Good Morning!

DB2 Users Group Meeting

Please …set cell phones to stun (vibrate)

Enterprise Data Management

Essentials of Test Data Management

Agenda

- Introductions
- EDM Overview
- Test Data Management Concepts
- TDM Solution
  - Subsetting
  - Data Privacy
- Q & A

Why?

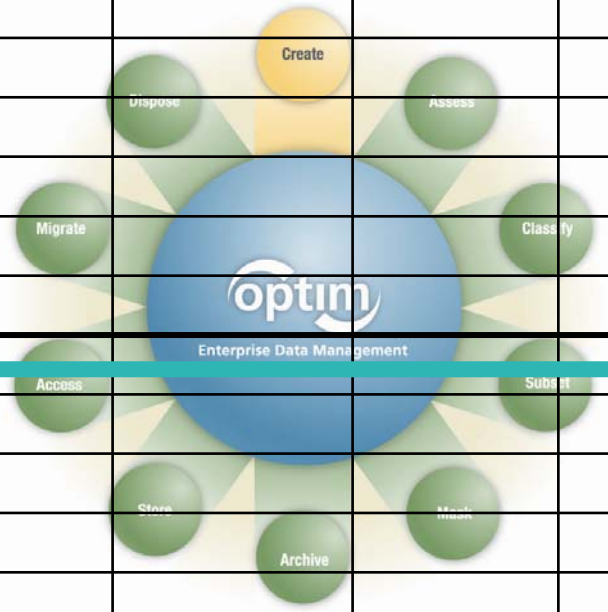
- Solve client pains surrounding their growing volumes of data
- Enable client to reduce application development costs while addressing security risks
- Why is data growing so fast?
  - Retention of data for auditing purposes
  - Growing compliance and data privacy requirements
  - High growth data warehouses
  - Why is this a difficult problem for customers to solve?
    - Disparate applications and infrastructure
    - Auditing of both “live” and “historical” data
    - Aligned IT and business compliance is required
  - What are the pains involved?
    - Data maintenance and storage costs
    - Inactive data degrades data server performance
    - Security breaches or compliance failures can have enormous costs to the business
  - Part of Information Management Software Group
    - Data Server Division

IBM Optim

- Market leading Enterprise Data Management (EDM) Solution and Platform:
  - Data Growth
  - Retention & Discovery
  - Data Privacy
  - Test Data Management
  - Application Upgrades
  - Application Retirement
- Solving complex data management issues since 1989
- 2500+ clients worldwide; c. 50% of Fortune 500
- Princeton Softech acquired by IBM, September 2007

Optim™ Enterprise Platform

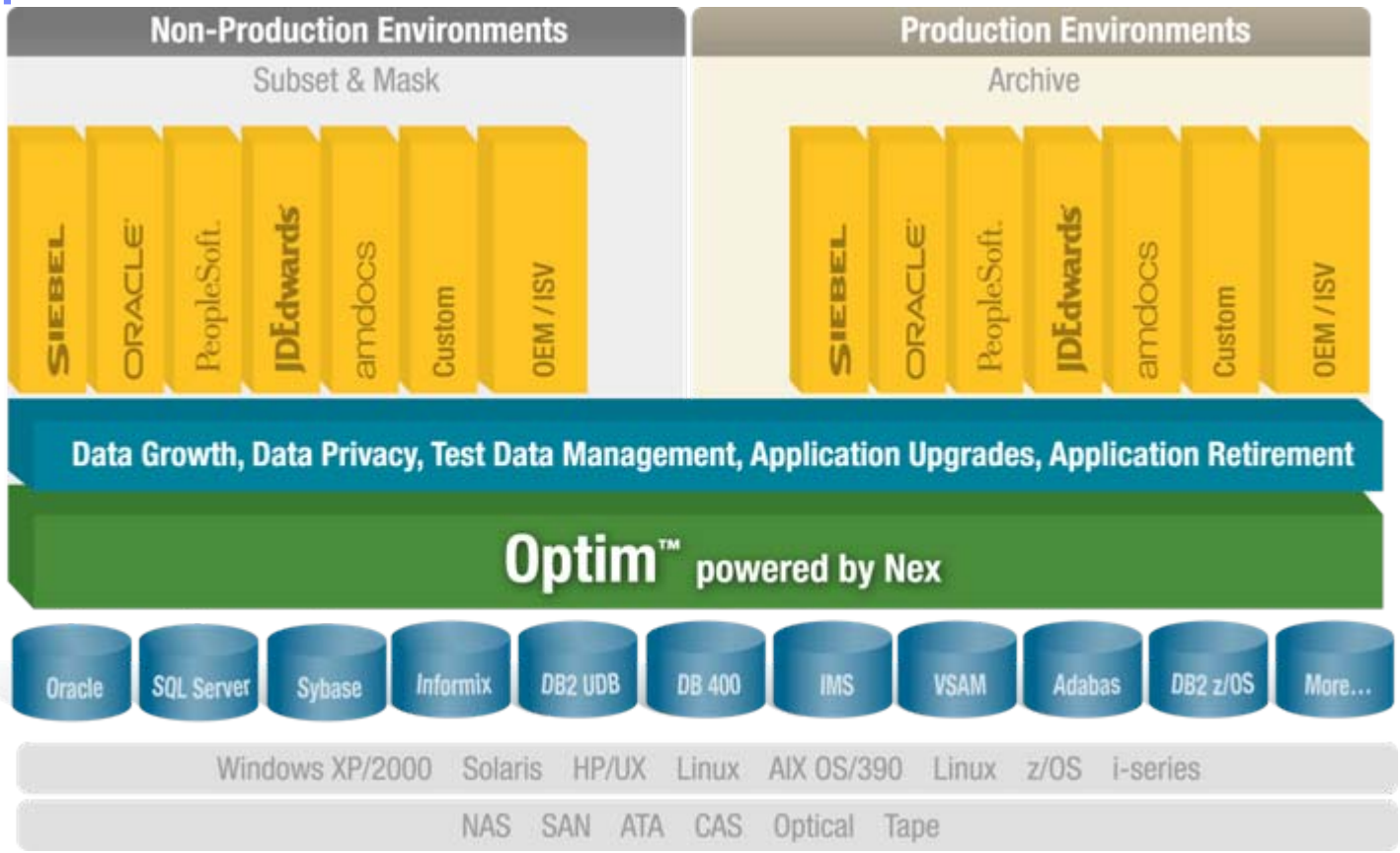
	Data Growth	Upgrades & Migrations	Application Retirement	Data Privacy	Retention & Discovery	Test Data Management
Custom Z-DB2						
Custom - UDB						
Custom - Oracle						
Custom – SQL SVR						
Custom- Sybase						
Custom – Informix						
PeopleSoft	X					
E-Business						
JDE						
Siebel						
Amdocs						
SAP	2008 Introduction					



