

5-9 November

Athens Hilton

Athens, Greece

Session: G01

IDUG® 2007

Europe

Better Faster Cheaper Using Princeton Softech Optim for Test Data Management



Olivier Jouannic
Princeton Softech

5 November 2007 • 11:30 – 12:30



© 2005 Princeton Softech, Inc.

Platform: Cross-Platform

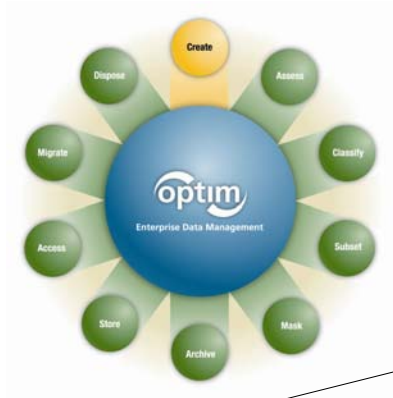
GoFurther

The application development cycle requires significant amounts of time and resources in order to effectively launch new and/or improve applications. When testing with structured data, many times a full clone of a production database is made for testing purposes, while a relationally intact “right-size” subset will give users the necessary data required for testing instead. Join us for a step by step demonstration that explores how a relationally intact subset can be created from your production environment as well as how that subset can be inserted into a previously created database or one that is new. See how forced errors can be used to test the quality of the application, and then compared against data not yet tested.

- Test Data Management in the Context of EDM*
- Enterprise Architecture of OPTIM
- What is a Complete Business Object in the context of Test Data Management?
- Optim Processes Overview
- Demos of Extract, Mask, Insert, Edit and Compare functions

* Enterprise Data Management

Optim™ Solves the EDM Challenge



- Archiving
 - Improve performance
 - Control data growth, save storage
 - Support retention compliance
 - Enable application retirement
 - Streamline upgrades
- Test Data Management
 - Create targeted, right sized test environments
 - Improve application quality
 - Speed iterative testing processes
- Data Privacy
 - Mask confidential data
 - Comply with privacy policies

Today's Discussion

3

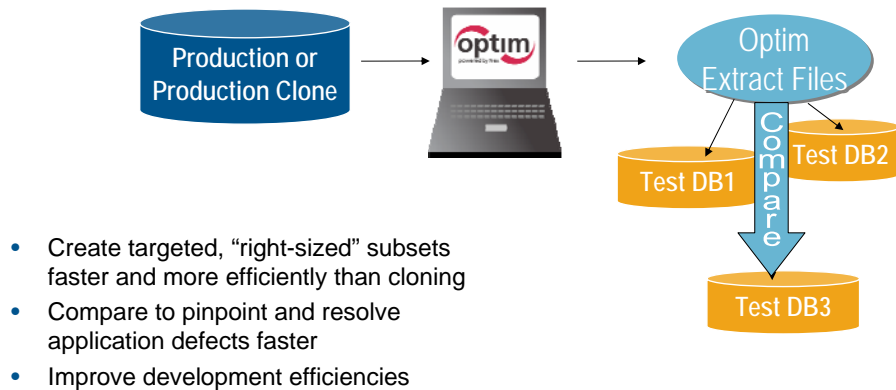
GoFurther

Today's discussion is on 2 of the Optim Features

- Test Data Management (TDM)

