Integrated Data Management – Data Studio

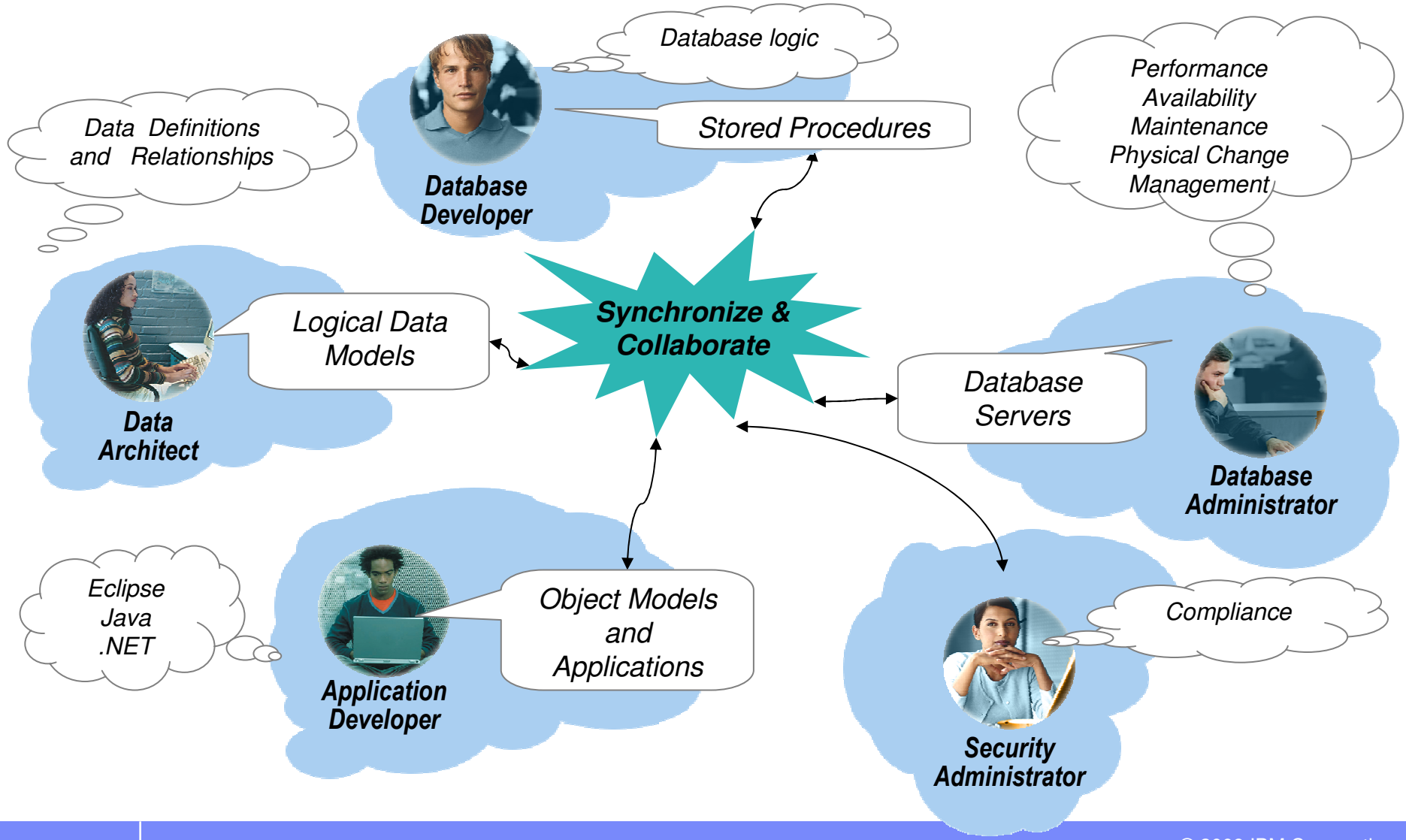
Alfred Horng

IBM SWG

kfhorng@tw.ibm.com

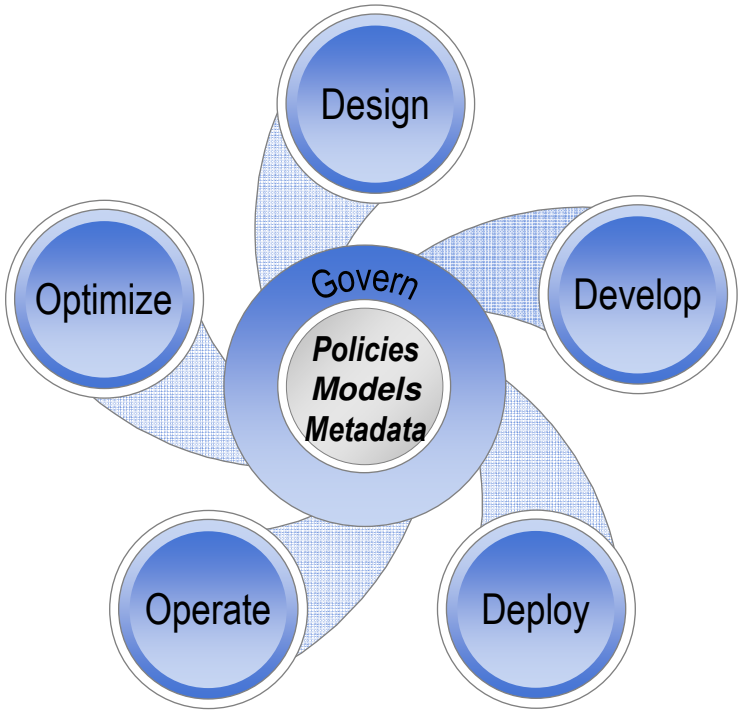
© 2009 IBM Corporation

Integrated Data Management



© 2009 IBM Corporation

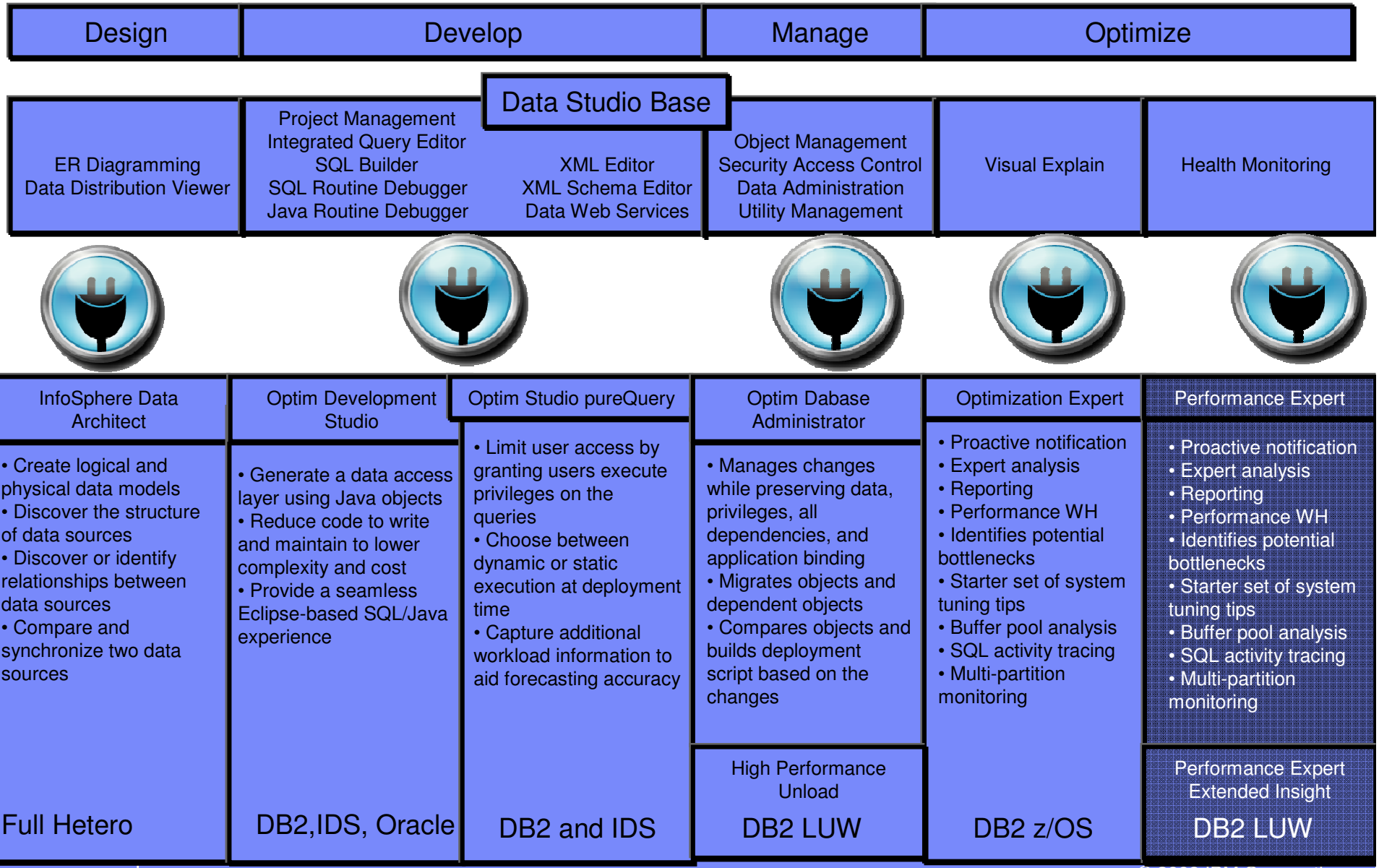
Integrated Data Management – What's Different?