EDM Solution Requirements – The Four Pillars

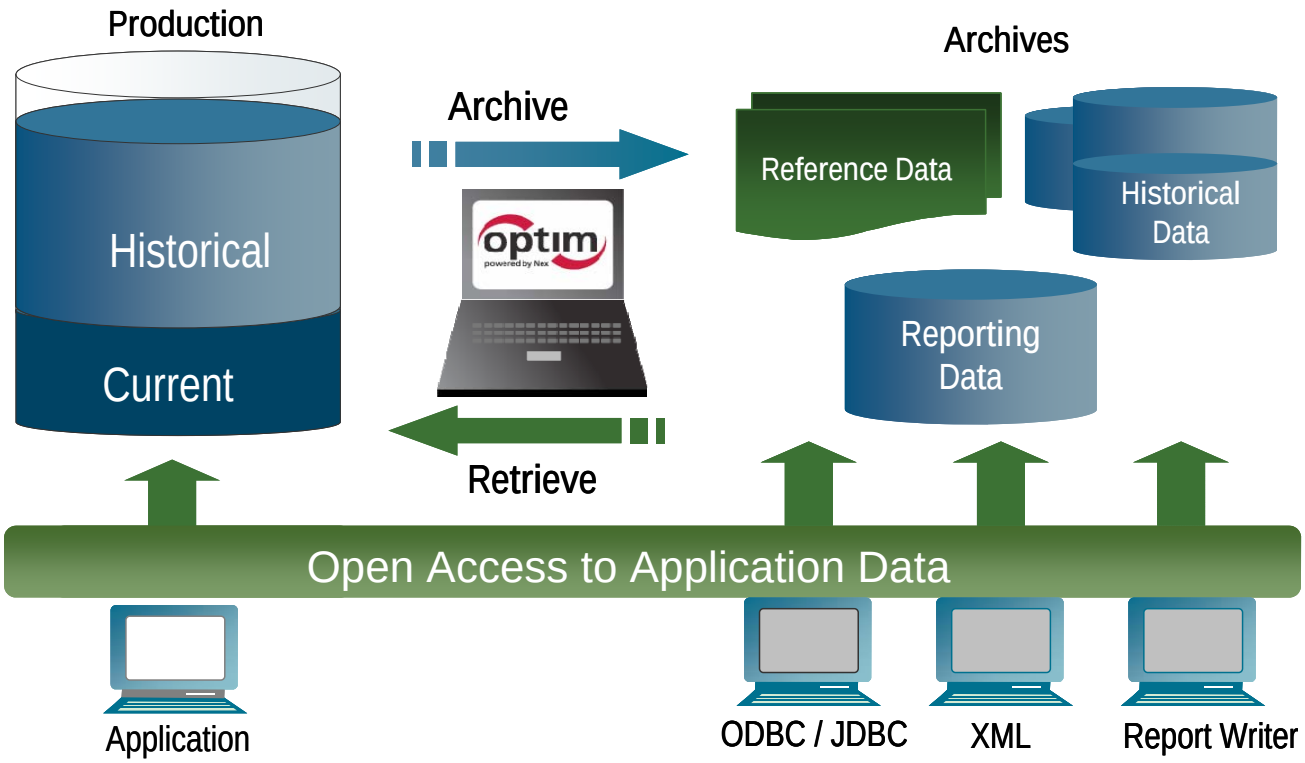
1. Enterprise Architecture
2. Complete Business Object
3. Extract, Store, Port and Protect
4. Universal Access

Enterprise Architecture



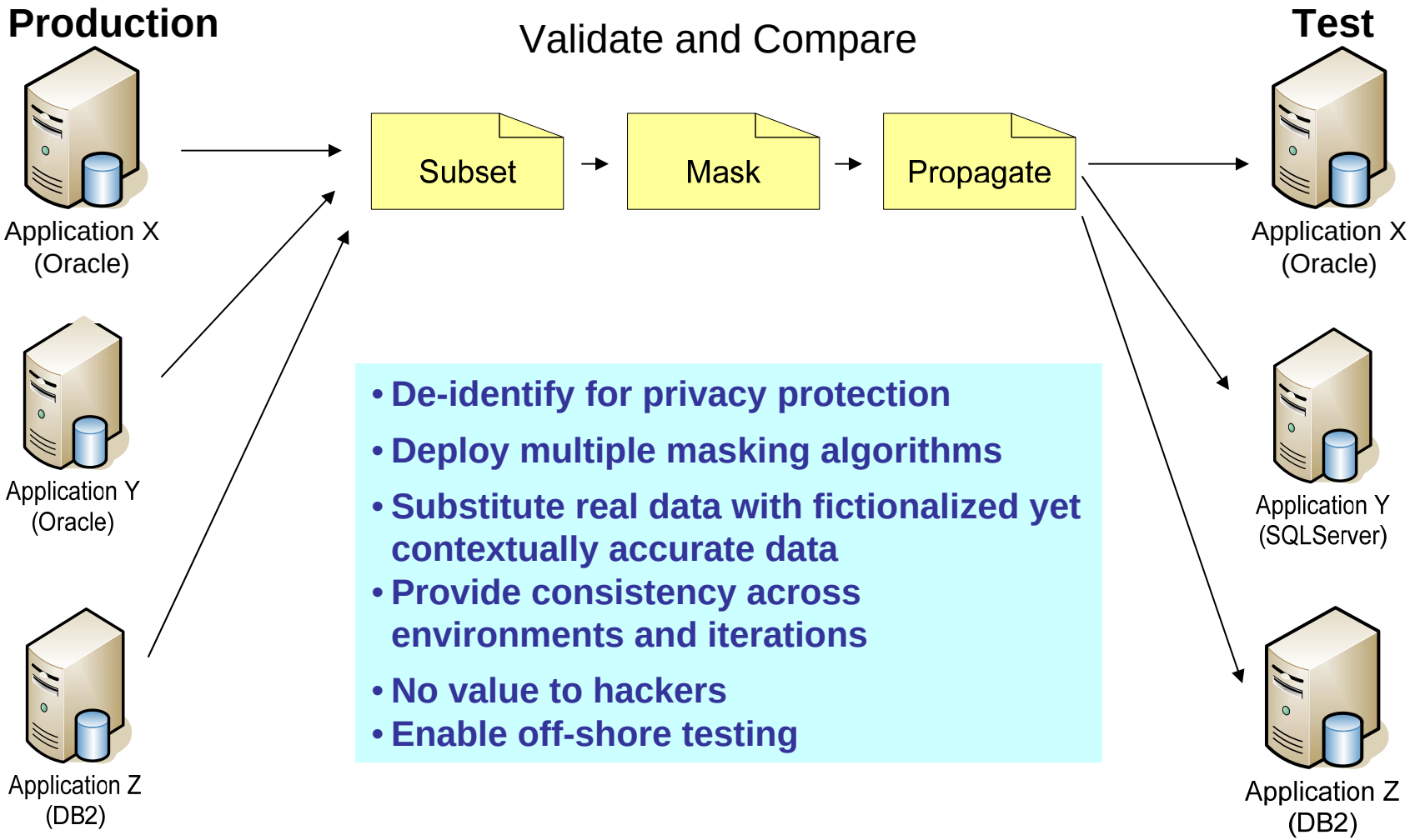
Single, scalable, interoperable EDM solution provides a central point to deploy policies to extract, store, port, and protect application data records from creation to deletion

Optim™ Data Growth Solution: Archiving



- Complete Business Object provides historical reference snapshot of business activity
- Storage device independence enables ILM
- Immutable archives enables data retention compliance

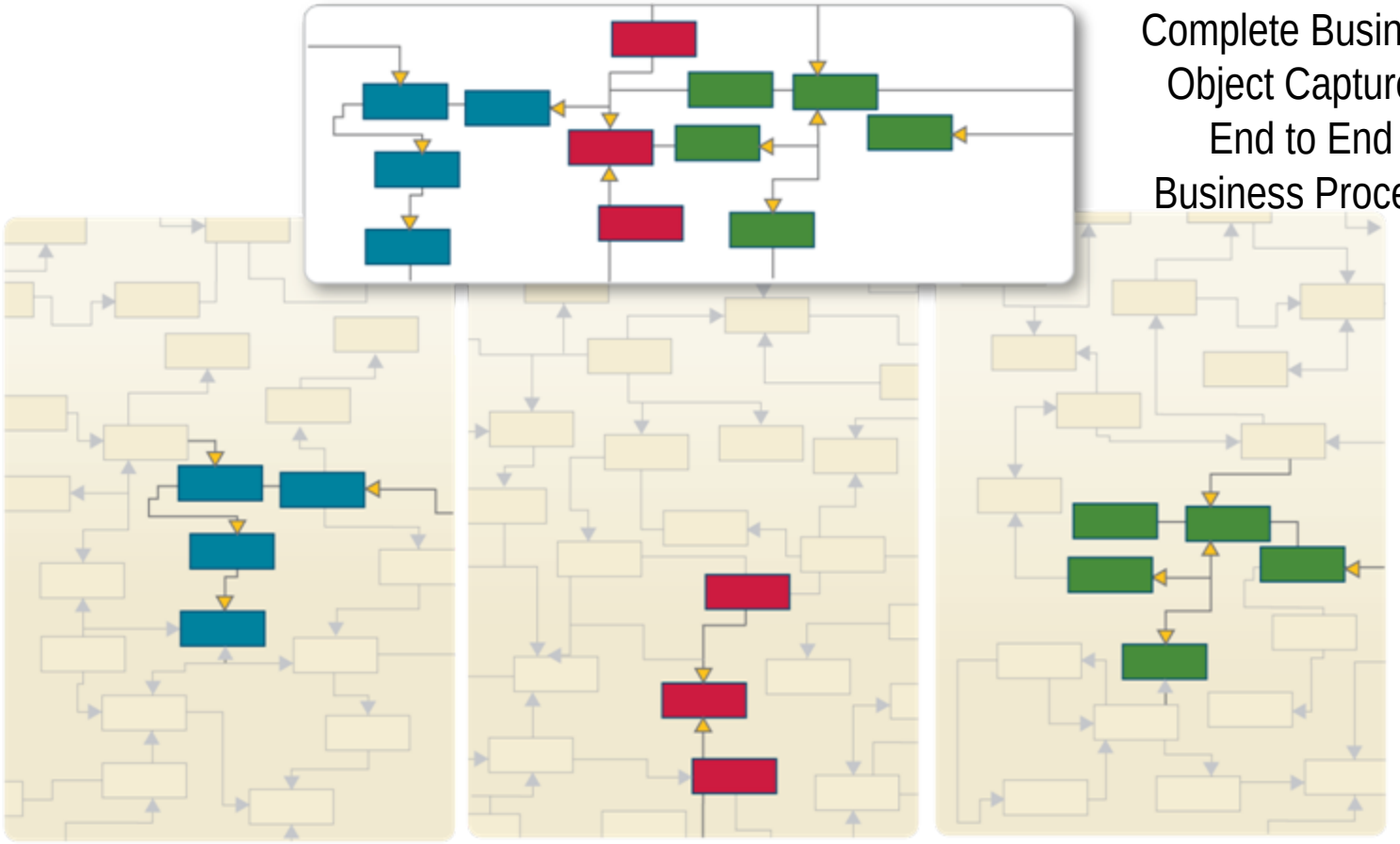
Test Data Management and Data Privacy Solution



