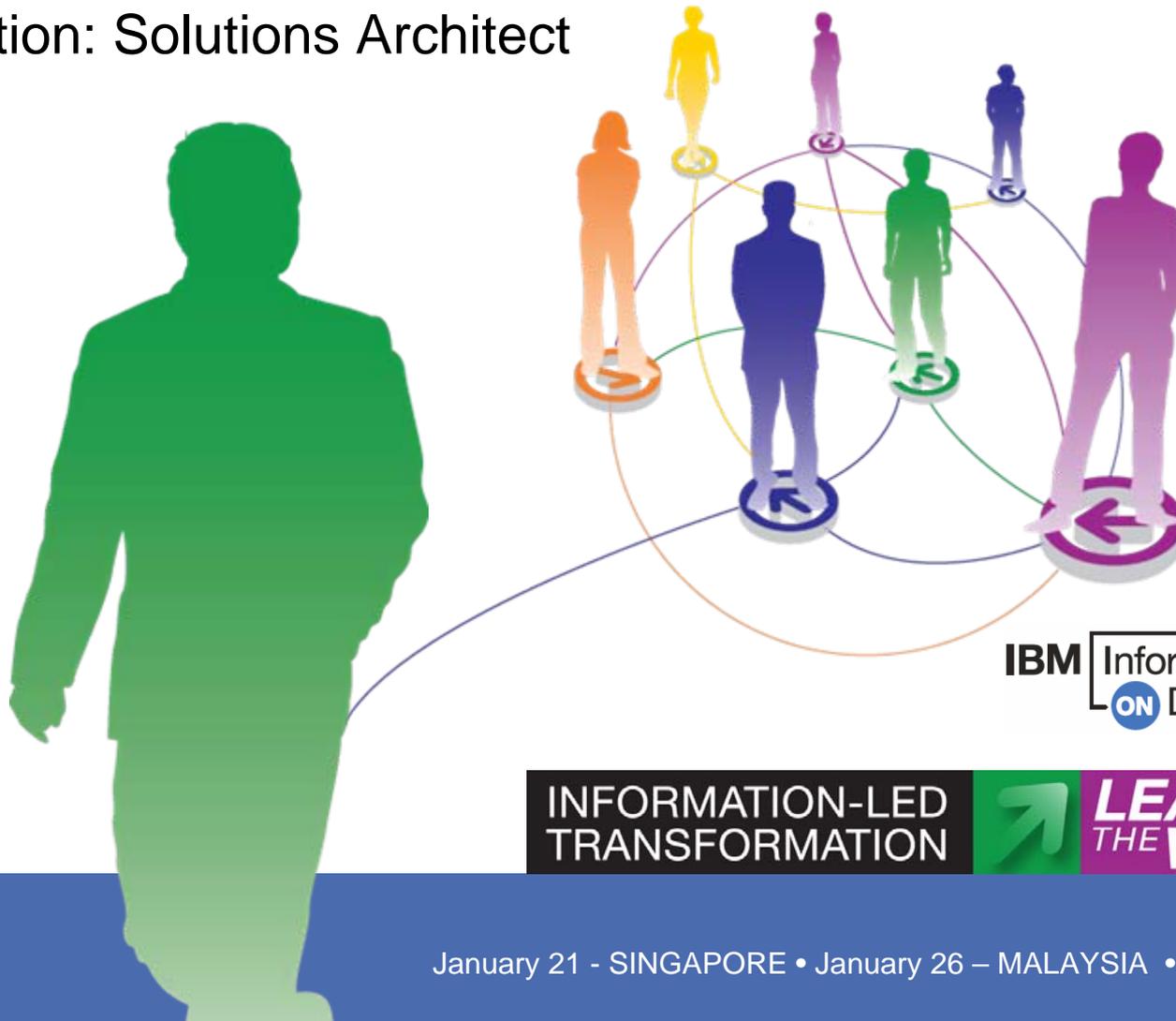


ASEAN INFORMATION ON DEMAND CONFERENCE 2010

Presenter: George Lapis

Designation: Solutions Architect



IBM Information
ON Demand 2010

INFORMATION-LED
TRANSFORMATION

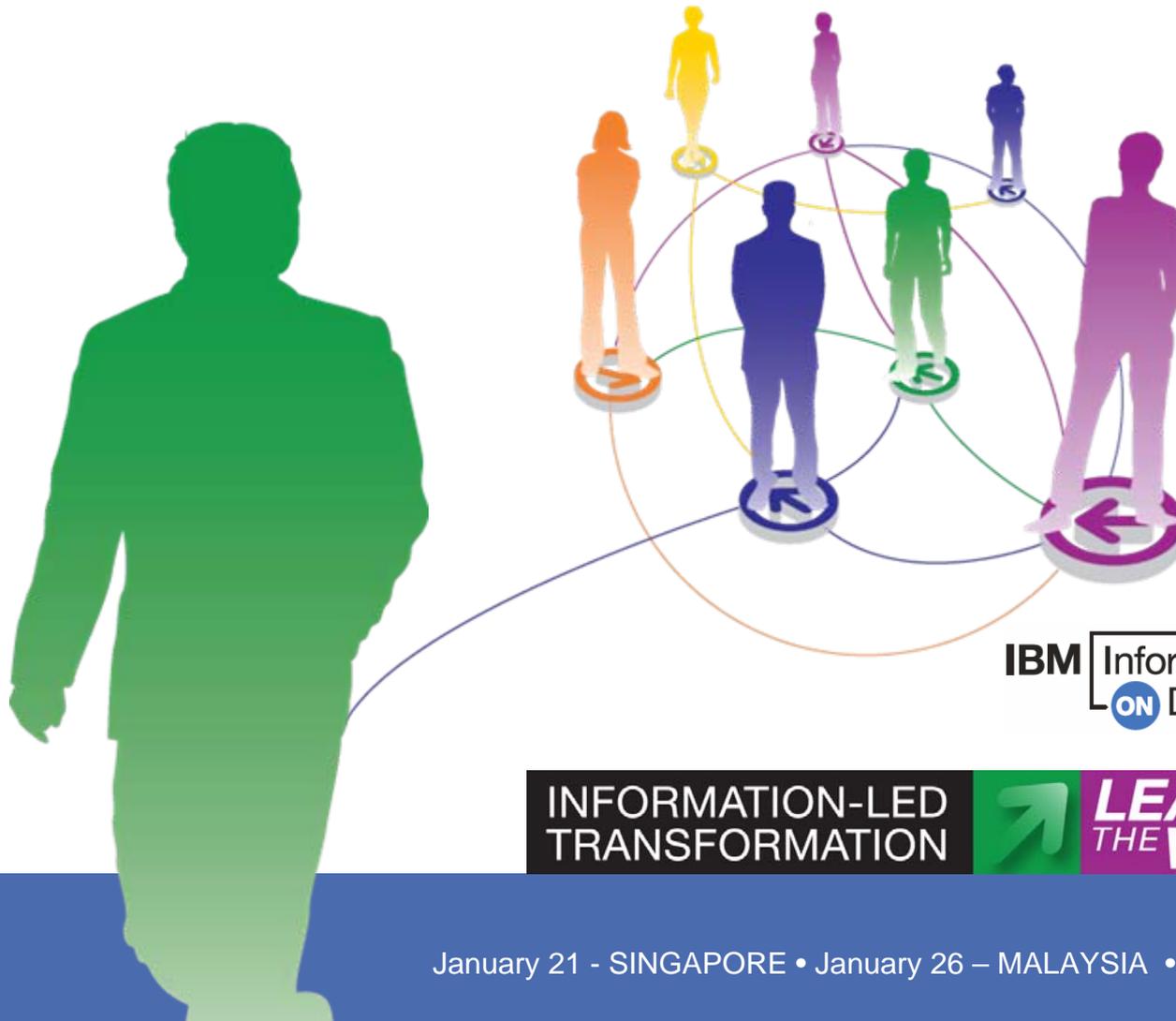
LEAD
THE WAY



January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

IBM Optim Solutions

Introduction to Integrated Data Management



IBM Information
ON Demand 2010

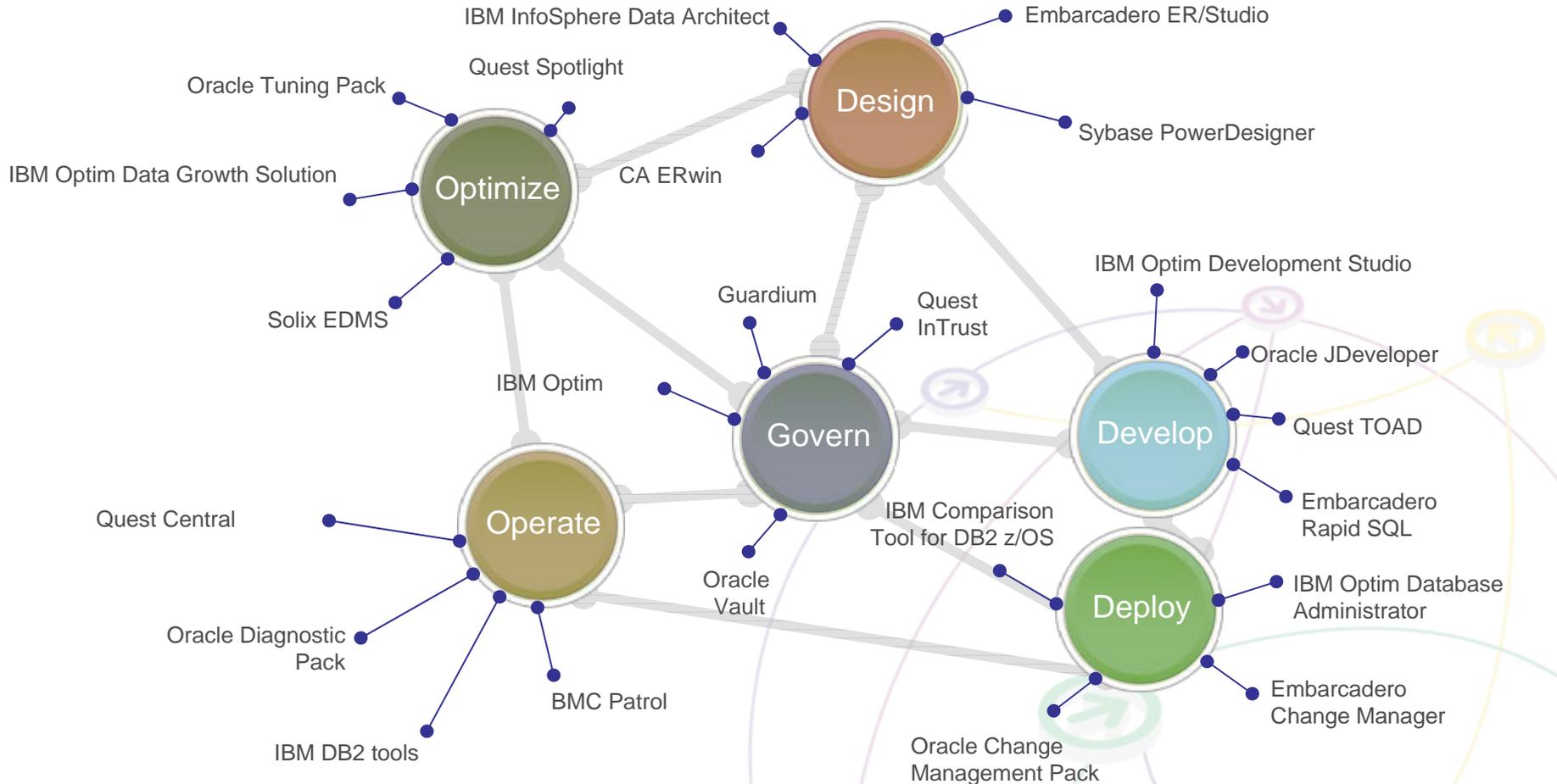
INFORMATION-LED
TRANSFORMATION



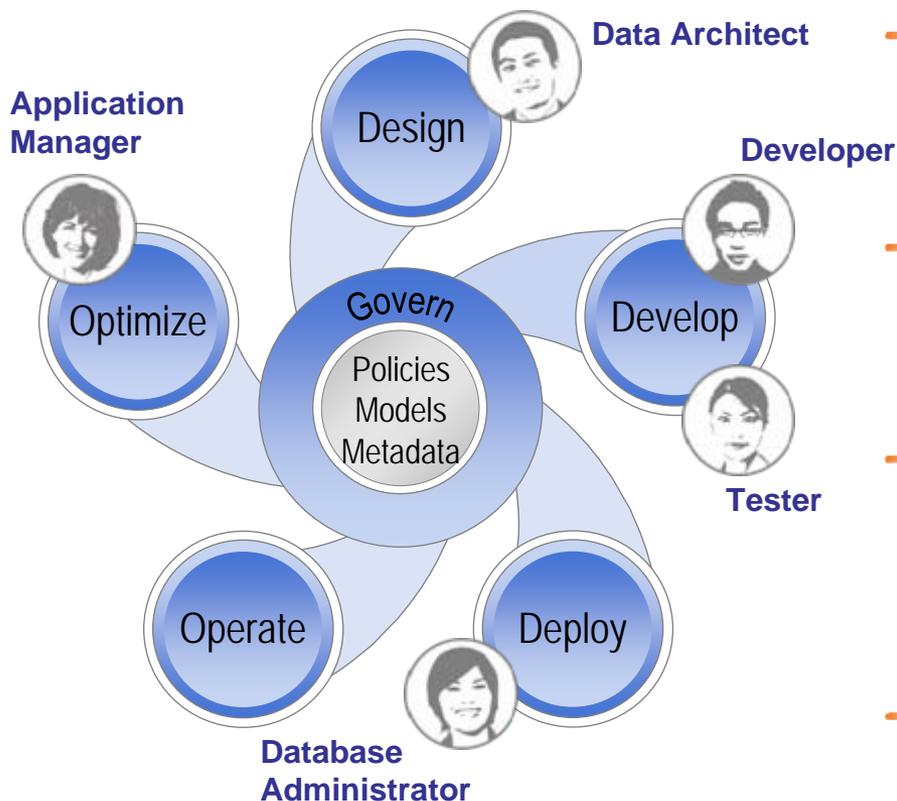
January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

What do Businesses Have?

A Collection of Disparate, Single-Purpose Products



Integrated Data Management – What’s Different?



→ Produce enterprise-ready applications faster

- Improve data access, speed iterative testing
- Empower collaboration between architects, developers & DBAs

→ Consistently achieve service level targets

- Automate and simplify operations
- Provide contextual intelligence across the solution stack

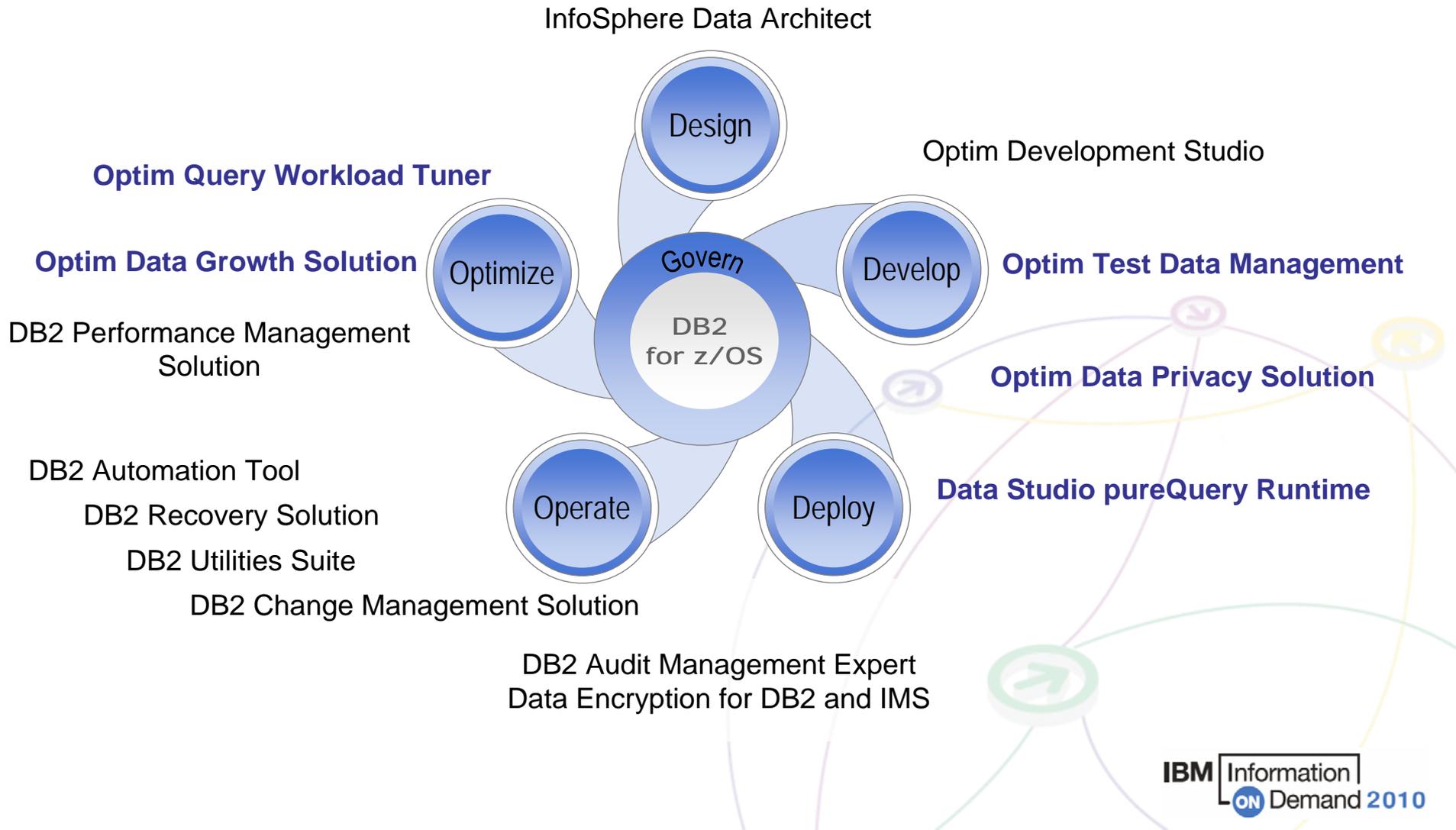
→ Support business growth

- Accommodate new initiatives without expanding infrastructure
- Simplify application upgrades, consolidation & retirement

