IBM Technical Summit

Stay ahead.

IBM

Test Data Management in the New Era of Computing

Vinod Khader

IBM InfoSphere Optim Development

Agenda

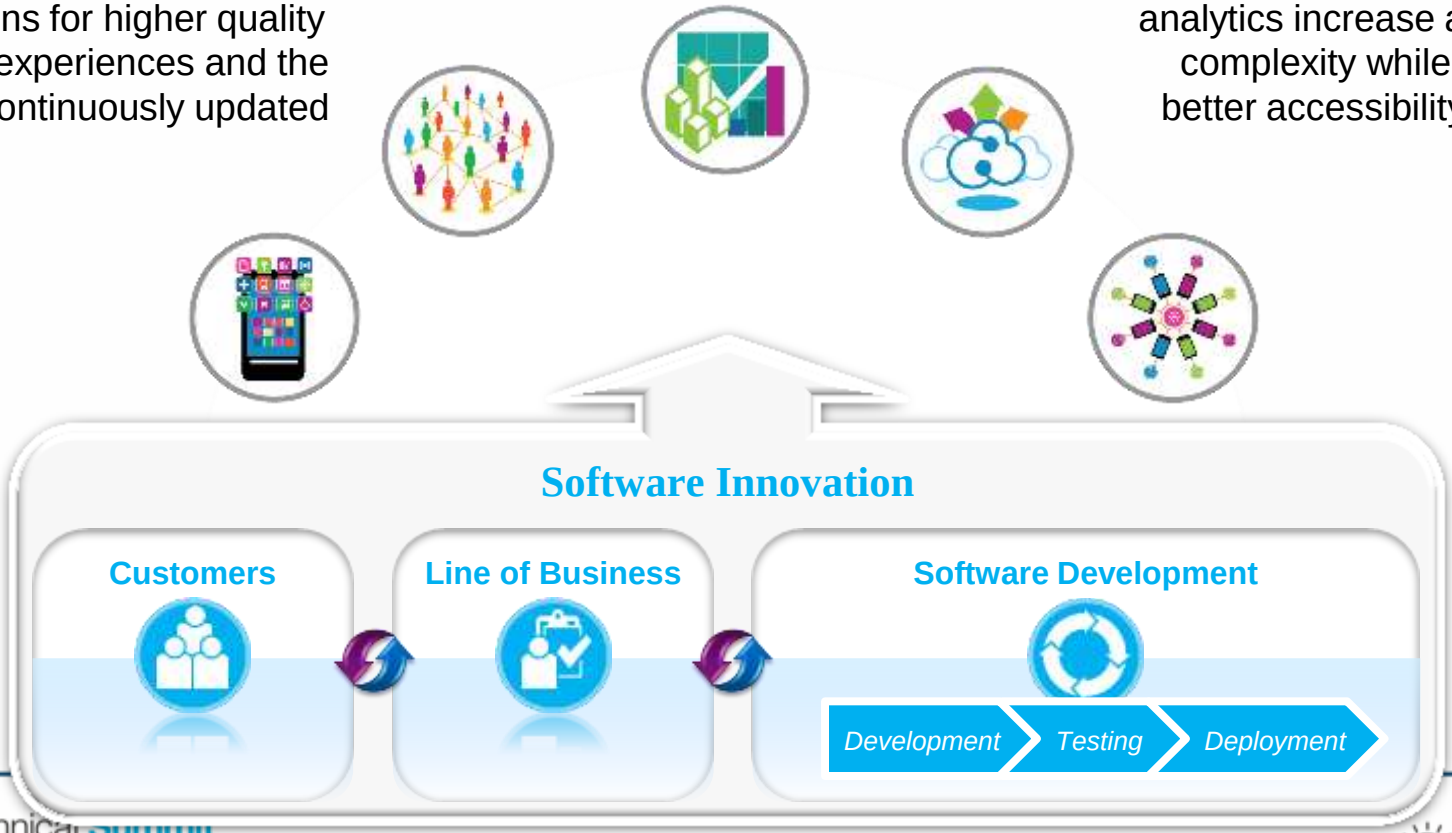
- **Changing Business Environment and Data Management Challenges**
- What is Test Data Management
- Best Practices in Test Data Management
- Solution for Test Data Management
- InfoSphere Optim Test Data Management with Rational Test Workbench & Rational Quality Manager
- Conclusion

Market shifts are fundamentally changing the way businesses approach software innovation

Empowered Users
The exponential increase in empowered users drives expectations for higher quality customer experiences and the need for continuously updated software

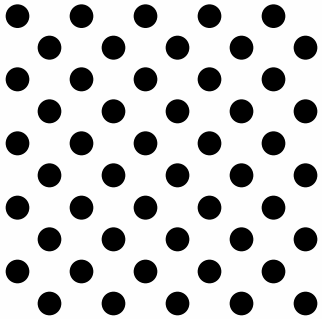
Macro Business Environment
Volatile economic and changing regulatory environments require businesses adapt quickly to changing market conditions

Technology Trends
Mobile, cloud, big data, social, agile and delivery to analytics increase application complexity while promising better accessibility, flexibility and agility



Big data is a phenomenon that results in increased complexity

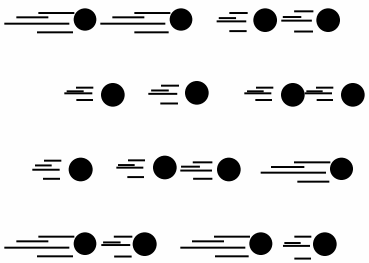
Volume



Data at rest

Terabytes to exabytes of existing data to process

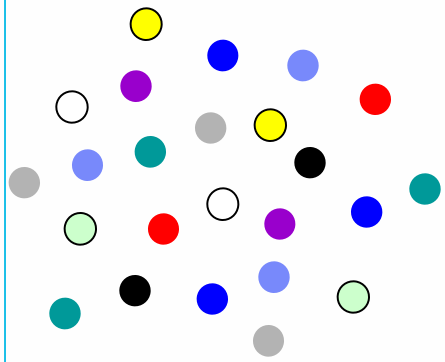
Velocity



Data in motion

Streaming data, milliseconds to seconds to respond

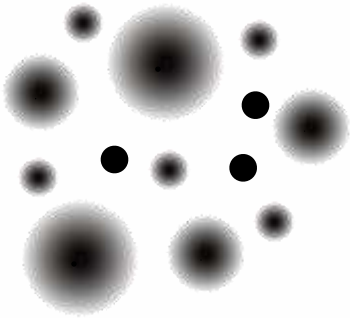
Variety



Data in many forms

Structured, unstructured, text, multimedia

Veracity*



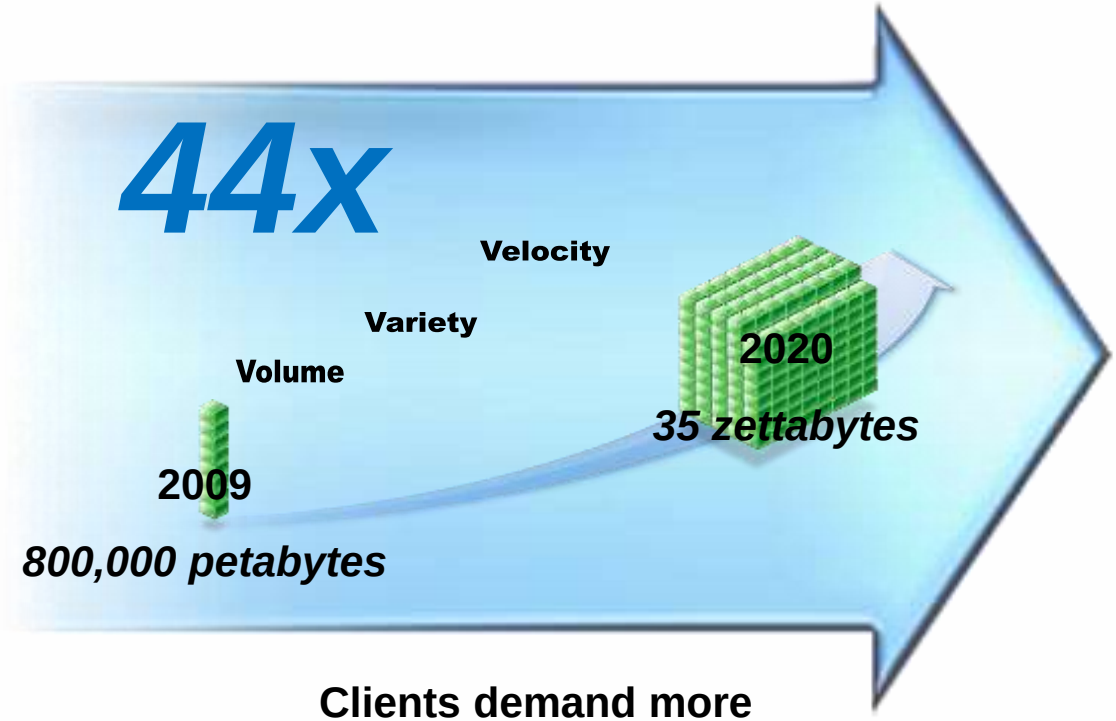
Data in doubt

Uncertainty due to data inconsistency & incompleteness, ambiguities, latency, deception, model approximations

* Truthfulness, accuracy or precision, correctness

Challenges often mean shortcuts during test & development

More pressure to deliver high quality transactional & analytical applications faster



Clients demand more functionality faster!

Client needs are growing.

Data sources are growing.

How do development & testing environments scale to handle the big data explosion?

How are different types of data protected as they speed across the enterprise?

In a rush to beat the competition, are you testing post deployment?