- **De-identify for privacy protection**
- **Deploy multiple masking algorithms**
- **Substitute real data with fictionalized yet contextually accurate data**
- **Provide consistency across environments and iterations**
- **No value to hackers**
- **Enable off-shore testing**

Ensure Data Privacy Across Non-Production Environments!

Key Capability - Federated Data Support



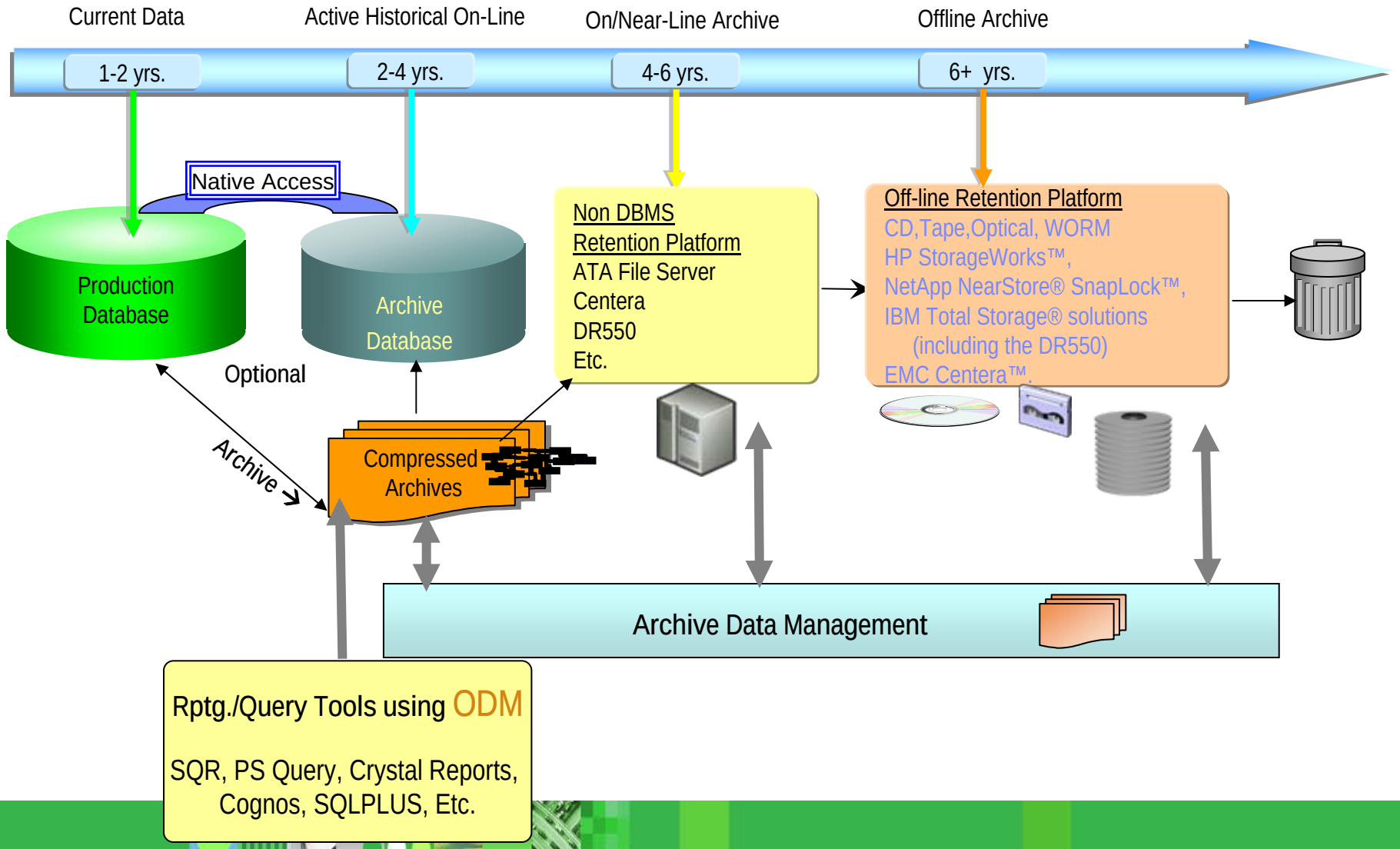
Complete Business Object Captures End to End Business Process