- Support business growth
  - Accommodate new initiatives without expanding infrastructure
  - Simplify application upgrades, consolidations & retirements
- 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
- Facilitate alignment, consistency & governance
  - Define business policies and standards up front; share, extend, and apply throughout the lifecycle

© 2009 IBM Corporation

Data Studio Packaging



© 2009 IBM Corporation

Core Users of the IBM Data Studio

© 2009 IBM Corporation

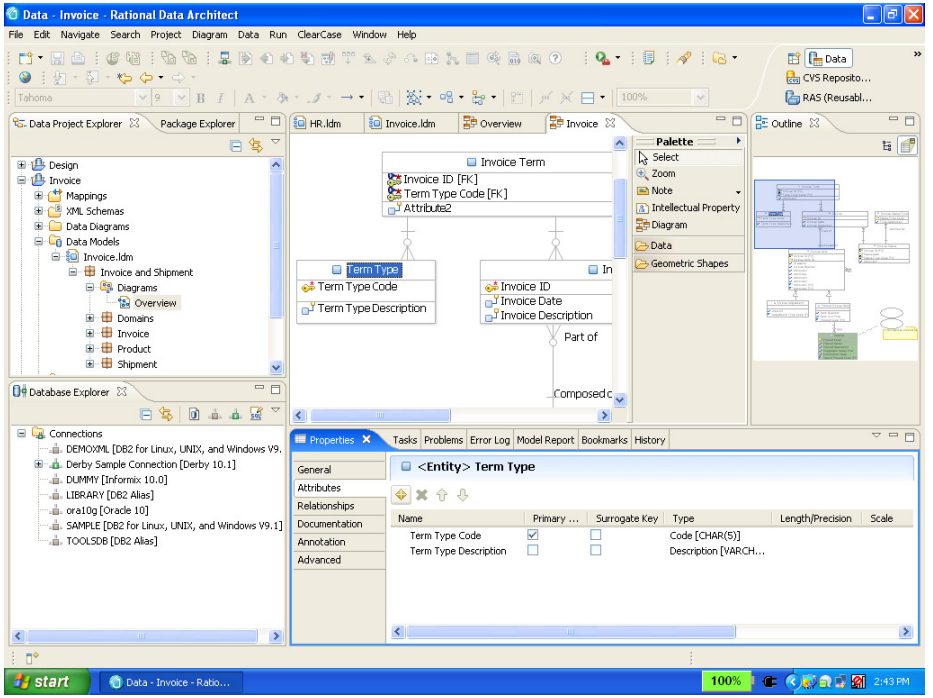
Users of the IBM Data Studio

© 2009 IBM Corporation

InfoSphere Data Architect

InfoSphere Data Architect is a collaborative, data design solution to discover, model, relate, and standardize diverse data assets.

- Increase Data Quality and Integrity
  - Analyze and enforce compliance to enterprise standards and data privatization
  - Support business and IT collaboration via a common business glossary
  - Use with IBM Industry Models for industry-specific best practices
  - Facilitate model-driven development via seamless integration with Rational Software Delivery Platform
  - Automate transformations between the application model and the data model



© 2009 IBM Corporation

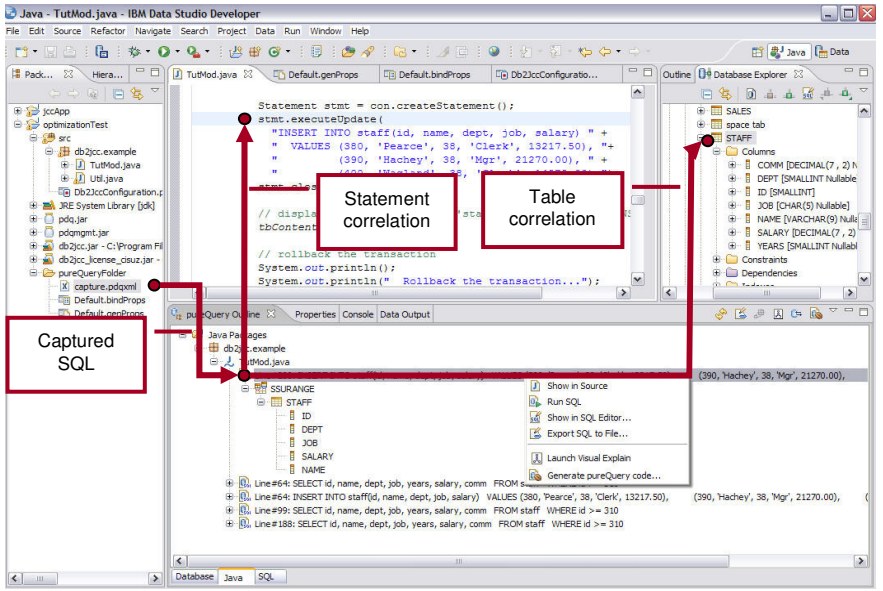
Users of the IBM Data Studio

© 2009 IBM Corporation

IBM Optim Development Studio and pureQuery Runtime

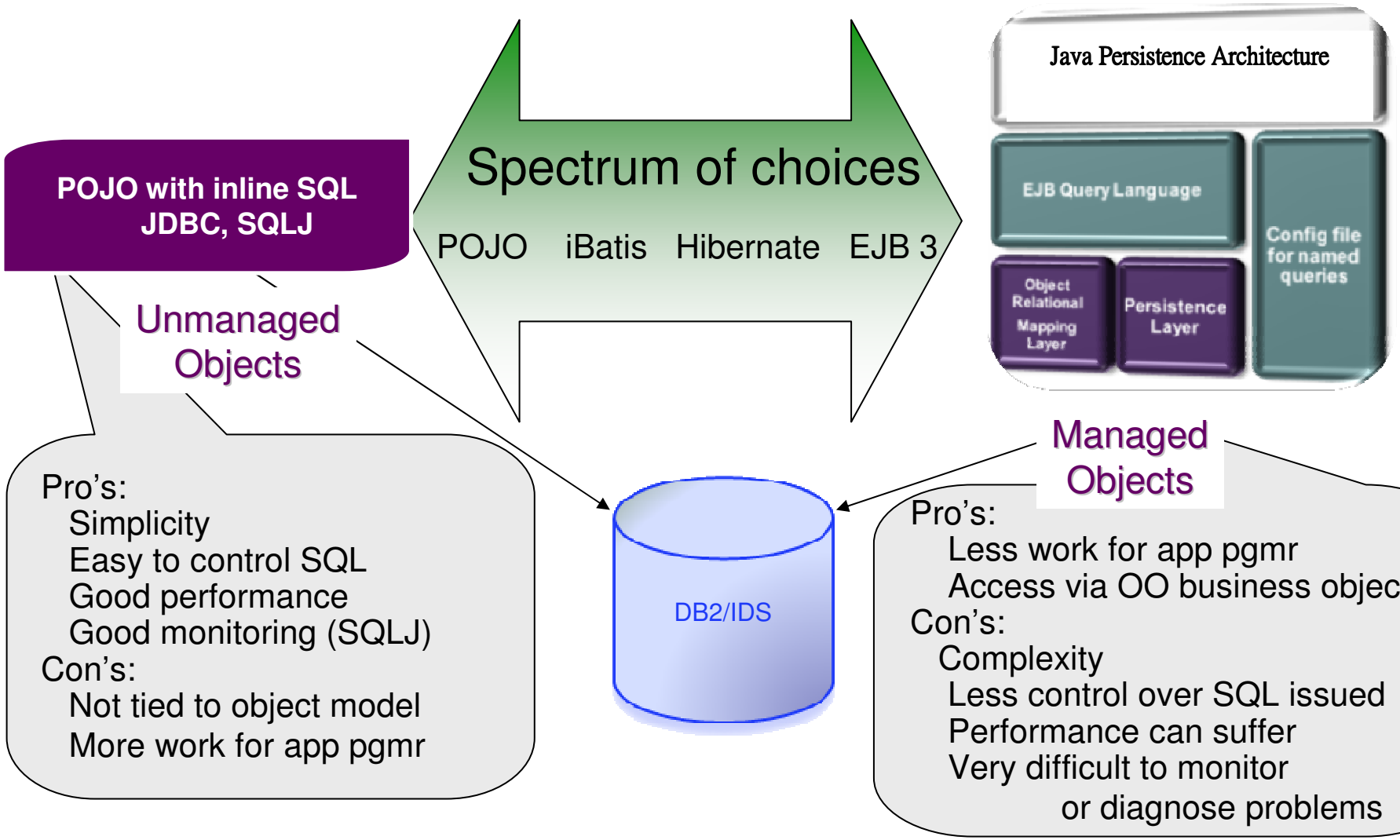
IBM Optim Development Studio is an integrated database development environment that speeds application design, development, and deployment while increasing data access performance and manageability.