→ Facilitate alignment, consistency & governance

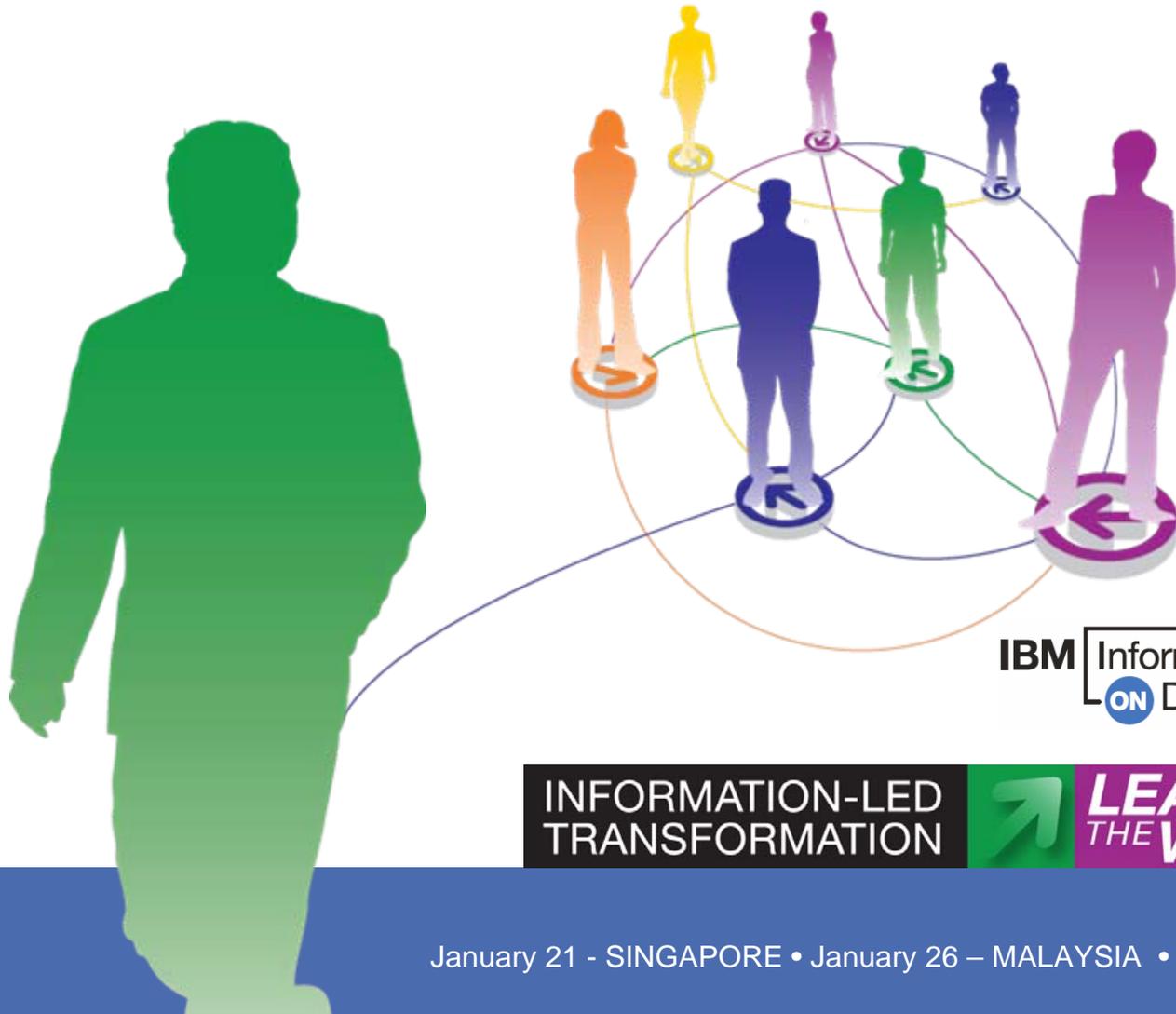
- Define business policies and standards up front; share, extend, and apply throughout the lifecycle

The broadest range of capabilities for managing the value of your data throughout its lifetime



IBM Optim Solutions

Challenges around data governance



IBM Information
ON Demand 2010

INFORMATION-LED
TRANSFORMATION



IBM

January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

Challenges Facing Customers Today

→ Mitigate Risk

- Effectively and securely manage archived data
- Protect data privacy
- Accurate, prompt responses to auditing requests

→ Maintain Performance in face of Data Growth

- Improve application performance by moving historical transaction records to a safe, secure archive
- Achieve Service Level Agreements (SLAs) consistently

→ Control Costs

- Reduce infrastructure costs; utilize cost effective tiered storage
- Minimize cost and time for compliance
- Improve productivity of development team

Challenges : Reduce Risk

→ Insiders and hackers are targeting data for profit

→ Data in and of itself has monetary value:

- Credit Card Number With PIN - \$500
- Drivers License - \$150
- Birth Certificate - \$150

Source: USA TODAY research 10/06

→ Average cost of a data breach in 2007 was **\$197 USD per customer record leaked** *Source: Ponemon Institute*

→ This has been a driving factor for creating data protection and privacy regulations

→ How to protect Personal Identifiable Information (PII)?

Challenges: Dealing with Data Growth

- Data is growing at a very rapid rate
 - Annual growth rates for databases exceed 125%
 - Mergers & acquisitions
 - Data Warehousing

- The Data Multiplier effect
 - OLAP cubes, data marts, and so on
 - Copies of data for test, development, quality assurance, disaster recovery, etc.

- Retention of data for compliance purposes

- Web 2.0 applications are verbose

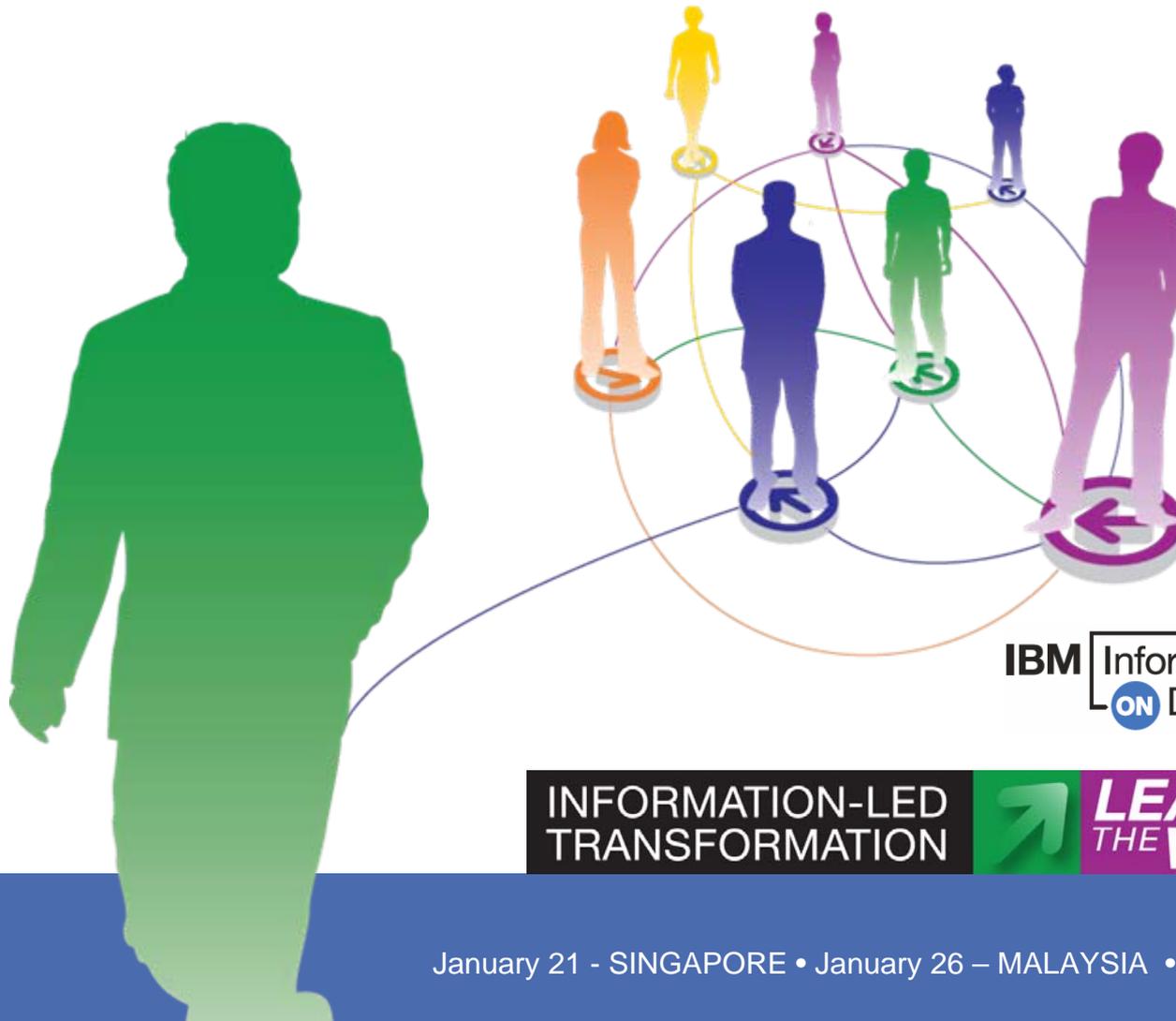
- How to manage data growth and aging effectively?

Challenges: Control Costs

- Growing storage costs due to rapid data growth
- Cost of storing and managing the many copies of your production data
- Cost of implementing data privacy measures for compliance across different databases and applications
- Cost of archive retrievals for compliance requests and e-discovery
- Growing development time costs
- How to control storage and data management costs?

IBM Optim Solutions

How IBM Optim Can Help



IBM Information
ON Demand 2010

INFORMATION-LED
TRANSFORMATION



LEAD
THE WAY

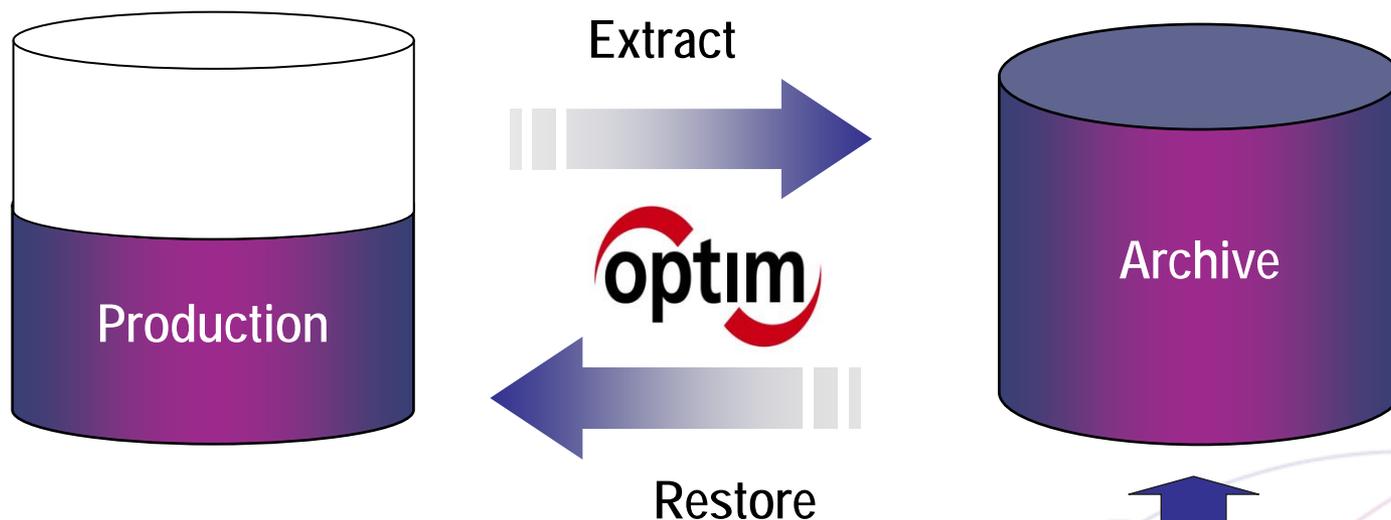


January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

How IBM Optim Helps

- Archive, Retention, and e-Discovery
- Test Data Management
- Application Decommissioning
- Data Privacy
- Information Lifecycle Management (ILM)

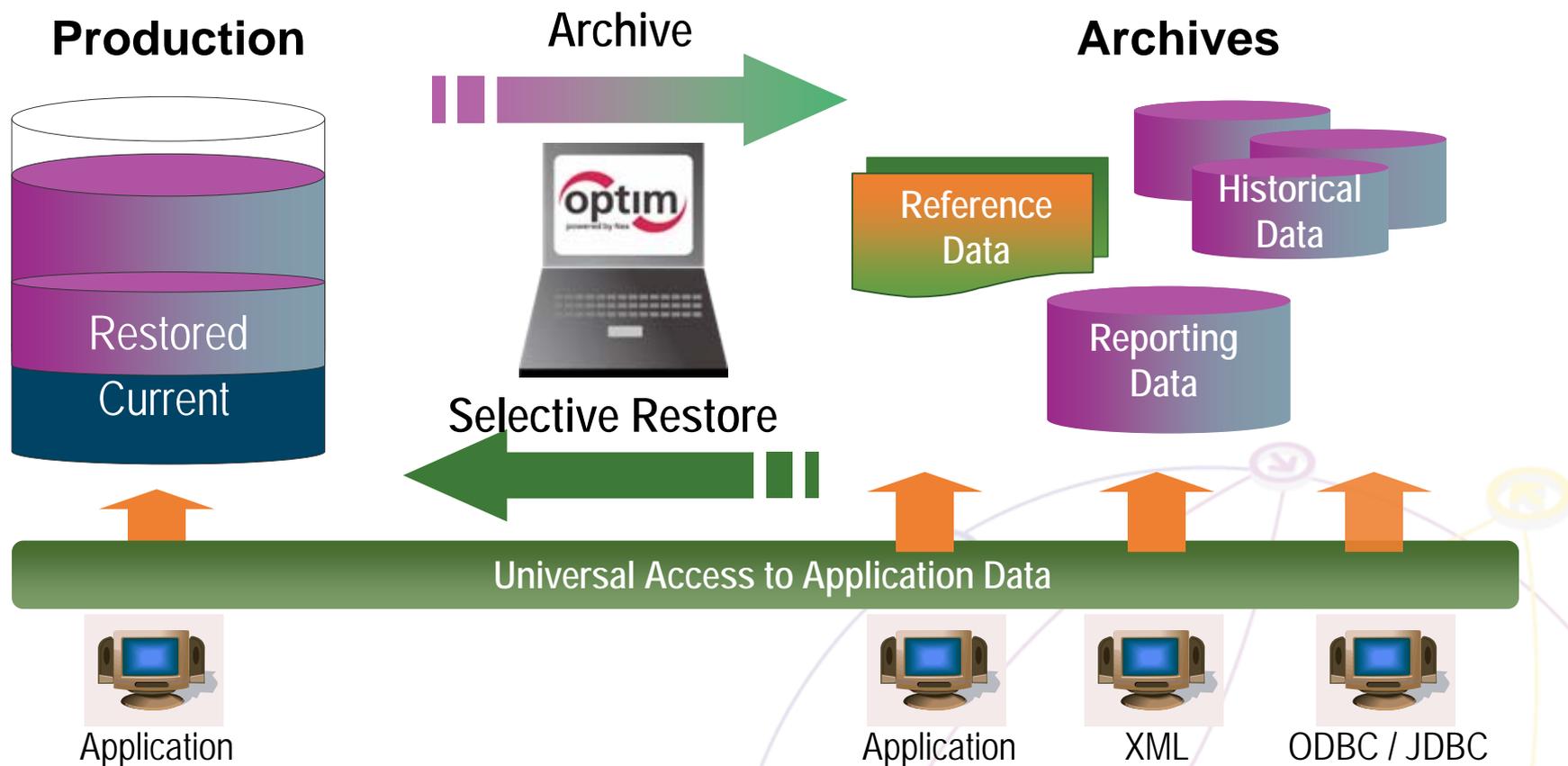
Archive, Retention, and e-Discovery



- **Optim safely moves inactive or historical data to an archive**
- **Archive can be accessed in many ways**



Optim Data Growth Solution: Archiving

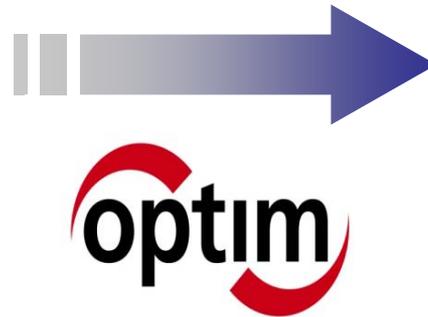


- Complete Business Object is historical snapshot of activity
- Storage device independence enables ILM
- Immutable file format enables data retention compliance