-Data masking

Optim™ Test Data Management Solution



4

GoFurther

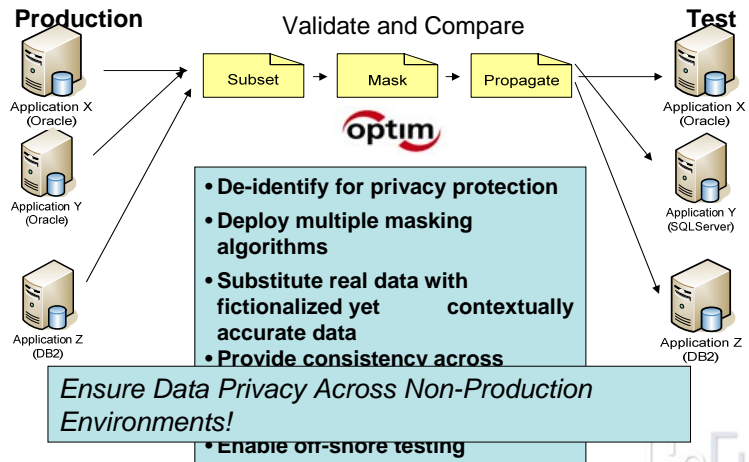
Principle

Extract data from production

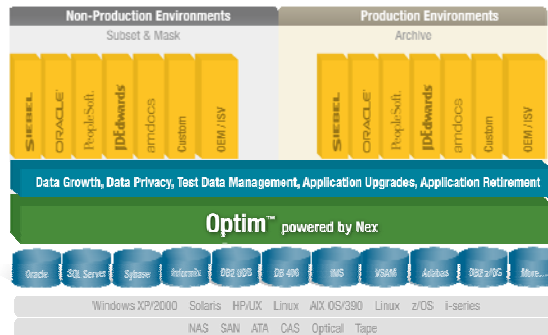
Insert to target (Test)

Compare snapshots for

The Optim™ Data Privacy Solution



Enterprise Architecture



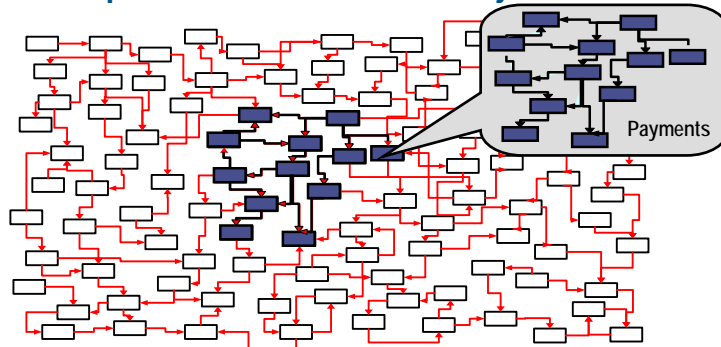
- Single, scalable, interoperable EDM solution provides a central point to deploy policies for extract, storing, porting, and protecting application data records from creation to deletion

GoFurther

Why is it important to have a single, scalable solution that supports all the applications, databases and environments in your enterprise? Because every company – whether large or small – has a mixed or “heterogeneous” infrastructure that includes some combination of mission critical applications, running on a variety of databases and operating systems. And these systems interoperate with one another to form your key business processes – like booking a sales order, shipping a product or posting a payment.

By leveraging a single, scalable and interoperable EDM solution, you have a central repository to define and store all the policies for managing, moving, copying, retaining and ultimately disposing of your application business records. You achieve the greatest level of consistency with the fewest resources at the lowest cost. You protect and leverage the your investment in enterprise applications and data assets.

Complete Business Object



- Represents application data record – payment, invoice, customer
 - Referentially-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

7

GoFurther

Data representing a business object is often stored in multiple tables, each table might be on a different DBMS, MACHINE or OS

Example : How can you collect coherent information to extract a ‘Payment’ from production ?