Peoplesoft

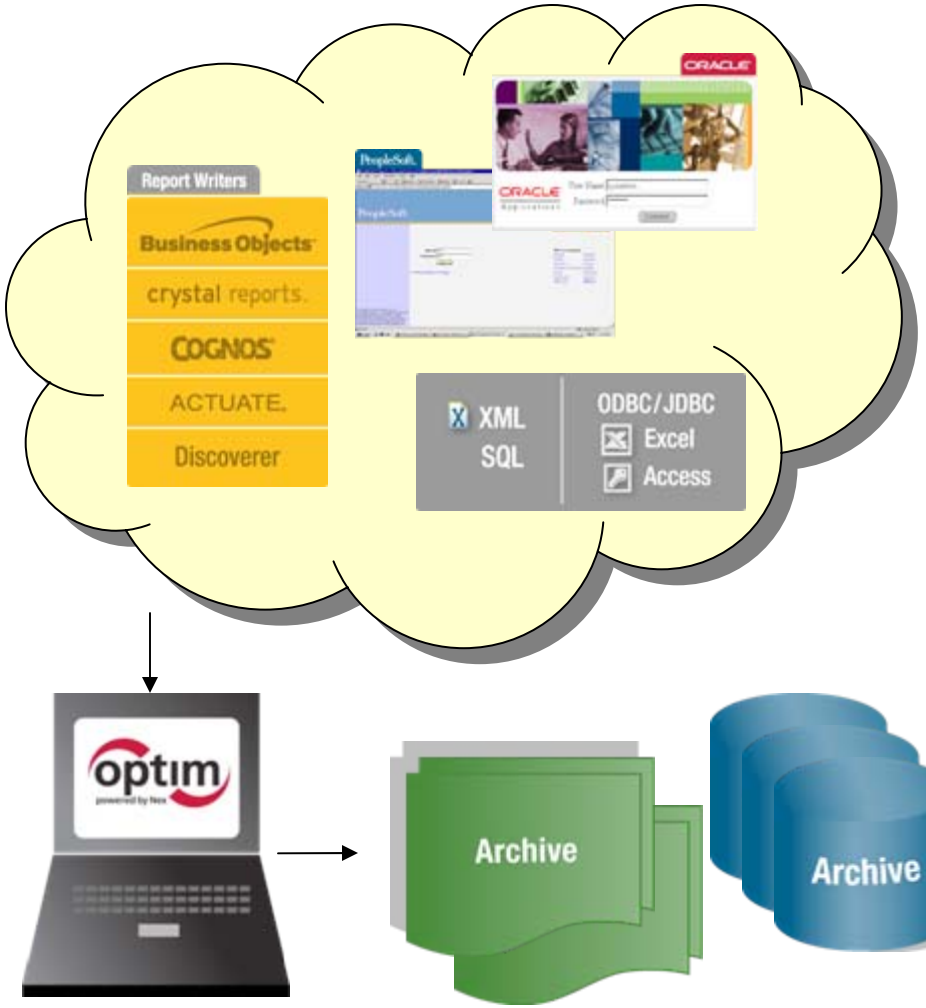
Custom Application

Other apps / any DBMS

Data Retention Strategies



Universal Access



- **Application independent access**
  - Industry standard methods: SQL, ODBC/JDBC, XML
  - Portals
  - Report writers: Crystal Reports, Cognos, Business Objects, Discoverer, Actuate
  - Desktop formats: Excel, CSV, MS Access
  - Database formats
- **Native application access**
  - Familiar screens and processes

Access Any Record, Anytime, Anywhere!

Benefits of Subsetting

- Maximize allocated disk space
- Increase number of test/dev environments
- Reduce infrastructure costs
- Realize development and test efficiencies
  - Reduce the cycle times for test upgrades
  - Reduces time and resources required to backup and maintain non-production environments
- Cap or eliminate future disk purchase

Solution

- Extract precise subsets of related data to build realistic, “right-sized” test databases
  - Create referentially intact subsets
  - Remove the bulk of production data
  - Minimize the load on testing and staging servers
- Speed iterative testing tasks with reusable processing definitions and Extract Files to ensure consistency
- Change Original Data with real-looking Masked Data

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

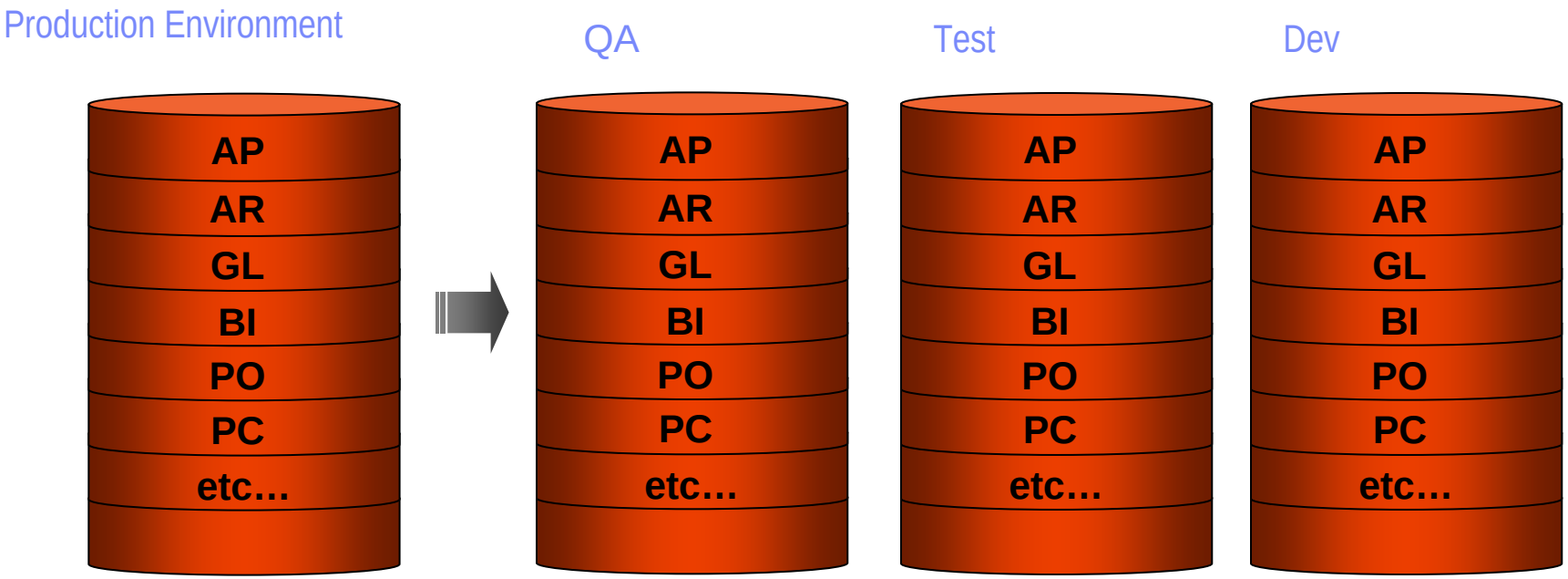
After

- RI Accuracy?
- Right Data?

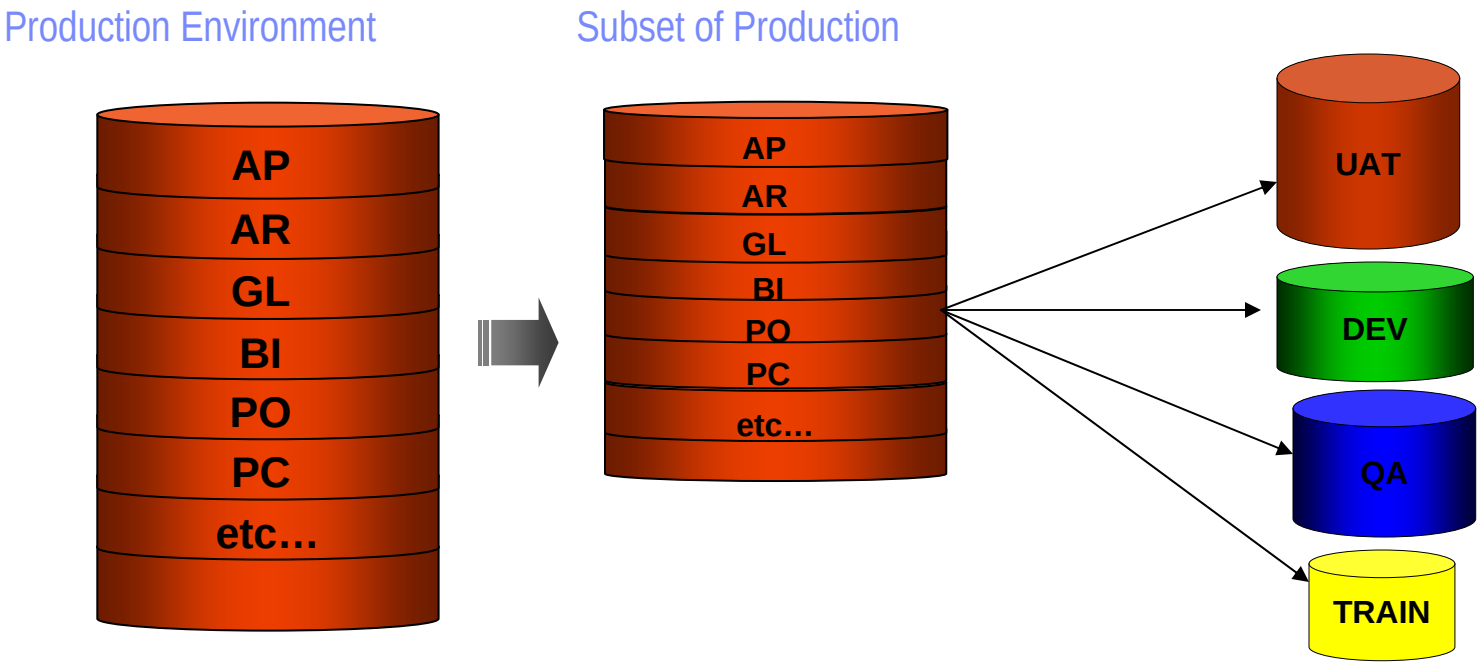
Expensive, Dedicated Staff, Ongoing Responsibility.

