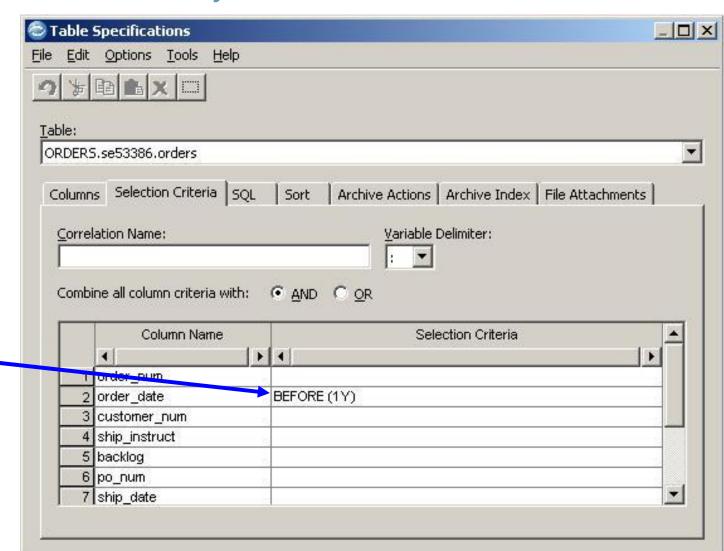


Relationship Editor - define <u>new</u> relationships





Access Definition – add your Selection Criteria



order rows older then 1 year -



Selection Criteria - For any column in a selected table

- Appropriate logical or relational operator and corresponding value or list of values.
- Can be input/specified at execution time of extraction
- Format selection criteria to comply with SQL syntax.
 - For example, to obtain data for all customers, named from M to Z, in New Jersey, you would specify:
 - CUSTNAME > 'M'
 - STATE = 'NJ'
 - Then, select:
 - AND to include all customers who satisfy both conditions, or
 - OR to include all customers who satisfy either condition.

Date Criteria

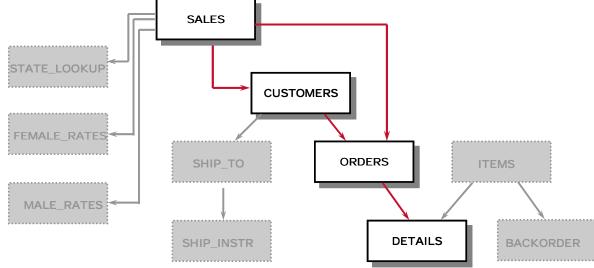
- If entering selection criteria for a column where the data type is Date, you can use the date format used by the database.
- Or, you can use the **BEFORE** operator which allows you to specify selection criteria based on a relative date. Specify the BEFORE command as follows:
 - ORDER_DATE BEFORE (1Y)
 - select all entries older than 1 year from today
 - ORDER_DATE BEFORE (1Y 4M)
 - select all entries from the ORDER_DATE columns that were older than eight months.



Normal Database Relationship Traversal Processing



During extract processing, normal traversal of relationships paths proceed like a waterfall through a data model. (Parent -> Child)



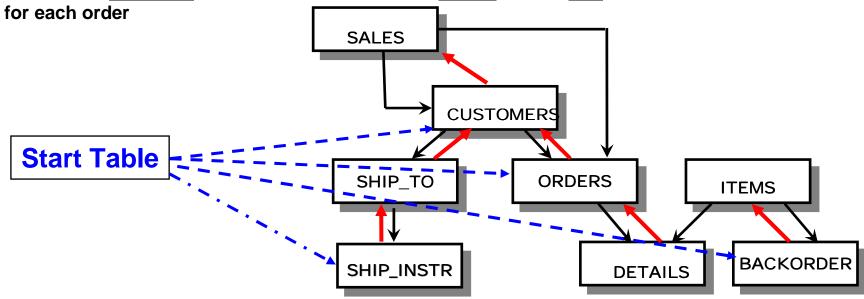


What if I need *more* than the direct parent-child data?

Examples

- 1. Get all <u>customers</u> from New Jersey, with <u>orders</u> placed by those customers, with related details and the salesman for each customer
- 2. Extract <u>customers</u> who have <u>orders</u> with <u>items</u> on <u>backorder</u>
- 3. Get all <u>customers</u> who have been shipped <u>orders</u> via *'Eventual Express'* and the <u>details</u> of those orders

4. Extract all <u>orders</u> for any <u>customer</u> with year-to-date <u>sales</u> greater than \$10,000.00, that customer's <u>salesman</u>, and that salesman's other <u>orders</u>, but *do <u>not</u>* extract the salesman





Optim Relationship Traversal: Option 1 and Option 2

Option 1:

- Causes processing to follow the relationship in the <u>opposite direction</u>
 - flow will proceed <u>from child table to parent</u>.