Introducing Optim

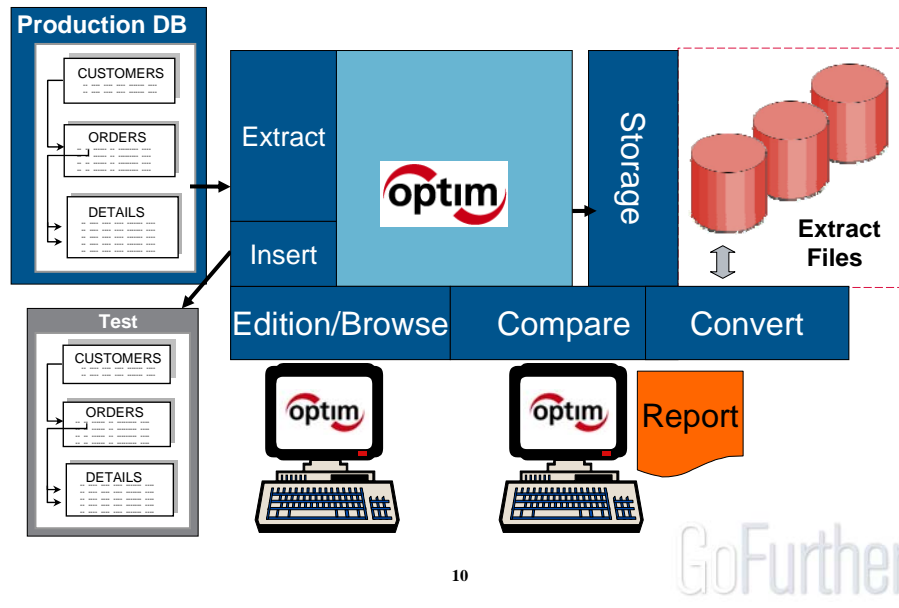
8

GoFurther

Optim - Test Data Management

- Subsetting capabilities to create realistic and manageable test databases
- De-identify (mask) data to protect privacy
- Edit data to create targeted test cases
- Compare baseline data across successive test runs

Test Data management: Processes



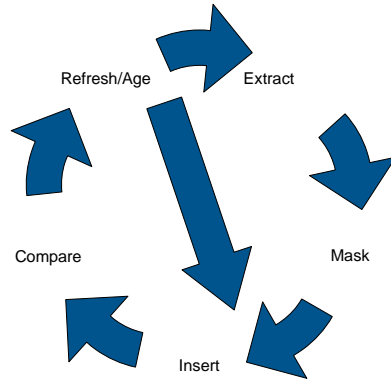
10

Main processes are

- Extract
- Insert
- Convert
- Edit / Browse
- Compare.

Storage options are multiple....

STEPS with High Value



- **Extract**
 - Realistic Test Cases
 - Intact Subsets including on complicated data models
 - With all forms of selection criteria including object lists
 - Data + DDL
- **Mask**
 - Sensitive Personal information Coherently
 - With Realistic values
- **Insert**
 - Replace all or add or refresh
 - Load
 - Map to target with different structure/content/business rules
- **Compare**
 - Image Before/After test
 - Image Result test V1/Result test V2
- **Refresh/Age**
 - Iterate tests
 - Iterate tests at different point in time aging all time related values

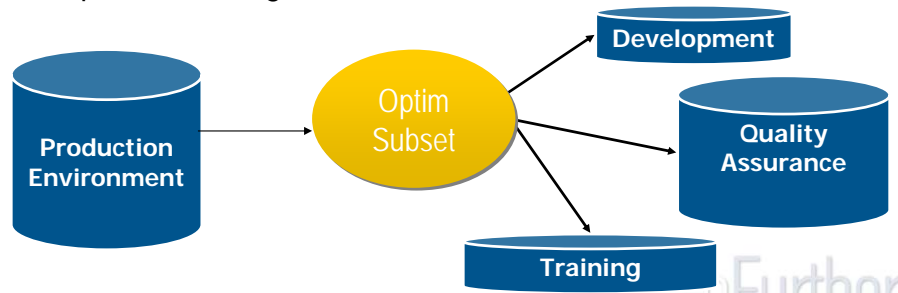
11

GoFurther

Value of each step

Subsetting Test Data

- Build intact, smaller instances
Development, Quality Assurance, Acceptance, Training
- Multiple subsets created from one source
- Varied selection criteria
- Optimizes storage utilization



12

GoFurther

Extract DEMO

Extract All customers from NJ

De-Identifying Test Data

- Transform and mask confidential data
- Protect privacy, support compliance initiatives
 - HIPAA, GLBA, PIPED, DDP, NPP, others
- Multiple transformation techniques
 - Random & sequential number generation, substrings, lookup values, date aging, numeric expressions, concatenated expressions
 - Key propagation
 - Custom routines to support site-specific requirements