Share test database with everyone else

Typical Cloning



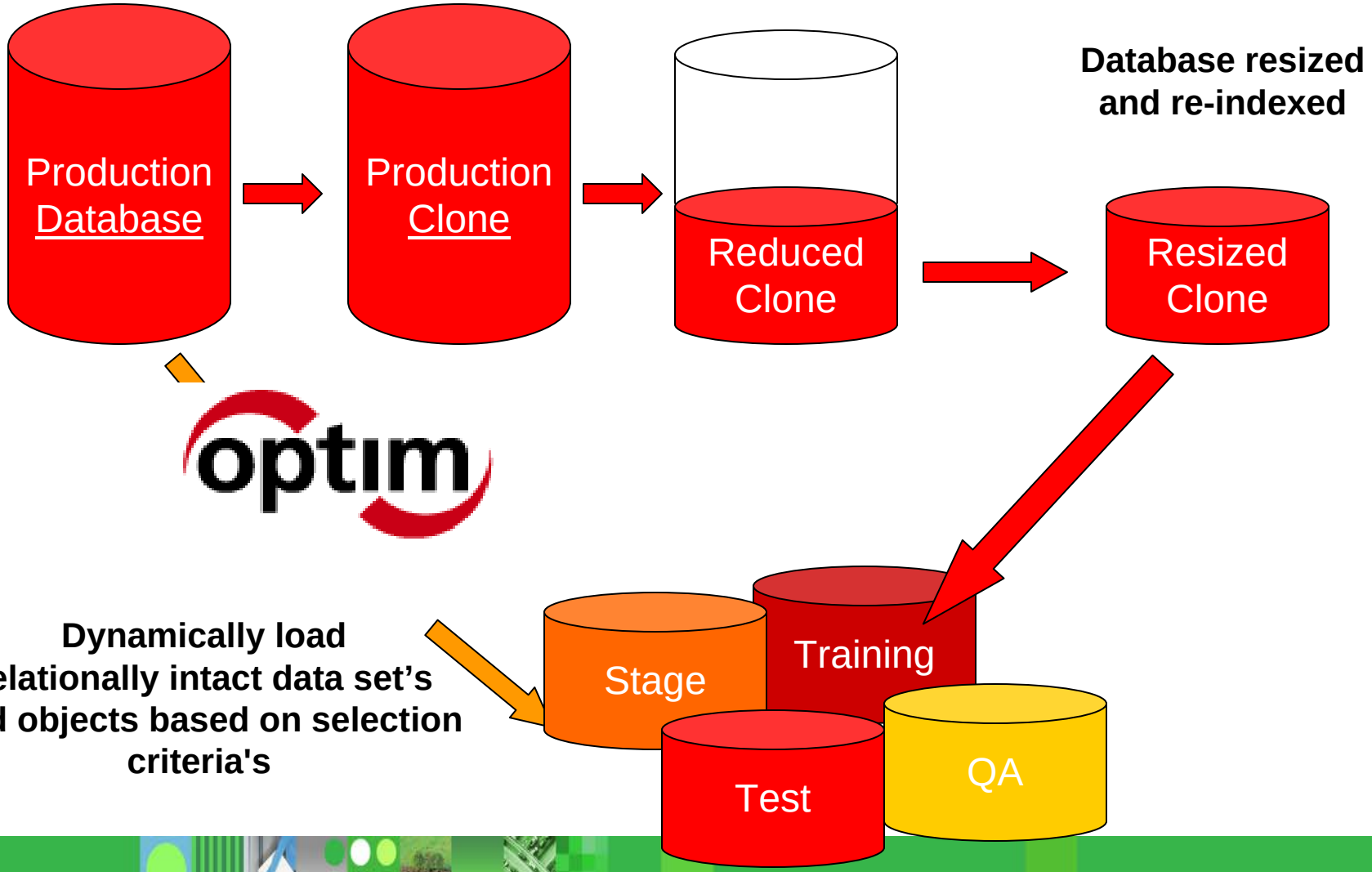
Subset Process



Conceptual Options

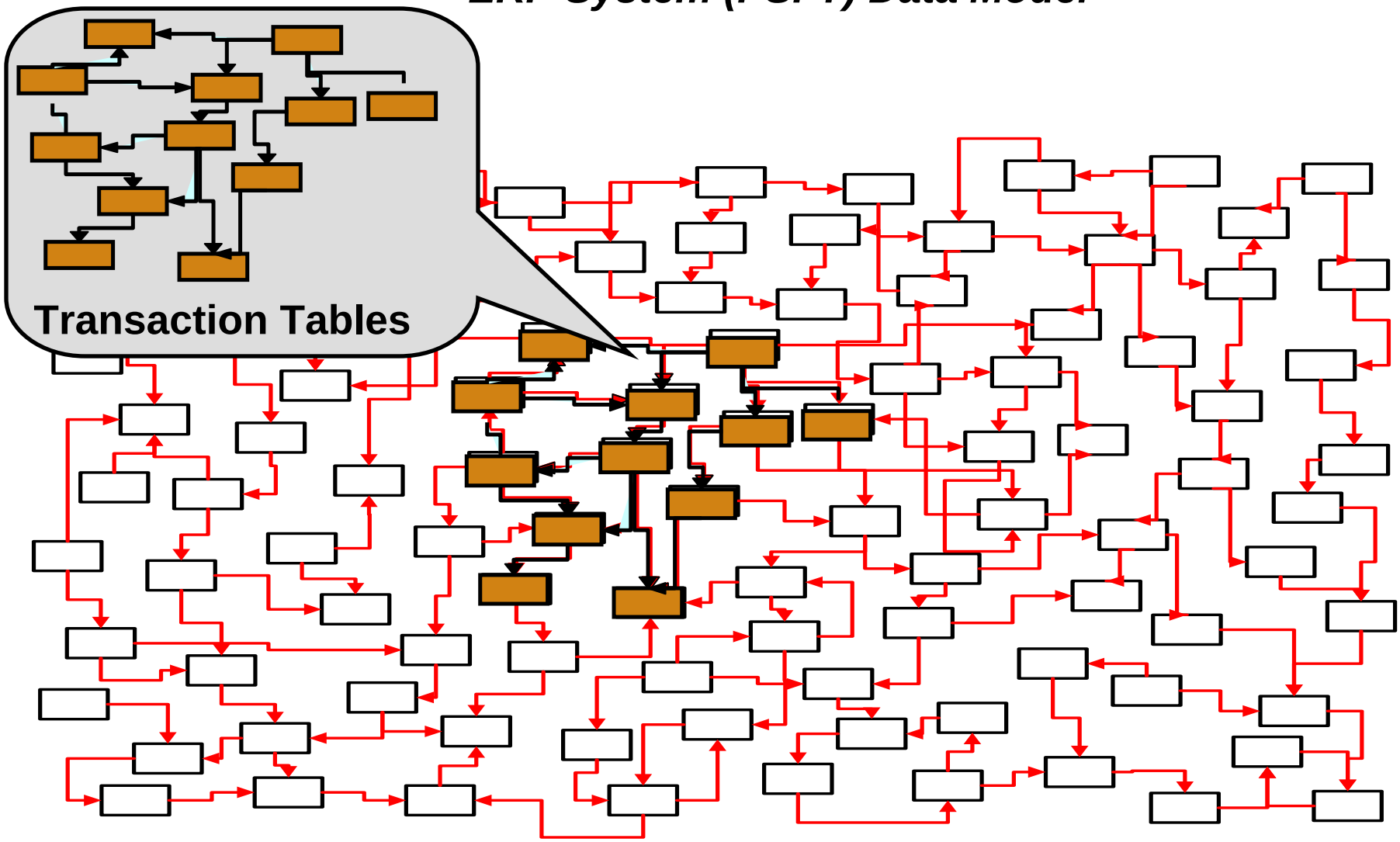
Tables are Truncated, but database footprint still the same