Building quality applications – Challenges Quantified

Increasing Risk

45,000+

Number of sensitive records exposed to 3rd party during testing [c]

62%

companies use actual customer data to test applications[a]

Time to Market

37%

Satisfied with speed of software development[f]

30-50%

Time testing teams spend on setting up test environments, instead of testing[b]

Increasing Costs

$300 billion

Annual costs of software-related downtime.[d]

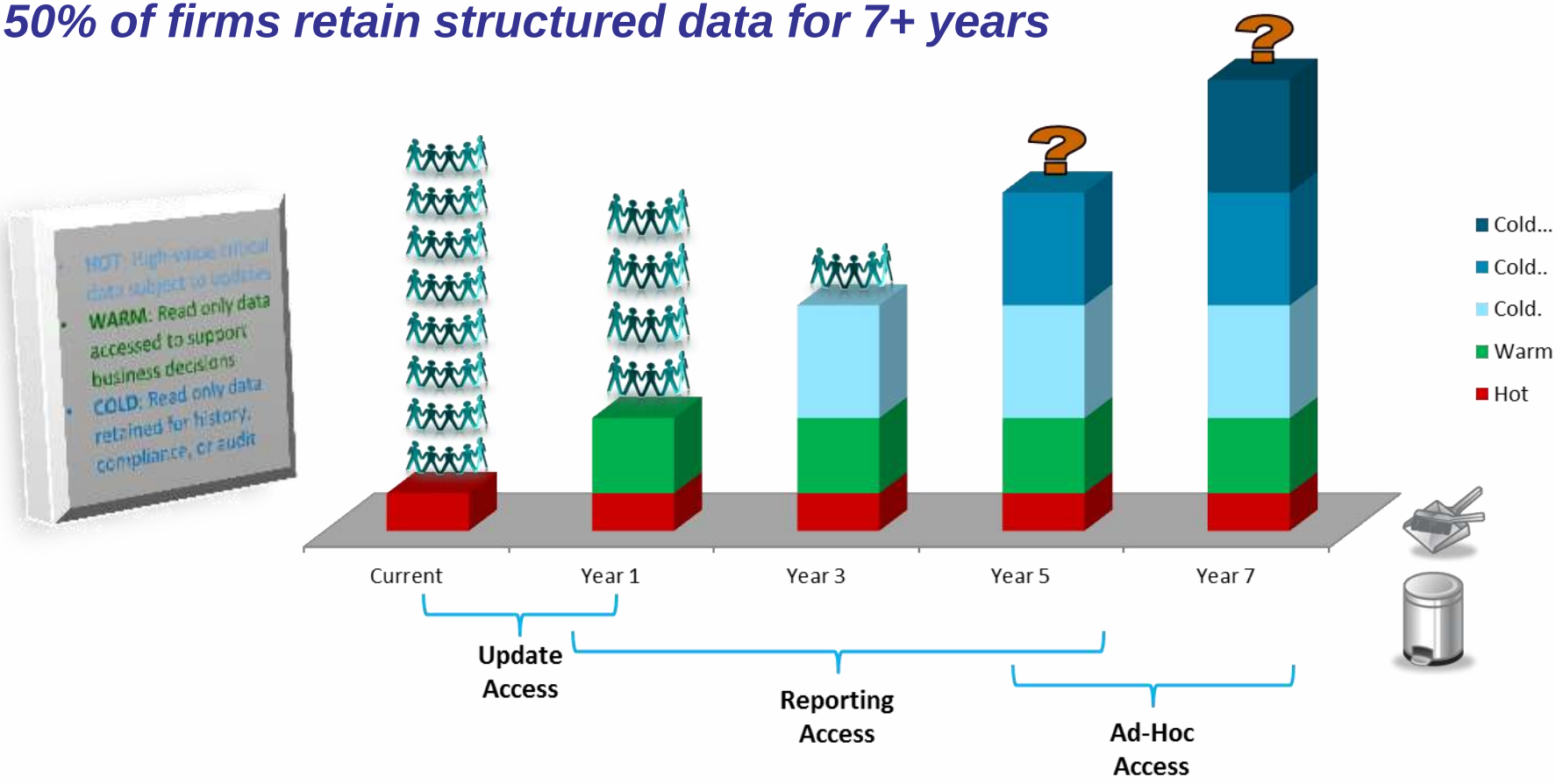
32%

Low success rate for software projects[e]

a. The Ponemon Institute. The Insecurity of Test Data: The Unseen Crisis
b. NIST, Planning Report. The Economic Impacts of Inadequate Infrastructure for Software Testing
c. Federal Aviation Administration: Exposes unprotected test data to a third party http://fcw.com/articles/2009/02/10/faa-data-breach.aspx
d. The Standish Group, *Comparative Economic Normalization Technology Study,* CHAOS Chronicles v12.3.9, June 30, 2008
e. The Standish Group, Chaos Report, April 2009
f. Forrester Research, *"Corporate Software Development Fails To Satisfy On Speed Or Quality",* 2005

Data classification

50% of firms retain structured data for 7+ years



Majority of your resources are deployed to manage COLD data

The impact of inefficient test practices

"We did not want to create 'fake' or unrealistic test data. All test data had to be created and set up manually. So it would take us a month to setup test data for 30 or more accounts."

-- Large US Healthcare Insurer

"We needed to improve efficiencies in development and testing environments, as well as our production environments. We can create realistic test environments that use much less disk space than we would by cloning the production database."

-- Allianz Seguros

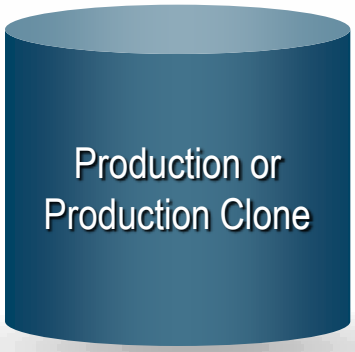
"Our staff wanted to implement more efficient and cost-effective testing processes that would shorten the time for creating and managing multiple test environments."

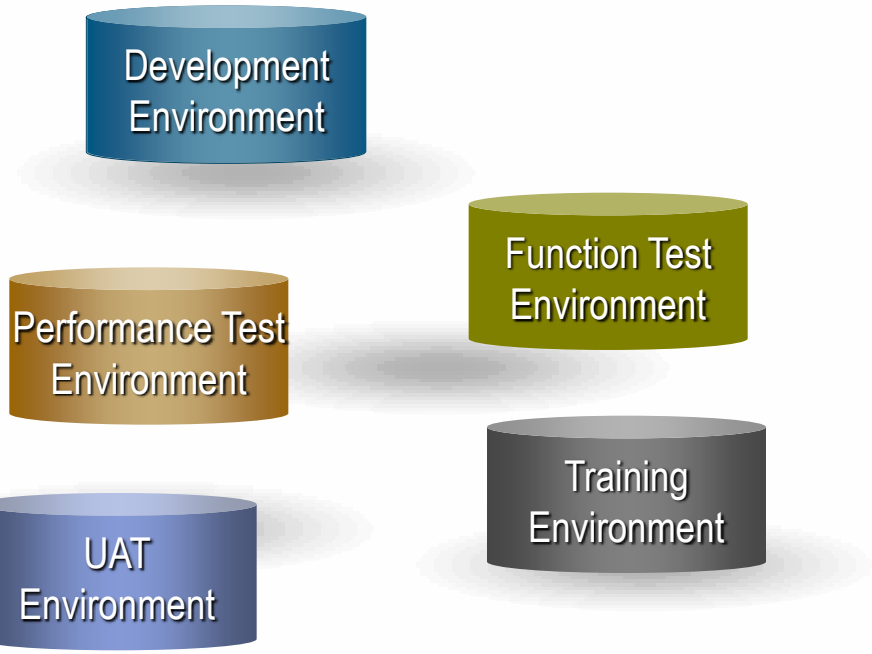
-- Cetelem

Agenda

- Changing Business Environment and Data Management Challenges
- **What is Test Data Management**
- Best Practices in Test Data Management
- Solution for Test Data Management
- InfoSphere Optim Test Data Management with Rational Test Workbench & Rational Quality Manager
- Conclusion

What is Test Data Management?





Create targeted, right-sized test environments instead of cloning entire production environments.

Development environments are then more manageable, speeding the testing process!

Agenda

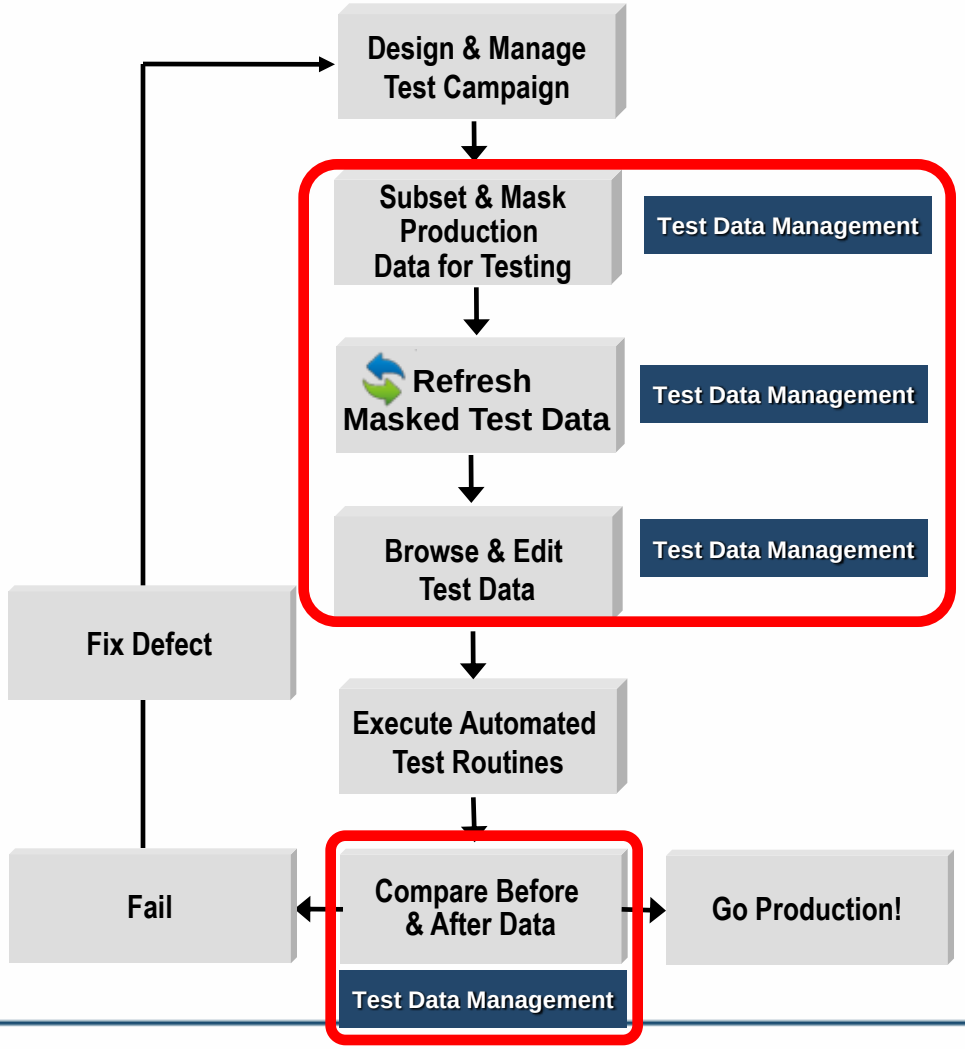
- Big data and software delivery
- What is Test Data Management
- **Best Practices in Test Data Management**
- Solution for Test Data Management
- InfoSphere Optim Test Data Management with Rational Test Workbench & Rational Quality Manager
- Conclusion

Where does test data management fit into the testing discipline?