- **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**



© 2009 IBM Corporation

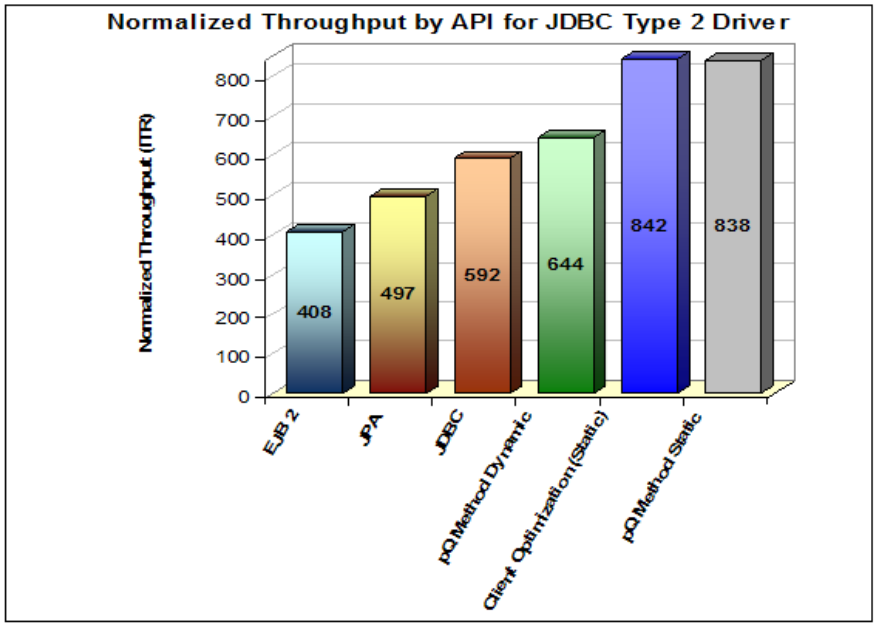
Optim Studio pureQuery – Best of Breed

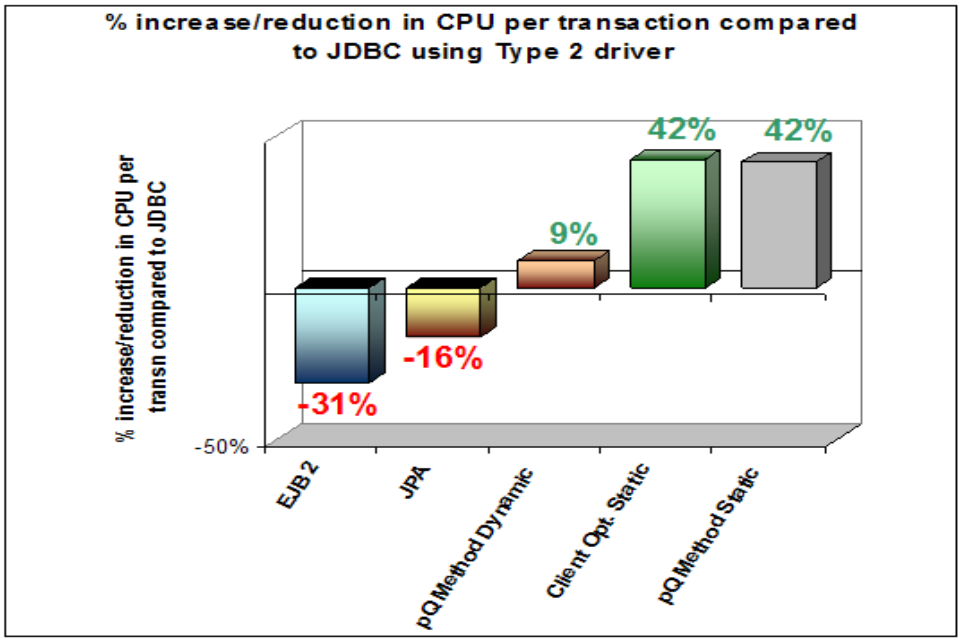


© 2009 IBM Corporation

Data Studio pureQuery Runtime for z/OS

- **In-house testing shows double-digit reduction in CPU costs over dynamic JDBC**



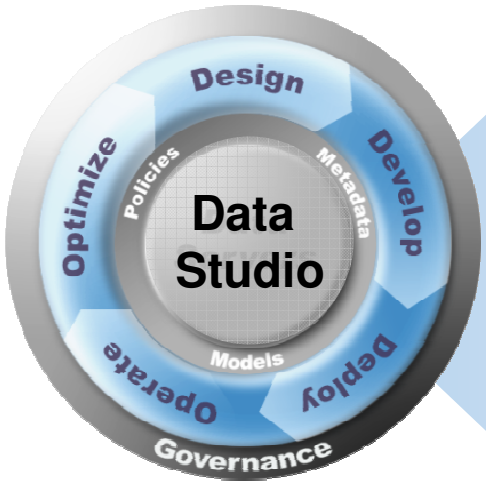


- **IRWW – an OLTP workload,** Type 2 driver (local call)
- **Cache hit ratio between 70 and 85%**
- 42% reduction in CPU per transaction **over dynamic JDBC**

© 2009 IBM Corporation

What is pureQuery

pureQuery is a high-performance, data access platform to simplify developing, managing, securing, and optimizing Java data access for new and existing applications.



Design and Develop

pureQuery Components:

- Development tools

  Integrated development environment with Java and SQL support delivered with IBM Data Studio Developer

- Simple and intuitive API

  Enables SQL access to databases or in-memory Java objects

- pureQuery Runtime

  Flexible deployment options with static SQL support with IBM Data Studio pureQuery Runtime for z/OS or for Linux, UNIX, and Windows

© 2009 IBM Corporation

pureQuery for IDS & DB2

- **For existing and new JDBC Applications**
  - **Capture**
    - See all SQL and Java for every database table & column
    - Find and Fix problems quickly
  - **SQL Injection Prevention and Rogue Query Control**
    - Lock down dynamic SQL to only approved SQL
  - **Tuned Query Replacement**
    - Replace SQL with optimized SQL without changing application
  - **Static SQL**
    - Lock access path and get performance and reliability of Static SQL against DB2
- **For new applications**
  - **Java Best Practices**
    - queryFirst, updateMany
    - Heterogenous Batch

© 2009 IBM Corporation

Retrieve a single row from Database

pureQuery API's:

```
Employee my_emp = db.queryFirst("SELECT Name, HomeAddress, HomePhone
FROM Employee WHERE Name=?", Employee.class, my_emp);
-or-
Employee my_emp = getEmployee(name);
```

XML file or Java annotation
SELECT * FROM EMPLOYEE
WHERE NAME=?1;

SQLJ:

```
#sql [con] { SELECT NAME, ADDRESS, PHONE_NUM INTO :name, :addr, :phone FROM EMP
        WHERE NAME=:name };
new Employee my_emp;
my_emp.setName(name);
my_emp.setHomeAddress(addr);
my_emp.setHomePhone(phone);
```

JDBC:

```
java.sql.PreparedStatement ps = con.prepareStatement(
    "SELECT NAME, ADDRESS, PHONE_NUM FROM EMP
        WHERE NAME=?");
ps.setString(1, name);
java.sql.ResultSet names = ps.executeQuery();
names.next();
new Employee my_emp;
my_emp.setName(names.getString(1));
my_emp.setHomeAddress(names.getString(2));
my_emp.setHomePhone(names.getString(3));
names.close();
```

Table	Column	Type
EMP	NAME	CHAR(64)
EMP	ADDRESS	CHAR(128)
EMP	PHONE_NUM	CHAR(10)

ad2

```
class Employee
{   public String Name;
    public String HomeAddress;
    public String HomePhone;
  …
}
```

ad2

change the color

Anshul Dawra, 2008/10/23

Users of the IBM Data Studio

© 2009 IBM Corporation

Optim Database Administrator

Future add-ons

Optim

Additional add-ons

Infosphere Data Architect

- Physical Database Model
- Logical Database Model
- Enforce Enterprise Standard for Data quality and consistency
- Integration with Industry Models
- Generate DDL for Various Data Servers

Optim Database Administrator

- Database Change Management
- Extended Alters
- Advanced Data Migration
- Enhanced Scripting
- Advanced Compare Multiple Objects
- Dependency Support
- Model Driven Change

ClearQuest / CVS

High Performance Unload

- Fast Unload
- Extraction from DB2 full backups
- Repartition feature

Data Studio base (no-charge capabilities)

Data Studio Administration Console

- At-a-Glance Health and Availability
- Problem Determination
- Replication Monitor
- Recommendation
- Limited Historical Information

Optim Development Studio

- ER Diagramming
- Integrated Query Editor
- SQL & Java Routine + Debugger
- XML, XML Schema Editors
- Data Web Services
- Object, Data and Security Management
- Visual Explain
- Export / Import Data
- Generate DDL