Database resized and re-indexed

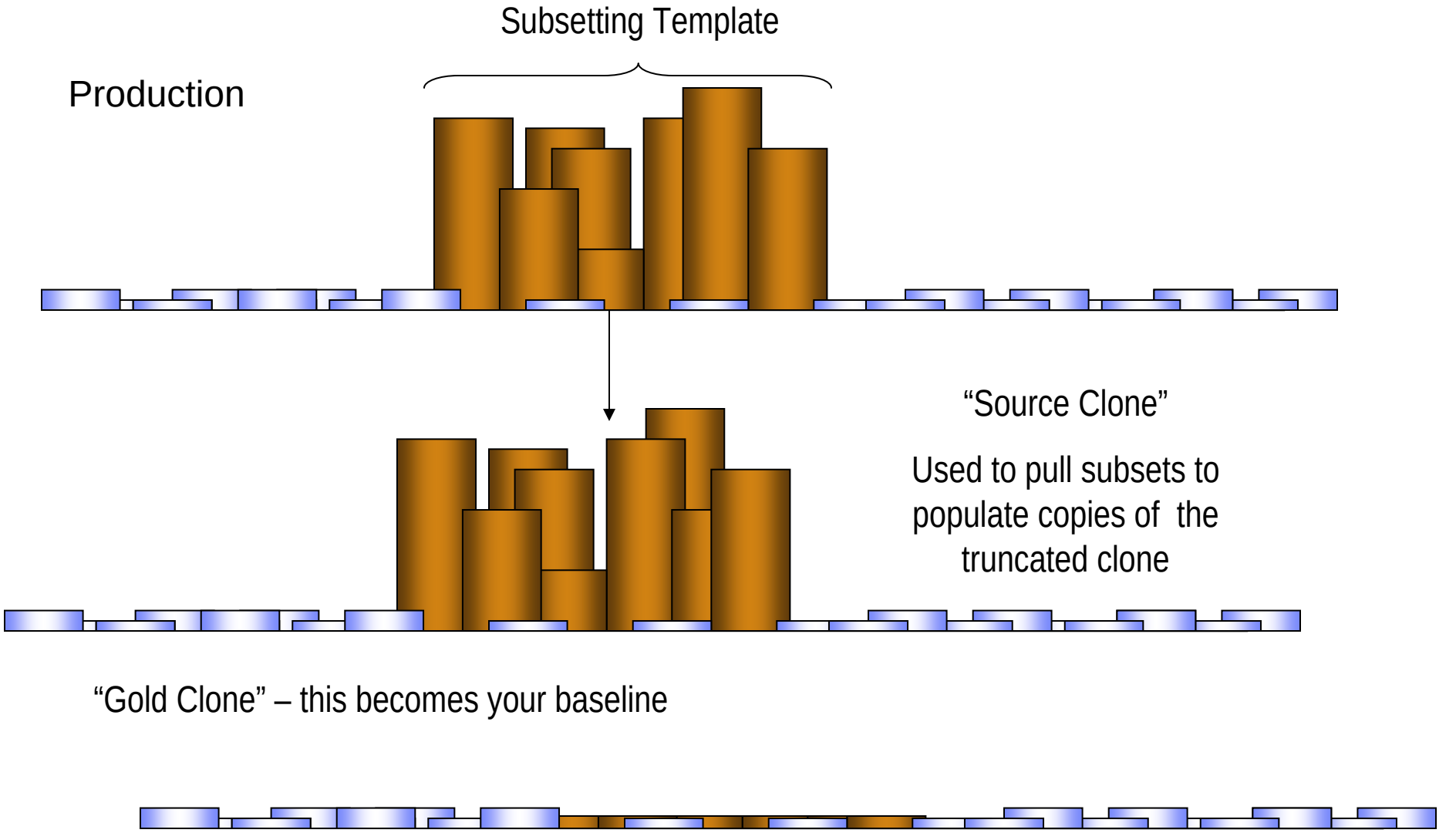


Dynamically load relationally intact data set's and objects based on selection criteria's