Option 2:

 Causes processing to follow the relationship in the parent to child direction, but is only needed when the parent table was processed as a result of following a relationship path upward (option 1) to get to it.

<u>If</u> a relationship path is followed upward to the parent as a result of Option 1 being selected on any path to the table,

then Option 2 must be selected to follow a relationship path to a child of that table



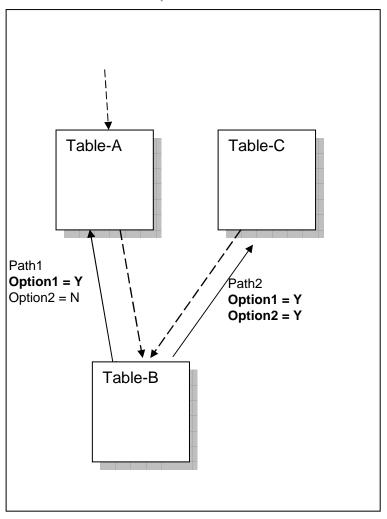
Examples

Optim Relationship traversal:

- Get selected rows from Table-A
- Get all related <u>children</u> rows from Table-B via Path1
- Get all related <u>parents</u> rows from Table-C via Path2.Option1
- Get additional children rows from Table-B via Path2.Option2
- Get all <u>new parent</u> rows from Table-A via Path1.Option1

The data extracted from Table-B is relationally intact because all related parent rows from both parent tables have also been extracted.

Example



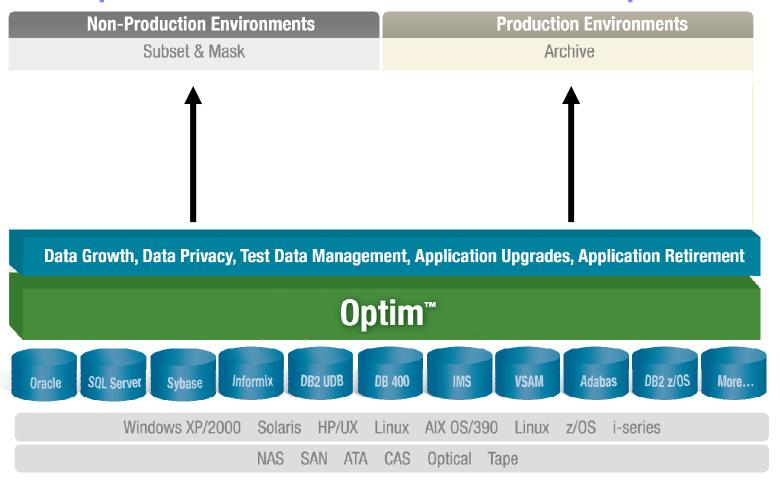


Defining logical relationships for TDM/Archiving

- Needed for subsetting data but ... how the heck can I find them in big complex ERP/CRM systems?
- Examples
 - AmdocsCRM
 - JD Edwards Enterprise One
 - Oracle E-Business Suite
 - Peoplesoft Enterprise
 - Siebel CRM
 - SAP
 - Microsoft Business Dynamics



Enterprise Architecture – across dbms/platforms



 Single, scalable, interoperable EDM solution provides a central framework to deploy policies to execute actions based on events within application data records from creation to deletion

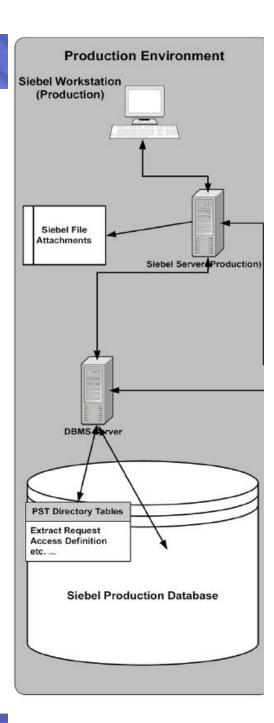


Enterprise Application Support

	Test Data Management	Data Privacy	Data Growth
Oracle E-Business Suite	✓	\checkmark	✓
PeopleSoft Enterprise	✓	✓	✓
JD Edwards Enterprise One	✓	✓	✓
Siebel CRM	✓	✓	✓
AmdocsCRM	✓	✓	✓
SAP		\checkmark	

Armonk, NY, May 19, 2008 – IBM (NYSE: IBM) today announced open client and partner testing of the new IBM® Optim™ Data Privacy Solution for use with SAP® solutions, which provides advanced enterprise data management and data privacy capabilities in support of SAP applications

Optim Siebel
Solution for Test
Data Management
System in Unix or
Windows
Environment





Optim Data Privacy Solution



Data Privacy

- Optim provides the fundamental components of test data management and enables organizations 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, Optim protects the privacy and security of confidential customer data, and supports compliance with local, state, national, international and industry-based privacy regulations



What are the capabilities?