- **Comprehensive** software quality process to minimize cost and shorten development cycles

- Streamline test data management processes and **deliver projects sooner and with fewer defects**

- **Define, enforce, reuse & extend policies and standards**; execute consistently



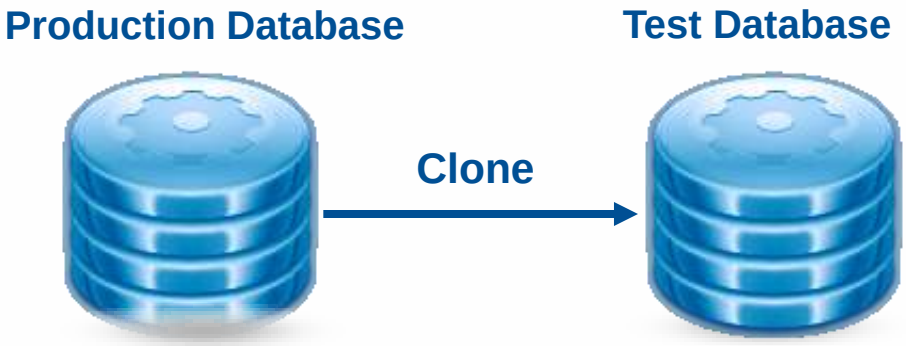
Cloning is often used as a short cut to proper test data management

Positives

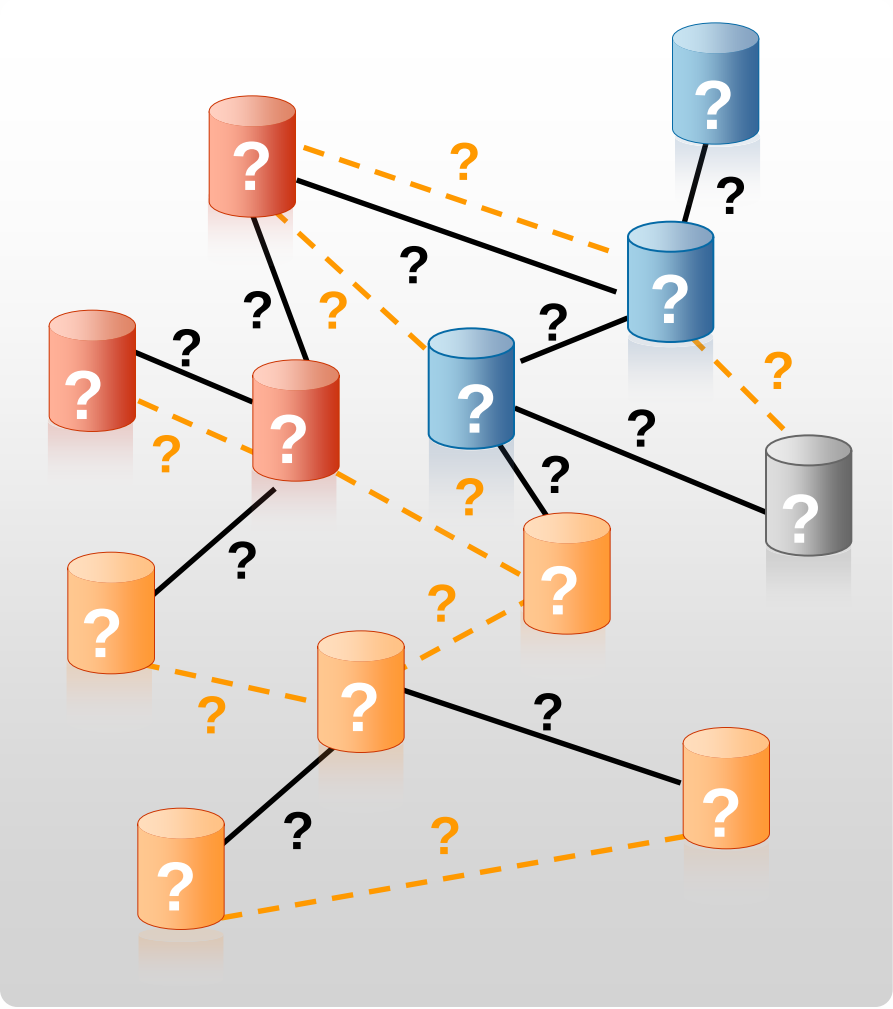
- **Simple** - requires little knowledge of technology
- **Realistic** - creates an exact copy of production

Negatives

- **Costly** - significant storage; more than required
- **Not precise** – no specific use cases or teams
- **Risky** - sensitive data used in test
- **Time consuming** - must copy all of production
- **Hard to use** - no way to analyze before/after test
- **Not scalable** - across sources/applications
- **Inefficient** – downtime waiting for test data

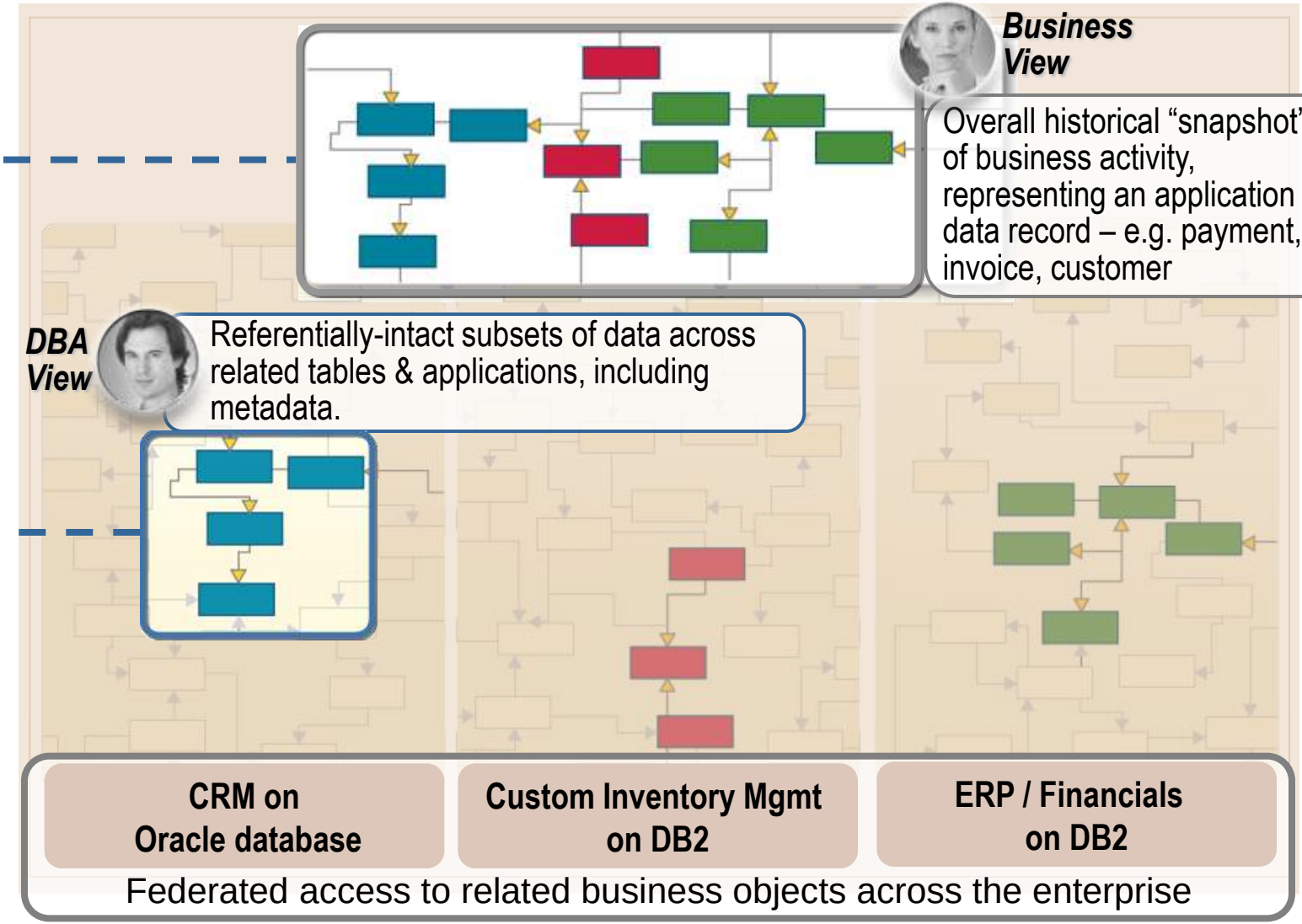


You can't govern what you don't understand



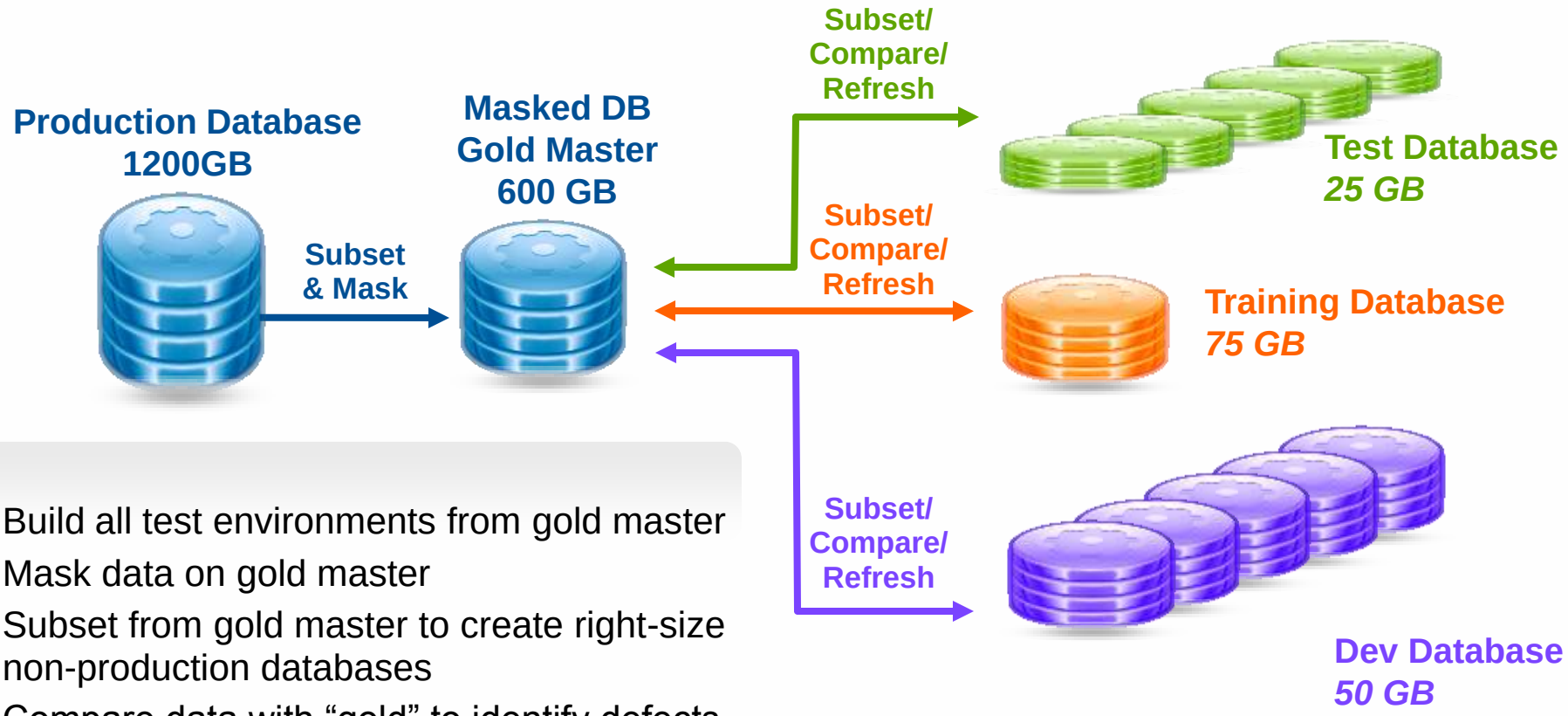
- Complex data relationships within and across sources
- Historical and reference data
- Test data needed to satisfy test cases
- Sensitive data identification