Application Decommissioning

Obsolete Application

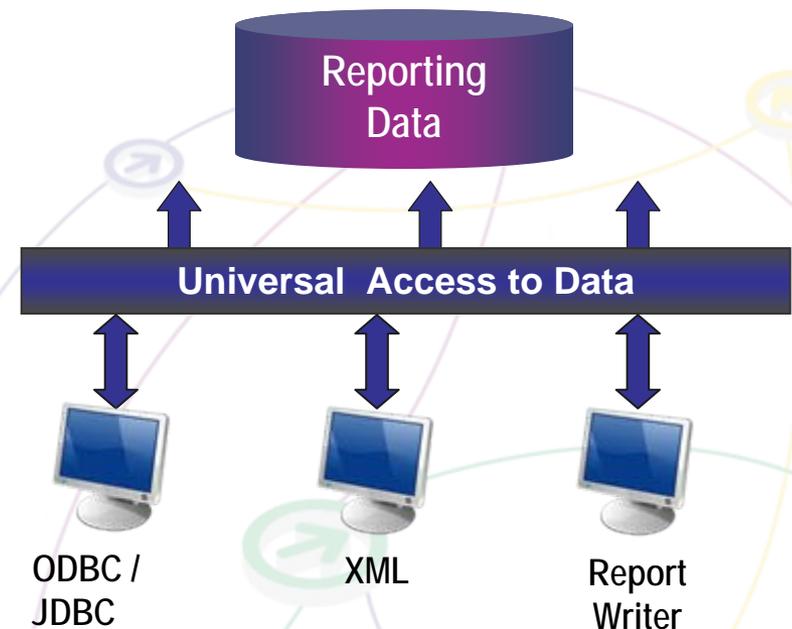
Archive



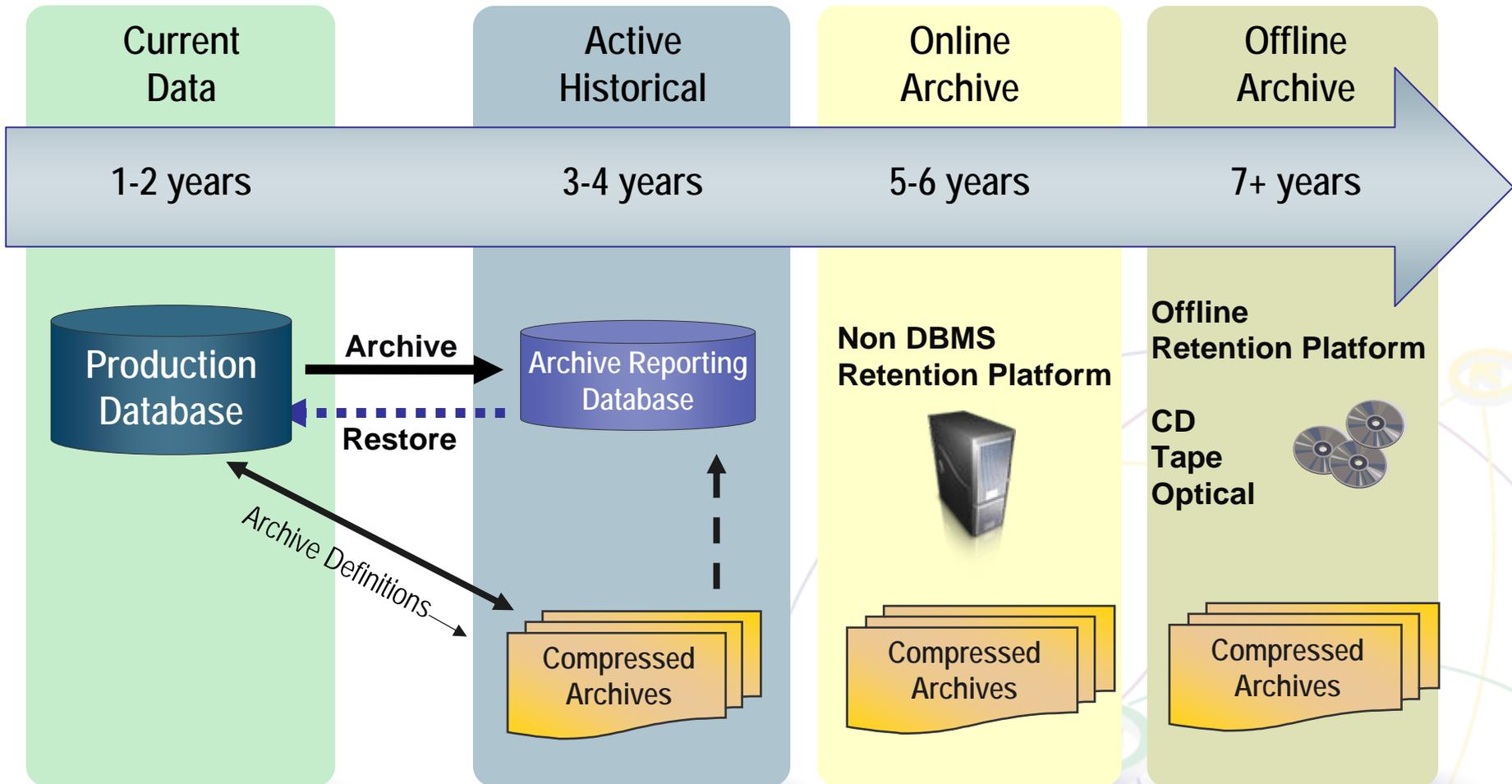
Archives



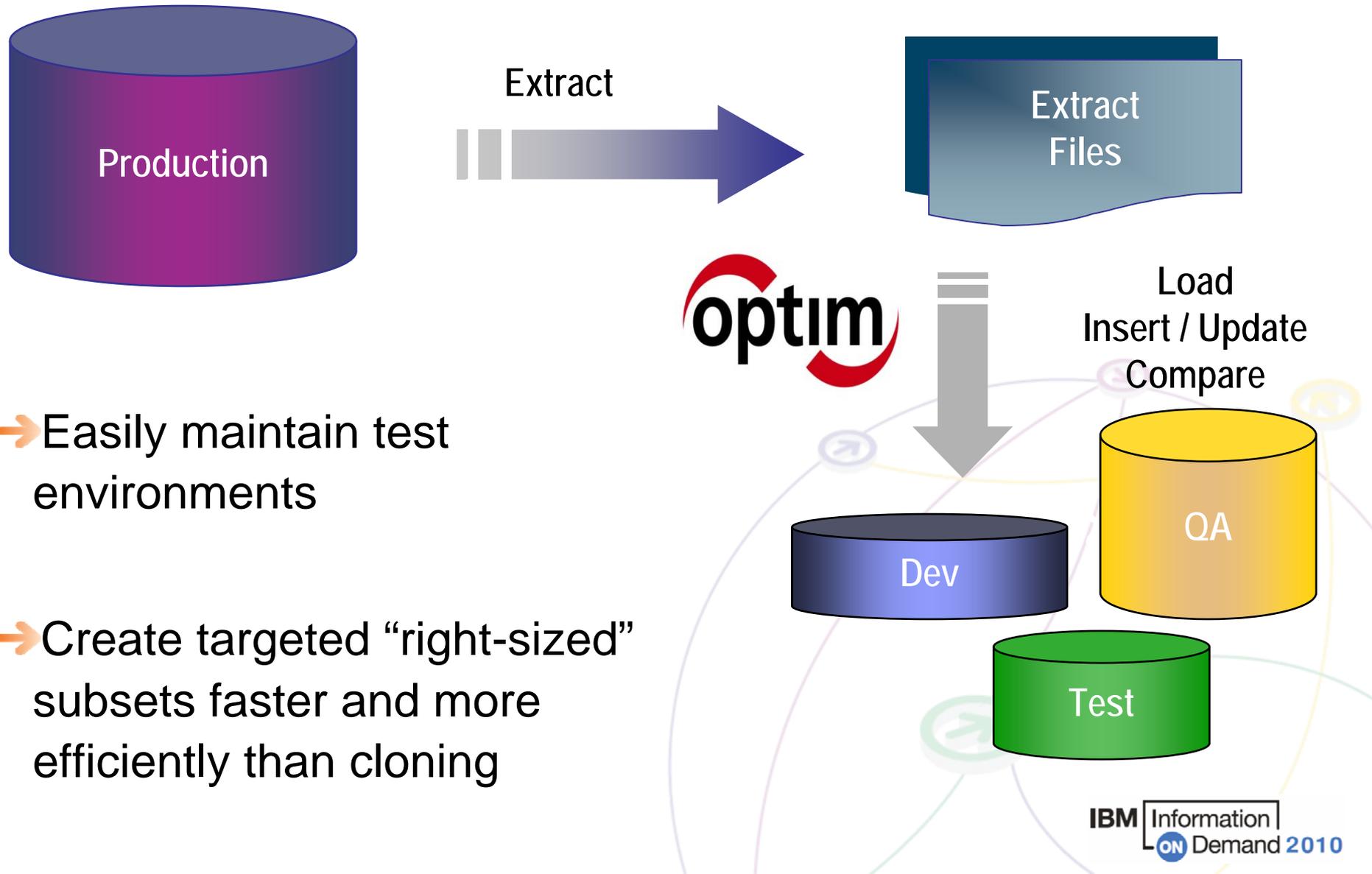
- Archive your application data before decommissioning
- Access this data **without** the application, or restore at any time



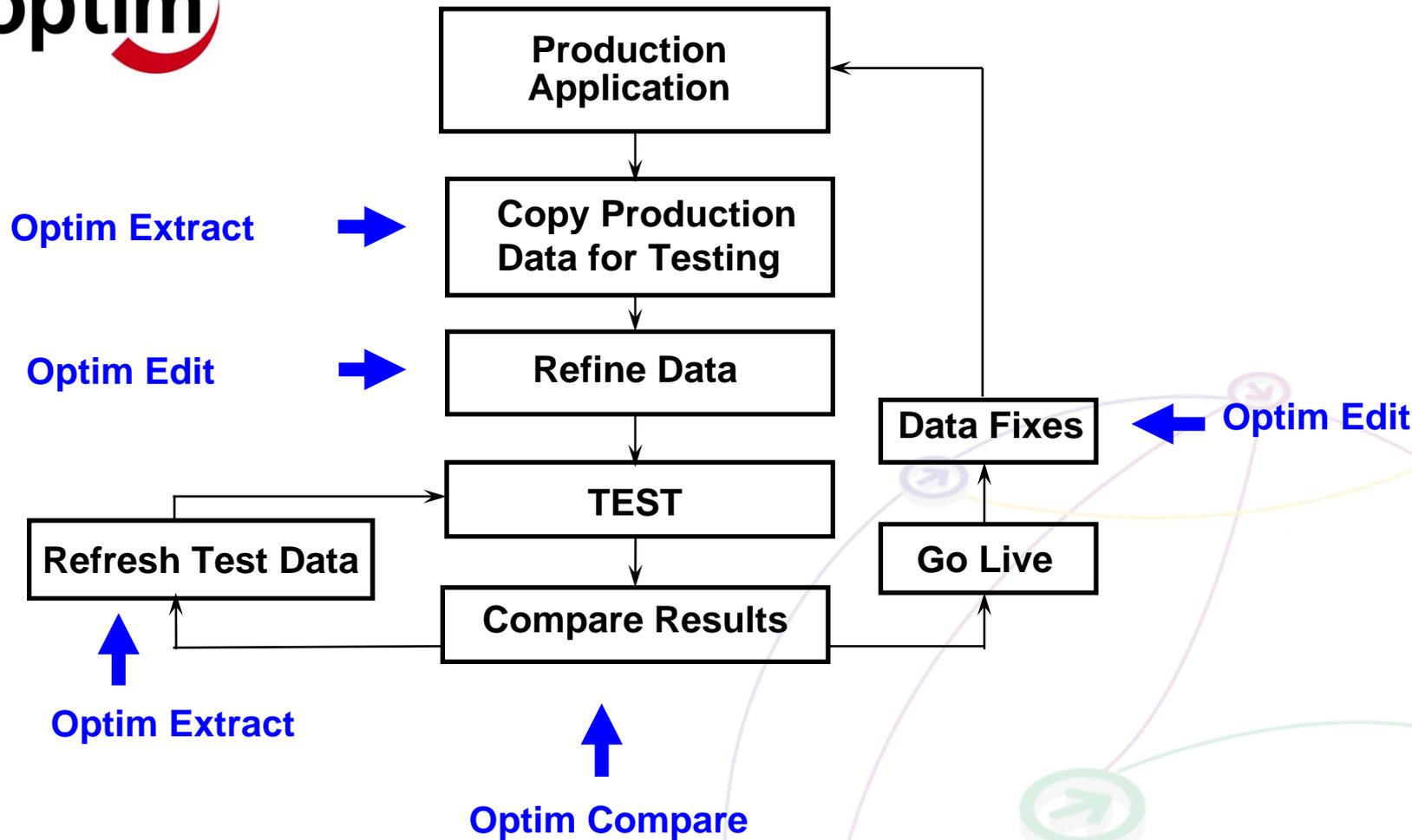
Optim Enables Information Lifecycle Management



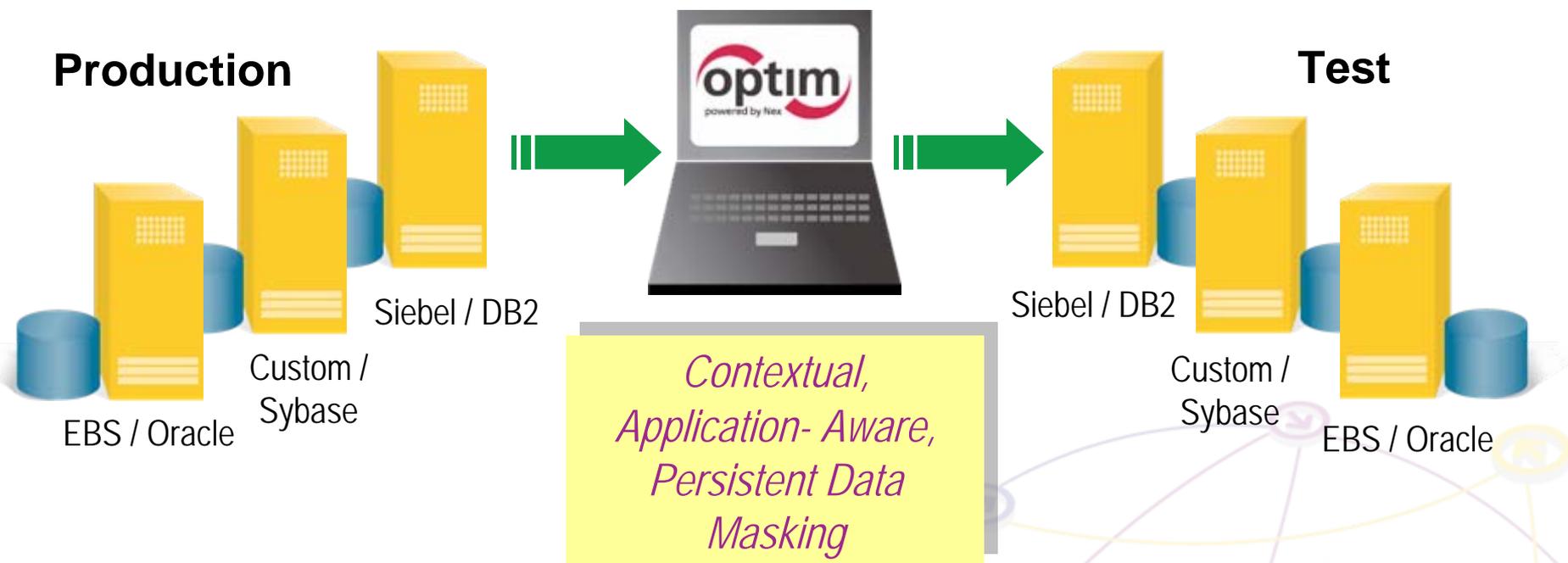
Test Data Management



Optim Test Data Management



Optim Data Privacy



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

Data Privacy



- A comprehensive set of data masking techniques to de-identify data
- Replaces (masks) confidential data with contextually accurate but fictionalized data

De-Identify test data

→ Can Be Performed

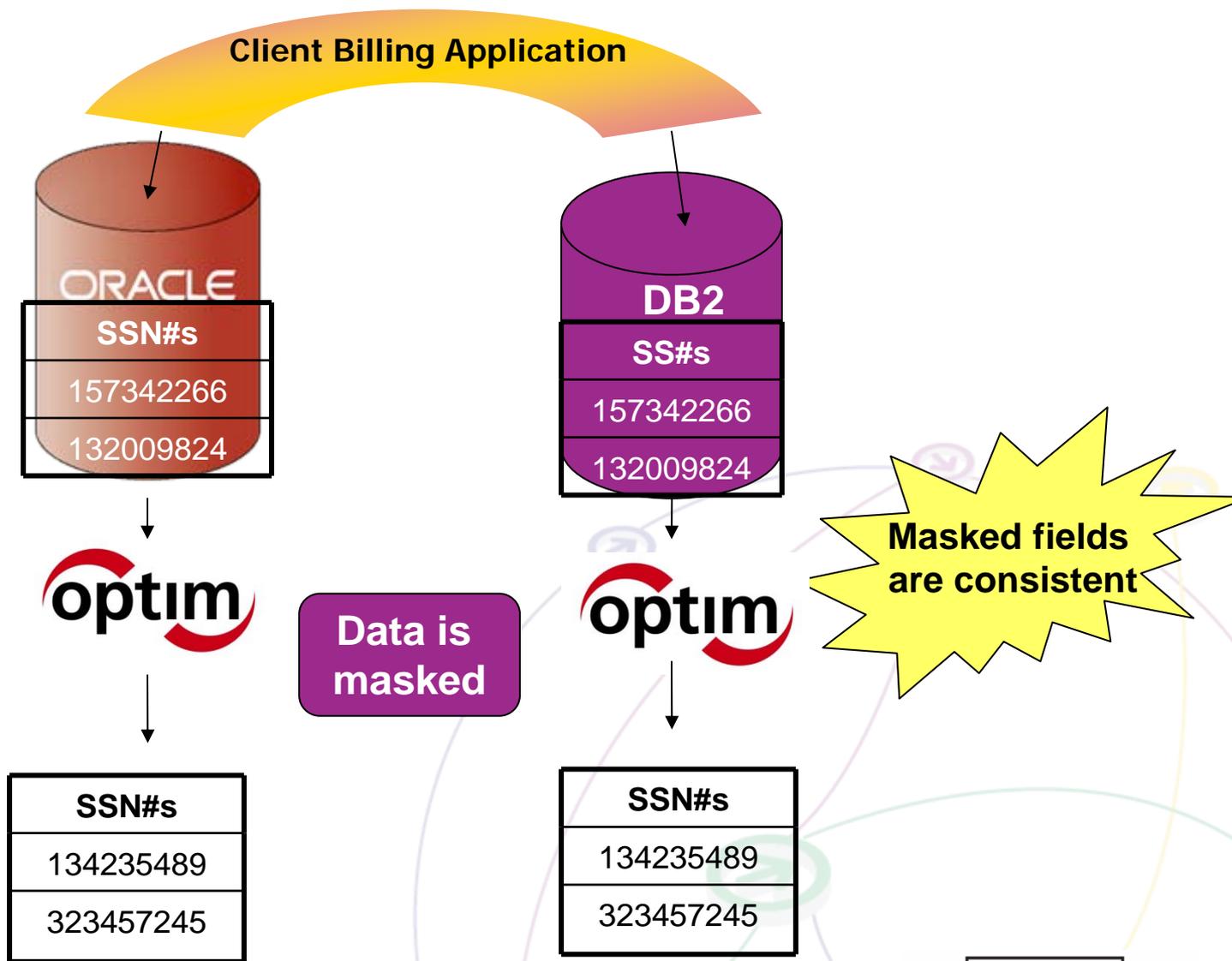
- During Extract Process from DB
- During Insert/Load Process to DB
- Or as a Standalone Convert Process

→ Transform or mask sensitive data using :

- Standard rules: Literals, Special Registers, Expressions, Default Values, Look-up tables
- Intelligent transformation rules: PCI, Addresses etc.
- Custom mapping rules: user exits

→ Converted extract file is safe to share – sanitized data

Consistent mapping Across the enterprise



Optim is Designed for the Enterprise

- Built on Enterprise Architecture
 - Support for all popular platforms and data sources
 - Built-in support for many popular applications
 - Scalable

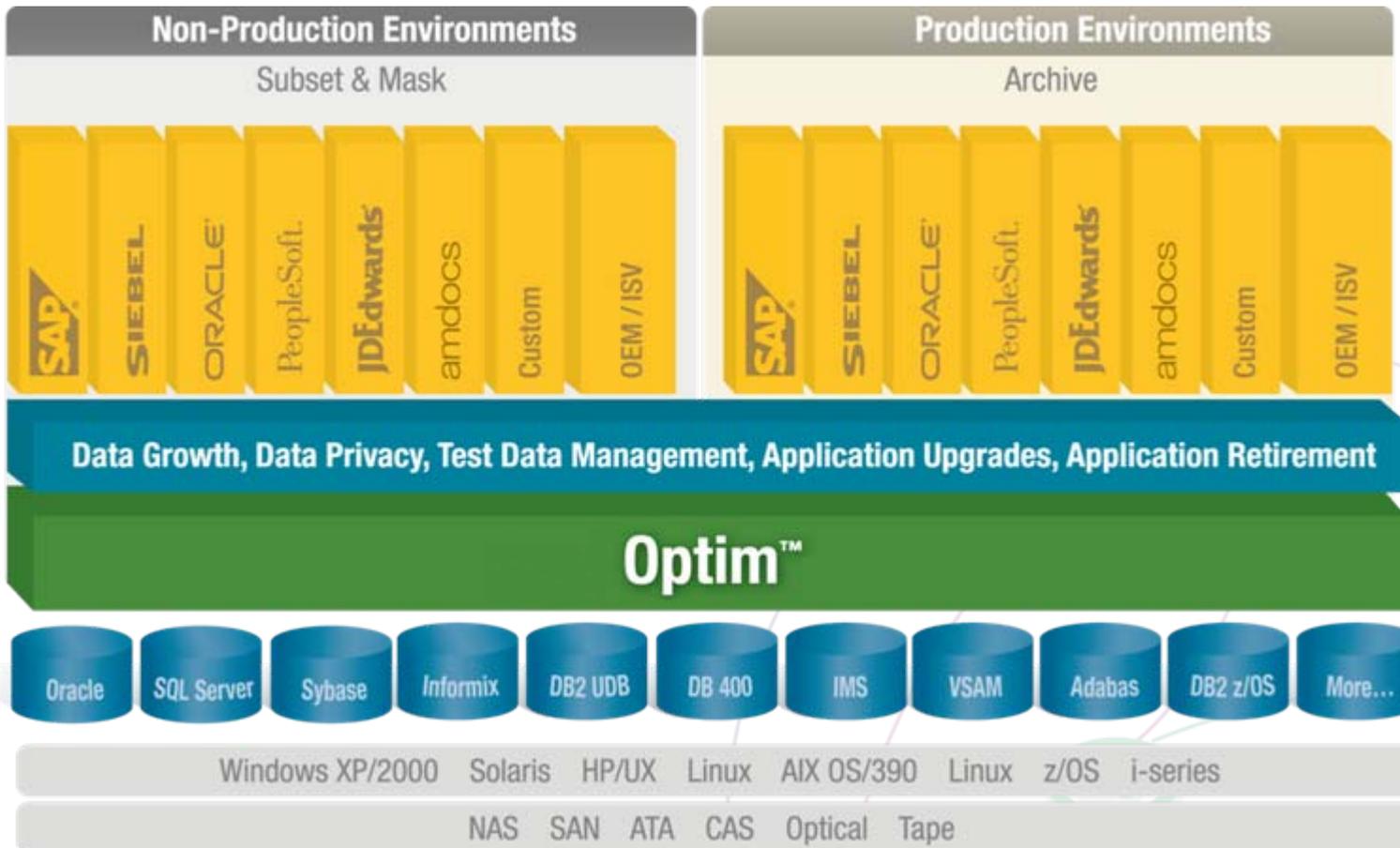
- Complete Business Object
 - A referentially intact subset of data
 - Maintains relationships across your tables, databases, and applications throughout its operations