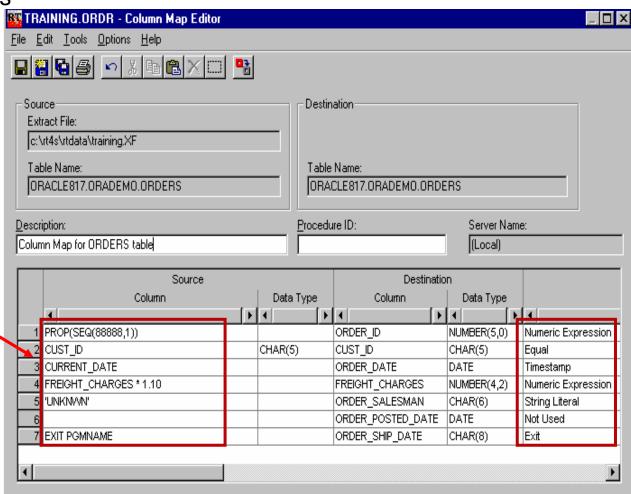
- Can occur from a standalone file or automatically after Extract
- Can be used prior to Insert/Update/Load of Data
- Implemented via Table Maps and Column Maps
- Transforms data at the column level
- Generates new values
- Semantic transformations
- Lookups
- Customer extensions
- Propagation



Column Map screen

Map unlike column names

- Transform/mask sensitive data
- Datatype conversions
- Column-level date aging
- Column-level currency conversion
- Source columns:
 - Literals
 - Special Registers
 - Expressions
 - Default Values
 - User exits





What are the techniques?

Expressions:

- Literals
- Alphanumeric
- Numeric
- Random or sequential numbers
- Character substrings
- Map special registers
- Generate random or sequential numbers

Lookups:

- Prepackaged
- Random
- Hash
- Multi-column

Functions:

- Aging
- Automatic data conversion
- Concatenation
- Sequential string function
- Character substrings
- Create an expression
- Create substrings from Char/Varchar
- Currency
- Identity or serial function
- Exits
- Propagation



Masking Techniques - Expressions

String literal example:

- Substitute Street_Address with string literal
 - '111 Campus Drive'

Random number generation example

- Replace Employee_Number with random nine-digit number
 - RAND(111111111,999999999)

Numeric Expressions

Operand1 (operator) Operand2

- Supports addition, subtraction, multiplication or division
- One value must be a column name
 - Example: FREIGHT_CHARGES * 1.1

Random or Sequential Numbers

- RAND(low, high)
 - Example: RAND(1000,10000)
- SEQ(start, step)
 - Example: SEQ(100,5)

Character Substrings

- Returns a substring of the contents of the named column
- SUBSTR(column_name, start, length)
 - Example: SUBSTR(CUST_ID,1,3)



Masking Techniques - Expressions

Map Special Registers

User ID Information

CURRENT SQLID CURRENT SQLID

USER WORKSTATION_ID

Date and/or Time

CURRENT_DATE CURRENT_TIMESTAMP GETDATE

CURRENT DATE CURRENT TIMESTAMP GETTIME()

CURRENT_TIME CURDATE NOW()

CURRENT TIME CURTIME() SYSDATE

Create an Expression

Concatenate any value or function (except special registers)

Example:

'Optim' concat SUBSTR(CUST_ID,4,3) concat SEQ(1,1)

Assume CUST_ID in ('000146', '003586'); generated values are:

Optim1461 and Optim5862



SHUFFLE

- Shuffle takes a column value and replaces it with a value for the same column from a different row.
 - SHUFFLE
- Multiple Columns: inserts shuffled values in the STATE and ZIP columns and refetches a replacement value that does not match the source up to twelve times.
 - SHUFFLE(DEST=(STATE,ZIP),RETRY=12)

Street	City	State	Zip
3526 Diamond Rd	Seattle	WA	98102
21 Street Rd	Las Vegas	NV	89101
After:			
3526 Diamond Rd	Seattle	NV	89101
21 Street Rd	Las Vegas	WA	98102
	_		



LOOKUP / RAND_LOOKUP / HASH_LOOKUP

- Assume the source column, STATE, contains state abbreviations (for example, NJ) and the destination column contains the full state name (for example, New Jersey).
- In the STATE_LOOKUP table, the abbreviations for state names are in the CODE column and the corresponding full state names are in the NAME column.
- LOOKUP (STATE, STATE_LOOKUP(CODE, NAME))

 Source col Lookup table

 Cols in lookup table
- To select a value at random from the CODE column in the first 50 rows of a table named STATE_LOOKUP and insert it in the destination column, specify
- RAND_LOOKUP(STATE_LOOKUP,CODE,50)