Discover & define business objects across heterogeneous databases & applications



More strategic approaches create optimized, secure environments

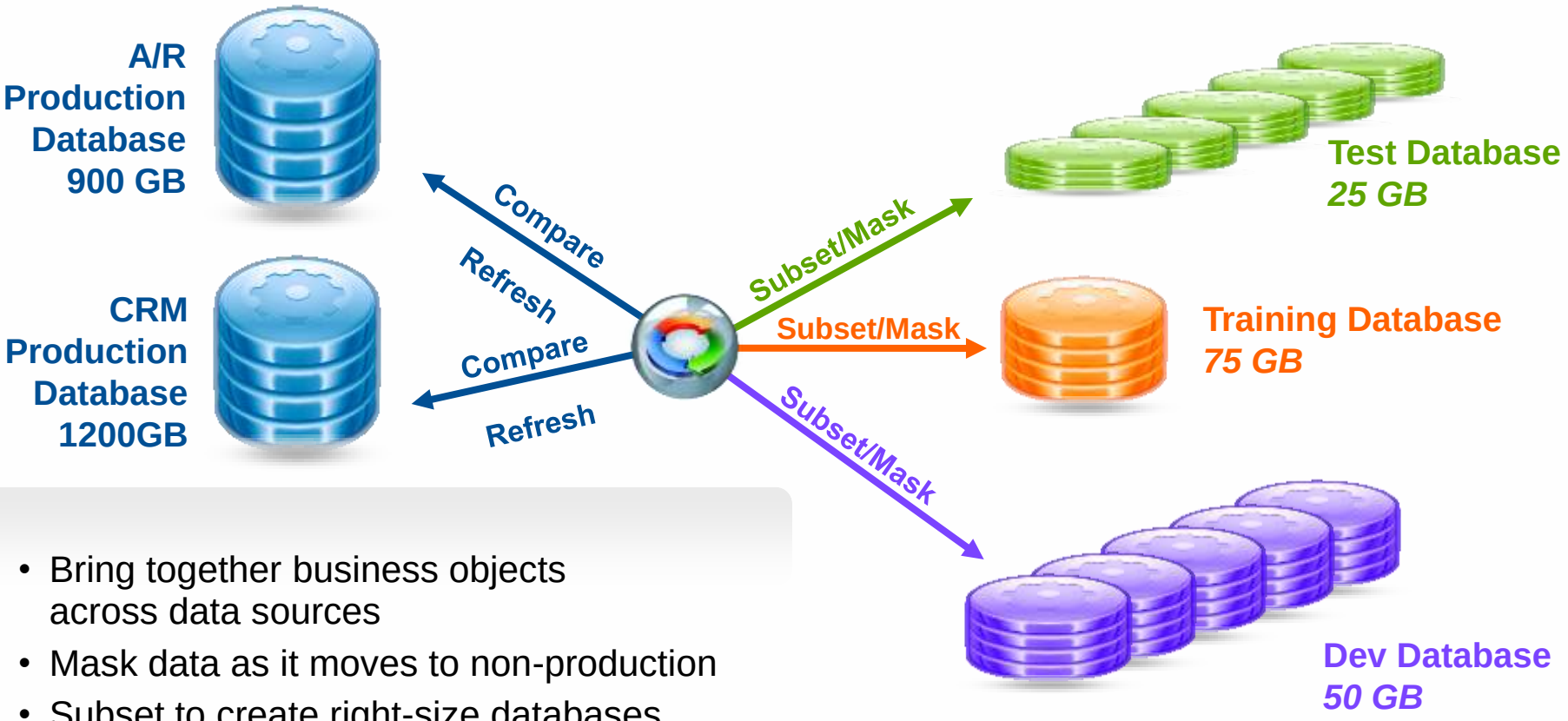
Example: Leverage a gold master



- Build all test environments from gold master
- Mask data on gold master
- Subset from gold master to create right-size non-production databases
- Compare data with “gold” to identify defects
- Refresh test data with “gold” to get latest data for testing

More strategic approaches make testing more agile

Example: Create test data directly from production

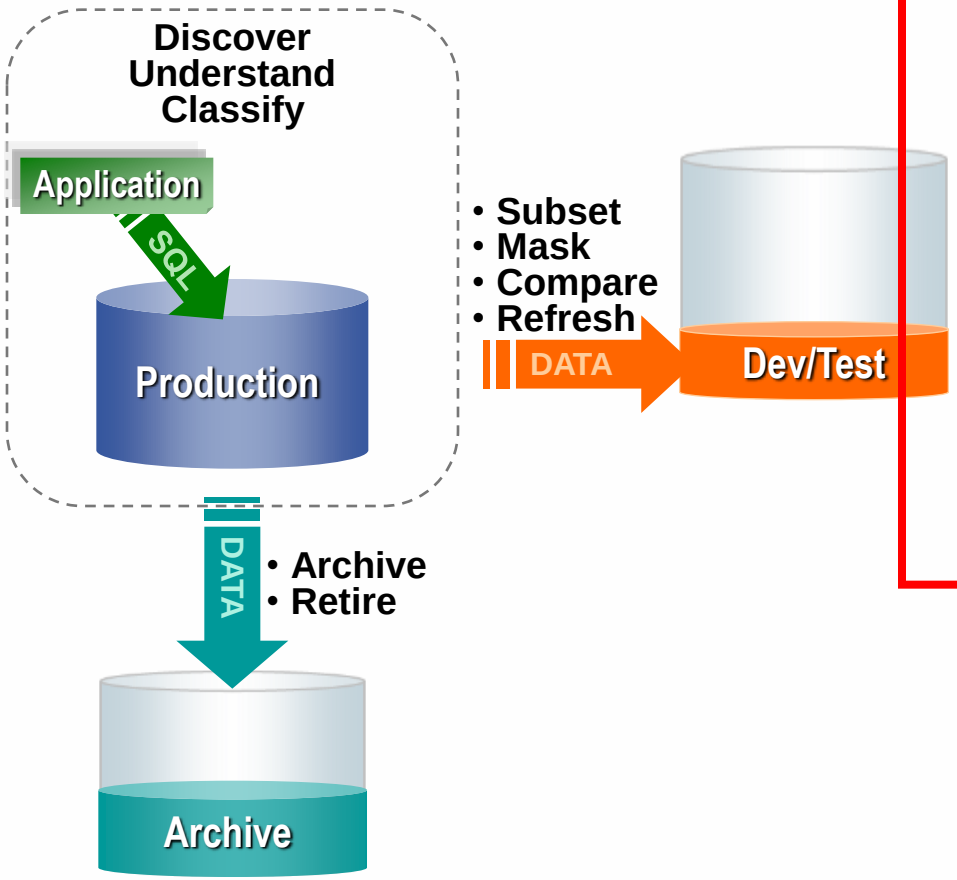


- Bring together business objects across data sources
- Mask data as it moves to non-production
- Subset to create right-size databases
- Compare data to production to identify defects
- Refresh test data to get latest for testing

Agenda

- Changing Business Environment and Data Management Challenges
- What is Test Data Management
- Best Practices in Test Data Management
- **Solution for Test Data Management**
- InfoSphere Optim Test Data Management with Rational Test Workbench & Rational Quality Manager
- Conclusion

IBM InfoSphere Optim Solutions



Discover

- Accelerate data management projects and reduce risk by understanding complex data relationships within & across apps / systems

Test

- Reduce cost, reduce risk and speed application delivery by maintaining right sized test environments
- Ensure compliance and privacy by masking
- Improve customer satisfaction, reduce the cost of change and meet SLAs by using production workloads for testing

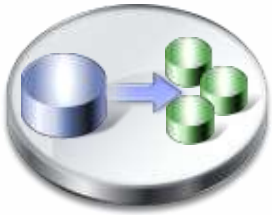
Mask

- Ensure compliance & privacy by masking data on demand

Archive

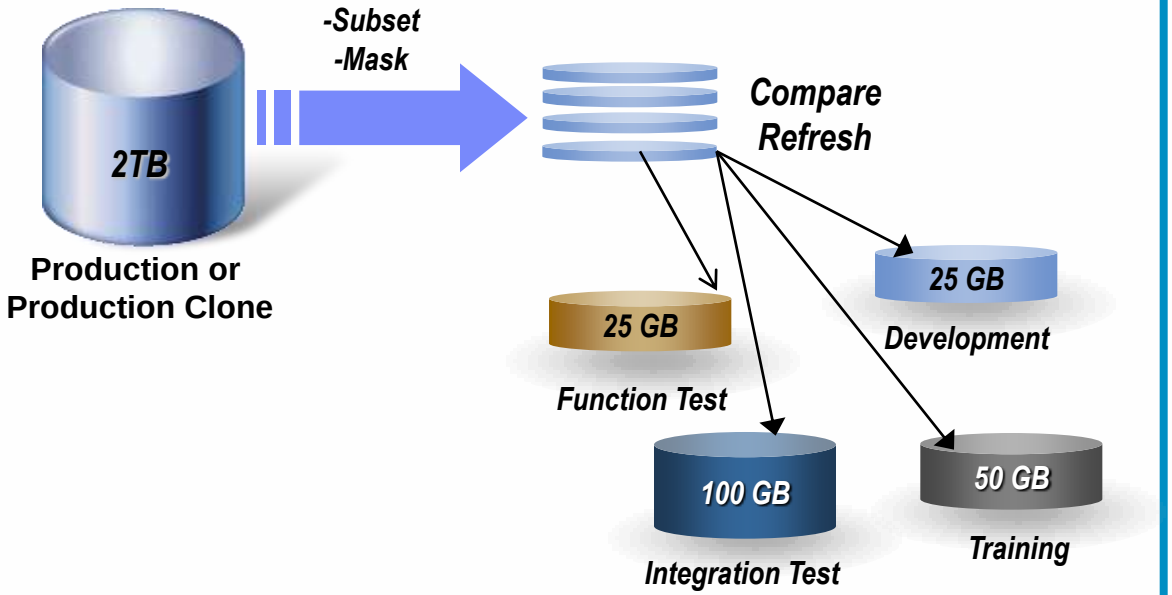
- Reduce hardware, software, storage & maintenance costs for enterprise applications
- Improve application performance & streamline back-ups and upgrades
- Support data retention regulations & safely retire legacy/redundant applications

IBM InfoSphere Optim Test Data Management Solution



Test Data Management

Create "right-size" production-like environments for application testing



InfoSphere Optim TDM supports data on distributed platforms (LUW) and z/OS.