Optim Database Administrator

- Enhanced Database Object Management
- Instance Management
- Database Management
- Schema Management
- Table Management
- Configuration Management
- Backup / Restore Database
- Run DB2 Commands

© 2009 IBM Corporation

Optim Database Administrator Features

- **Manage and Administer systems**
  - Manage and quickly access database objects
  - Manage database connections
  - Manage DB2 LUW server at instance level
  - Perform preventive maintenance by backing up and restoring databases or table spaces.
  - View / edit configuration parameters
- **Database Change Management**
  - Initiate database object modification from data source explorer
  - External Data Management
  - Auto generate undo
  - Auto generate delta DDL
  - Auto generated maintenance commands
  - Enhance audibility with integrated reporting and documentation of changes and their impact

© 2009 IBM Corporation

Optim Database Administration Features

- **Easily find your database objects**
- **Group over a list of connections**
- **Stop or start your DB2 instances or databases**
- **Backup, restore or recover database or table spaces**
- **Unload and load data into tables**
- **Manage user privileges**

© 2009 IBM Corporation

Scenario

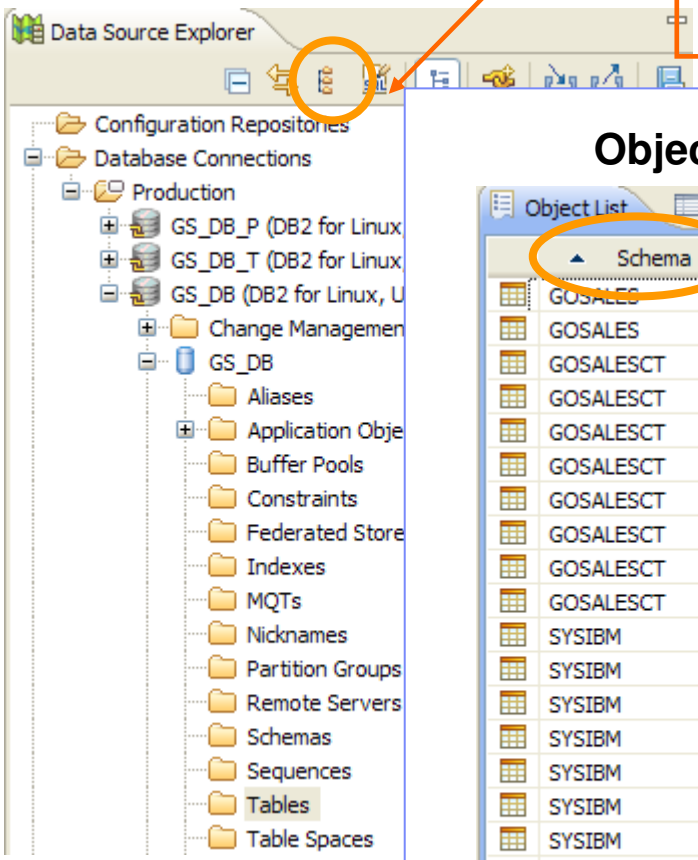
DBA has 100+ schemas and each schema has 20+ tables. He needs to only edit one column in the PROD table.

How is he going to quickly find that table?

© 2009 IBM Corporation

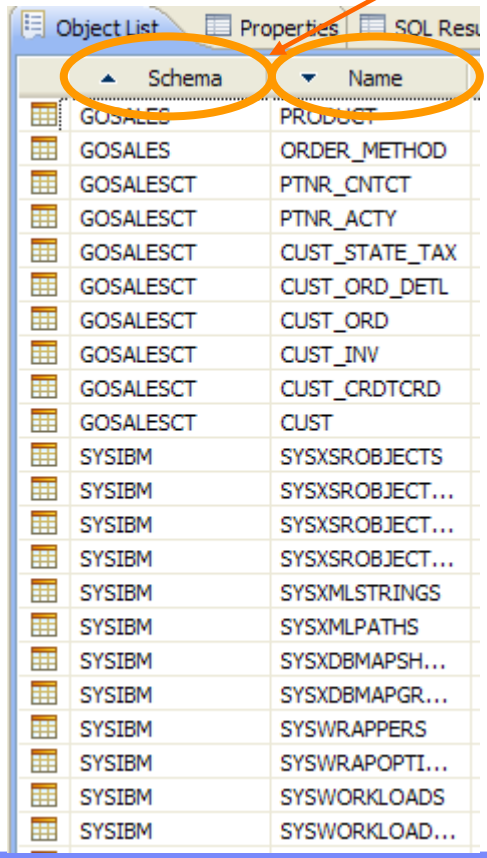
Easily Find Database Objects

Flat presentation



One Click

Object List



Sort by clicking on the columns

© 2009 IBM Corporation

Scenario

DBA normally has to manage 10-15 database…but he current project only requires him to use a subset of the databases

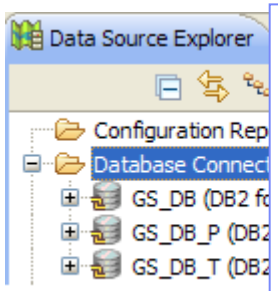
How is he going to quickly organize the databases?