Lookup table Col in lookup table

- To select values from the CITY, STATE, and ZIP columns in a random row of a table named STATE_LOOKUP and insert them in the corresponding destination columns, specify



Masking Techniques - Exits and Column Map Procedures

Invoke custom masking routines: Exits

- Exit routinename
- Useful for special processing, conditional data manipulation, and literals that exceed the length provided for Source Column
- Example:
 - EXIT CUSTPGM
- The function of a <u>Column Map Procedure</u> is generally the same as that of an exit routine.
 - An <u>exit routine</u> must exist in a DLL external to Optim, and must conform to calling conventions used in the **C** programming language.
 - A <u>Column Map Procedure</u>, however, is stored in the Optim Directory and is written in **Optim Basic**, which is distributed with Optim



Referential integrity and data masking

Original Data CUSTOMERS

ORDERS

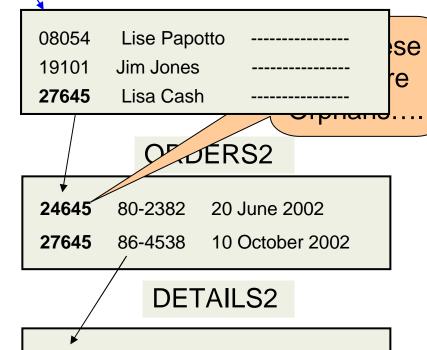
27645 80-2382 20 June 2002 **27645** 86-4538 10 October 2002

DETAILS

86-4538 DR1001 Lord of the Rings 86-4538 CM2010 Dude, Where's My Car?

De-Identified Data

mask CUSTOMERS2



86-4538 DR1001 Lord of the Rings 86-4538 CM2010 Dude, Where's My Car?



Propagating Keys

Mask and propagate

CUSTOMERS

ORDERS

27645 80-2382 20 June 2002 **27645** 86-4538 10 October 2002

DETAILS

86-4538 DR1001 Lord of the Rings 86-4538 CM2010 Dude, Where's My Car?

CUSTOMERS2

55555 Lise Papotto Referential 33333 Jim Lee integrity is 88888 Lisa Cash maintained! ORDERS2 88888 80-2382 20 June 2002 88888 86-4538 10 October 2002 **DETAILS2**

86-4538 DR1001 Lord of the Rings 86-4538 CM2010 Dude, Where's My Car?



Propagate Keys

- Propagates/cascades <u>new</u> values for
 - the primary key or
 - foreign key to
- all related tables
- Necessary to retain referential integrity of masked data
- How?
 - Assign a value (literal, function, expression) to a key column in a parent table
 - Propagate key value to child tables at all levels
 - Enables you to "manufacture" new test data



© 2008 IBM Corporation

Propagation - Cascade Primary/Foreign Key Values

- PROP(value,column-name)
 - Value: literal, expression, special register, exit routine
 - Column-name: specifies source column containing value that is subject of the function

Example: PROP('SMITH')

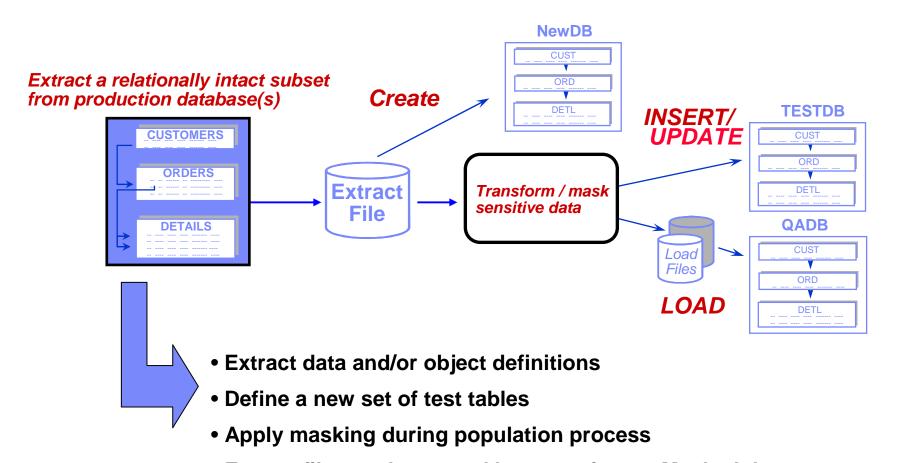
PROP(RAND(10000,99999),CUST_ID)

Example: PROP(TRANS SSN)

PROP(SEQ(10000,1),CUST_ID)