Out-of-the-box subset support for packaged applications ERP/CRM solutions as well as :

Requirements

- Create referentially intact, "right-sized" test databases
- Automate test result comparisons to identify hidden errors
- Protect confidential data used in test, training & development
- Shorten iterative testing cycles and accelerate time to market

Benefits

- Deploy new functionality more quickly and with improved quality
- Easily refresh & maintain test environments
- Protect sensitive information from misuse & fraud with data masking
- Accelerate delivery of test data through refresh

ORACLE e-businesssuite | PeopleSoft | SIEBEL | JDEdwards | amdocs | IBM PureData for Analytics | Teradata® | Other | IBM

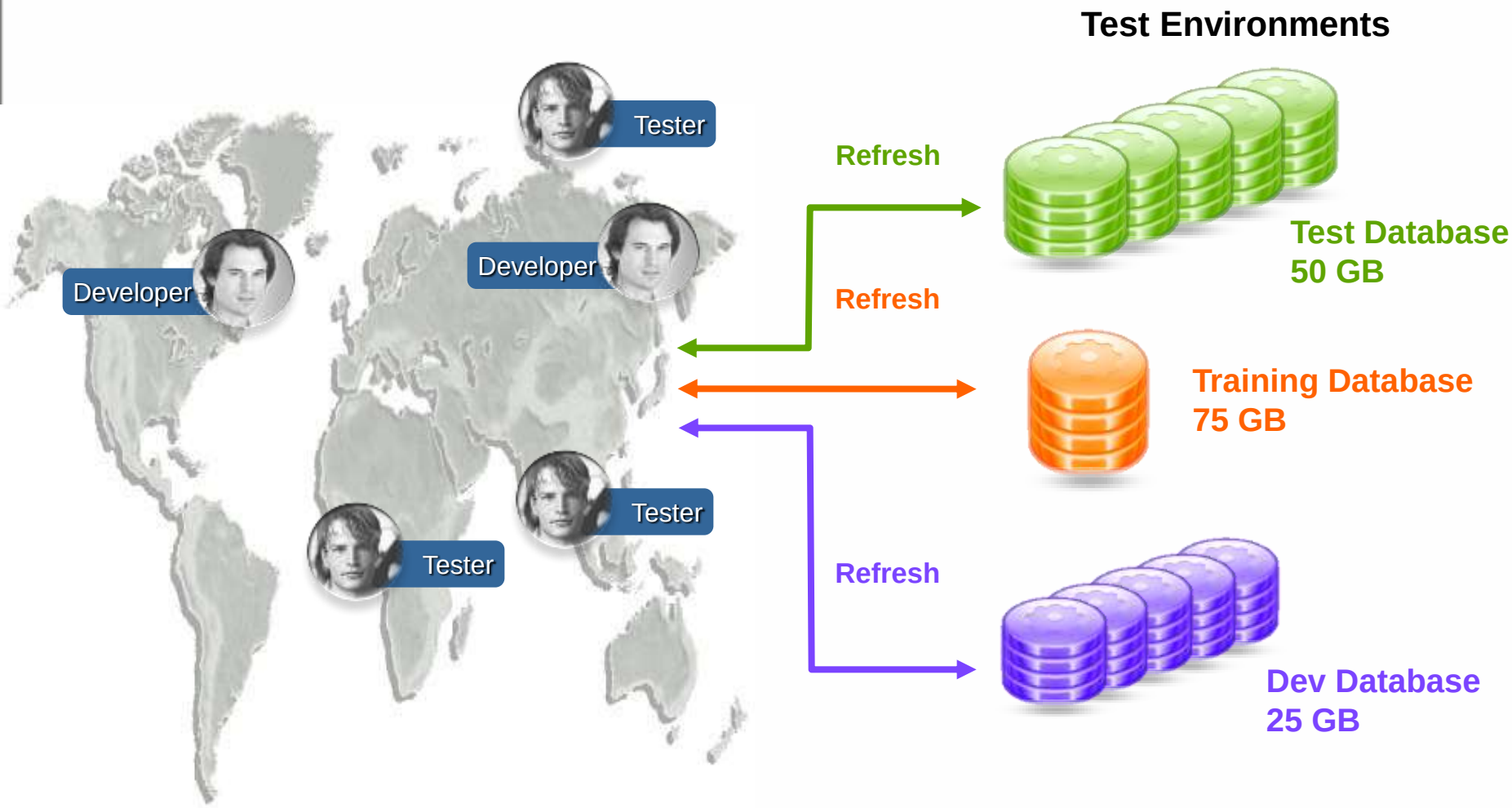
InfoSphere Optim protects Data Privacy during development & testing

Sensitive Data

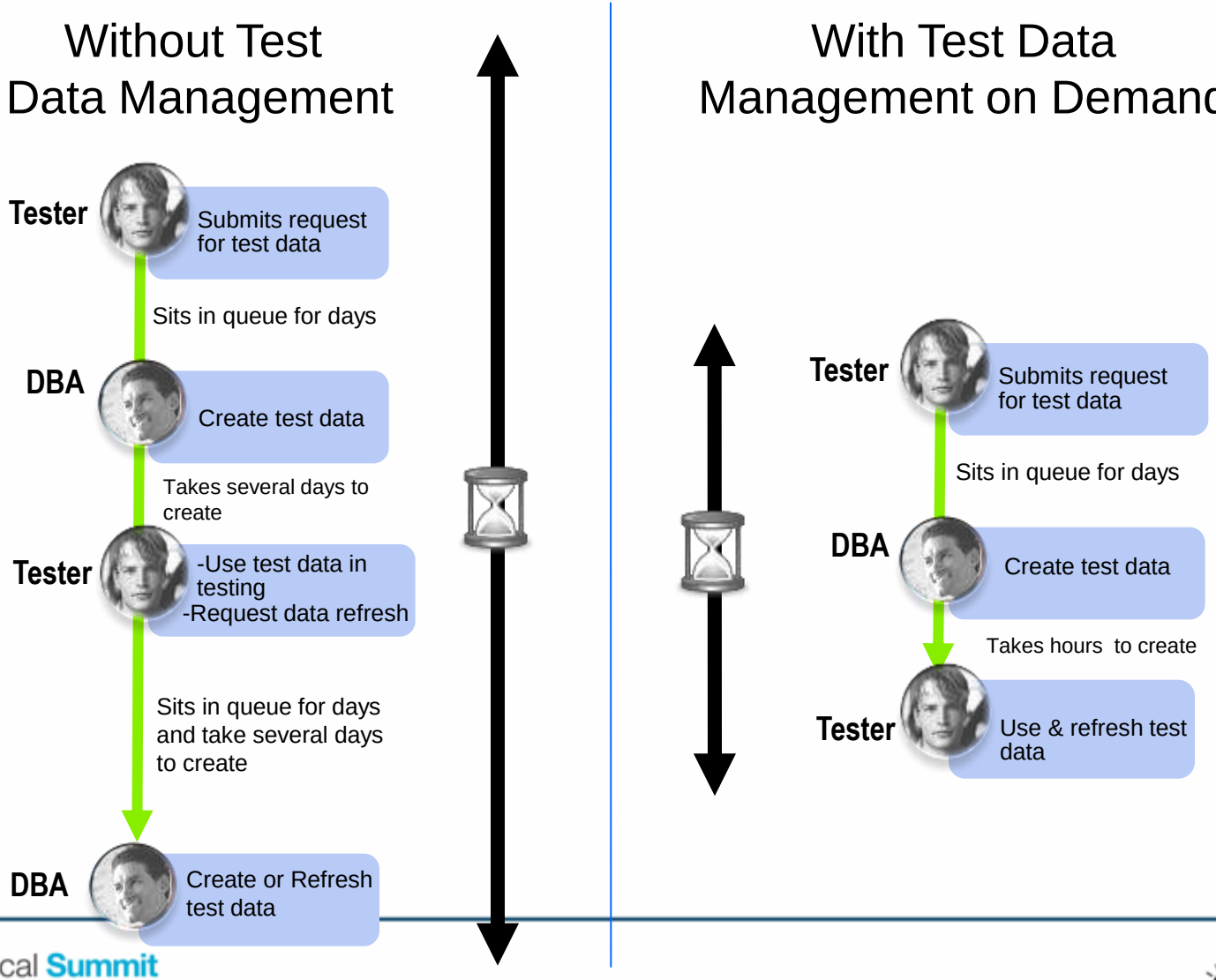
- Names
- Geography
- Credit Card Numbers
- Telephone numbers
- Email addresses
- Social Security numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers numbers
- Web URL's
- IP Addresses
- Business Data
- Corporate intelligence

- ✓ **Mask data in non-production databases such as test and development**
- ✓ **Improve privacy of non-production environments**
- ✓ **Facilitate faster testing processes with accurate test data**
- ✓ **Support referential integrity**
- ✓ **Mask custom & packaged ERP/CRM applications, BI/Analytics platforms & warehouses**