Mask Demo

De-identify NJ Customers

De-Identifying Test Data for Privacy Protection

Original Data

Customers Table

Cust ID	Name	Street
08054	Alice Bennett	2 Park Blvd
19101	Carl Davis	258 Main
27645	Elliot Flynn	96 Avenue

Orders Table

Cust ID	Item #	Order Date
27645	80-2382	20 June 2004
27645	86-4538	10 October 2005

De-Identified Data

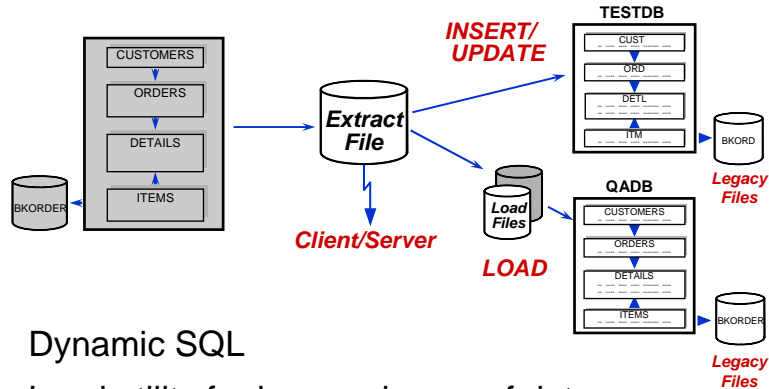
Customers Table

Cust ID	Name	Street
10000	Auguste Renoir	Mars23
10001	Claude Monet	Venus24
10002	Pablo Picasso	Saturn25

Orders Table

Cust ID	Item #	Order Date
10002	80-2382	20 June 2004
10002	86-4538	10 October 2005

Populate Destination Tables



- Dynamic SQL
- Load utility for large volumes of data
- Download to Client/Server RDBMS from MVS

17

GoFurther

Once you have extracted the data you want, the next step is to populate the destination tables. Move gives you several options for the population process.

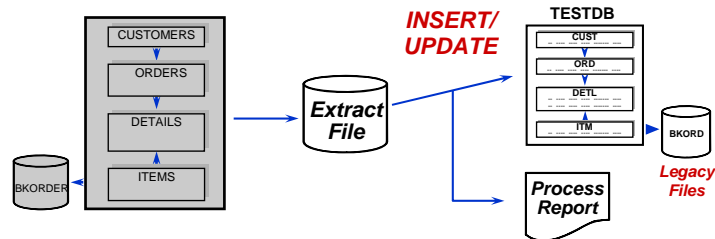
Move can populate your destination tables using dynamic SQL.

The Load utility is the more logical choice for populating the destination tables when you have a large volume of data.

Or, you may want to download to a Client/Server platform and use Move for Servers or Import for Servers to populate a database on your workstation.

Or, if using Move for Servers, upload the extract file from your workstation and use it to populate a mainframe DB2 database.

Populate Destination Tables - Insert



- Insert new rows
- Update existing rows (insert others)
- No impact on other users

18

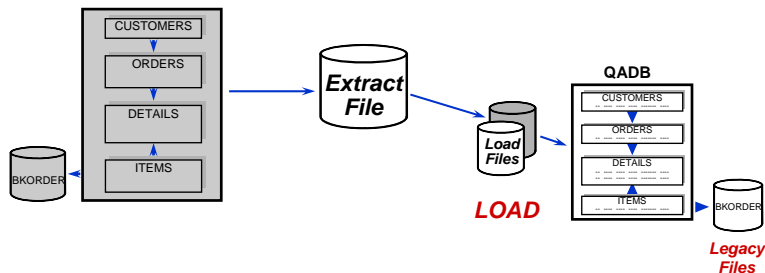
GoFurther

When you execute either of the SQL processes (Insert or Update), Move allows you to control whether locks will be held on any tables during the population process. This allows you to create and maintain your test data without having to either schedule the task or interfere with anyone else who might be using the tables.

While the advantage to the SQL process is that utility (tablespace) locks are not held, sometimes you may want to keep others from accessing the tables during the Insert or Update process. Move also allows you to control when COMMITs will be issued during the Insert process. By default, a COMMIT is issued every 1000 rows and at the end of each table.