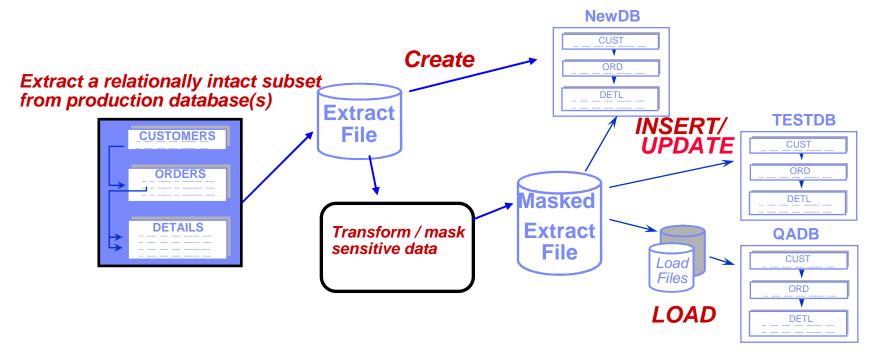
Data Privacy in Application Testing – 1(3)



- Extract file may be reused but contains un-Masked data
- Good practice for testing masks



Data Privacy in Application Testing – 2(3)

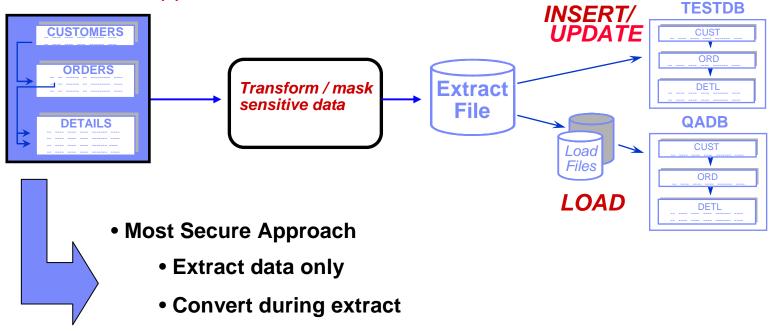


- Extract data and/or object definitions in pre-masked file
- Use pre-masked Extract file to create new set of tables
- Convert Pre-masked extract file data into second masked extract file
- Share masked extract file to be reused for population step
- Good practice for testing masks using COMPARE



Data Privacy in Application Testing – 3(3)

Extract a relationally intact subset from production database(s)



- Extract file already contains masked data
 - •Can be shared with testers to reuse



Edit, Browse and Compare



Why a Relational Data Editor in a test data environment?

Because you need a fast, flexible way to:

- Examine, edit, and understand related data from multiple tables
- Add data to test application logic
- Correct bad data

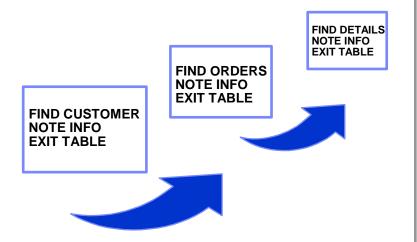


Approach

Traditional Tools

Single Table Editors

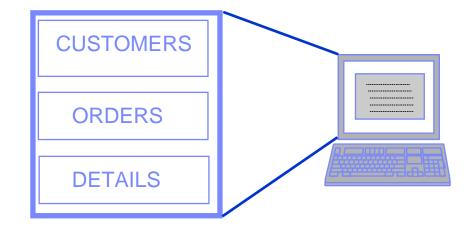
- One table/view at a time
- No edit of related data from multiple tables



Optim[™] Test Data Management

The Relational Editor

 Simultaneous browse/edit of related data from multiple tables



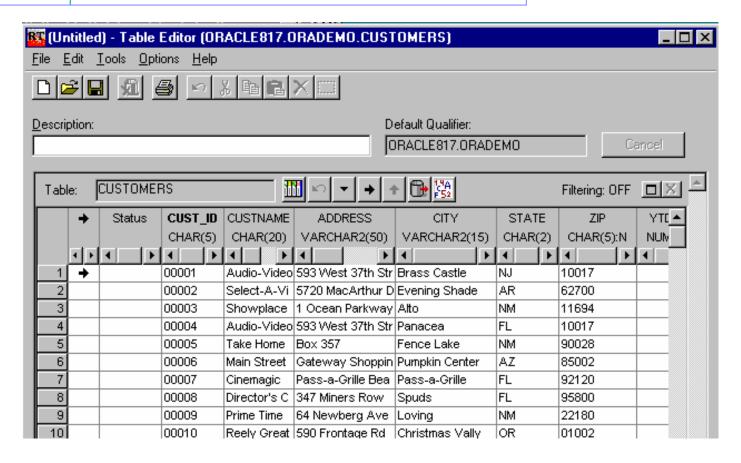


Relational Edit

- Edit data while respecting the original RI within its relational context
- Dynamic join related tables into a single view
- Synchronously scroll through related data
- Use a single editing and viewing tool across multiple database types (DB2 on mainframe or server, Oracle, Sybase, SQL Server, Informix)
- Simultaneous browse, update, insert, and delete of related data from multiple tables
- Extended backout capabilities
- Audit Trail of updates