Refresh test data – Centralized solution

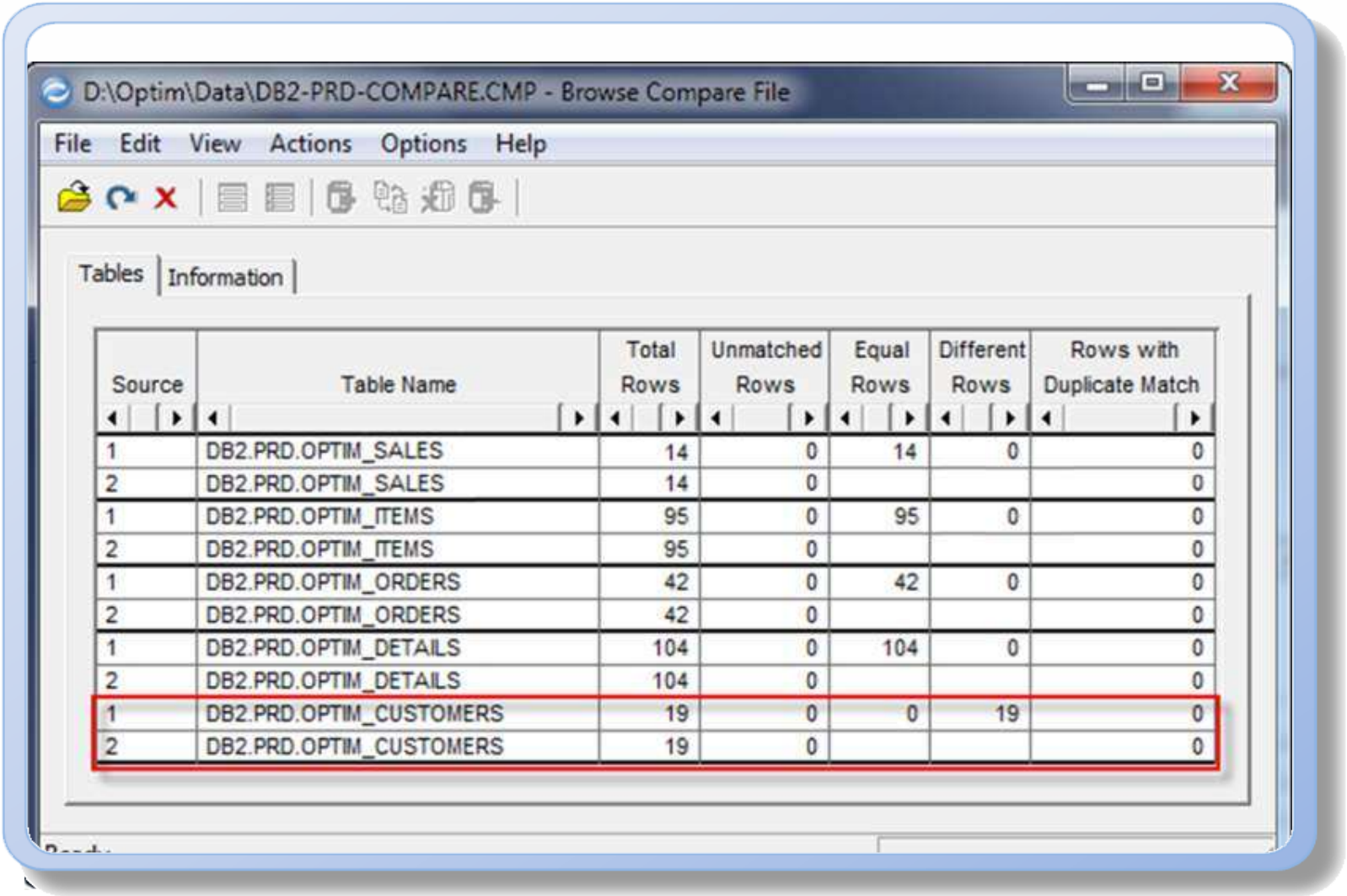


InfoSphere Optim Services on Demand create agile test environments



Analyzing test data results

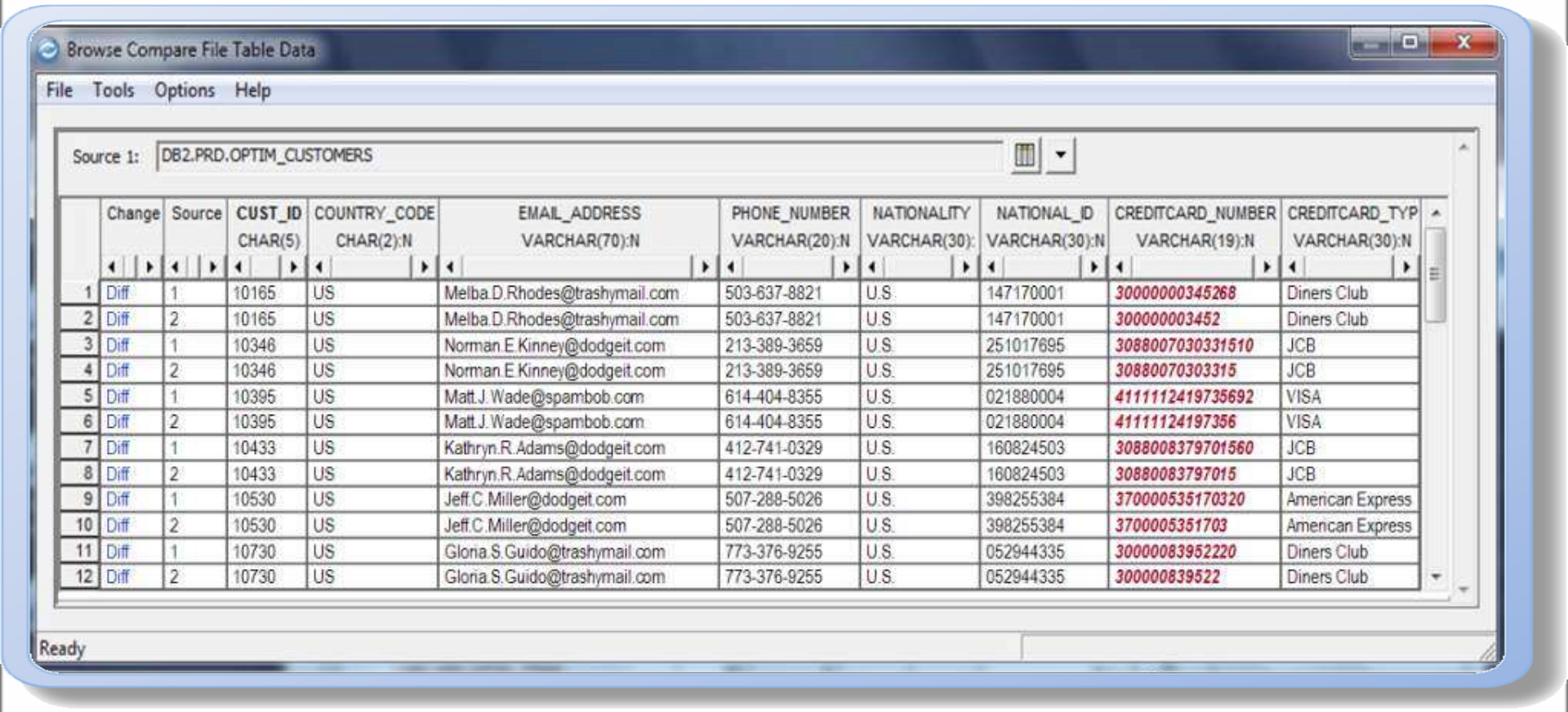
Compare 'before' and 'after' data to assess passed or failed



Source	Table Name	Total Rows	Unmatched Rows	Equal Rows	Different Rows	Rows with Duplicate Match
1	DB2.PRD.OPTIM_SALES	14	0	14	0	0
2	DB2.PRD.OPTIM_SALES	14	0			0
1	DB2.PRD.OPTIM_ITEMS	95	0	95	0	0
2	DB2.PRD.OPTIM_ITEMS	95	0			0
1	DB2.PRD.OPTIM_ORDERS	42	0	42	0	0
2	DB2.PRD.OPTIM_ORDERS	42	0			0
1	DB2.PRD.OPTIM_DETAILS	104	0	104	0	0
2	DB2.PRD.OPTIM_DETAILS	104	0			0
1	DB2.PRD.OPTIM_CUSTOMERS	19	0	0	19	0
2	DB2.PRD.OPTIM_CUSTOMERS	19	0			0

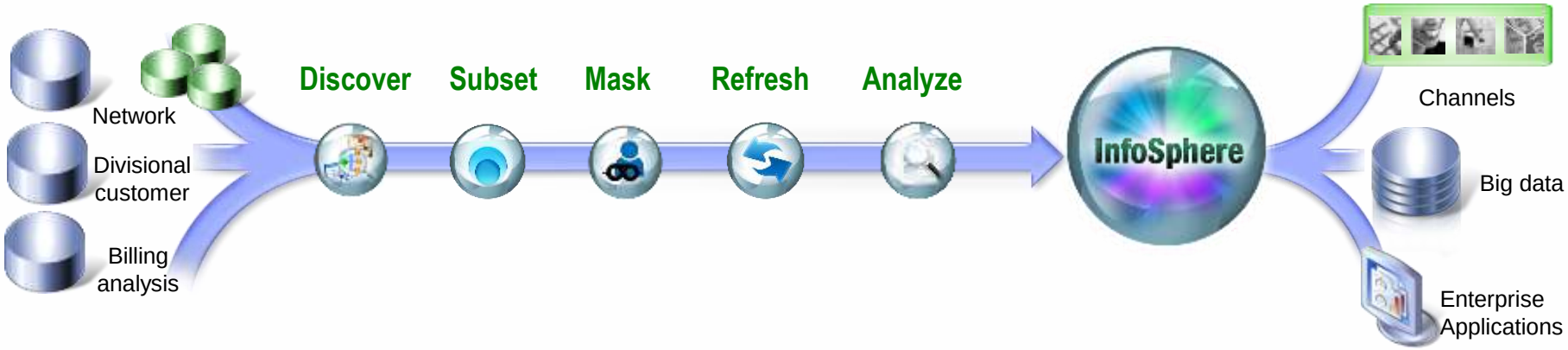
Analyzing test data results

Compare 'before' and 'after' data to assess passed or failed



	Change	Source	CUST_ID CHAR(5)	COUNTRY_CODE CHAR(2):N	EMAIL_ADDRESS VARCHAR(70):N	PHONE_NUMBER VARCHAR(20):N	NATIONALITY VARCHAR(30):	NATIONAL_ID VARCHAR(30):N	CREDITCARD_NUMBER VARCHAR(19):N	CREDITCARD_TYP VARCHAR(30):N
1	Diff	1	10165	US	Melba.D.Rhodes@trashymail.com	503-637-8821	U.S.	147170001	*30000000345268*	Diners Club
2	Diff	2	10165	US	Melba.D.Rhodes@trashymail.com	503-637-8821	U.S.	147170001	*300000003452*	Diners Club
3	Diff	1	10346	US	Norman.E.Kinney@dodgeit.com	213-389-3659	U.S.	251017695	*3088007030331510*	JCB
4	Diff	2	10346	US	Norman.E.Kinney@dodgeit.com	213-389-3659	U.S.	251017695	*30880070303315*	JCB
5	Diff	1	10395	US	Matt.J.Wade@spambob.com	614-404-8355	U.S.	021880004	*4111112419735692*	VISA
6	Diff	2	10395	US	Matt.J.Wade@spambob.com	614-404-8355	U.S.	021880004	*41111124197356*	VISA
7	Diff	1	10433	US	Kathryn.R.Adams@dodgeit.com	412-741-0329	U.S.	160824503	*3088008379701560*	JCB
8	Diff	2	10433	US	Kathryn.R.Adams@dodgeit.com	412-741-0329	U.S.	160824503	*30880083797015*	JCB
9	Diff	1	10530	US	Jeff.C.Miller@dodgeit.com	507-288-5026	U.S.	398255384	*370000535170320*	American Express
10	Diff	2	10530	US	Jeff.C.Miller@dodgeit.com	507-288-5026	U.S.	398255384	*3700005351703*	American Express
11	Diff	1	10730	US	Gloria.S.Guido@trashymail.com	773-376-9255	U.S.	052944335	*30000083952220*	Diners Club
12	Diff	2	10730	US	Gloria.S.Guido@trashymail.com	773-376-9255	U.S.	052944335	*300000839522*	Diners Club