© 2009 IBM Corporation

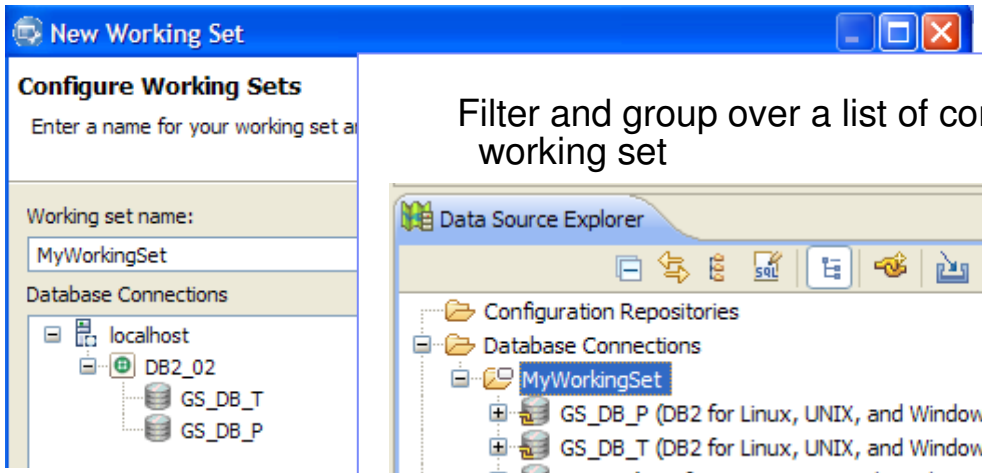
Manage Database Working Sets

Group Database Together

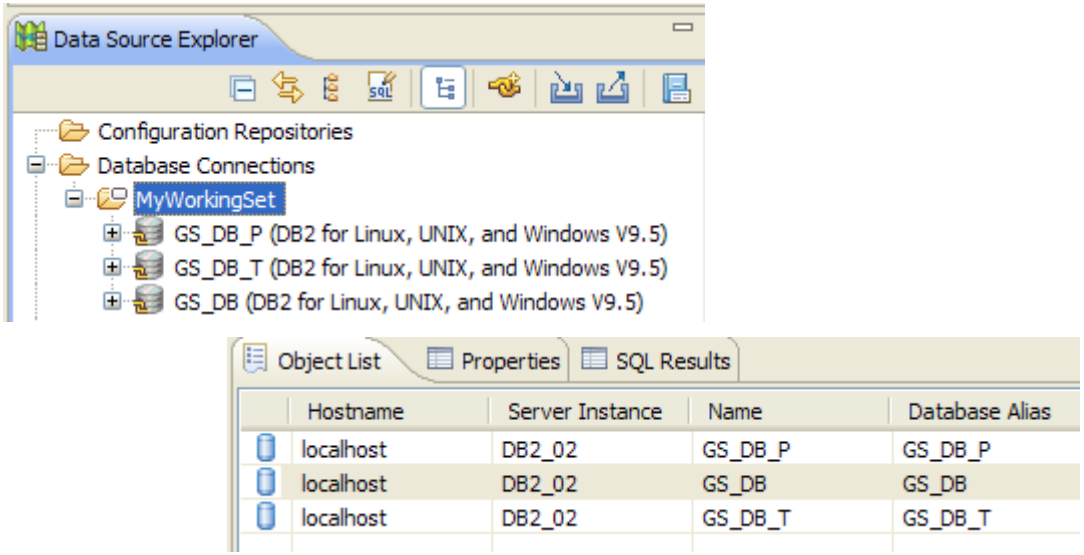
Configure a connection



Identify a set of database connections you will be working with



Filter and group over a list of connections to create a working set



© 2009 IBM Corporation

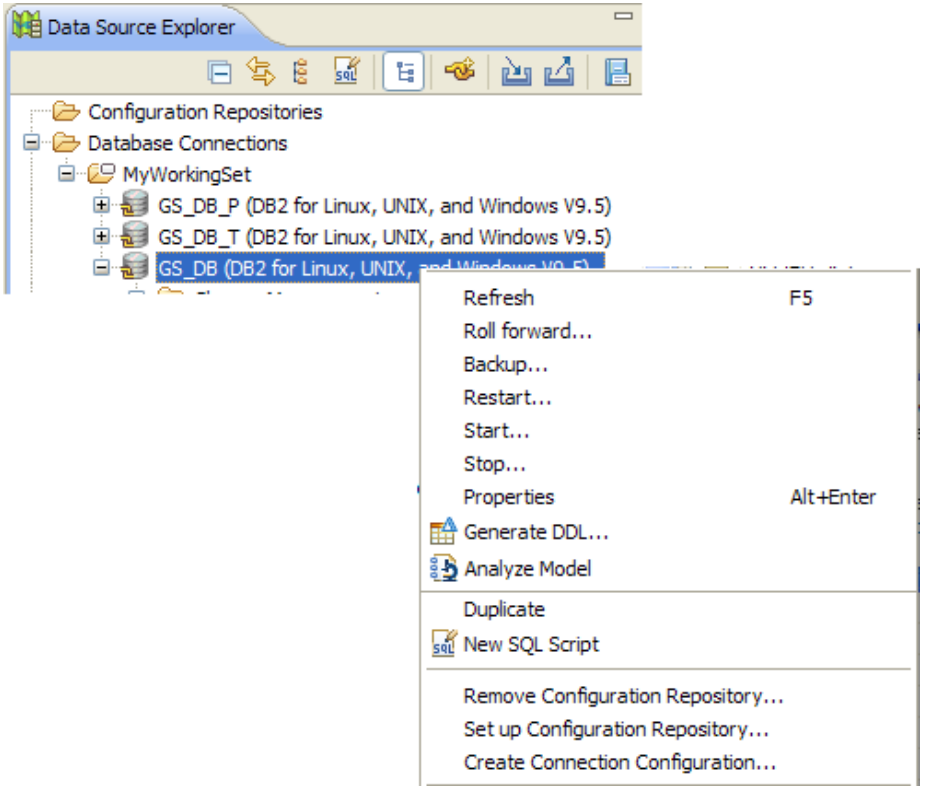
Scenario

What about my other day to day database administration task?

© 2009 IBM Corporation

Simplified Database Administration

- **Using the Data Source Explorer**
  - Stop or start DB2 instances or database
  - Create or drop database
  - Backup, restore or recover database or table spaces
  - Unload and load data into tables
  - Reorganize tables and indexes
  - Manager user privileges



© 2009 IBM Corporation

Optim Database Administrator Features

- **Manage and Administer systems**
  - Manage and quickly access database objects
  - Manage database connections
  - Manage DB2 LUW server at instance level
  - Perform preventive maintenance by backing up and restoring databases or table spaces.
  - View / edit configuration parameters
- **Database Change Management**
  - Initiate database object modification from data source explorer
  - External Data Management
  - Auto generate undo
  - Auto generate delta DDL
  - Auto generated maintenance commands
  - Enhance audibility with integrated reporting and documentation of changes and their impact

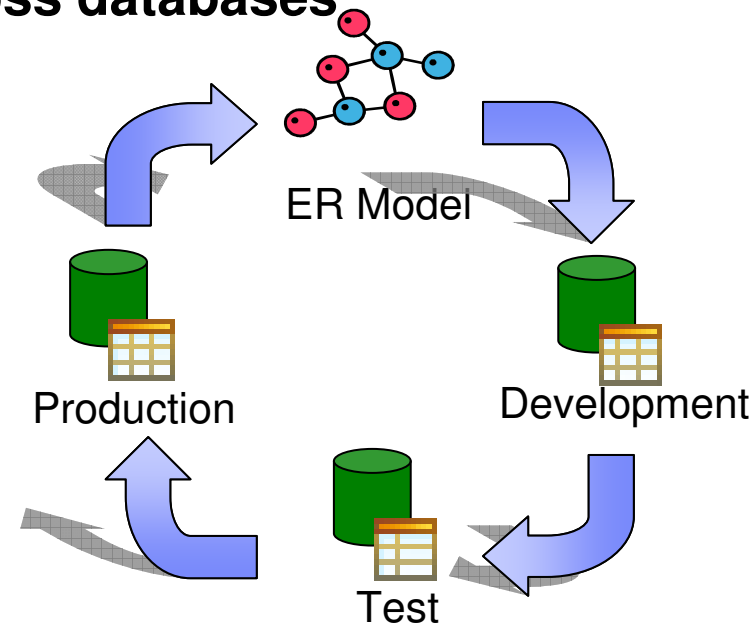
© 2009 IBM Corporation

3 Database Change Scenarios

1. **Change driven by data modeling**

2. **Compare and Synchronize from source to target**

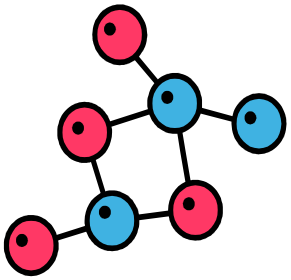
3. **Simple object copy and paste across databases**



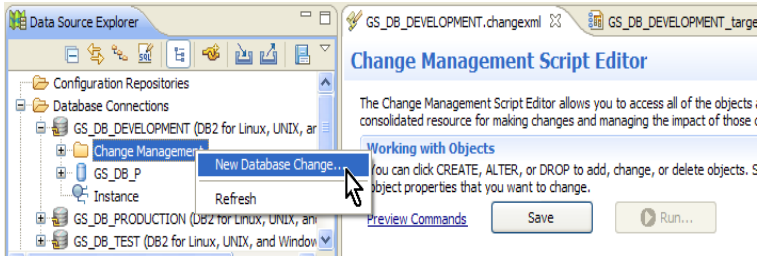
© 2009 IBM Corporation

1 - Change driven by data modeling

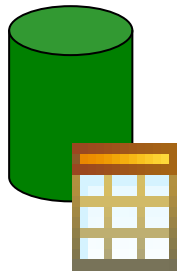
- **Business requirements modeled using Infosphere Data Architect**
- **Logical data model transformed to physical data model**
- **Physical data model drives Data Studio Administrator to deploy changes to target database**