- Extract, Store, and Restore
 - Clone, subset, move, and restore your data throughout your enterprise

- Universal Access

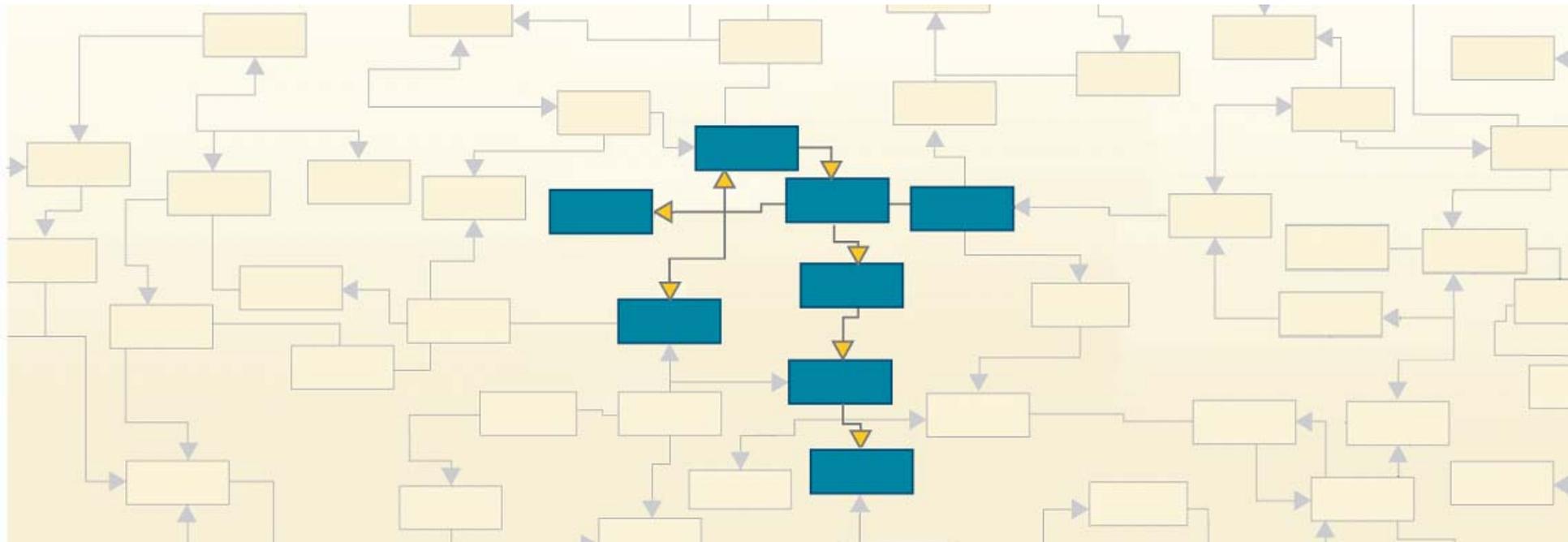
Optim's Enterprise Architecture

- Single scalable architecture provides central point to extract, store, restore, and protect (mask) application data records



Complete Business Object

- ➔ Referentially-intact subset of data across related tables and applications; includes metadata
- ➔ Provides “historical reference snapshot” of business activity



Extract, Store, and Restore

→ Extract

- Identify and extract business objects across multiple related applications, databases, and platforms

→ Store

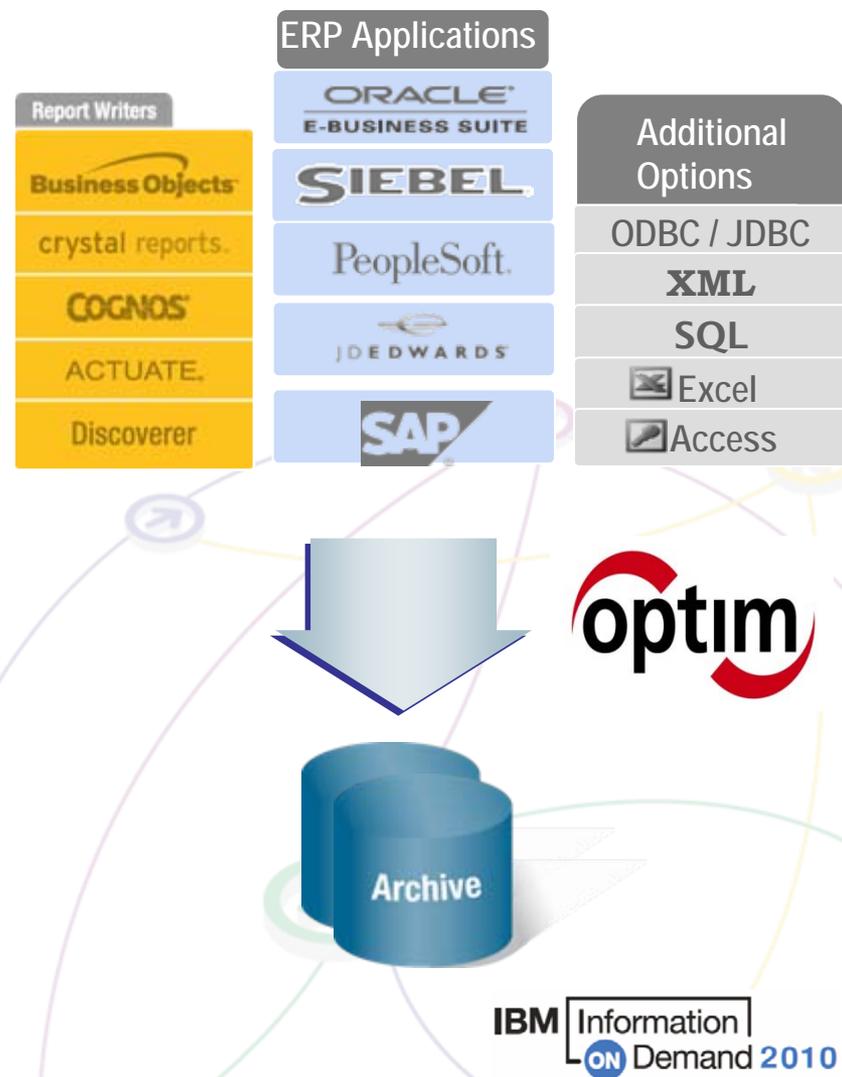
- Store immutable business objects independent of infrastructure – any hardware device or platform

→ Restore

- Restore complete business objects from any database, application or version to any other database, application, or version
- Choose exactly what items to restore – a single transaction record, a year's worth of data, or any other range

Universal Access

- Access any record, any time, anywhere!
- Native application access
- Application independent access
 - Industry standard methods (e.g. SQL, ODBC/JDBC, XML)
 - Portals
 - Report writers (e.g. Cognos)
 - Desktop formats (e.g. Excel, CSV, MS Access)
 - Database formats



Optim Products

→ IBM Optim Test Data Management

→ IBM Optim Data Growth

→ IBM Optim Data Privacy

→ For Custom Applications or Packaged Application Solutions

- Amdocs CRM
- JD Edwards Enterprise One
- Oracle E-Business Suite
- PeopleSoft Enterprise
- Siebel Applications

→ IBM Optim Data Privacy for SAP



Optim Terminology

→ Referentially-Complete

Optim extracts data based on primary/foreign key relationships (“parent/child”) between tables.

Handling data this way reduces errors and allows data to be moved without breaking application software

→ Subsetting

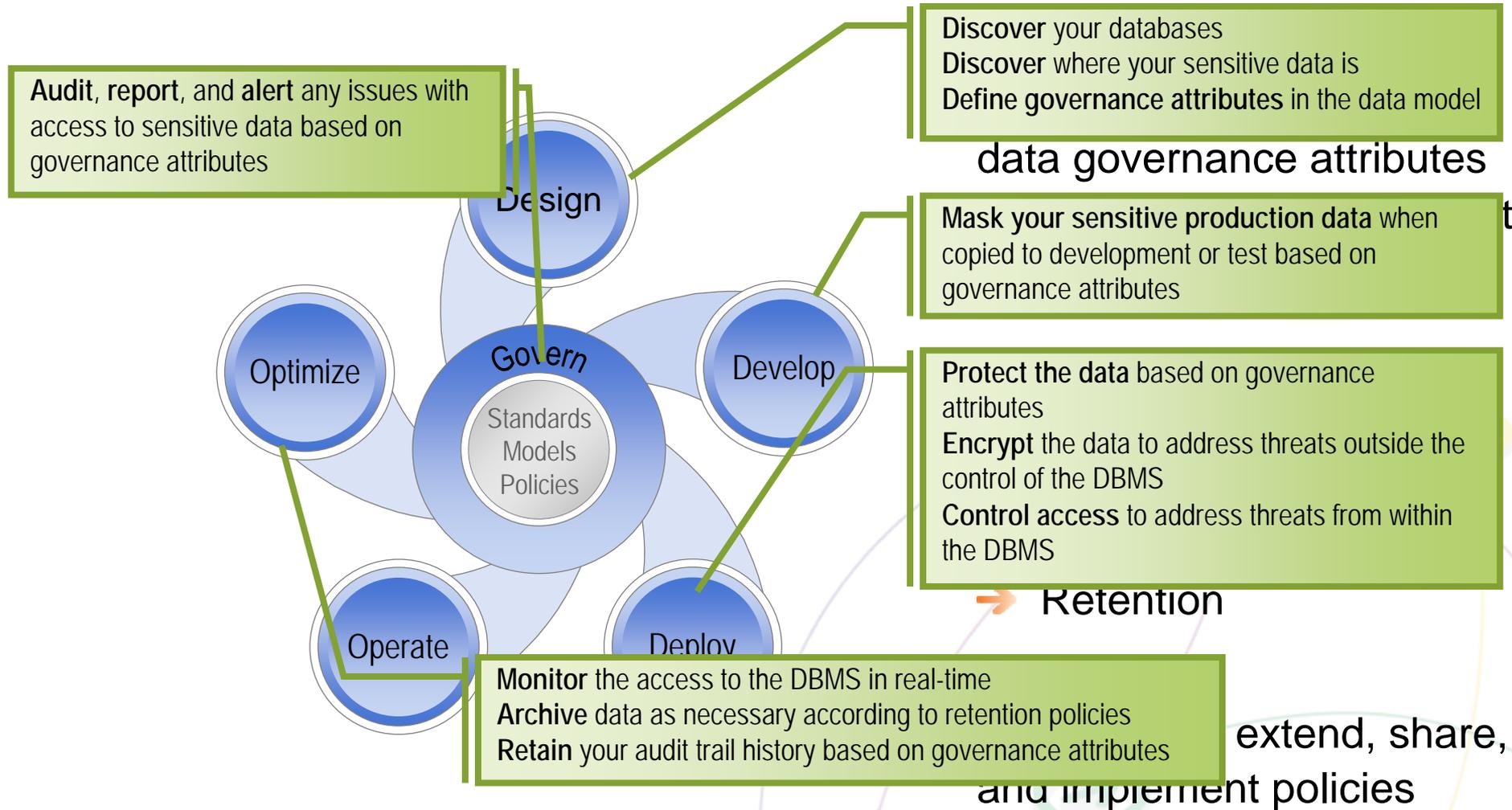
Using Optim to create a reduced size but referentially complete copy of a database for development or test.

→ Masking

Changing sensitive data before testing by replacing it with false but equally valid data.

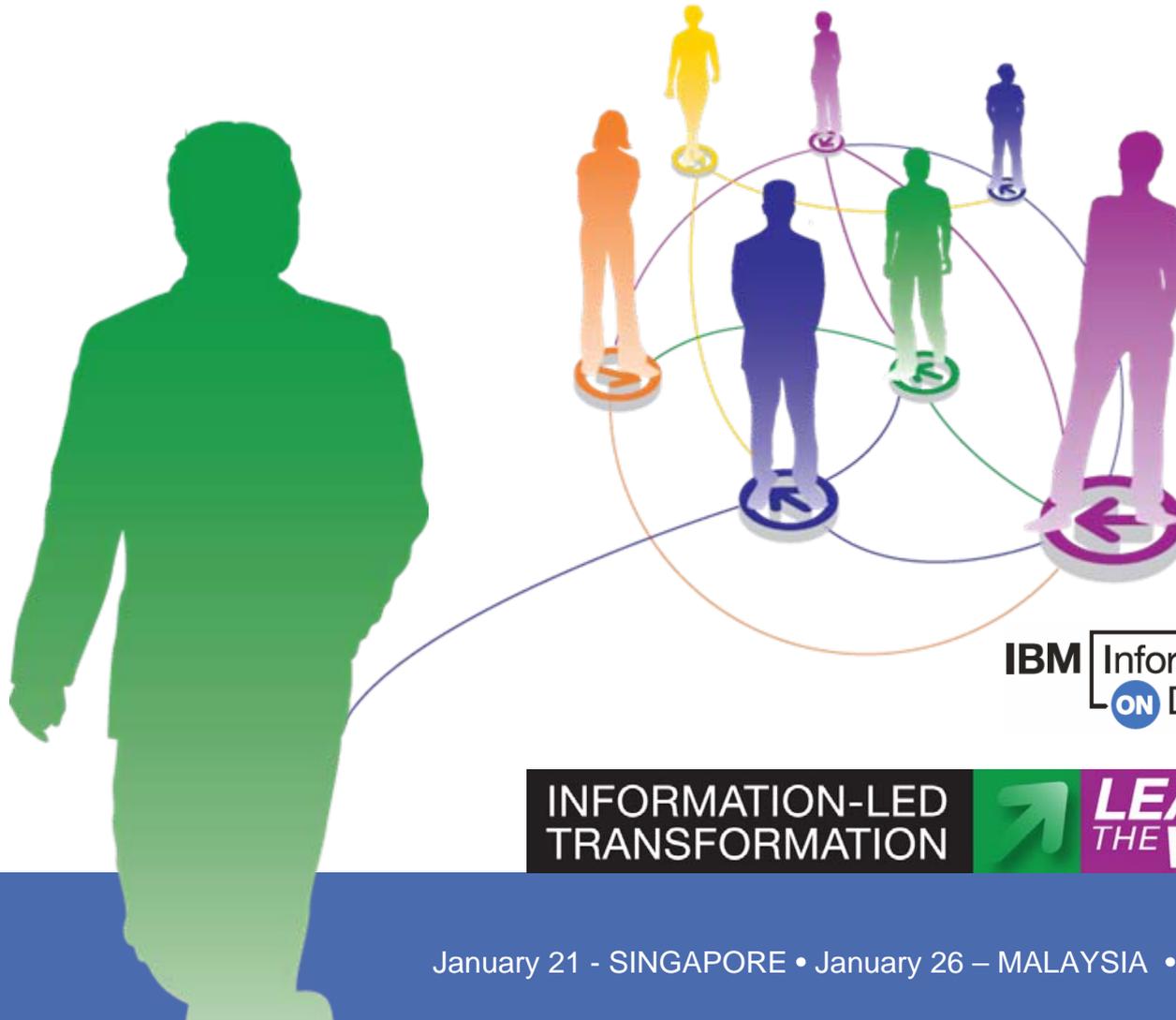
Integrated Data Management

Data Governance Direction



IBM Optim Solutions

Optim Development Studio and PureQuery



IBM Information
ON Demand 2010

INFORMATION-LED
TRANSFORMATION



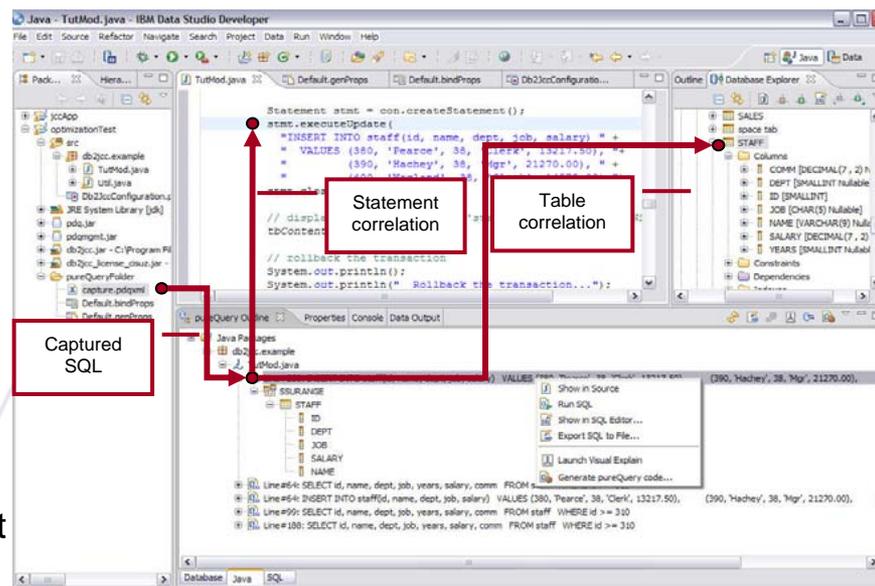
IBM

January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

Optim Development Studio and pureQuery Runtime

Development Studio is an integrated database development environment that speeds application design, development, and deployment while increasing data access performance and manageability.