In MVS, you also have the choice of an online or batch process.

Populate Destination Tables - Load Processing



- User-specified load options
- Transforms extract file into Load format files
- Efficient for very large amounts of data

At some point in your development cycle, you will want to do volume testing to make sure your application can process the volume of production data. For large volumes of data, LOAD processing is best for populating the test database. Load format files are built from the extract file data at runtime. Because the extract file is not modified, this means the same extract file can be used at any time for Insert, Update or Load processing.

If you choose to use the load utility to populate the destination tables, a panel will be displayed where you can choose which options you want for the load process. In MVS you can choose either DB2 Load or BMC LOADPLUS on the user options panel, and the appropriate load parameters screen is presented.

Insert Demo

Insert de-identified NJ customers
into a target test database

Editing Test Data

Table: CUSTOMERS

	Status	CUST_ID CHAR(5)	CUSTNAME CHAR(20)	ADDRESS VARCHAR2(50)	CITY VARCHAR2(15)	STATE CHAR(2)	ZIP CHAR(5):N
1	Updated	00001	Audio-Video	593 West 37th Str	Brass Castle	NC	10017
2		00002	Select-A-Vi	5720 MacArthur D	Evening Shade	AR	62700
3		00003	Showplace	1 Ocean Parkway	Alto	NM	11694
4	Inserted	66666	New Custo	408 Parkway Plac	New York City	NY	10002
5		00004	Audio-Video	593 West 37th Str	Panacea	FL	10017
6	Deleted	00005	Take Home	Box 357	Fence Lake	NM	90028
7		00006	Main Street	Gateway Shoppin	Pumpkin Center	AZ	85002

Edit data to:

- Insert, delete, update rows
- Create error and boundary conditions

GoFurther

Joining Related Tables

Table: CUSTOMERS

Status	CUST_ID CHAR(5)	CUSTNAME CHAR(20)	ADDRESS VARCHAR2(50)	CITY VARCHAR2(15)	STATE CHAR(2)	ZIP CHAR(5)N	YTT NUM
1	00001	Audio-Video	593 West 37th Str	Brass Castle	NJ	10017	
2	00002	Select-A-Vi	5720 MacArthur D	Evening Shade	AR	62700	
3	00003	Showplace	1 Ocean Parkway	Alto	NM	11694	

Table: ORDERS

Status	ORDER_ID NUMBER(5,0)	CUST_ID CHAR(5)	ORDER_DATE DATE	FREIGHT_CHARGES NUMBER(4,2)N	ORDER_SALESMAN CHAR(6)N	ORDEF
1	20	00001	2/12/98 00:00	14.80	NE005	1/27/98
2	229	00001	9/24/98 00:00	19.05	NE005	1/27/98
3	275	00001	10/19/98 00:00	21.97	NE005	1/27/98
4	30023	00001	4/2/98 00:00.0	33.85	NE005	1/27/98

- Browse or edit related rows
- Scroll of higher-level table automatically synchronizes all lower-joined tables

Editor demo

Edit a customer in test database

Verifying Test Results

Change	Source	CUST_ID CHAR(5)	CUSTNAME CHAR(20)	ADDRESS VARCHAR2(50)	CITY VARCHAR2(15)	STATE CHAR(2)	ZIP CHAR(5):N
1 Equal	Both	00001	Audio-Video	593 West 37th Street	Brass Castle	NC	10017
2 Only	1	00002	Select-R-VI	5720 MacArthur Drive	Evening Shade	AR	62700
3 Only	1	00003	Showplace	1 Ocean Parkway	Alto	NM	11694
4 Diff	1	00004	Audio-Video	593 West 37th Street	Panacea	FL	10017
5 Diff	2	00004	Audio-Video	593 West 37th Street	Hollywood	CA	10017
6 Equal	Both	00006	Main Street	Gateway Shopping C	Pumpkin Center	AZ	85002
7 Diff	4	00008	Director's C	347 Miners Row	Tatars	FL	95800
8 Diff	2	00008	Director's C	1986 Panhandlers	Tatars	FL	95800
9 Equal	Both	00009	Prime Time	64 Newberg Avenue	Loving	NM	22180
10 Only	1	00010	Reely's Great	590 Frontage Rd	Christmas Vally	OR	01002
11 Diff	2	00010	Ted's Excel	590 Frontage Rd	Christmas Vally	OR	01002
12 Diff	1	00011	Director's C	347 Miners Row	Kiester	MN	95800
13 Diff	2	00011	Director's C	347 Miners Row	Kiester	MN	70455