Improve efficiency of application development & test with InfoSphere



Discover

- Intelligently analyze enterprise data
- Uncover relationships across heterogeneous sources
- Identify hidden sensitive data
- Identify proper test data

Subset

- Automatically extract test data required for each test case
- Test only on the required values to keep environments efficient
- Exercise outliers or usual test cases

Mask

- Enforce data integrity while masking
- Support context & application aware masking
- Use compliance packs
- Rely on scalable solution to protect privacy

Refresh & Analyze

- On demand access & refresh of test data
- Automate test result comparisons to reduce errors
- Easily maintain test & dev environments

Improve application quality and delivery efficiency with InfoSphere Optim Test Data Management Solution

Reduce Cost

Automate creation of realistic “right sized” test data to reduce the size of test environments

Reduce Risk

Mask sensitive information for compliance and protection

Speed Delivery

Refresh test data speeding testing and application delivery

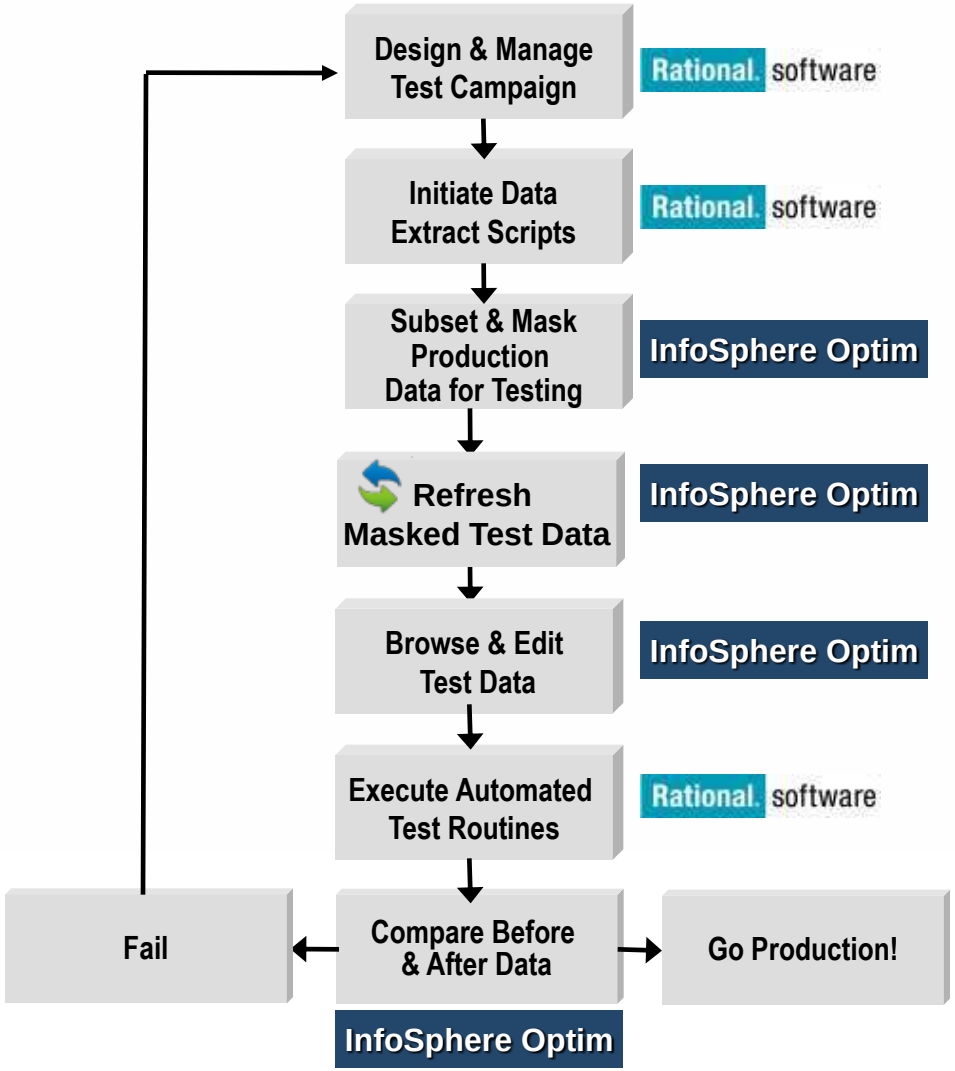
- Understand what test data is needed for test cases
- Create “right-sized” test data by sub-setting
- Ensure masked data is contextually appropriate to the data it replaced, so as not to impede testing
- Easily refresh & maintain test environments by developers and testers
- Automate test result comparisons to identify hidden errors
- Support for custom & packaged ERP applications in heterogeneous environments

Agenda

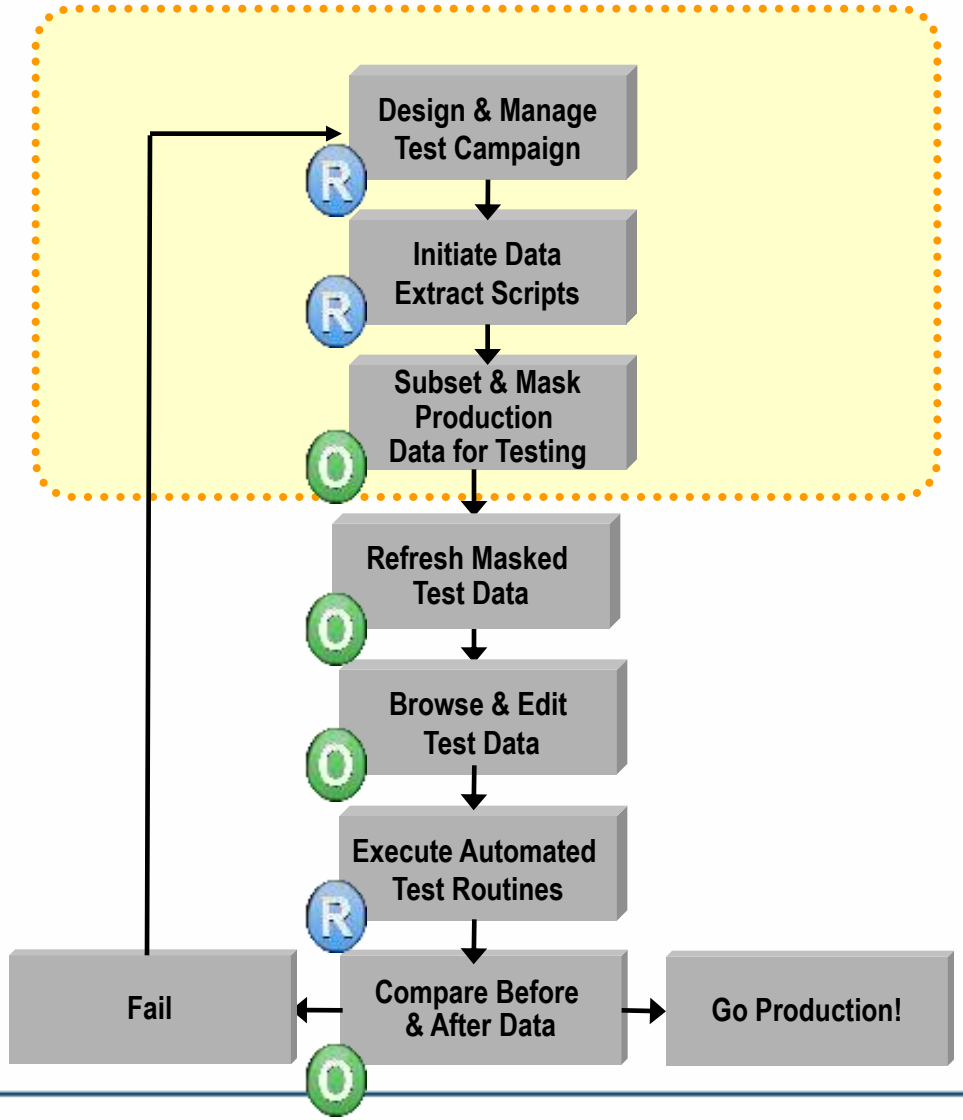
- Big data and software delivery
- What is Test Data Management
- Best Practices in Test Data Management
- Solution for Test Data Management
- **InfoSphere Optim Test Data Management with Rational Test Workbench & Rational Quality Manager**
- Conclusion

Enterprise Testing Solution with Rational and InfoSphere Optim
Building better quality applications for continuous testing

- **Comprehensive software quality process to minimize cost and shorten development cycles**
  - Manage test labs
  - Create realistic test environments from production data
  - Ensure protection of sensitive data
  - Manage unit, functional and performance testing and quality test cases
  - Simulate virtual services
- **Streamline your test data management processes** and deliver your project sooner and with fewer defects

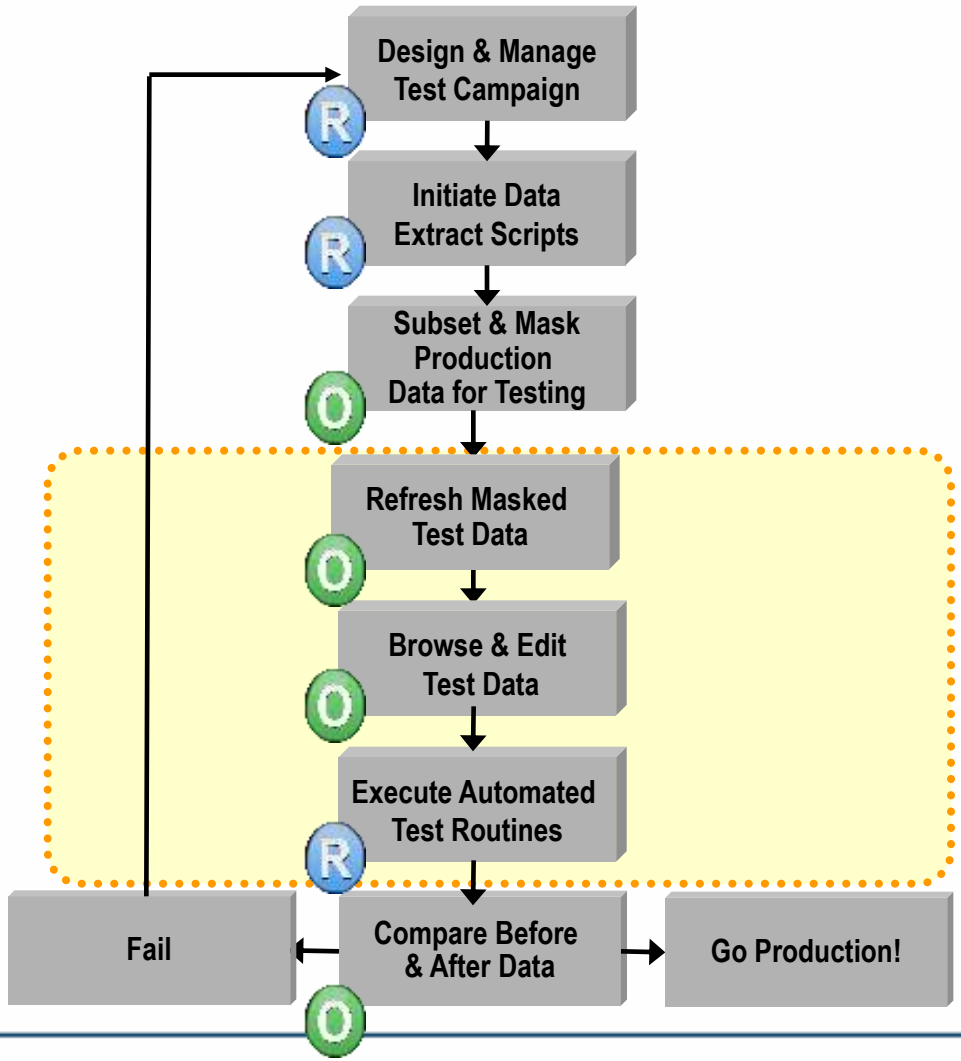


How does Rational and Optim work together?