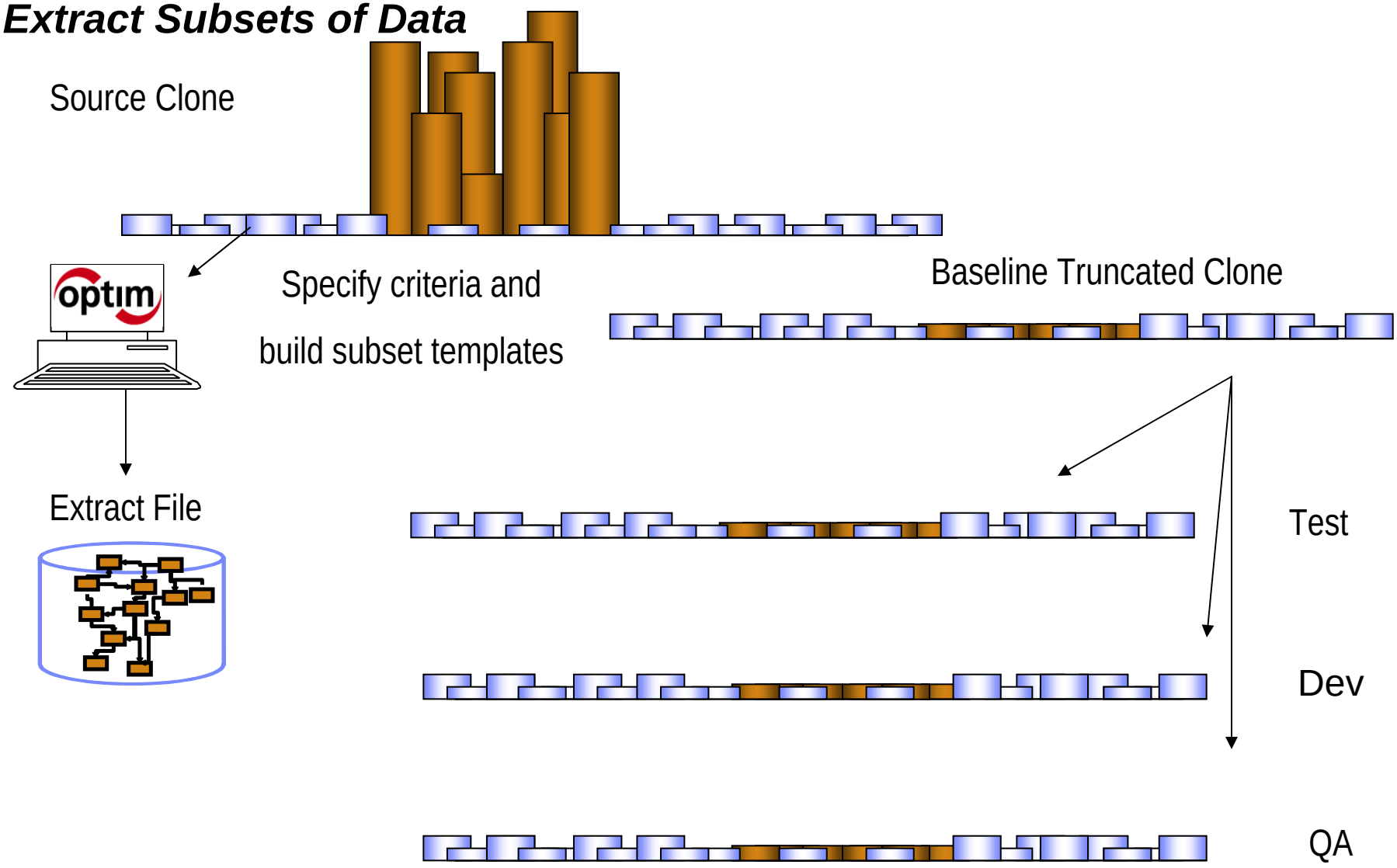
ERP System (PSFT) Data Model



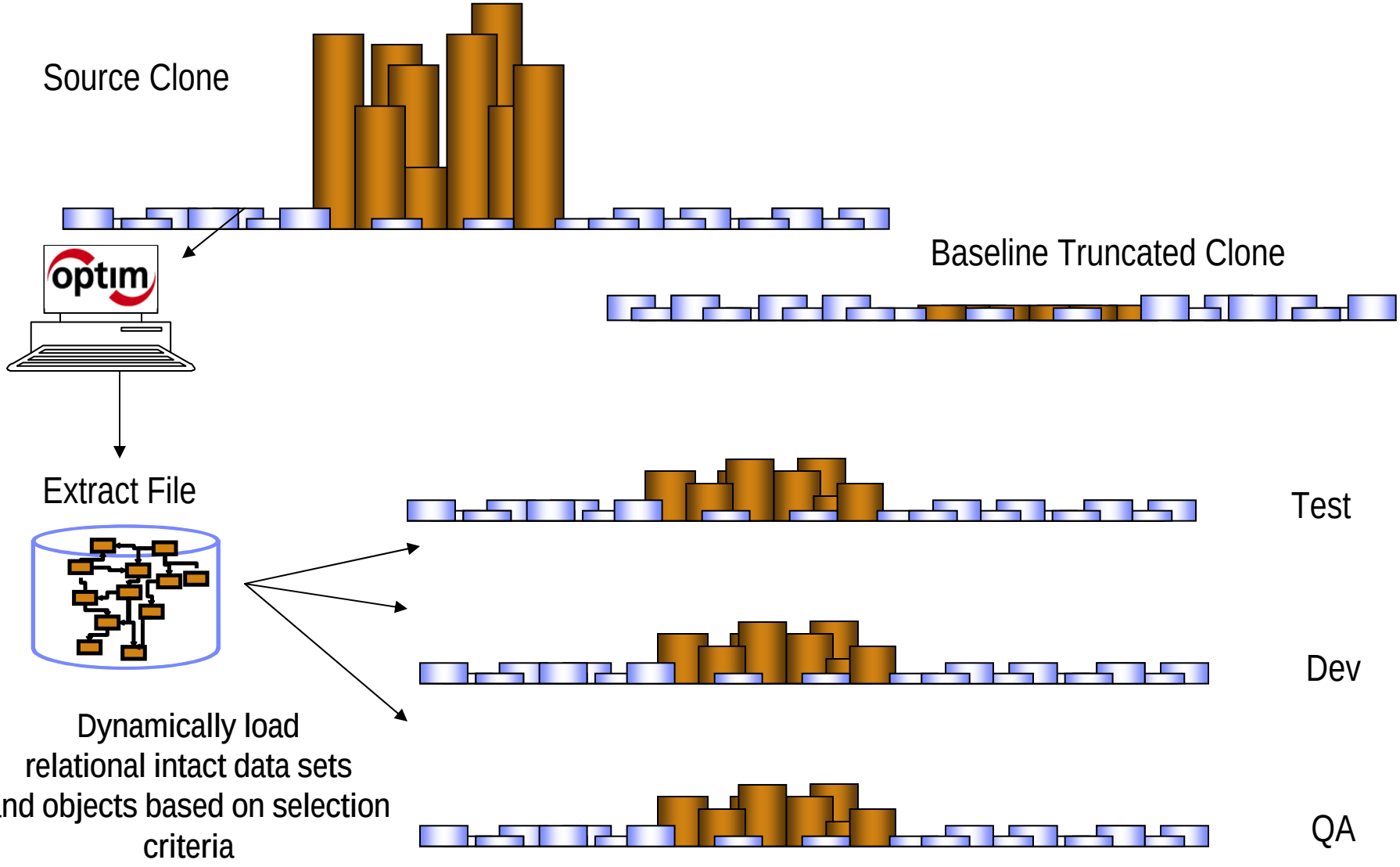
Create Gold Clone



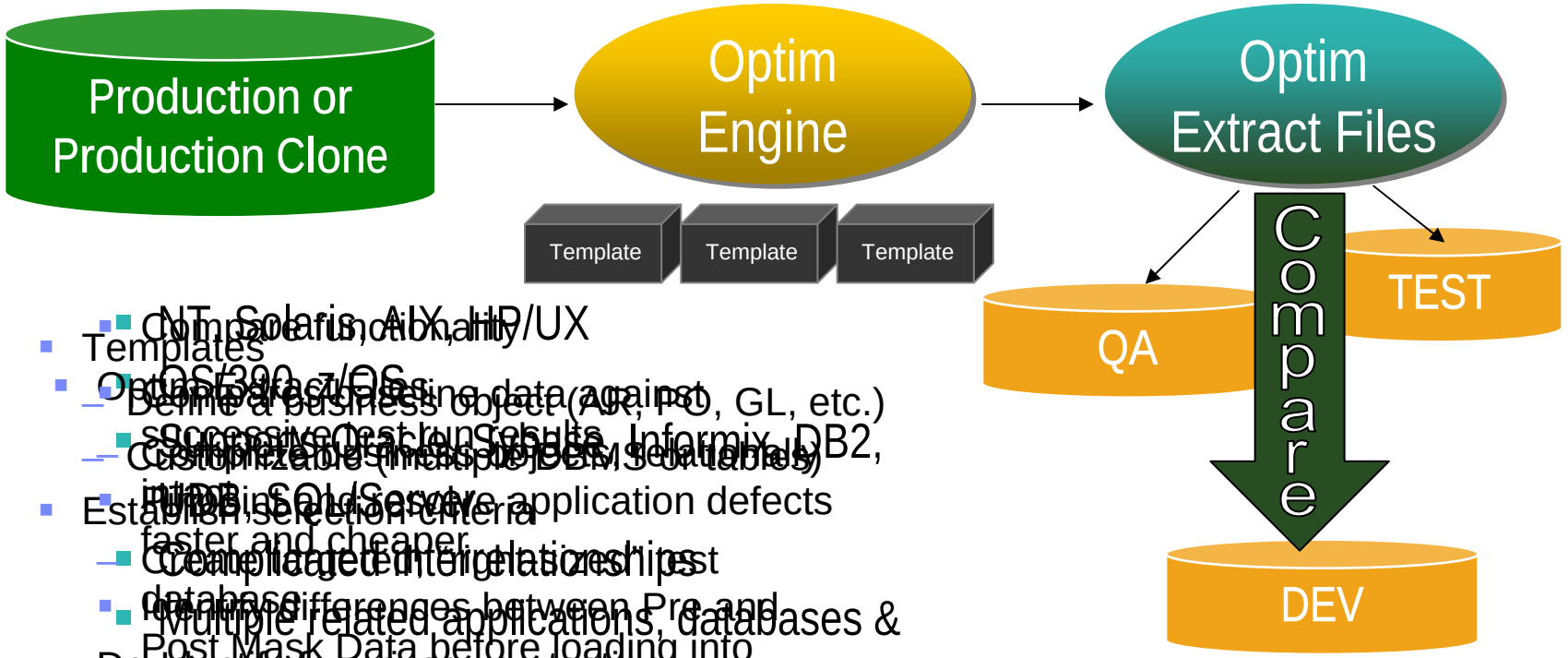
Extract Subsets of Data



Load Subset Data into Different Environments



Optim Subset Creation

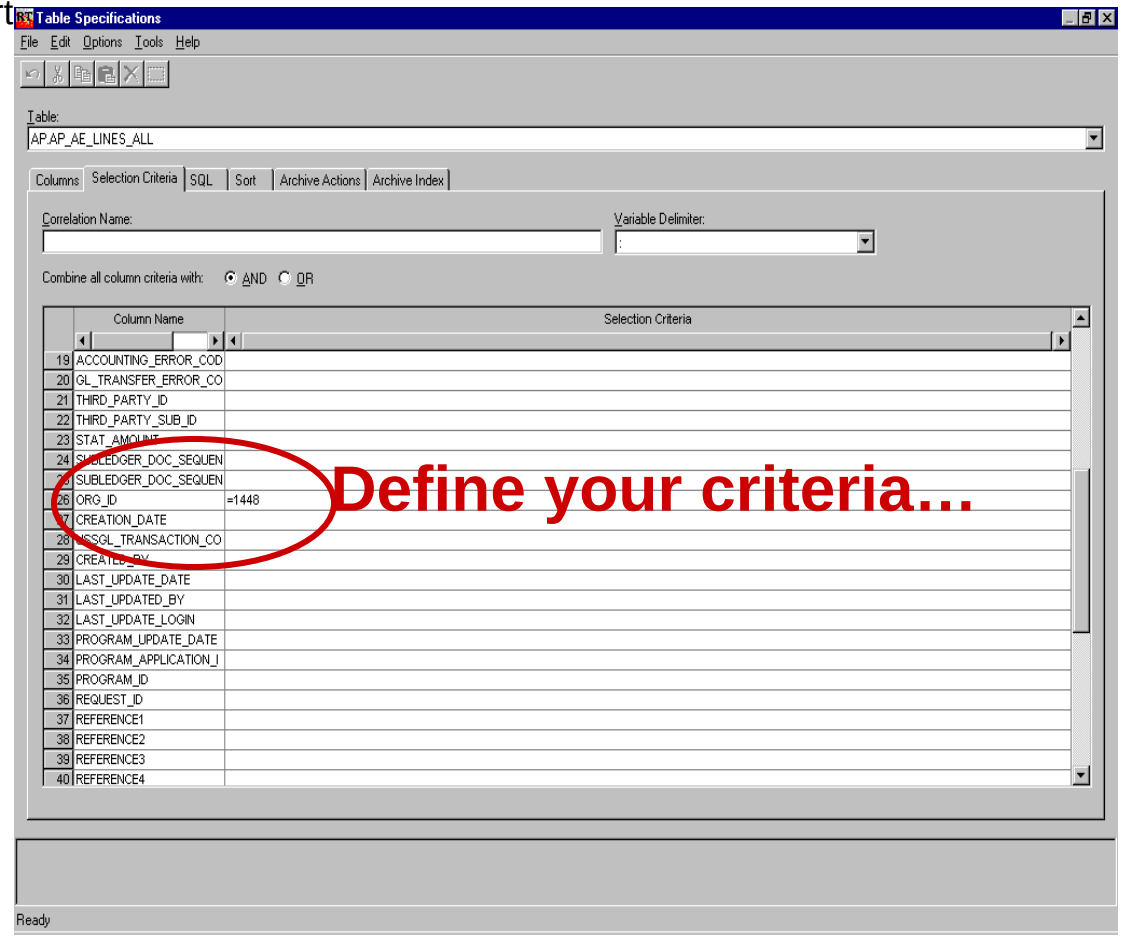


- Templates
  - Define a business object (AR, PO, GL, etc.)
  - Customizable (multiple DBMS tables)
  - Portable (NT, Solaris, AIX, HP/UX, OS/390, z/OS, Oracle, Informix, DB2, IMS DB, SQL/Server, Sybase, Teradata, Datacom, Adabas)
- Establish selection criteria
  - Complex table interrelationships
  - Multiple related applications, databases & platforms
- De-Identify for privacy protection
  - Deploy one or more masking algorithms
  - Substitute real data with fictionalized but contextually accurate data
  - No value to hackers

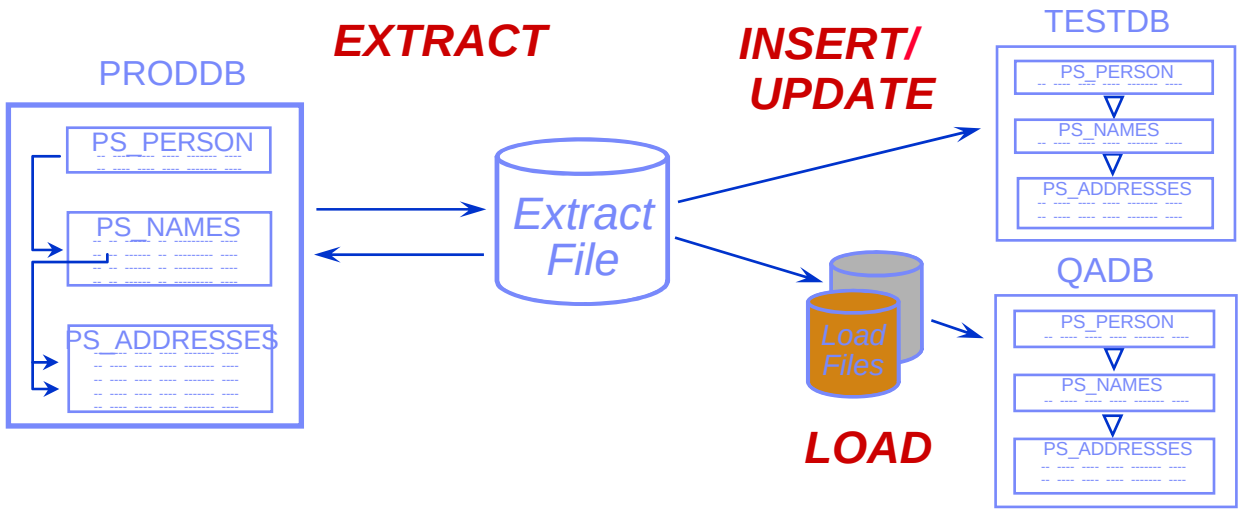
- Compare functionality
  - Validate testing data against successive test runs
  - Identify application defects faster and cheaper
  - Compare Pre and Post Mask Data before loading into target database
  - Audit changes to a database
  - Single-table or multi-table compare

Creating Subset Templates

- Specify % of data to be used for subsetting
- Specify Max # of rows to be used from the Start Table
- Add tables including any custom tables
- Add relationships
  - Same instance
  - across instances and database platforms
- Define Relationship Traversals
- Criteria can be based on one or more modules
- All Date Values
  - Accounting Date
  - Fiscal Year
- Business Unit, Ledger Groups Status
- “And/Or” combinations
- More….



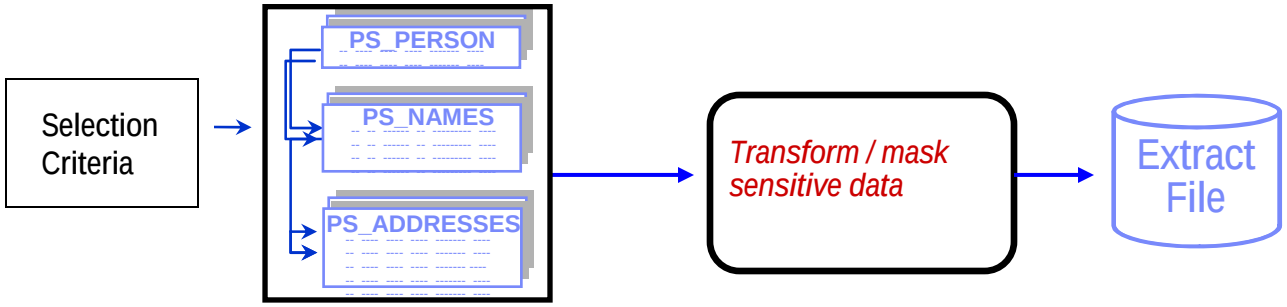
The Relational Extract Facility



- GUI based - saves Programmer/DBA Time, Disk Space