- Compare baseline data against successive test runs
Table to table, or multiple sets of related tables
- Data differences are highlighted

GoFurther

24

The change column indicates one of several statuses:

DIFF - the row has differences in one or more columns as shown by the highlighted fonts.

EQUAL - the rows of data matched exactly in the two input sources.

ONLY - the row exists only in the source number shown

The source number column will indicate if the row originated from Source file 1 or Source file 2.

Compare Demo

Compare current updated Database
with the original values
(De-identified NJ customers)

Optim's Test Data Management's Most Important Features

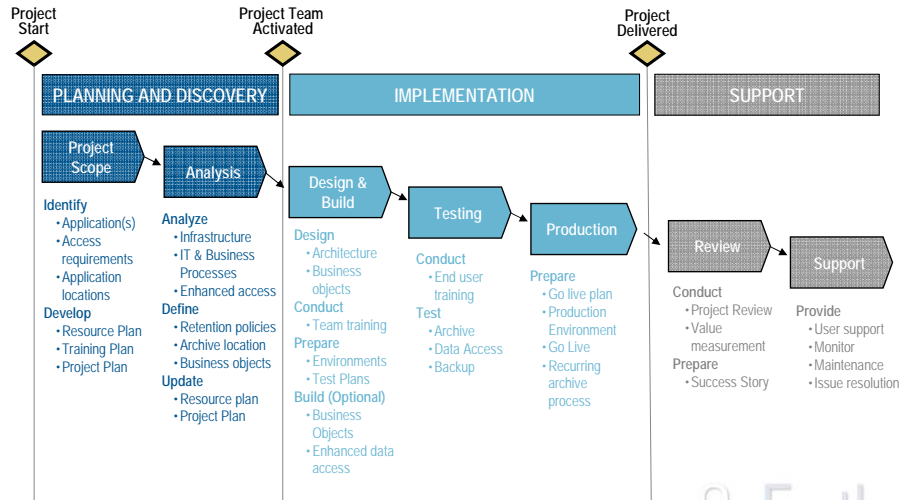
- Referentially Intact Subsets Guarantee
 - 100% Integrity
 - Cycle Management
 - Dynamic Model Maintenance
 - Say What not How
- Performance
 - Dynamic optimisation / Options / Index Analysis
- Exploitability
 - Share / Operate from production transparently / Schedule
- Full Test Data Management
 - Subsetting
 - Data Privacy
 - Unified Editor
 - Regression testing (compare)
 - Current Date Management
- Multi Database / Multi Platform
- Predefined Templates for Siebel
- DDL re-generation
- ...

Optim Implementation



- Bundled services speed time to value
- Define your business objectives for enterprise data management
- Apply Optim processes and technology
- Implement Optim in your enterprise application environments
- Benefit from ongoing optimization and technical support

Implementation Time Line



28

GoFurther

Roles Tasks

Project Manager Establishes project plan with Princeton Softech

- Identify internal resources, their roles & assumptions
- Capture business feedback and identify additional business needs
- Identify and resolve issues
- Secure internal resources

Business Define what business transactions to archive

- Define retention time frame
- Conduct data access training (optional)

Technical Install Software

- Configure test environment
- Configure archive parameters
- Test configurations and parameters
- Enhance user data access (optional)
- Business training for data access (optional)

Session G01

Better Faster Cheaper

Using Princeton Softech Optim for Test Data Management

Olivier Jouannic

Princeton Softech

ojouannic@princetonsoftech.com