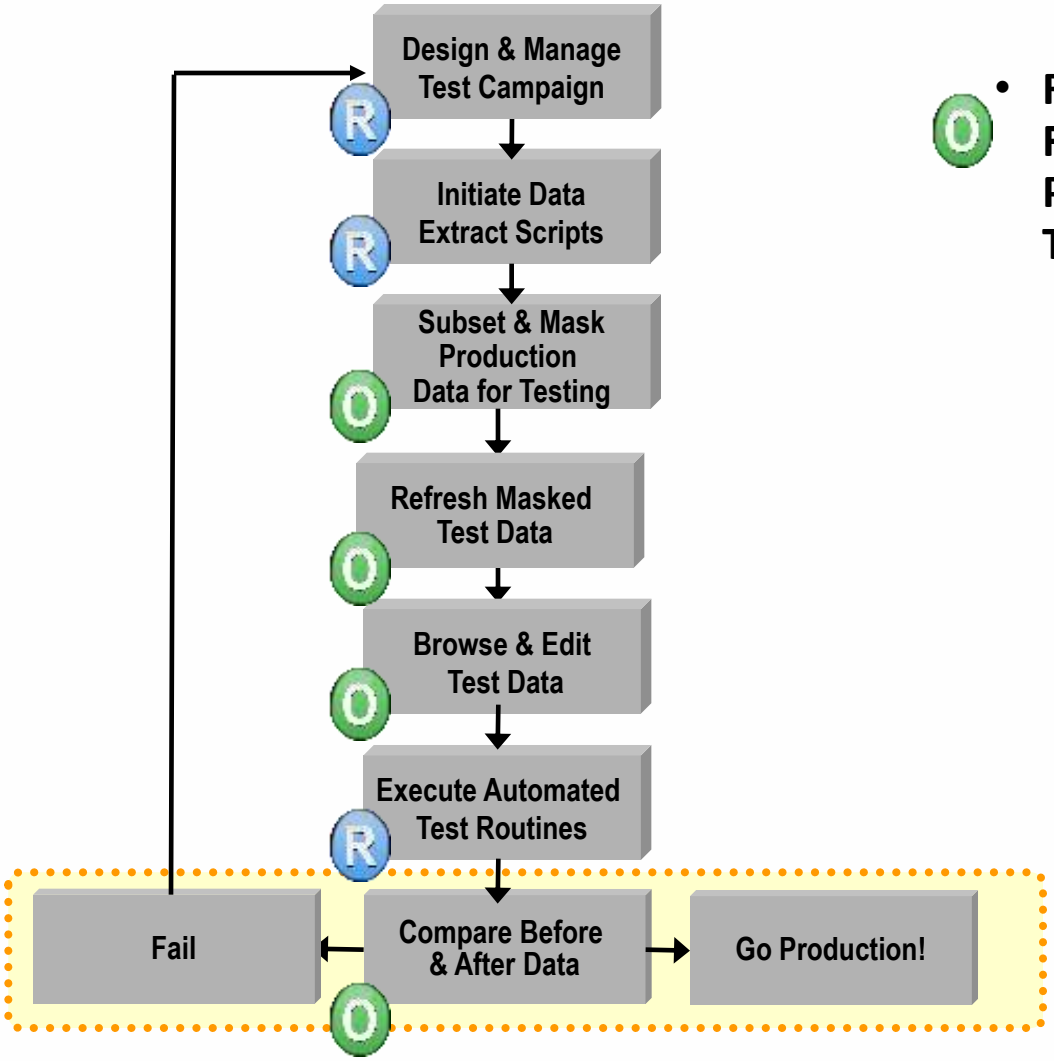
- **Design test cases in Rational Quality Manager**
- **Rational Test Workbench (includes RFT & RPT) uses automated test process to invoke creation of test data**
- **Optim Test Data Management accesses production DB**
  - Subsets
  - Masks
  - Propagates data to test environment

How does Rational and Optim work together?



- O • **Refresh test data via Optim Services On Demand**
- O • **Browse and edit test data e.g. for boundary conditions, unusual combinations of data**
- R • **Rational Functional Tester executes automated test scripts, using test data as input**
- R • **Rational Performance Tester executes load capacity testing, using test data as input**
- R • **Rational Test Workbench simulates virtual services, using test data as input**
- R • **Rational Quality Manager executes manual tests, using test data as input**

How does Rational and Optim work together?



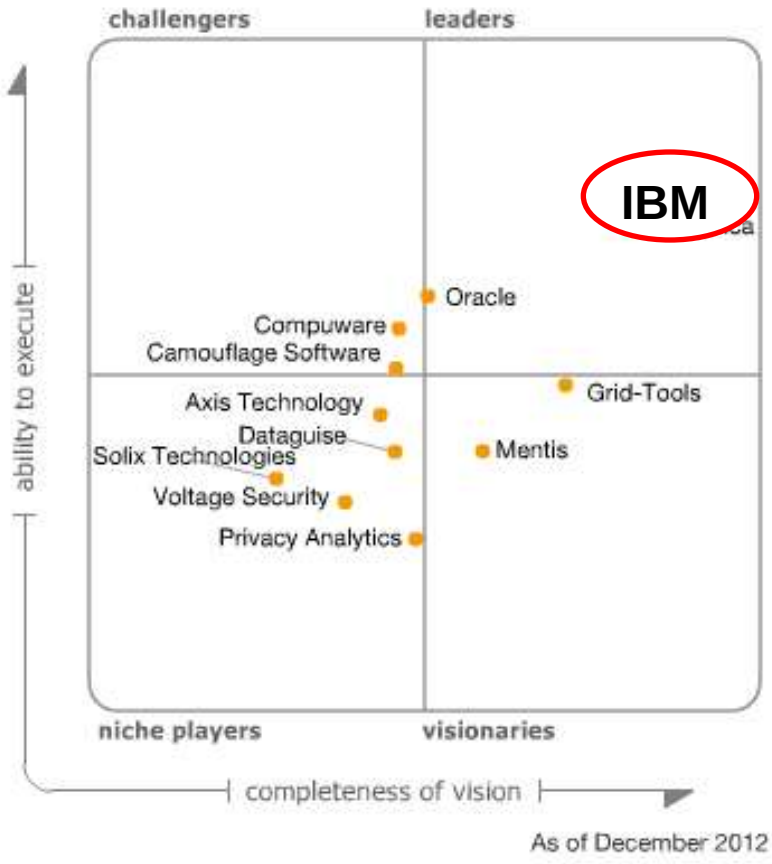
- **Rational Test Workbench, Rational Functional Tester & Rational Performance Tester invokes Optim Test Data Management**
  - Analyze differences between before and after test
    - Did everything change that was supposed to?
    - Are there any unexpected changes?

Agenda

- Big data and software delivery
- What is Test Data Management
- Best Practices in Test Data Management
- Solution for Test Data Management
- InfoSphere Optim Test Data Management with Rational Test Workbench & Rational Quality Manager
- **Conclusion**

InfoSphere Optim recognized leader according to analysts

Gartner MQ for Data Masking Technology



IBM Leader

IBM's Optim product line led the ILM segment with a 52.3% share, and showed nearly 18% growth.

---IDC 2012

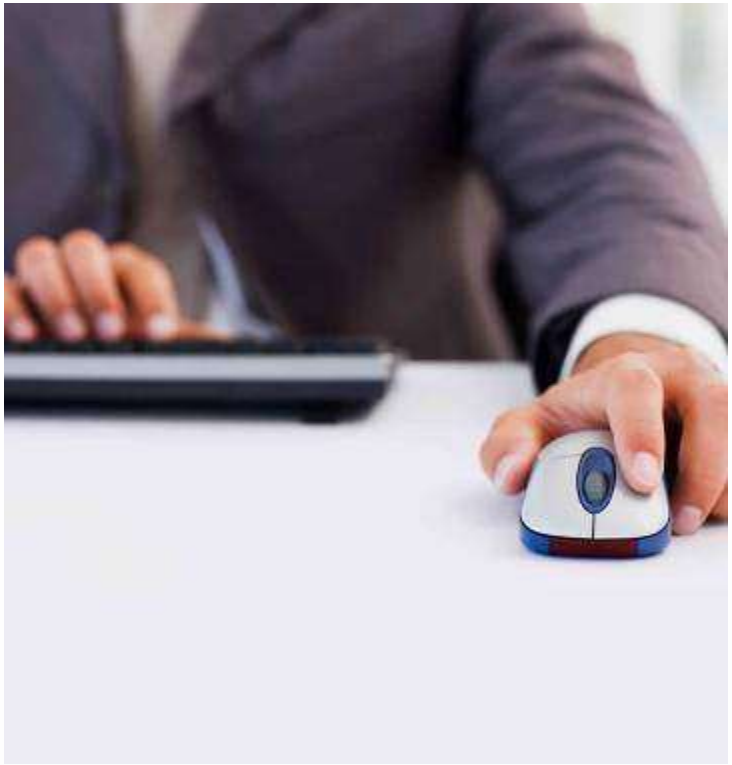
Reduces capacity requirements for testing by **40%**

- Implemented InfoSphere Optim Test Data Management to reduce the size of 10 development, testing and training database environments
- Anticipates cost savings of about US$240,000/year
- Trained and enabled the client to be self-sufficient

Travelport

Resources to Learn More!

- Analyst report: Successful approaches to test data management
- ebook: Fundamentals of Test Data Management
- Solution Sheet: InfoSphere Optim Test Data Management
- Whitepaper: Integrated Strategies to Improve Application Testing
- Case Study: InfoSphere Optim Test Data Management
- Whitepaper: Deploy Enterprise Changes with Confidence



got questions?