- ➔ Ease JAVA coding
 - pureQuery code assistance
- ➔ Improve predictability and manageability with static SQL
 - Switch dynamic to static SQL without changing code
 - Eliminate SQL Injection Risk by approving which statements can be executed
 - Replace existing SQL with more efficient SQL without changing the source code
- ➔ Visualization of JAVA request to SQL code
 - SQL Outline correlates SQL to JAVA code and the associated objects
 - Provide developers with: No. of executions, elapsed time, min./max executions
- ➔ Impact Analysis
 - Display lines of code associated to a DB2 object to understand the impact of an object change.
- ➔ PureQuery for JAVA, Open JPA and .NET



Problem Statement

→ High cost of database application development

- No support for SQL development within Java IDE
- No way to generate and customize optimized SQL code and test cases

Improve developer productivity by bridging the gap between data and Java

→ Unexpected and uncontrolled application behavior

- Unpredictable performance due to runaway, ad-hoc and unexpected queries
- No link between application and executed SQL makes troubleshooting time consuming
- Difficulty meeting Service Level Agreements (SLA) and availability targets

Improved Quality of Service using static SQL

Introducing pureQuery

A high-performance, data access platform to simplify developing, managing, securing, and optimizing data access.

pureQuery Components:

→ Simple and intuitive API

- Enables SQL access to databases or in-memory Java objects
- Facilitates best practices

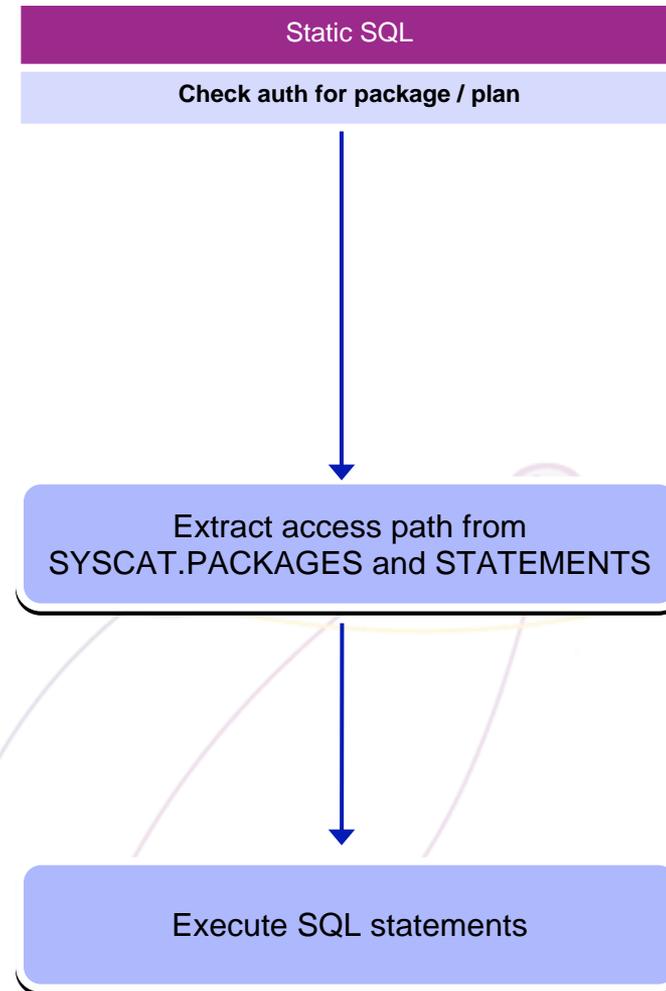
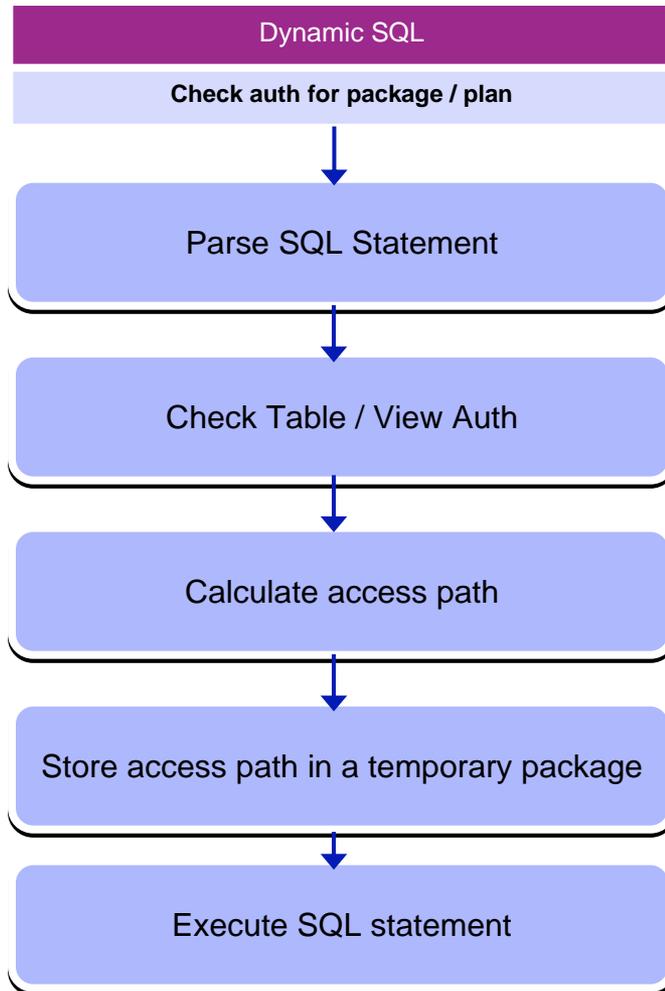
→ Optim Development Studio

- Integrated development environment with Java and SQL support
- Improve problem isolation and impact analysis

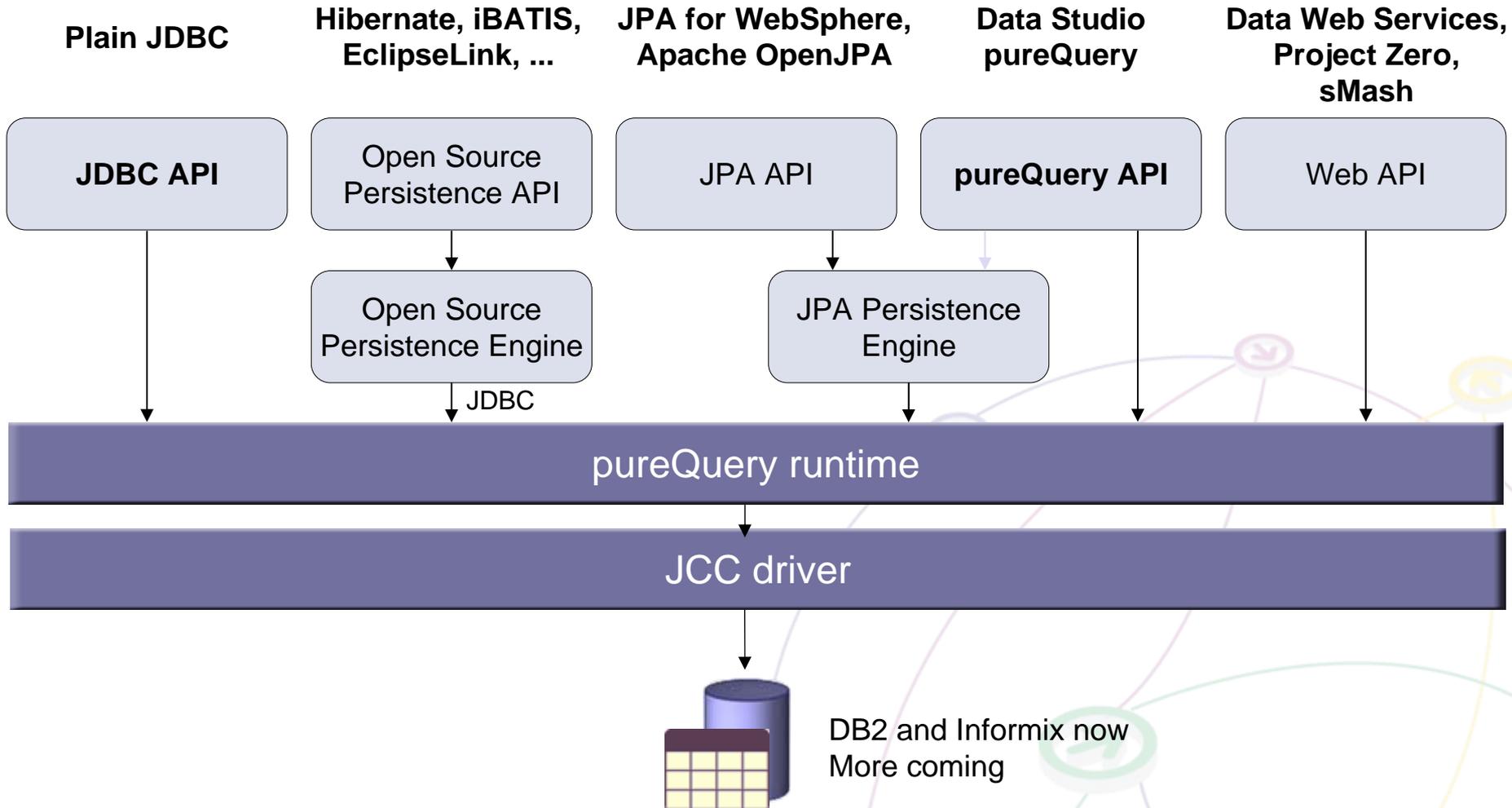
→ Optim pureQuery Runtime

- Flexible static SQL deployment for DB2

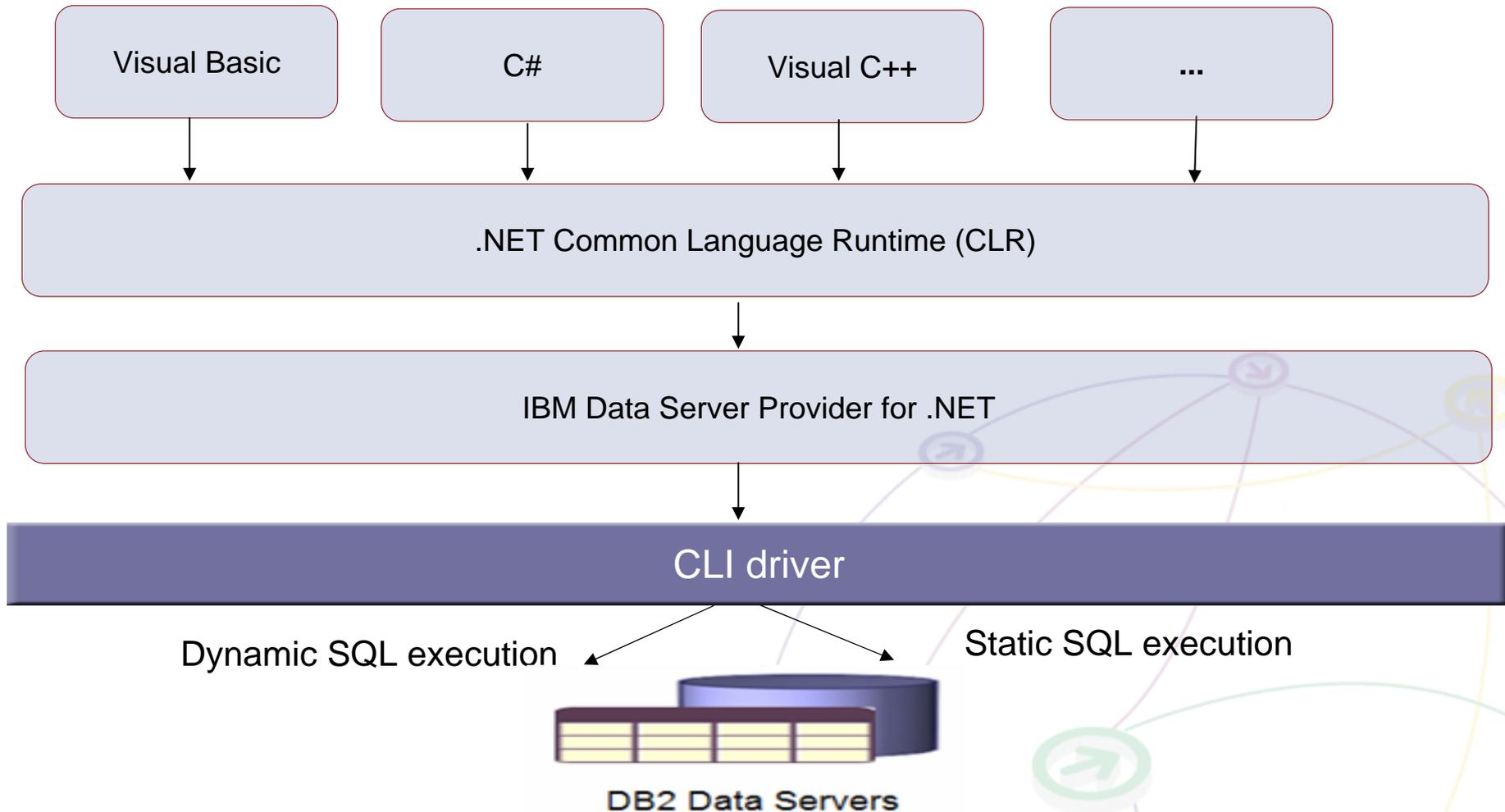
SQL Execution – Dynamic vs. Static



Java Database Access and pureQuery



Optimizing .NET Database access using pureQuery



Step 1: Browse the captured metadata

The screenshot displays the IBM Data Studio Developer interface with the following components:

- Package Explorer:** Shows the project structure for 'CustomerAnalysis', including 'src', 'JRE System Library [jdk]', 'Referenced Libraries', 'pureQueryFolder', and 'analysis'. A blue arrow points from the 'analysis' folder to the 'Captured SQL' label.
- Source Code Editor:** Displays the Java file 'AnalyzePurchasePatterns.java'. The code includes a SQL query:


```
String sql = "SELECT ORDER_DETAIL_CODE, ORDER_NUMBER, SHIP_DATE, PRODUCT_NUMBER, PROMOTION_CODE, UNIT_COST, UNIT_PRICE, UNIT_SALE_PRICE FROM GOSALES.ORDER_DETAILS WHERE ...";
```

 A blue arrow points from this SQL query to the 'Source code correlation' label.
- Data Source Explorer:** Shows the database schema for 'GOSALES', including tables like 'ORDER_DETAILS' and 'CUSTOMER'. A blue arrow points from the 'ORDER_DETAILS' table to the 'Database object correlation' label.
- Properties View:** Shows the 'pureQuery Outline' with a list of SQL queries. A context menu is open over the third query, showing options like 'Show in Source', 'Run SQL', and 'Export SQL to File...'. A blue arrow points from the 'Source code correlation' label to this context menu.

Captured SQL

Source code correlation

Database object correlation

Step 2: Configure database package

1. Identify which SQL should be excluded from the database package

2. Replace captured suboptimal SQL with better performing SQL

3. Configure database package characteristics

4. Preview database package in pureQuery Outline view

Database Packages

- [-] CUSTAN
 - [+] SELECT CUST_FRST_NAME, CUST_CITY FROM GOSALESC.T.CUST WHERE CUST_CITY = ?
 - [+] SELECT CUST_FRST_NAME, CUST_CITY FROM GOSALESC.T.CUST WHERE CUST_CITY = ? FETCH FIRST 10 ROWS ONLY
 - [+] CustomerAnalysis
 - [+] GOSALESC.T
 - [+] SELECT COUNT(CUST_CODE) AS NUMBER_OF_CUST FROM GOSALESC.T.CUST
 - [+] SELECT ORDER_DETAIL_CODE, ORDER_NUMBER, SHIP_DATE, PRODUCT_NUMBER, PROMOTION_CODE, QUANTITY, UNIT_COST
 - [+] SELECT COUNT(PRODUCT_NUMBER) AS TOTAL_PRODUCT FROM GOSALES.PRODUCT