- Minimizes Testing Interference
- Facilitates:
  - Creating and Maintaining Test Databases
  - Migrating Referentially Intact Subsets of Data

Defining the Extract

Extract a relationally intact subset from production database(s)



- Extract data and/or object definitions
- Extract from source tables using dynamic SQL
- Select start table or define a set of tables
- Enter selection criteria
  - SQL where-clause
  - random rows, data partitioning & sampling
- Control traversal; Show steps

Why do something with Optim

Optim solves the market need…

- Extract precise subsets of related data to build realistic, “right-sized” test databases
  - Create referentially intact subsets
  - Flexible criteria for determining record sets
- Speed iterative testing tasks with reusable processing definitions and Extract Files to ensure consistency
- De-identify sensitive data in the test environment to ensure compliance with regulatory requirements for data privacy.
- Compare the test data before and after

What Clients can do with Optim:

- Effectively Manage Data Growth
  - Improve application performance
  - Control data growth, save HW & storage
  - Delay CPU/MIPS upgrades
  - Streamline application upgrades
- Perform True and Optimized Test Data Management
  - Create targeted, right sized test environments, save HW & storage
  - Improve application quality, Compare to pinpoint app defects faster
  - Speed iterative testing processes
- Facilitate Data Privacy
  - Mask confidential data, Mitigate Risk
  - Comply with privacy policies, enable offshore testing with privatized data
  - Exceed PCI standards
- Enhance Compliance
  - Decommissioning/Sunset unused applications
  - Low cost of Data retention compliance

Thank You