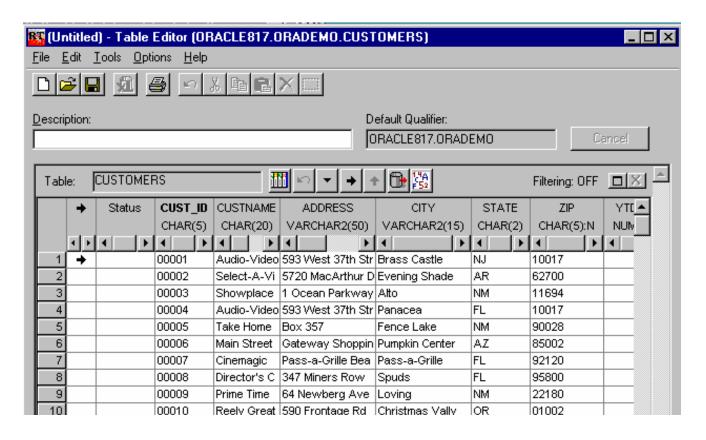
Browse/Edit a table



- User can define how data is displayed
 - SORT, Sidelabel/Columnar format, Attributes, Lock
- All database access authority enforced



Joining to another table(s)

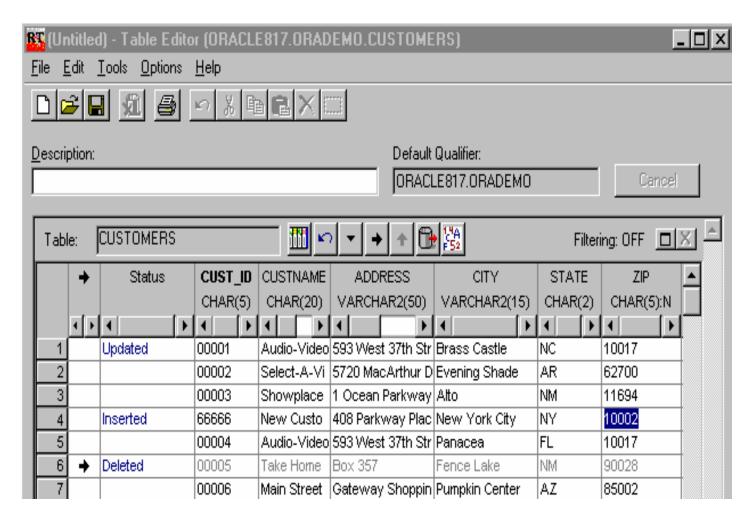


Browse or edit related rows

Scroll of
higher-level
table
automatically
synchronizes
all lowerjoined tables



Editing Data

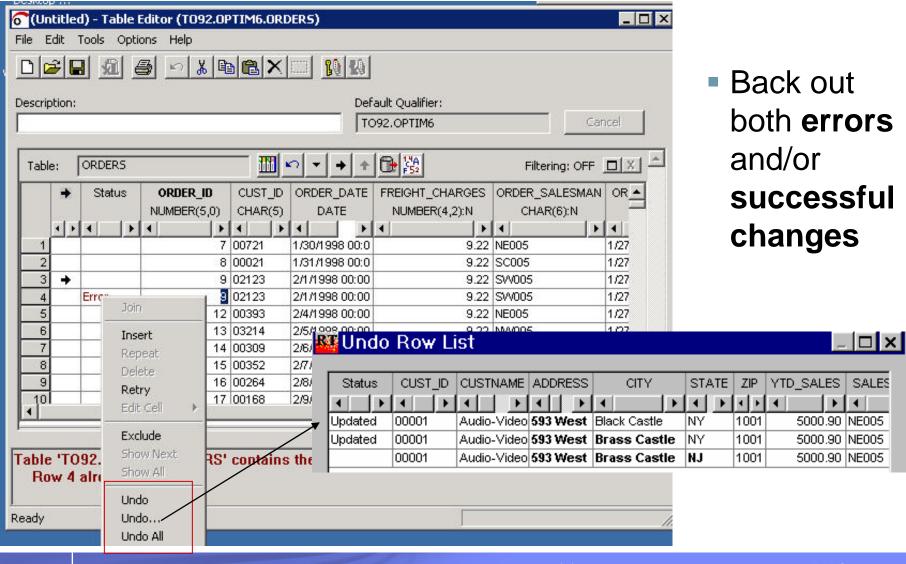


Edit data to:

- Insert rows
- Delete rows
- Update rows



Flexible Undo feature



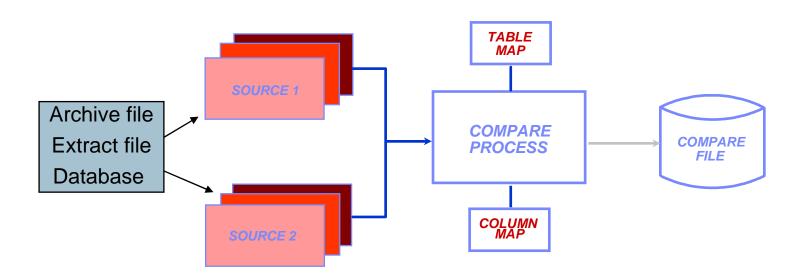


Compare: What are the capabilities?