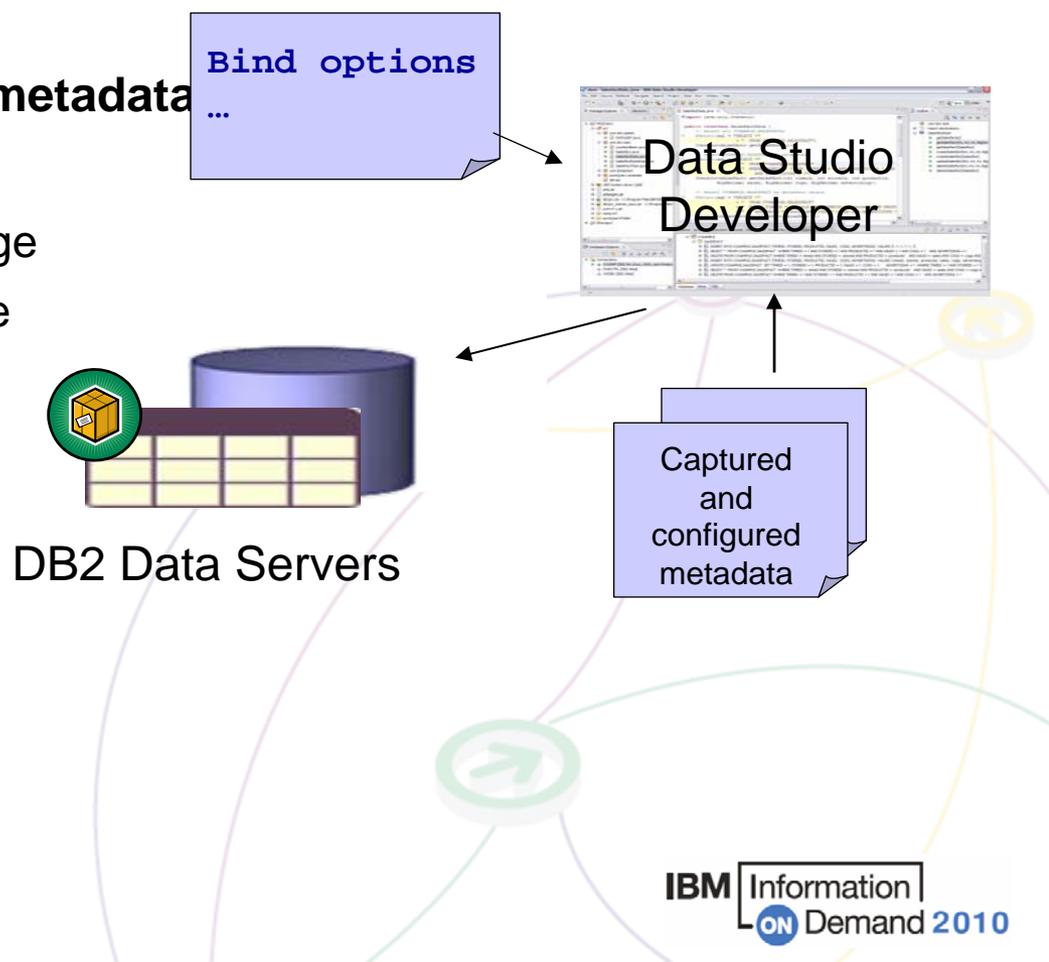
Database Java SQL

Optimize an existing application in four easy steps

Step 3: Bind database package to target database

→ Bind captured and configured metadata

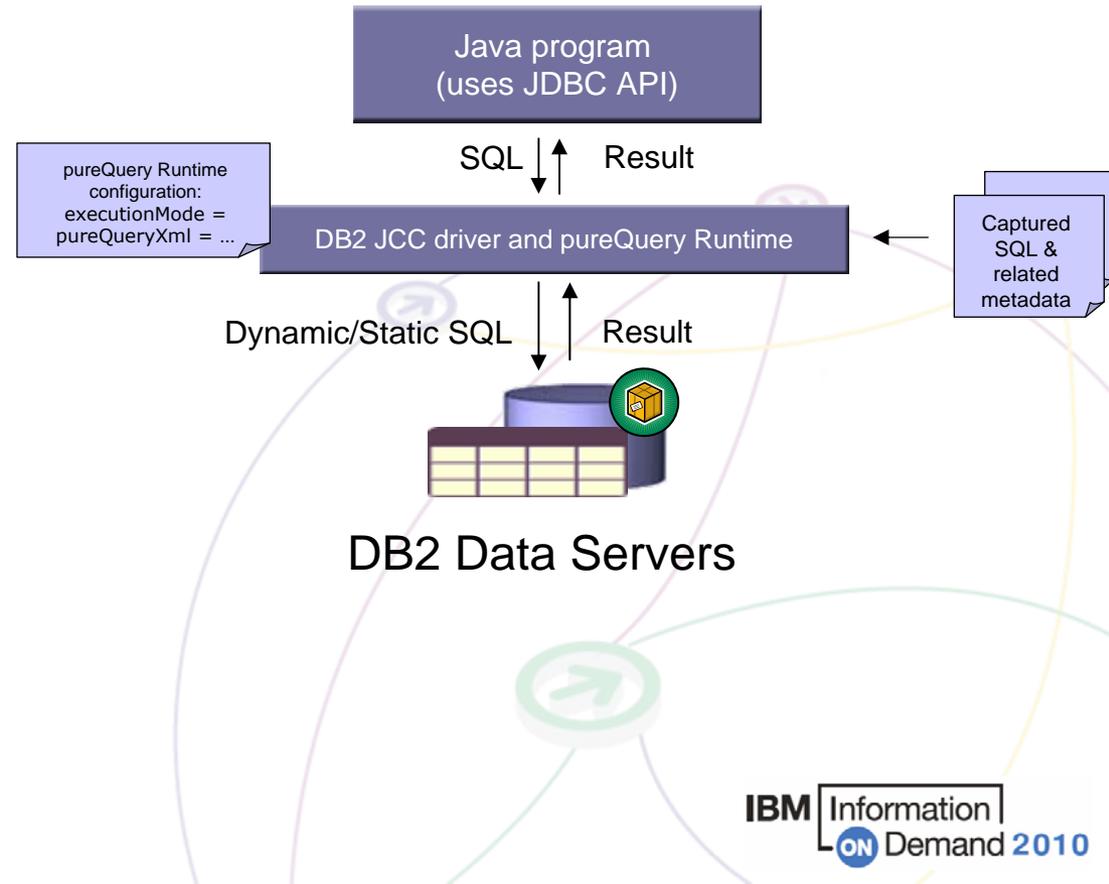
- Optionally, grant execute privilege for database package to one or more users



Optimize an existing application in four easy steps

Step 4: Execute application

- Application runtime environment configuration determines SQL execution behavior
- Configured metadata file can be loaded via class loader to simplify application deployment
- Run SQL statically, dynamically or in mixed mode
 - Choose execution mode based on life-cycle phase
- Prevent execution of SQL that has not been approved
 - Improve security by preventing SQL injection
- Run applications original SQL or (equivalent) optimized SQL
- Improve performance of problem queries without source code changes



SQL Concentrator

- SELECT employee FROM hr WHERE serial = 123456
- SELECT employee FROM hr WHERE serial = 563293
- SELECT employee FROM hr WHERE serial = 432345
- SELECT employee FROM hr WHERE serial = 243223



- SELECT employee FROM hr WHERE serial = ?
- Lock SQL and protect against SQL Injection
- Convert to static and save CPU cycles
- Save space in dynamic statement cache

pureQuery Improves Performance, Security, and Manageability for DB2 ...Without Changing a Line of Code

Three steps

1. Capture the SQL

- Use pureQuery API, generate from WebSphere JPA, or capture while executing
- Use with custom-developed, framework-based, or packaged applications

2. Bind SQL to DB2

- Use tooling in Data Studio Developer, WAS console, or command line

3. Choose execution mode

- Dynamic or static
- Choose at deployment time instead of development time

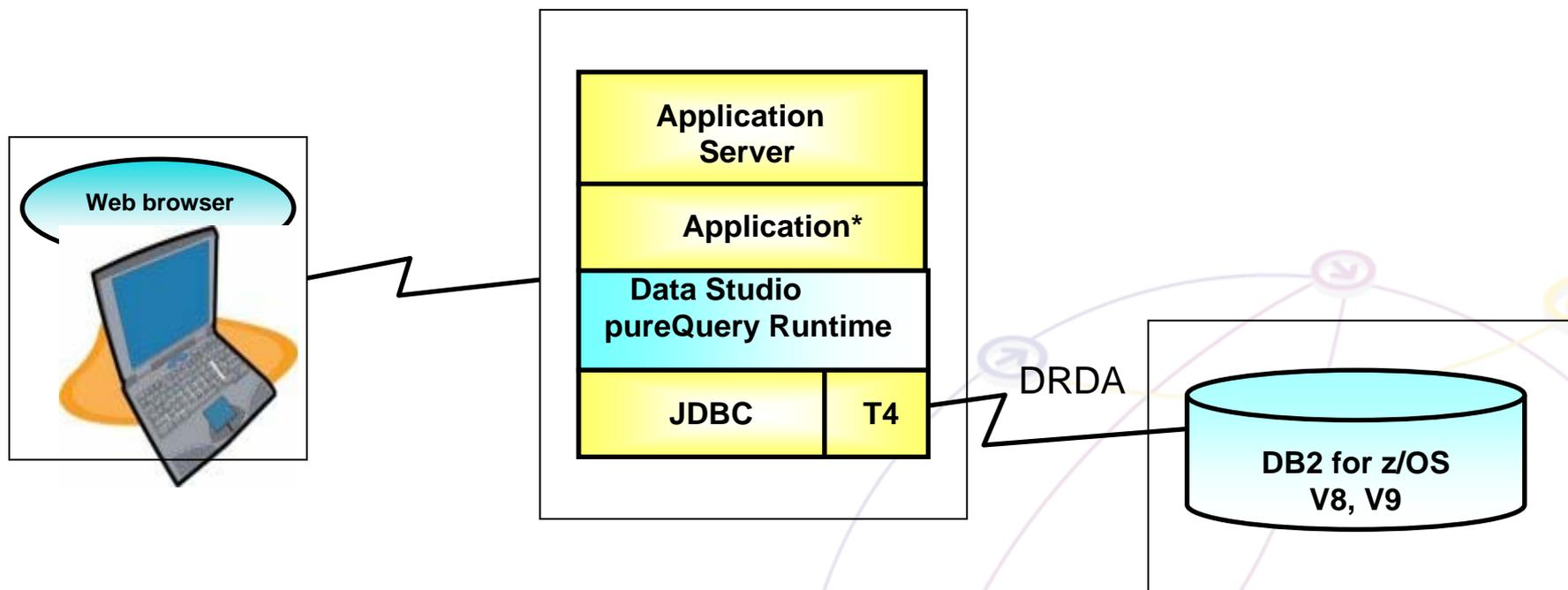
Static SQL value

- Make response time predictable and stable
 - By locking in the SQL access path pre-execution, rather than re-computing at access time
- Limit user access and reduce SQL injection
 - By granting execute privileges on the query packages rather than access privileges on the table
- Accelerate problem resolution
 - By tracing SQL execution to a specific package and the originating source
- Increase system capacity
 - By driving down DB cycles

80-90% of pure Query - Reference Architecture

Configurations Supported:

Application on Linux, UNIX, Windows, or z/Linux accessing DB2 for z/OS

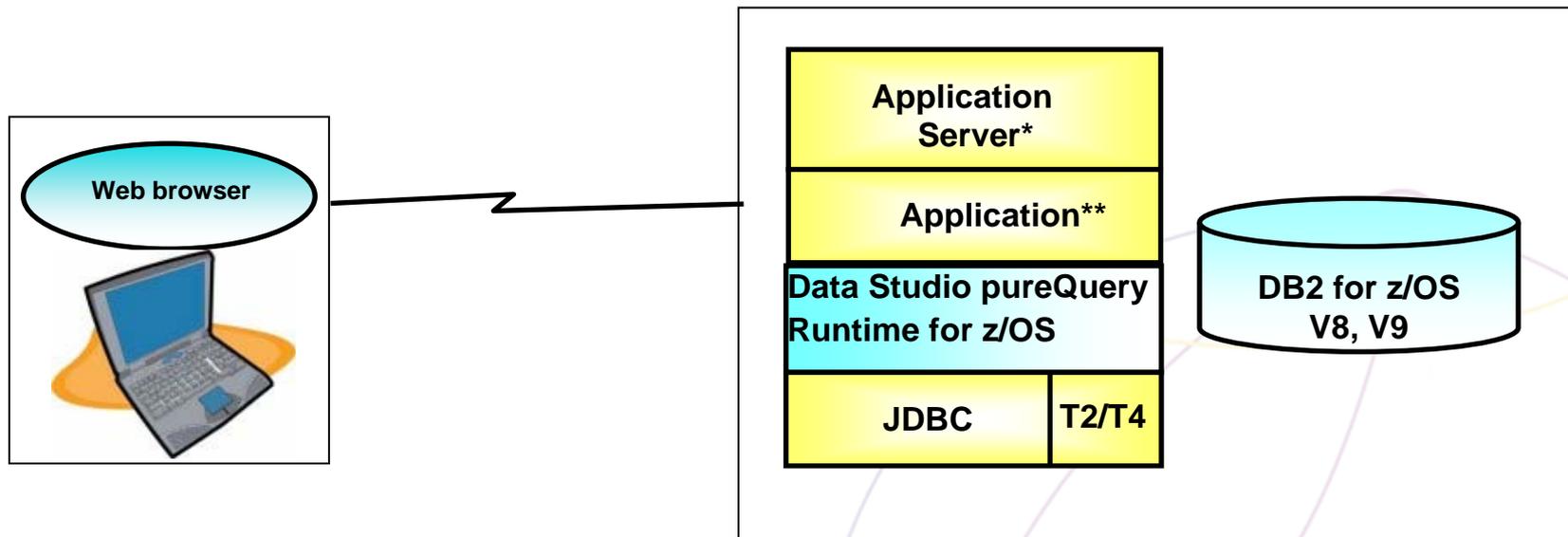


* - Application running on AIX, Linux, Solaris, Windows and z/Linux

5-10% of pureQuery - Reference Architecture

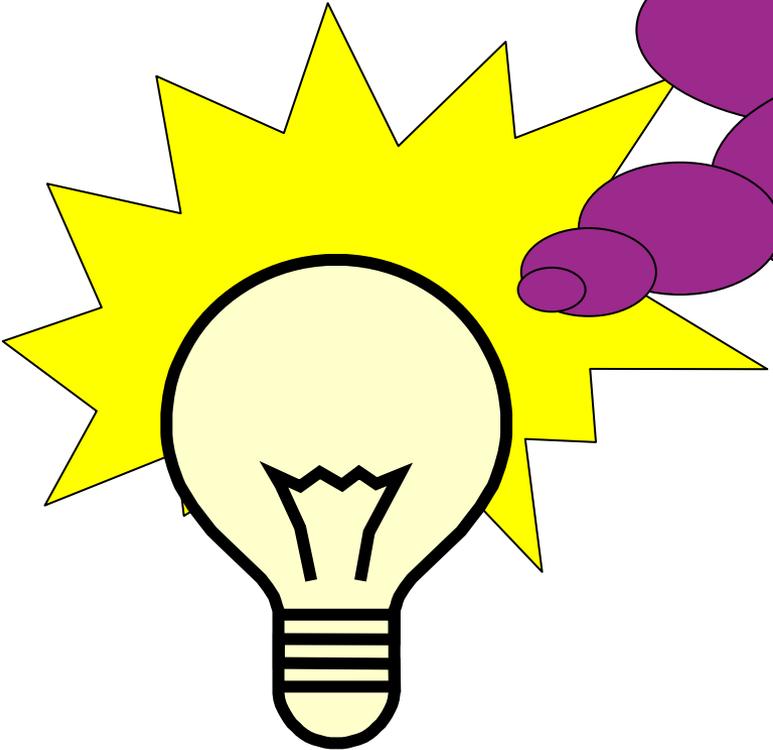
Configurations Supported:

Application running natively on z/OS accessing DB2 for z/OS



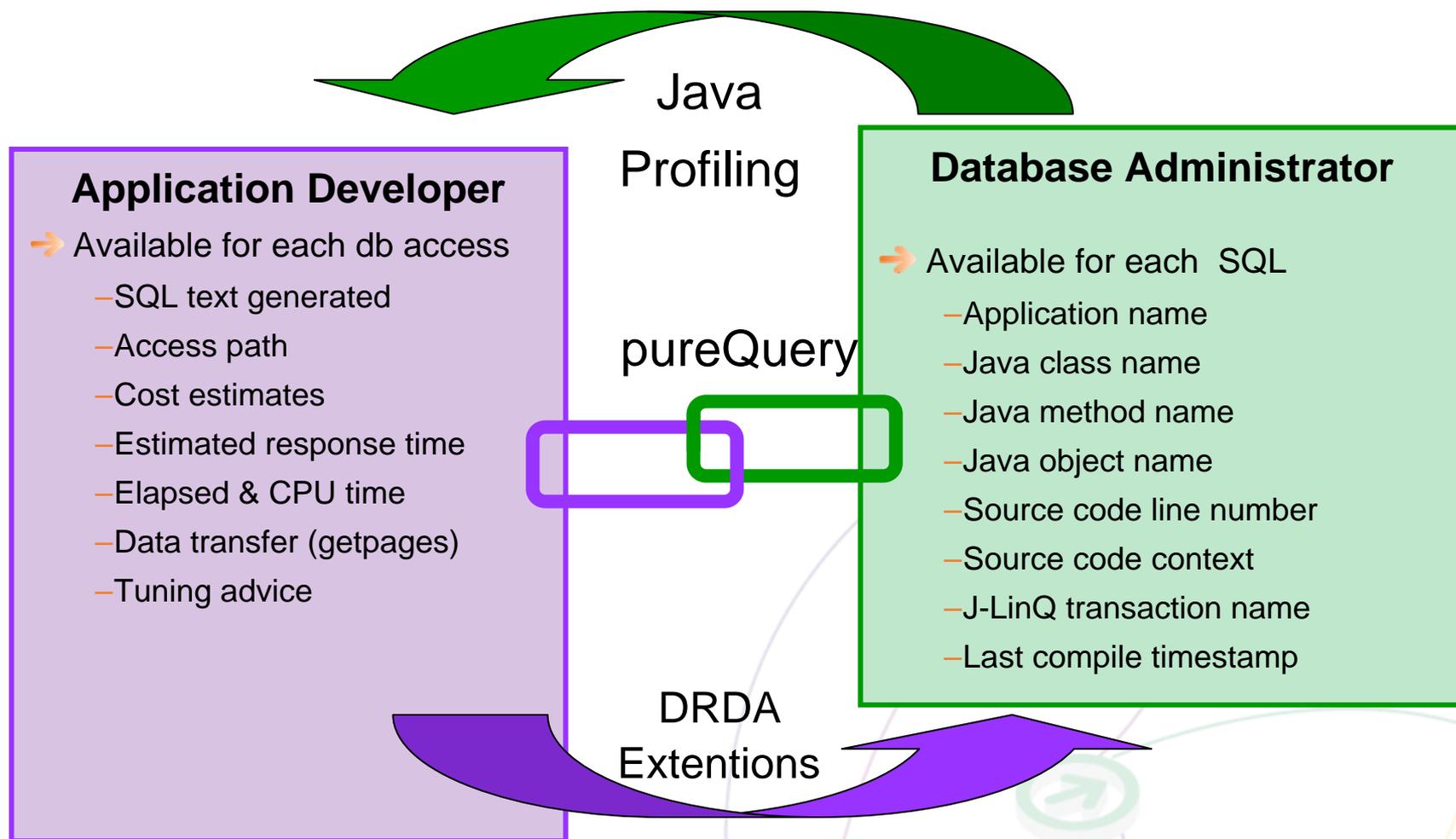
* - App. Server on z/OS

** - Application running on z/OS



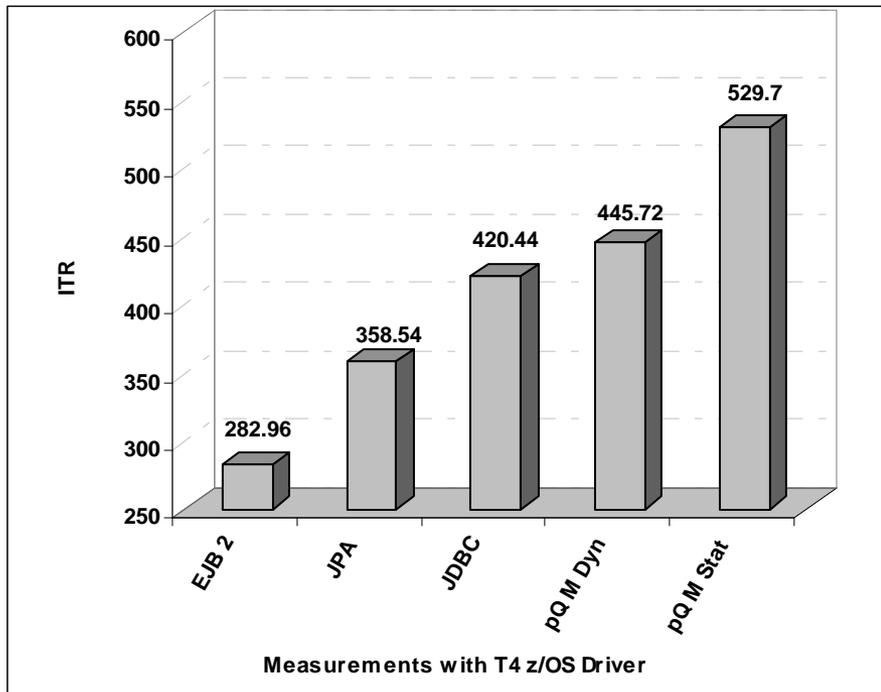
If you think about it;
pureQuery makes Java
web applications behave a
lot like CICS COBOL
applications...