ER Model



Optim Database Administrator

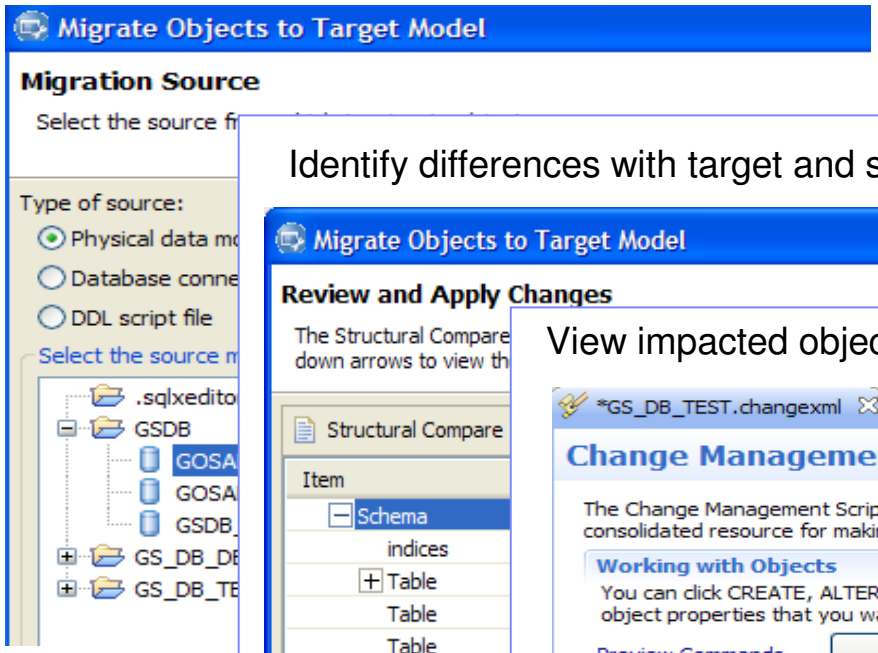


Target Database

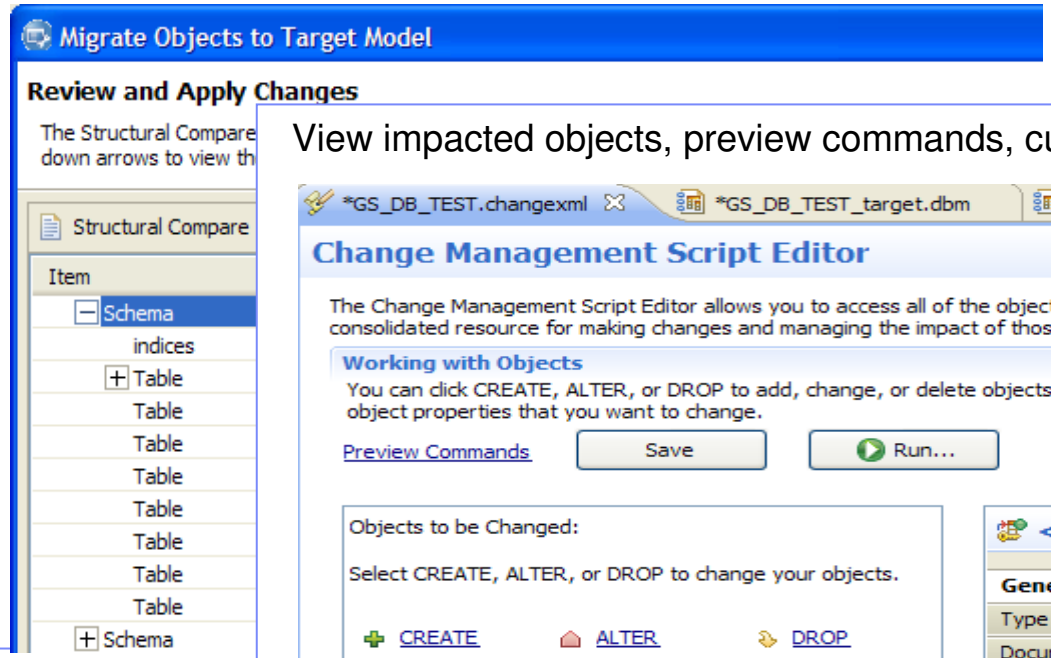
© 2009 IBM Corporation

Change driven by data modeling

Select the physical database model as the source



Identify differences with target and select for deployment

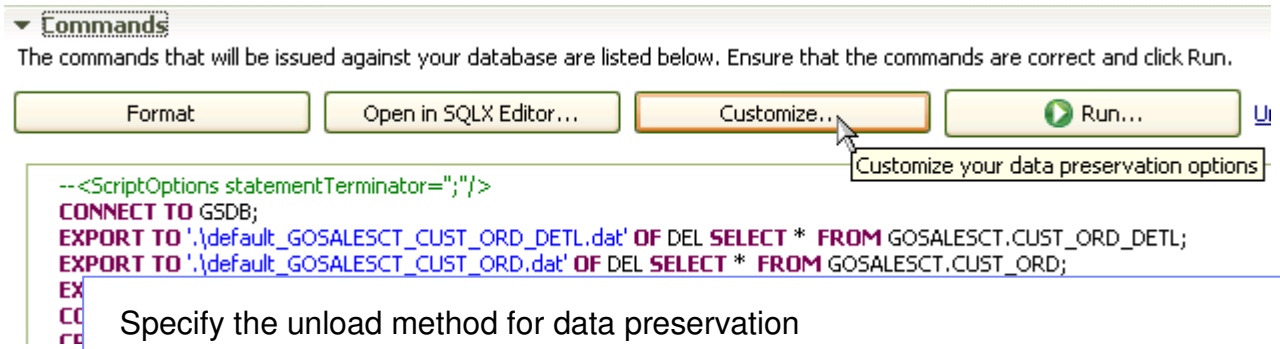


View impacted objects, preview commands, customize commands and run

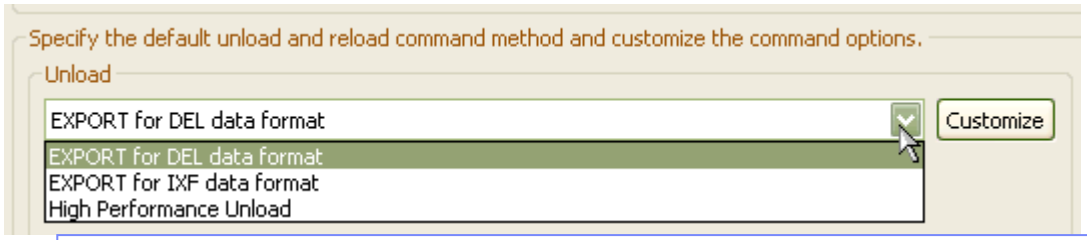
© 2009 IBM Corporation

Data Preservation

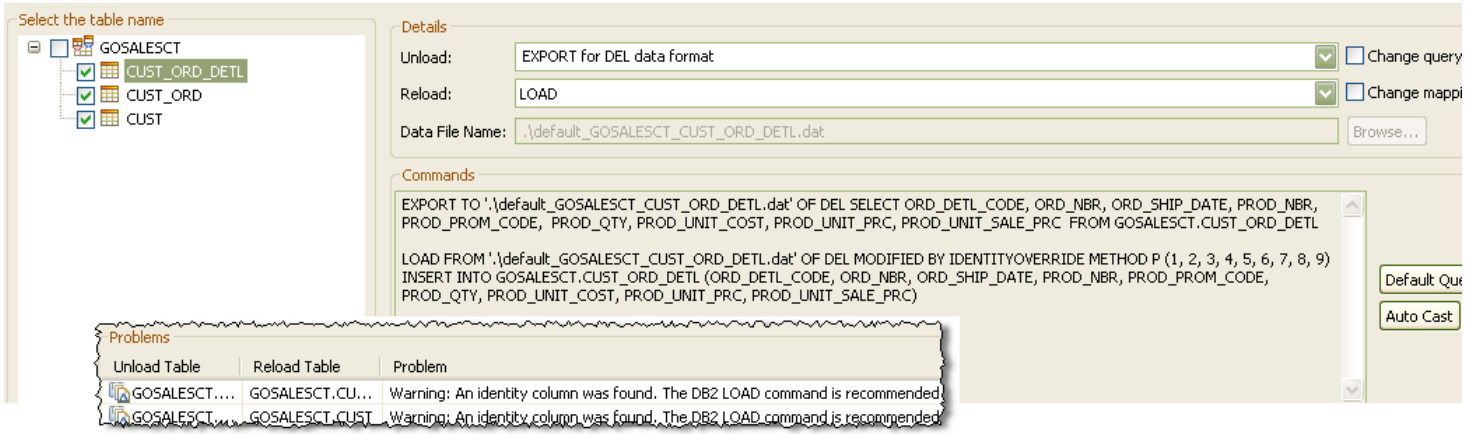
Customize DDL generated by Data Studio Administrator



Specify the unload method for data preservation

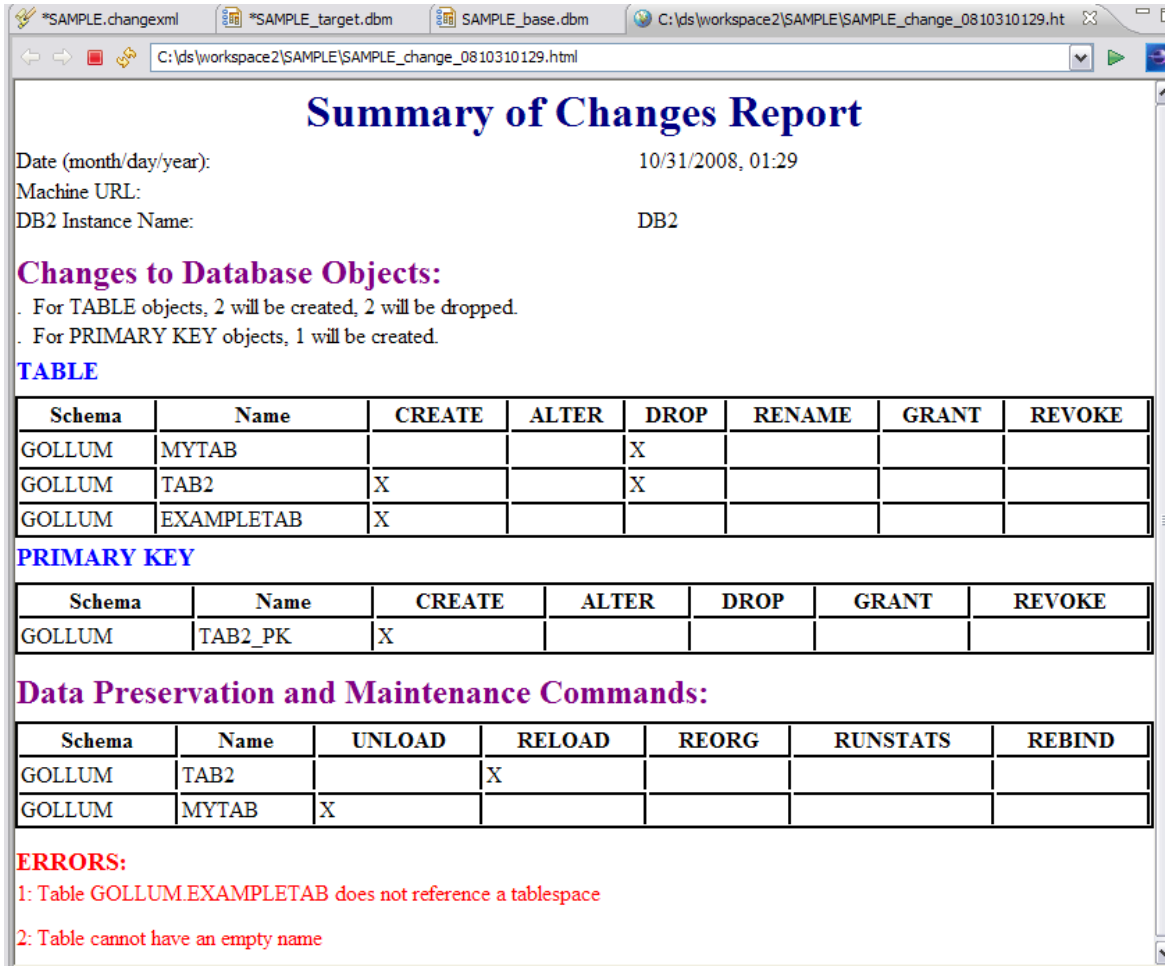


Verify generated commands and resolve errors



Reviews and Change Management

- **Reports**
  - Summary of changes
  - Deployment
  - Database unit test



Summary of Changes Report

Date (month/day/year): 10/31/2008, 01:29
Machine URL:
DB2 Instance Name: DB2

Changes to Database Objects:

. For TABLE objects, 2 will be created, 2 will be dropped.
. For PRIMARY KEY objects, 1 will be created.