- Compare the "before" and "after" data from a converted Extract
- Compare results after running modified application during regression testing
- Identify differences between separate databases
- Identify expected changes as well as unanticipated results
- View complete before and after images of the database or highlight only the differences



Compare Options



- Single-table or multi-table compare
 - If source is a database, rows are extracted into intermediate file to compare
- Creates compare file of results
- Displays results on screen



Different Table Names?

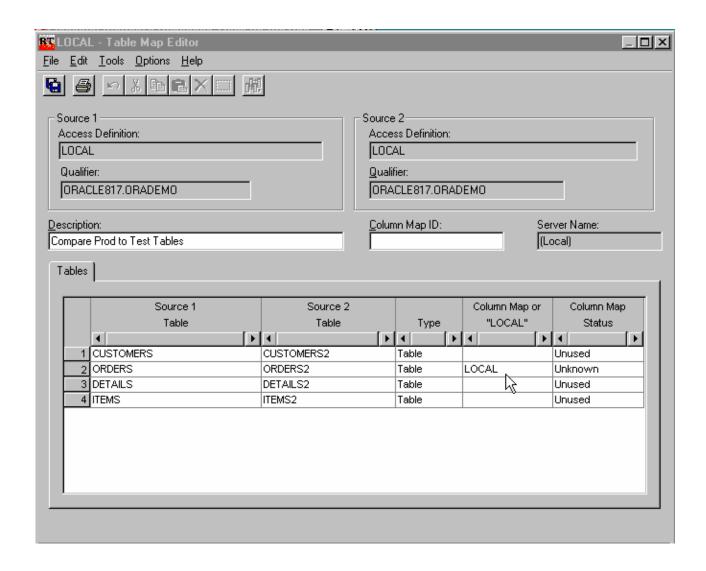
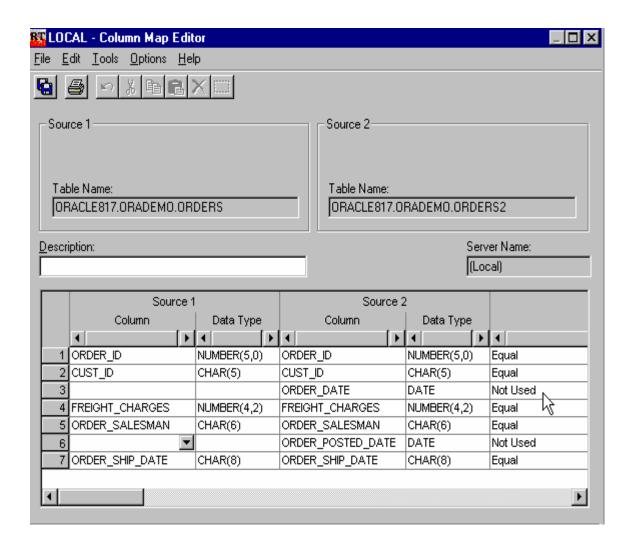


Table Map

- Allows mapping of different table names or creator-IDs
- Can be 'local' to this compare or saved in the Optim Directory for future use



Different Column Names?

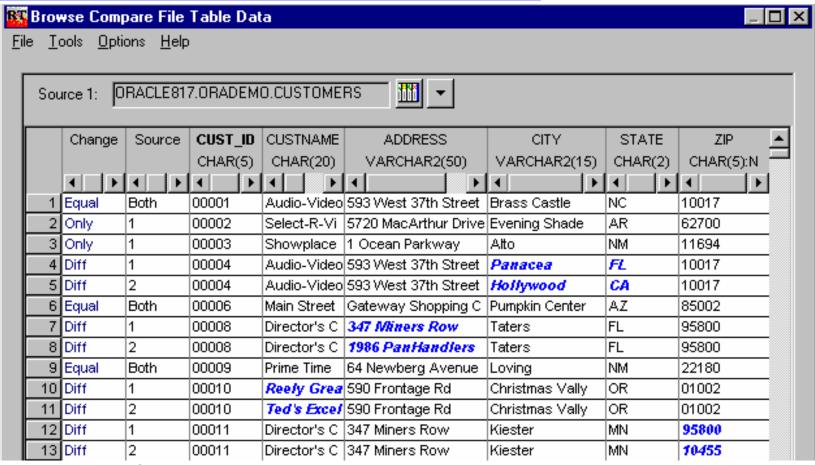


Column Map

- Correlates columns with unlike names
- Eliminates columns from the compare
- Can be 'local' to this compare or saved in the Optim Directory for future use