Simplifying Problem Determination Scenario

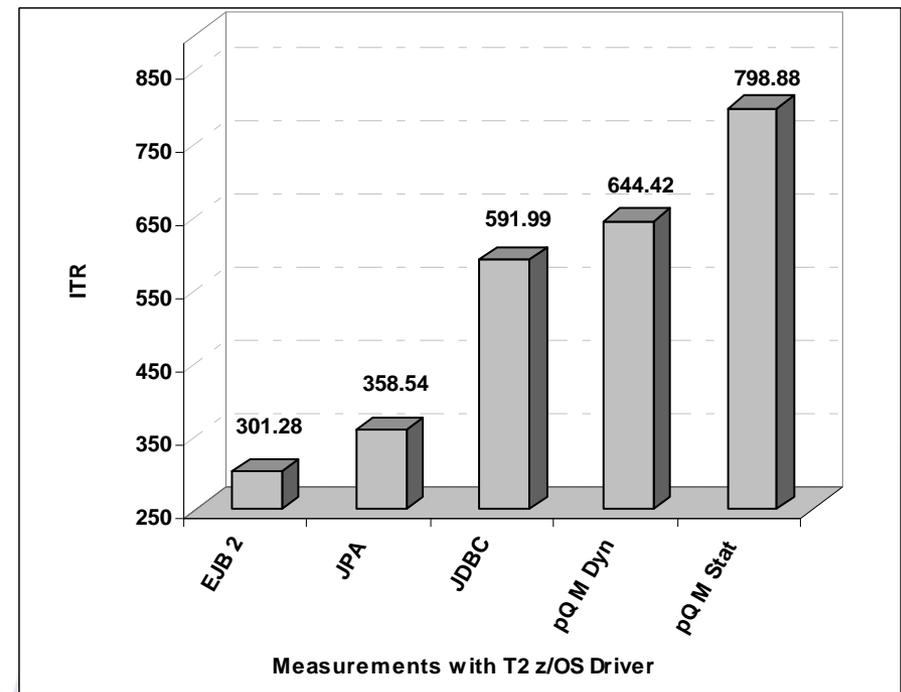


pureQuery improves throughput

(pre-release numbers for v1.2 on IRWW benchmark, ~75% Dynamic cache hit)



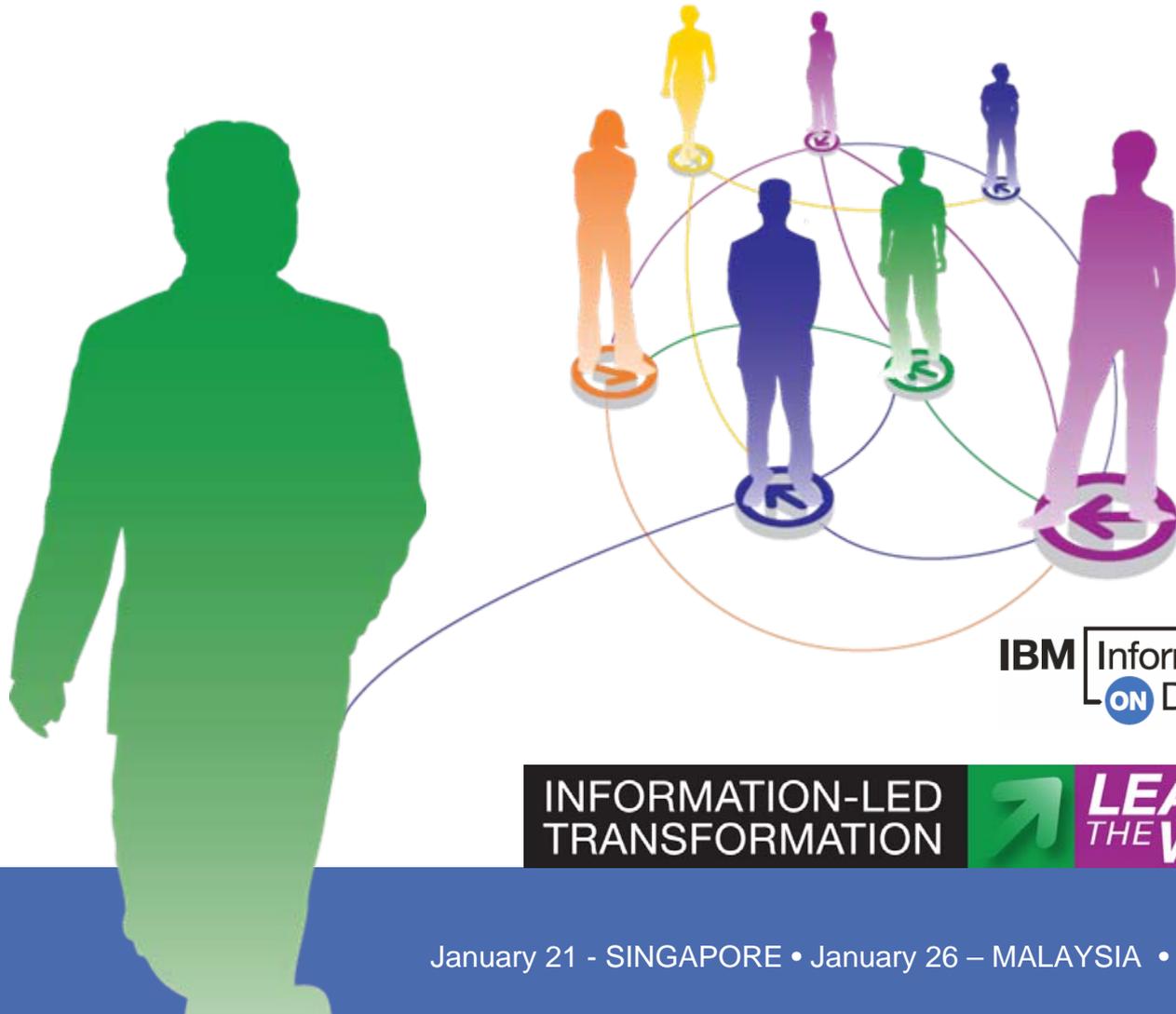
pQ M Static 26% faster than JDBC



pQ M Static 35% faster than JDBC

IBM Optim Solutions

Optim Query Tuner



IBM Information
ON Demand 2010

INFORMATION-LED
TRANSFORMATION



January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

Why is Query Tuning Such a Pain?

→ Development

- Lack of query tuning skills
- Use of frameworks that generate SQL
- Inadequate testing environments to drive the workload and data scale
- Disconnect between application and data groups

→ Production

- Performance problems appear without warning
- Aggregate information from multiple places
- Complexity of analysis requires specialized skills
- Need to consider entire workloads, not just single queries

IBM Optim Query Tuner Offerings

Optimize performance and reduce costs



→ Optimize performance

- Improve customer satisfaction
- Improve organizational productivity
- Satisfy service level objectives
- Fit workloads into maintenance windows

→ Reduce costs

- Increase overall system capacity
- Reduce chargebacks
- Simplify analysis for DBAs, developers, designers, and others
- Improve statistics collection efficiency
- Reduce time to respond

Expert-Enabled Query Tuning Out of the Box

→ Identify query candidates

- DB2 catalog
- Dynamic statement cache
- Development Studio hot spots
- Query or performance monitors

→ Facilitate analysis

- Query formatting
- Query annotation
- Access path visualization and annotation

→ Get expert tuning advice

- Improve query design
- Improve statistics quality
- Improve database design

The screenshot displays the IBM Query Tuner interface with the following components:

- Query Annotation:** Shows the parsed and formatted SQL query. The formatted query includes a subquery with a CASE statement and joins between TPCDS.CATALOG.REPAIRS and TPCDS.CATALOG.RETURNS.
- Access Plan Graph:** A visual representation of the query execution plan, showing nodes for RETURN, GRPBY, TBSCAN, and SORT.
- Advisor Recommendation Overview:** A table listing various advisors and their recommendations.

Advisor	Priority	Description
Statistics Advisor	HIGH	Gather and recollect all of relevant statistics
Query Advisor	MEDIUM	Consider adding join predicates between columns R_ITEM_SK in table REPAIR(R) and columns CR_ITEM_SK in table CATALOG.RETURNS(C) which use the referential constraints between table REPAIR(R) and table CATALOG.RETURNS(C) to avoid a potentially costly Cartesian join. Check the explanation for this warning for more details about possible impact and examples.
Query Advisor	LOW	Consider adding the following predicate to the TPCDS.CATALOG.RETURNS table is at
Access Path Advisor	LOW	The TPCDS.CATALOG.RETURNS table is at
- Selected Recommendation:** A detailed view of the MEDIUM priority recommendation, including a description and an explanation with a sample SQL query.


```
SELECT T1.C2, T2.C2
FROM T1, T2
WHERE T1.C2 = :charHV3
AND T2.C2 = :charHV4
```

Streamlined Analysis

Define or select workload

Select a workload and then click an item in the Open menu to view information about the workload.

New Workload Open Refresh Remove Tools

Wizard... Statement Cache... Catalog... QMF... QMF HPO... File... Category... Other Workloads... Import... Export...

Summary Status	Owner
ENABLED/STARTED	SYSADM
ENABLED/STARTED	SYSADM
ANALYZING	B3OSC07
ANALYZING	B3OSC12
ANALYZING	B3OSC12

Execute advisors

Workload Statements

Immediately capture statements or multiple sources to this workload, launch workload advisors, use tools capture, consolidation, and analysis.

Capture Workload Tools Query Tools Schedule History

Filter name: DEF

Run All Advisors
Run Workload Statistics Advisor
Run Workload Index Advisor
Run Workload Query Advisor
Generate Workload
Workload Report
Workload Environment Capture

Execution Count	Average Elapsed Time (sec)	Accumulated Elapsed Time (sec)
5	21.151768	0.0694
5	23.351143	2.2789
5	18.033714	10.322
5	5.907151	0.0459
5	20.778751	5.5212

Drill into advice

Advisor Recommendation Overview

These recommendations are part of a sample project. [Learn More](#)

Advisor	Priority	Description
Recommendations		
Statistics Advisor	HIGH	Repair statistics problems for this query.
Query Advisor	MEDIUM	Consider adding join predicates between tables.
Query Advisor	MEDIUM	Consider adding the following predicate to the query.
Access Path Advisor	LOW	The DSN8910.EMP table is accessed by a full table scan.
Index Advisor	LOW	Index recommendations found.
Statistics Advisor	MAINTENANCE	Gather and recollect all of relevant statistics.

Context

Query 1 Initial Analysis

Validate improvement

Workloads List

Select a workload and then click an item in the Open menu to view information about the workload. To create a workload, click New Workload.

New Workload Open Refresh Remove Tools

Name	Summary Status	Owner	Execution Time
WorkloadWithTypicalStats	ANALYZING	B3OSC12	CPU time: 97.32 (second...)
WorkloadTunedWithStatsAdvisor	ANALYZING	B3OSC12	CPU time: 53.19 (second...)
WorkloadTunedWithIndexAdvisor	ANALYZING	B3OSC07	CPU time: 40.67 (second...)
AbsoluteCPUTimeExceptionMonitor	ENABLED/STARTED	SYSADM	N/A
NormalMonitor	ENABLED/STARTED	SYSADM	N/A

Gather High Cost Queries and Workloads

DB2

- Plan table
- Statement table
- Catalog plan or package
- Statement cache
- Profile monitor

Optim Development Studio

QMF and QMF HPO

DB2 Query Monitor

OMEGAMON XE for DB2

File, text, or exported workload

User defined category

Define or select workload

Workloads List

Select a workload and then click an item in the Open menu to view information about the

New Workload Open Refresh Remove Tools

	Summary Status	Owner
Wizard...		
Statement Cache...	DISABLED/STARTED	SYSADM
Catalog...	DISABLED/STARTED	SYSADM
QMF...	ANALYZING	B3OSC07
QMF HPO...	ANALYZING	B3OSC12
File...	ANALYZING	B3OSC12
Category...		
Other Workloads...		
Import...		
Export...		

Visualize Queries and Costs to Speed Analysis

Formatted Query	Annotation	Additional Information
<pre>SELECT A.EMPNO , A.FIRSTNME , A.LASTNAME , A.JOB , A.SALARY , A.BONUS , A.COMM , B.LOCATION , C.PROJNAME FROM DSN8910.DEPT AS B , DSN8910.EMP AS A , DSN8910.EPROJ AS C WHERE (A.EMPNO IN (SELECT DSN8910.DEPT.MGRNO FROM DSN8910.DEPT WHERE DSN8910.DEPT.MGRNO IS NOT NULL) AND A.WORKDEPT = B.DEPTNO AND B.DEPTNO = C.DEPTNO) ORDER BY A.EMPNO ASC , A.FIRSTNME ASC , A.LASTNAME ASC</pre>	<pre>CARDF=14 QUALIFIED_ROWS= CARDF=42 QUALIFIED_ROWS= CARDF=(missing) QUALIFIED_ROWS= COLCARDF=42 MAX_FREQ=(r CARDF=14 QUALIFIED_ROWS= COLCARDF=9 MAX_FREQ=42</pre>	<pre>DSN8910.DEPT.MGRNO contain(s) skewed data DSN8910.EMP.WORKDEPT contain(s) skewed data</pre>

Easily see tables, sections, join predicates, etc.

Examine table statistics and additional information

→ Accelerate analysis, reduce downtime

- Spot human errors
- Identify where filtering should occur

View Optimizer Transformations

Original Transformed

Annotation to display: All

Expand All Collapse All Customize Save Print

Formatted Query

```
SELECT A.EMPNO
      , A.FIRSTNME
      , A.LASTNAME
      , A.JOB
      , A.SALARY
      , A.BONUS
      , A.COMM
      , B.LOCATION
      , C.PROJNAME
FROM DSN8910.DEPT AS B
      , DSN8910.EMP AS A
      , DSN8910.EPROJ AS C
WHERE ( A.EMPNO IN ( SELECT DSN8910.DEPT.MGRNO
                    FROM DSN8910.DEPT
                    WHERE DSN8910.DEPT.MGRNO IS NOT NULL
                  )
      AND A.WORKDEPT = B.DEPTNO
      AND B.DEPTNO = C.DEPTNO
      )
ORDER BY A.EMPNO ASC
      , A.FIRSTNME ASC
      , A.LASTNAME ASC
```



Original Transformed

Annotation to display: All

Expand All Collapse All Customize Save Print Clear Highlights

Formatted Query

```
SELECT A.EMPNO
FROM DSN8910.DEPT AS B
      , DSN8910.EMP AS A
      , DSN8910.EPROJ AS C
WHERE ( A.WORKDEPT = B.DEPTNO
      AND A.WORKDEPT = C.DEPTNO
      AND B.DEPTNO = C.DEPTNO
      AND A.EMPNO = SYSADM."DSNWFQB(02)".MGRNO
      DB2 creates a virtual table, SYSADM.DSNWFQB(02), to process the following correlated subquery
          SELECT DSN8910.DEPT.MGRNO
          FROM DSN8910.DEPT
          WHERE ( DSN8910.DEPT.MGRNO IS NOT NULL
                AND A.EMPNO = DSN8910.DEPT.MGRNO
                )
      End of the subquery for the SYSADM.DSNWFQB(02) virtual table.
      )
ORDER BY A.EMPNO ASC
      , A.FIRSTNME ASC
      , A.LASTNAME ASC
```

- Accelerate analysis, reduce downtime
- Spot transformations which will occur

Execute Advisors

Workload Statements

Immediately capture statements or multiple sources to this workload, launch workload advisors, use tools capture, consolidation, and analysis.

Capture Workload Tools Query Tools Schedule History

Filter name: DEF

< Previous

Run All Advisors

Run Workload Statistics Advisor

Run Workload Index Advisor

Run Workload Query Advisor

Generate Workload

Workload Report

Workload Environment Capture

Execution Count	Cache	Average Elapsed Time (sec)	Accum
5		21.151768	0.0696
5		23.351143	2.2789
5	CACHE	18.033714	10.322
5	CACHE	5.907151	0.0459
5	CACHE	20.778751	5.5212
5	CACHE	99.665344	0.5974
5	CACHE	13.440651	2.4417
9	CACHE	9.309578	0.3156
9	CACHE	18.914537	0.1205
11	CACHE	17.629927	14.428
11	CACHE	22.592033	5.0555
11	CACHE	18.943417	0.1524
11	CACHE	18.937361	0.0849
11	CACHE	3.410327	2.4663
11	CACHE	18.706799	0.0596
13	CACHE	18.79925	1.2846
13	CACHE	27.179655	6.4435
16	CACHE	5.250328	0.0872
16	CACHE	0.970876	0.3689

→ Statistics

- Get recommendations on the best statistics to capture to influence access path selection

→ Query

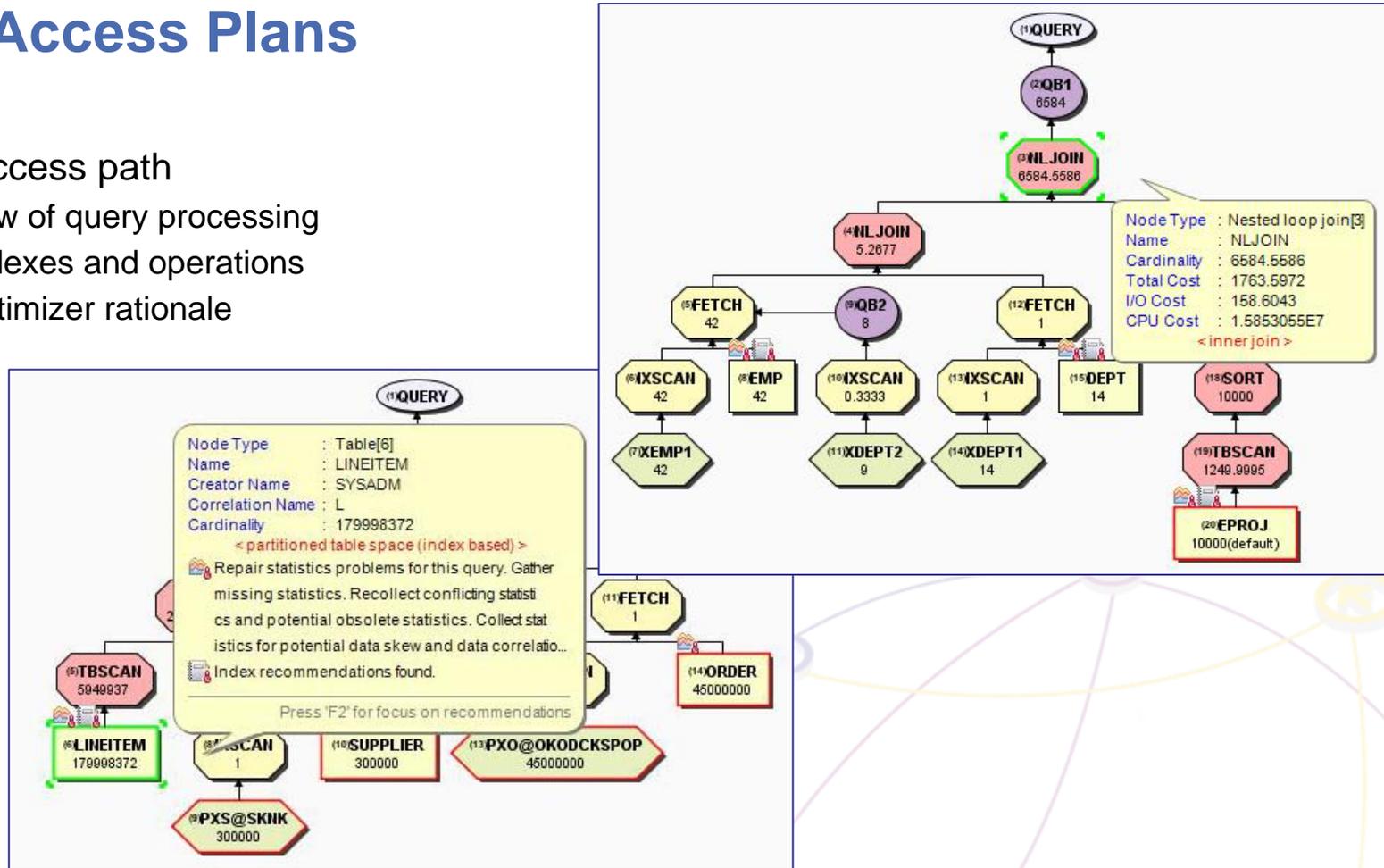
- Get recommendations regarding how to rewrite the query for better efficiency

→ Index

- Get recommendations on additional indexes that can reduce database scans

Analyze Access Plans

- Visualize access path
 - See flow of query processing
 - See indexes and operations
 - See optimizer rationale



- Assess access path stability to reduce risk of performance regression
 - Is the optimizer able to apply the filtering early?
 - Are there indexes that support an efficient path?
 - Do statistics allow distinction between the choices?