TABLE

Schema	Name	CREATE	ALTER	DROP	RENAME	GRANT	REVOKE
GOLLUM	MYTAB			X			
GOLLUM	TAB2	X		X			
GOLLUM	EXAMPLETAB	X					

PRIMARY KEY

Schema	Name	CREATE	ALTER	DROP	GRANT	REVOKE
GOLLUM	TAB2_PK	X				

Data Preservation and Maintenance Commands:

Schema	Name	UNLOAD	RELOAD	REORG	RUNSTATS	REBIND
GOLLUM	TAB2		X			
GOLLUM	MYTAB	X				

ERRORS:

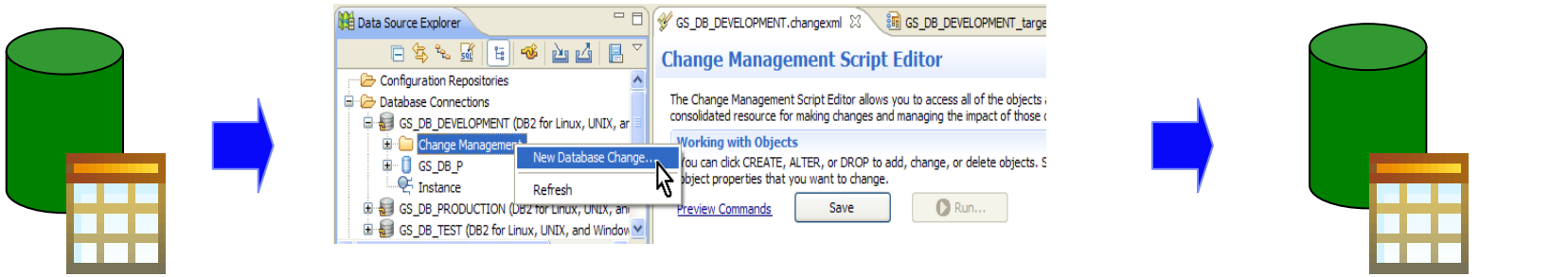
1: Table GOLLUM.EXAMPLETAB does not reference a tablespace

2: Table cannot have an empty name

© 2009 IBM Corporation

2 – Compare and Synchronize

- **Comprehensive solution for migration between database environments**



Source Database

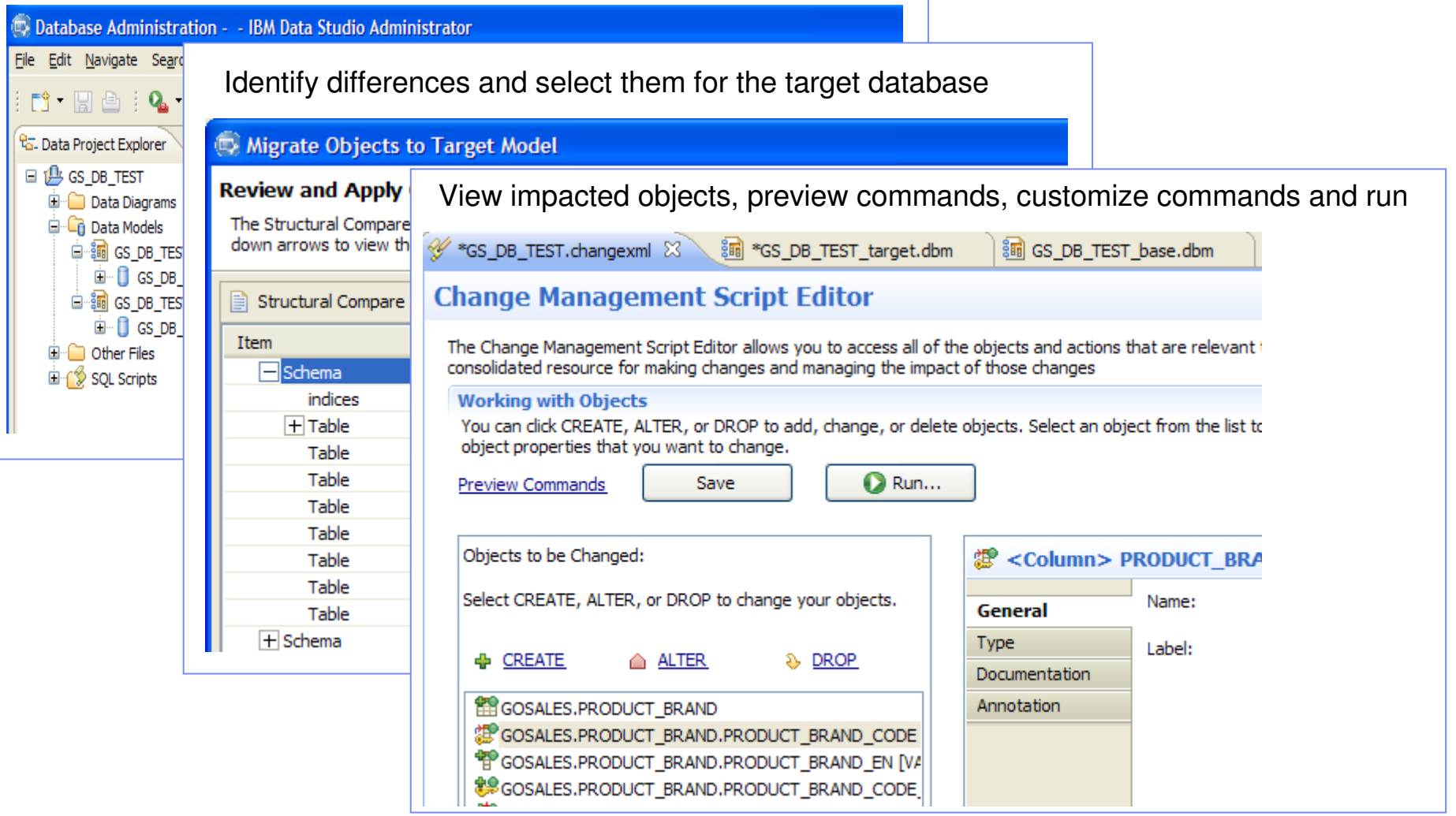
Data Studio Administrator

Target Database

© 2009 IBM Corporation

Compare and Synchronize Databases

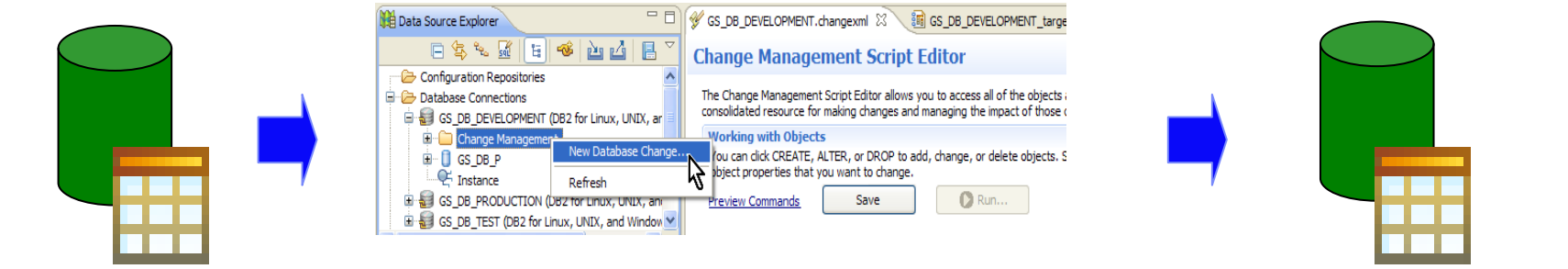
Invoke compare and migrate objects



© 2009 IBM Corporation

3 – Simple object copy and paste

- **Quick, easy and intuitive way to copy and paste or drag and drop objects from system to system**