Browsing the Compare File



- Change column identifies the type of change
- Source column identifies input source row
- Data differences are highlighted



Summary: Optim Capabilities for Test Data Management

- Manage test data across enterprise
 - all related applications and platforms
 - Extracts and migrates data between different types of databases (Informix, DB2 on mainframe or server, VSAM/sequential files, IMS, Oracle, Sybase, SQL Server, ...)
- Extract referentially intact subsets of data (including complete business objects)
 - Supports all data relationships, whether defined to the DBMS or defined to the application
- Dynamically create destination environment
 - insert or load data to target, create tables if needed
- De-identify or mask data in non-production environments
- Edit test data to create error and boundary conditions
- Compare baseline data against successive test run results to identify errors that would have otherwise gone undetected
- No need to write and maintain custom extract programs.



Optim Data Growth Solution

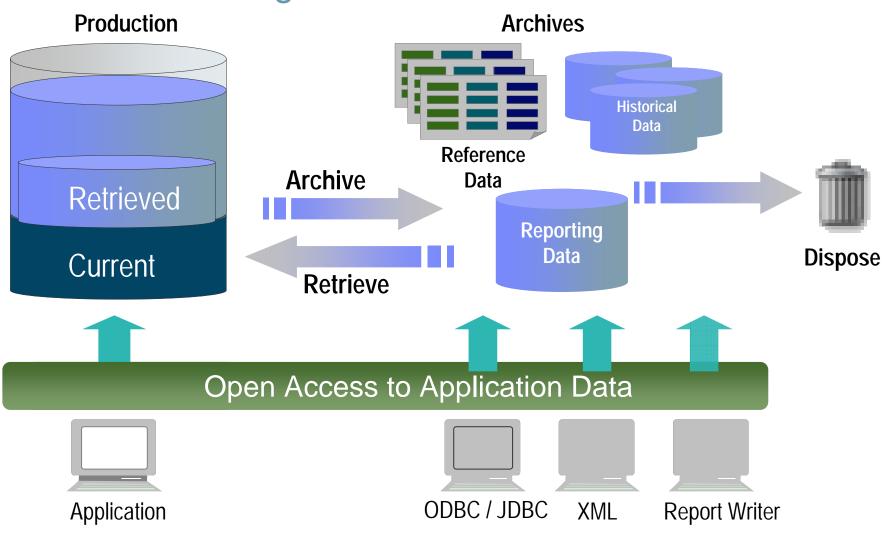


Database Archiving

- Improve Performance reduce size of production database
 - Faster Application Performance
 - Quicker Backup and Recovery
 - Improve IT Productivity
- Control Costs
 - Reduce infrastructure costs of servers/storage
 - Move data to different tiers based on current value and access requirements
- Mitigate Risks
 - Secure access to data
 - Preserve Data integrity
 - Maintain compliance with national and industry specific data retention regulations
 - Preparation for E-discovery requests



How does Archiving Work?





Optim Archive

An OPTIM Archive is:

- A self-describing relationally intact subset of data
- that retains its business context
- even when extracted from its original database
- that can be stored and accessed based on value
- and deleted in accordance to a retention policy.

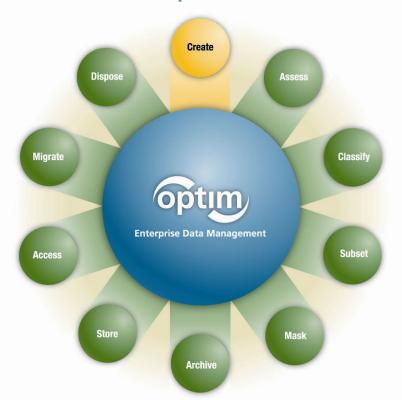


Optim Open Data Manager (ODM)

- Provides access to data in <u>Optim Archive Files</u> for programs that use the Open Data Base Connectivity (ODBC) and Java Data Base Connectivity (JDBC).
- You can use ODM on the same Windows, Solaris, AIX, HP-UX, and Linux platforms that are supported by *Optim*.
- ODM can be installed during the *Optim* installation process.
- ODM is implemented using the Attunity Connect® product in concert with a custom driver that provides access to Archive Files.



That's it! - Optim[™] Solutions





Optim Test Data Management Solution (Move, Edit, Compare)

- Create targeted, right sized test environments
- Improve application quality
- Speed iterative testing processes

Optim Data Privacy Solution

- Mask confidential data
- Comply with privacy policies

Optim Data Growth Solution (Archiving)

- Improve performance, Operational efficiencies
- Control data growth, save storage
- Support retention compliance