Improve Statistics Quality and Collection

- Provides advice on
 - Missing statistics
 - Conflicting statistics
 - Out-of-date statistics

- Results
 - Accurate estimated costs
 - Better query performance
 - Less CPU consumption
 - Improved maintenance window throughput

“Half of access path PMRs could be resolved by statistics advisor before calling IBM support.”
– IBM Support

Statistics Recommendation Detail

Recommendations

RUNSTATS Control Statements

```
RUNSTATS INDEX(SYSADM.PXS@SKNK FREQVAL NUMCOLS 1 COUNT 15)
SHRLEVEL CHANGE REPORT YES UPDATE ALL HISTORY NONE

RUNSTATS TABLESPACE DB4LINEI.TSLINEI
TABLE(SYSADM.LINEITEM) SAMPLE 40
COLUMN(L_SHIPMODE)
COLGROUP(L_SHIPDATE) HISTOGRAM NUMQUANTILES 25
COLGROUP(L_RECEIPTDATE) FREQVAL COUNT 15 HISTOGRAM NUMQUANTILES 25
COLGROUP(L_SHIPMODE) FREQVAL COUNT 15
COLGROUP(L_RETURNFLAG) FREQVAL COUNT 15
SORTDEVT SYSDA SORTNUM 4
INDEX(SYSADM.SXL@PKSKOKEPDSQN KEYCARD,
```

Generates RUNSTATS control statements

Table, index, column, and column group details

O_ORDERKEY

Cardinality: 4.5E7

Uniform statistics collection time: 0001-01-01 00:00:00.0

Uniform statistics status: **conflicting**

Frequency statistics collection time: 2001-10-04 13:01:28.076475

Frequency statistics status: conflicting

Histogram statistics collection time: null

Histogram statistics status: **missing**

Possibly point skewed: No

Possibly range skewed: No

O_ORDERPRIORITY

Cardinality: 5.0

Uniform statistics collection time: 0001-01-01 00:00:00.0

Uniform statistics status: OK

Frequency statistics collection time: null

Frequency statistics status: missing

Histogram statistics collection time: null

Histogram statistics status: missing

Possibly point skewed: Yes

Possibly range skewed: No

Indicates conflicting and missing statistics

Conflicting statistics explanation

Conflicts detail

TABLE SYSADM.LINEITEM

One of the frequency records (-1.0) of the L_ORDERKEY column group is out of range [0,1]
Tolerance: 0.0010

The maximum frequency of the column group or column (L_ORDERKEY), (0,0), is less than the average frequency, or 1 divided by the greater than the average unless only least-frequently occurring values are being collected.
Tolerance: 0.0010

Improve Query Design

Workload Query Advisor Recommendations Summary

The following is a summary of the queries analyzed in the workload. Use this criteria to filter the view for specific statements.

Statements Sorted by	Number
Statements Analyzed Successfully	22
Statements with Warnings	4
Number of High Severity Warnings	0
Number of Medium Severity Warnings	0
Number of Low Severity Warnings	7
Statements with High Severity Warnings	0
Statements with Medium Severity Warnings	0
Statements with Low Severity Warnings	0

View statements that meet the following:

Degree of warning severity: High severity
 Medium severity
 Low severity
 Show statements that do not contain warnings

Restore Defaults Save as Defaults

Statements | Index Advisor | Query Advisor Summary | Advisors

View analysis summary

Filter recommendations by severity

Guard against errors and oversights:
Further constrain query, increase index utilization, and reduce data reads

Query Recommendation Detail

SQL Text

```
SELECT A.EMPNO
, A.FIRSTNAME
, A.LASTNAME
, A.JOB
, A.SALARY
, A.BONUS
, A.COMM
, B.LOCATION
, C.PROJNAME
FROM DSN8910.DEPT AS B
, DSN8910.EMP AS A
, DSN8910.EPROJ AS C
WHERE ( A.EMPNO IN ( SELECT DSN8910.DEPT.MGRNO
FROM DSN8910.DEPT
WHERE DSN8910.DEPT.MGRNO IS NOT NULL
)
```

Selected Recommendation:

Description
Consider adding join predicates between columns EMPNO in table DSN8910.EMP and columns MGRNO in table DSN8910.DEPT which use the referential constraints between table DSN8910.EMP and table DSN8910.DEPT to avoid a potentially costly Cartesian join. Check the explanation for this warning for more details about possible impact and examples.
If a referential constraint is defined between corresponding join predicates that map exact RI1 is defined between table T1 and T2, in which an SQL statement joins the two tables as follows: SELECT T1.C2, T2.C2 FROM T1, T2 WHERE T1.C2 = :charHV3 AND T2.C2 = :charHV4 As written, the SQL statement does not include the join predicate between the two tables. The result is likely to be incorrect. Rewriting the SQL statement as follows: SELECT T1.C2, T2.C2 FROM T1, T2 WHERE T1.C2 = :charHV3

Highlights relevant components of the query

Recommendation and rationale

Query Advisor checks for

- Missing join predicate for referential constraint
- Predicates that can be rewritten as indexable
- Stage 2 predicates that can be rewritten as stage 1 predicates

Indexing Advice to Improve Query Efficiency

→ Improve query efficiency

- Indexing foreign keys in queries that do not have indexes defined
- Identifying index filtering and screening
- Support for index only access
- Indexing to avoid sorts

→ Simplify use

- Consolidate indexes and provide a single recommendation
- Enables what-if analysis
- Provides DDL to create indexes
- Run immediately or save

The screenshot displays the IBM Optim Query Tuner Client interface, specifically the Workload Tuning Editor. The main window shows a table of recommendations for index creation. A 'DDL Details' dialog box is open, showing the SQL DDL for creating two indexes on the SYSADM.LINEITEM table.

Feature Details	Action	Object Name	Columns
[-] PART			
<input checked="" type="checkbox"/> Index	Create	PART_VIR...	P_PARTKEY(ASC), P_NAM...
[-] PARTSUPP			
<input checked="" type="checkbox"/> Index	Create	PARTSUP...	PS_SUPPKEY(ASC)
<input checked="" type="checkbox"/> Index	Create	PARTSUP...	PS_SUPPLYCOST(ASC), PS...
[-] LINEITEM			
<input checked="" type="checkbox"/> Index	Create	LINEITEM...	L_SHIPMODE(ASC), L_OR...
<input checked="" type="checkbox"/> Index	Create	LINEITEM...	L_QUANTITY(ASC), L_EXT...
<input checked="" type="checkbox"/> Index	Create	LINEITEM...	L_RETURNFLAG(ASC), L_L...
<input checked="" type="checkbox"/> Index	Create	LINEITEM...	L_SUPPKEY(ASC), L_EXT...
[-] ORDER			
<input checked="" type="checkbox"/> Index	Create	ORDER_V...	O_CUSTKEY(ASC), O_ORD...
<input checked="" type="checkbox"/> Index	Create	ORDER_V...	O_ORDERKEY(ASC), O_O...

```
CREATE INDEX
DB20E.LINEITEM_VIRT_IDX_1159535554645 ON
SYSADM.LINEITEM (L_SHIPMODE ASC, L_ORDERKEY
ASC) FREEPAGE 0 PCTFREE 10;

CREATE INDEX
DB20E.LINEITEM_VIRT_IDX_1159535563807 ON
SYSADM.LINEITEM (L_QUANTITY ASC,
L_EXTENDEDPRICE ASC, L_TAX ASC) FREEPAGE 0
```

Workload Tuning vs. Query Tuning

→ Optimization decisions are based on tradeoffs

- Statistics – CPU costs vs. query savings
- Indexing – query speed vs resource and transaction

→ Workload tuning speeds up analysis

- Analyzes multiple queries at once

→ Workload tuning consolidates and optimizes recommendation for overall workload

- Statistics recommendations
- Index recommendations

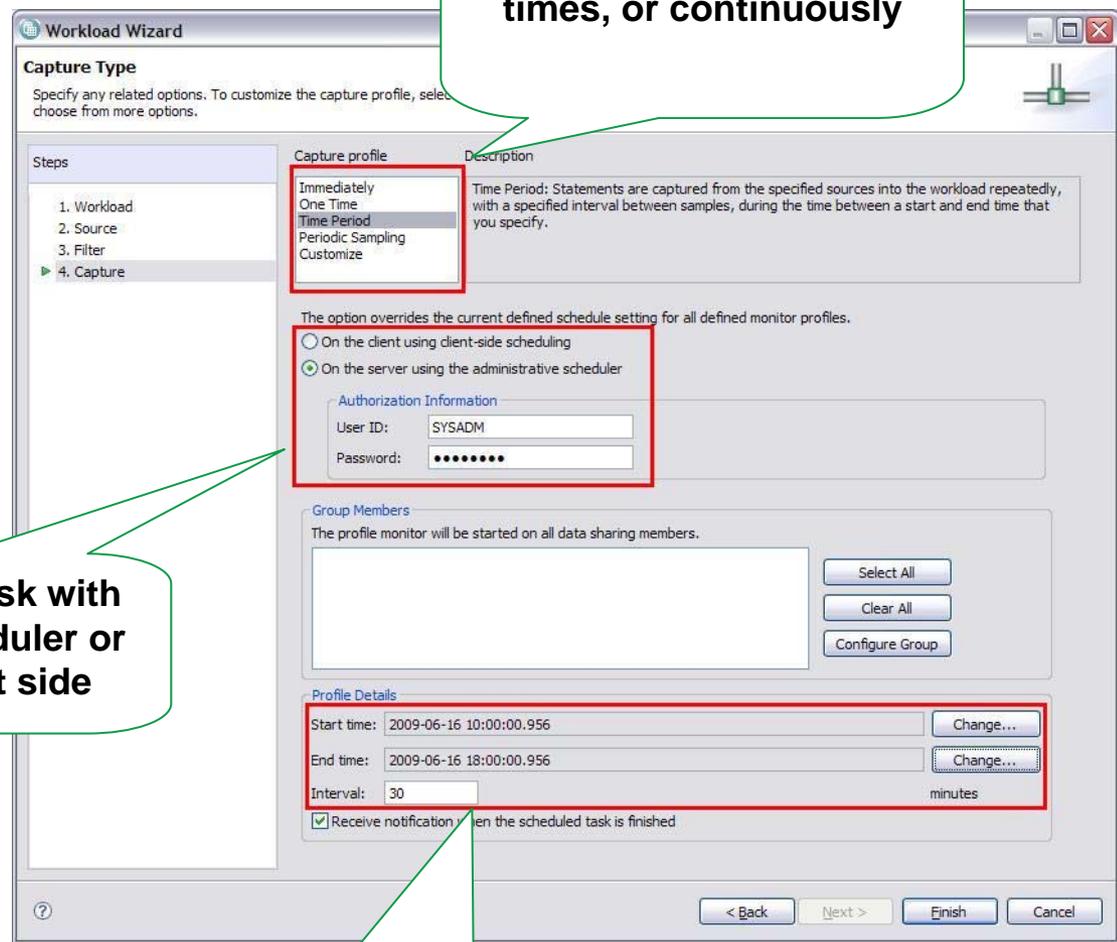
Speed up analysis, optimize design, and balance resource usage

Scheduling

- Easy task scheduling
 - Capture of workloads
 - Profile monitor activity
 - Explain of workloads
 - Advisors for workloads

Schedule the task to be executed once, multiple times, or continuously

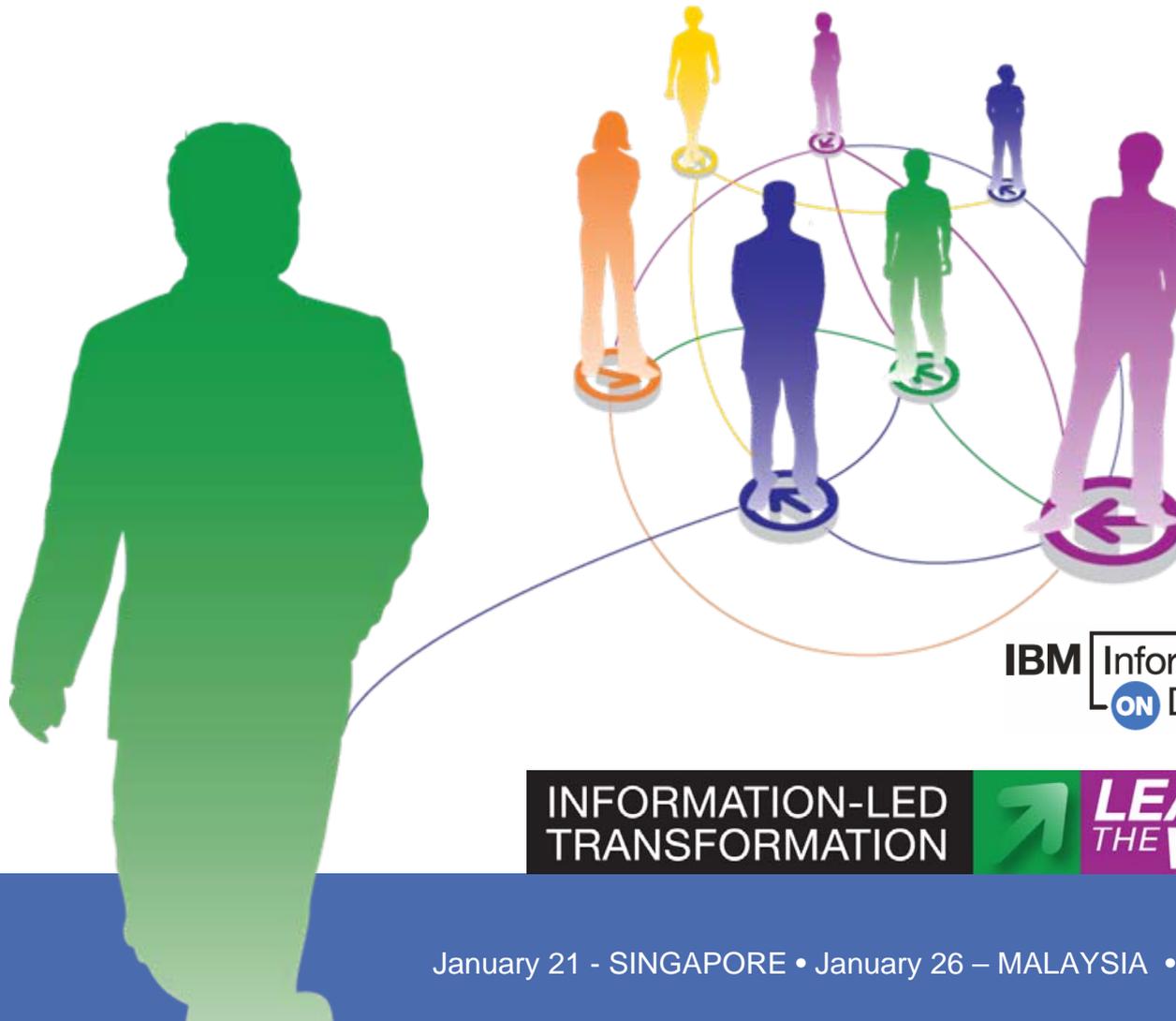
Schedule the task with the admin scheduler or from the client side



Specify the start time, end time, and interval

IBM Optim Solutions

What Customer are Saying



IBM Information
ON Demand 2010

INFORMATION-LED
TRANSFORMATION



January 21 - SINGAPORE • January 26 – MALAYSIA • January 28 - THAILAND

“Optim has already yielded the returns we were looking for, and we have just touched the tip of the iceberg...”



“We can easily comply with data privacy regulations without incurring additional costs...”



“The bottom line is that Optim's archiving capabilities will help us exceed our SLA...”



...And Many Other Happy Optim Customers

CORNING

TOSHIBA

BOEING

HONDA
The Power of Dreams

UCSF

DIRECTV

Rabobank

Advance
Auto Parts

Countrywide
HOME LOANS

Safeco Insurance

NAUTILUS Inc

newgistics.

CHIVAS REGAL
12

TD
Bank
Financial
Group

BlueCross
BlueShield
Association

AEP
AMERICAN
ELECTRIC
POWER

BOSTONIA
CONSTITUTA AD
1630
CIVITATIS REIPUBLICANAE DORATA ALIBI
City of Boston

ART IN MOTION

ROGERS
Your World Right Now

Limited brands

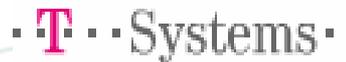
EOS

Puerto Rico
Federal Affairs
Administration

EAGLE
Family Foods, Inc.

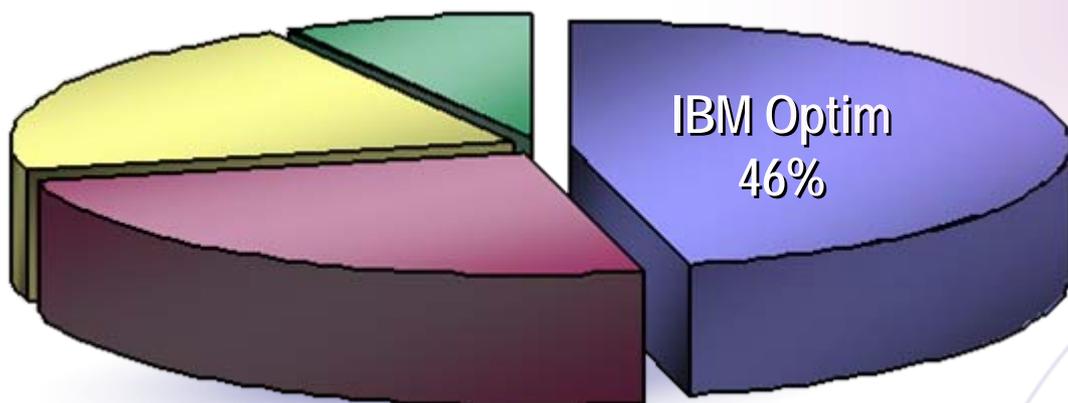
Bausch & Lomb

...And More



Customers Have Made Optim The Market Leader

“[Optim] customers spoke highly of the reliability and high performance of its archiving solution, which sets it apart from other vendors in this space.”



Gartner Research

-- Forrester Research

Sources:

Gartner, "Archiving Software Market to Experience Strong Growth Through 2010" 2006;

Forrester Research, "Database Archiving Remains An Important Part Of Enterprise DBMS Strategy," 2007

The broadest range of capabilities for managing the value of your data throughout its lifetime