Source Database

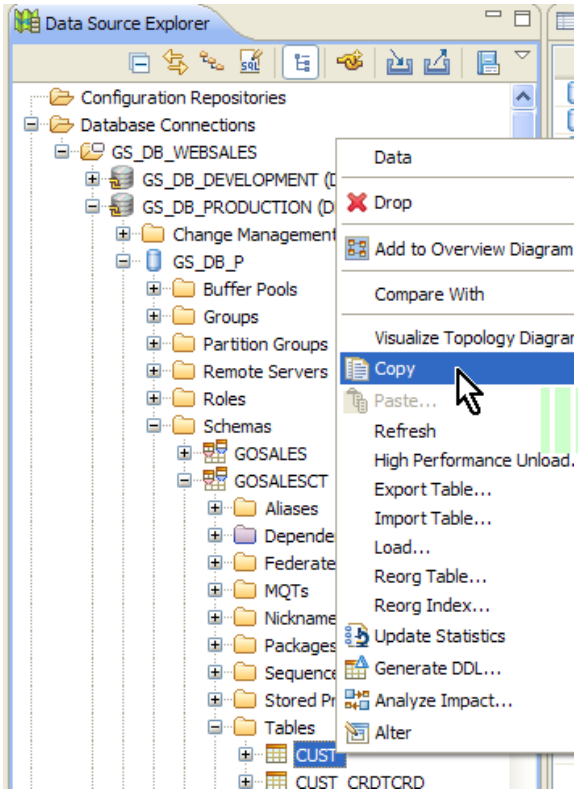
Data Studio Administrator

Target Database

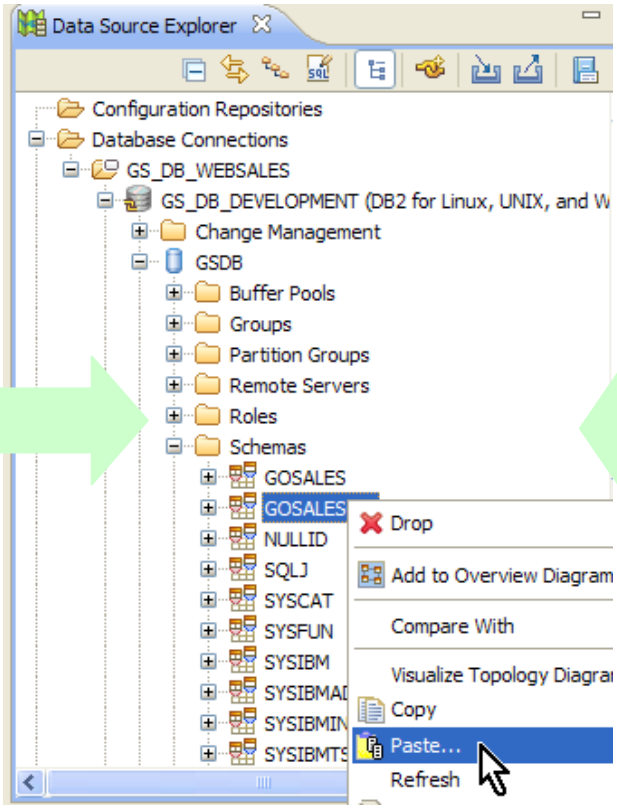
© 2009 IBM Corporation

Simple Copy and Paste

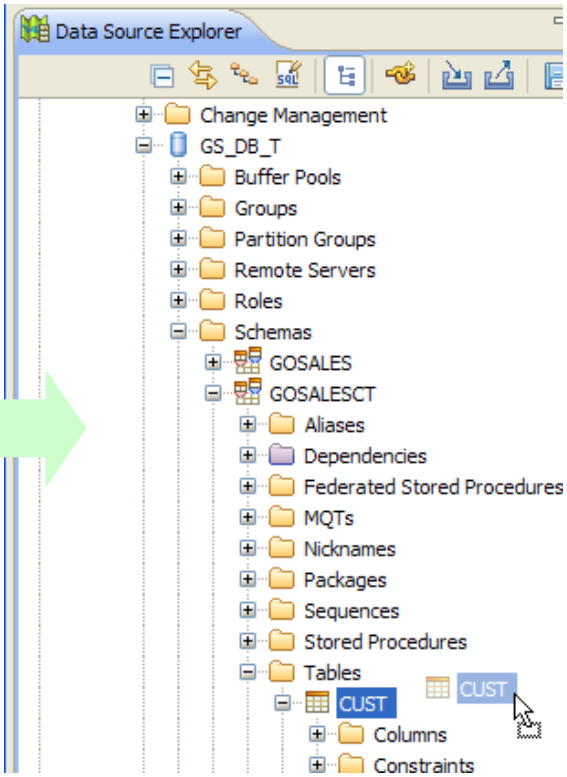
Copy the source object



Paste the source object to the target database

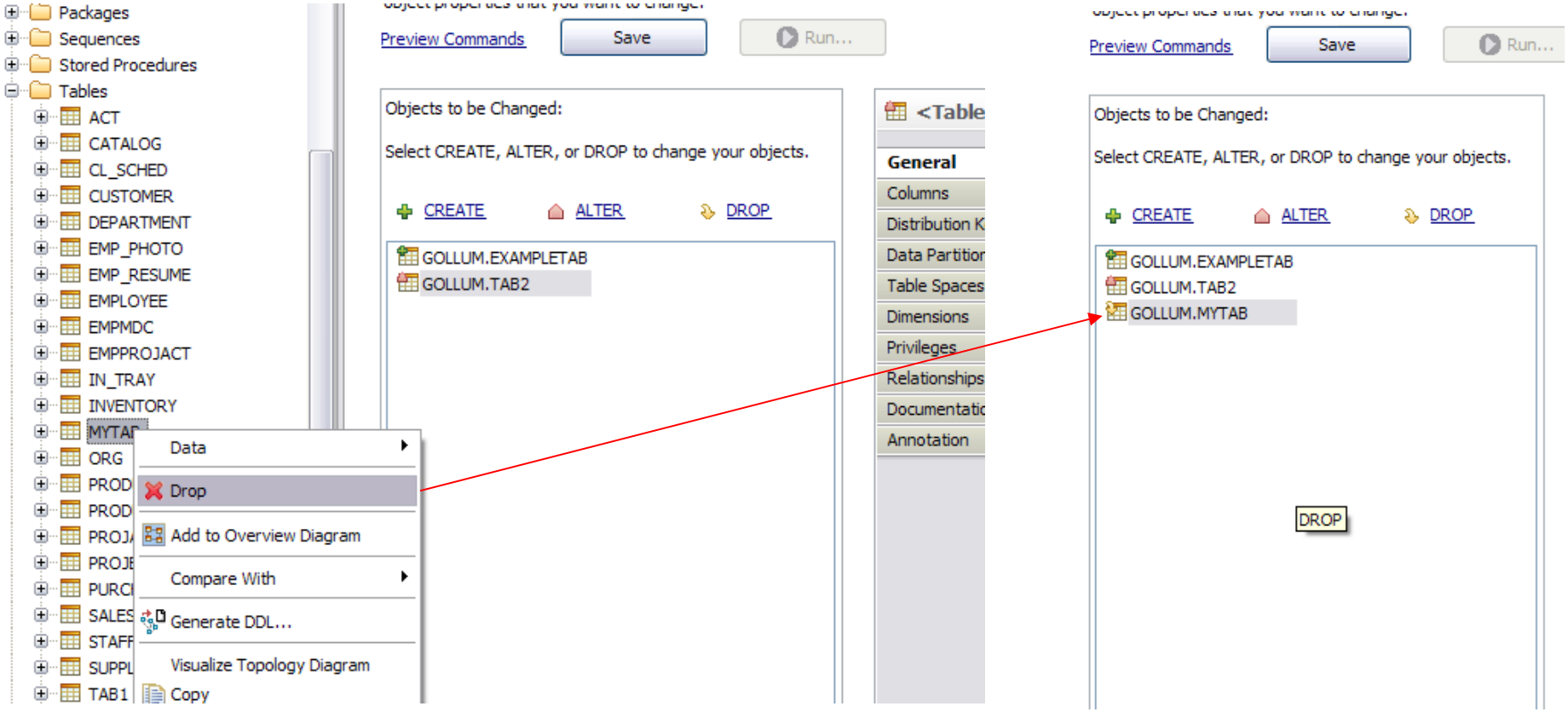


Drag-and-drop objects



© 2009 IBM Corporation

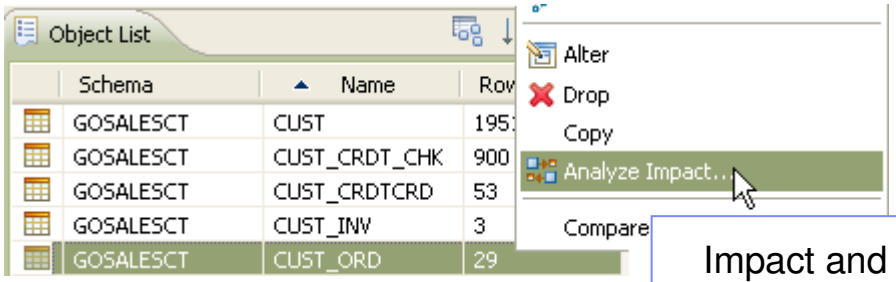
Drop Objects From Data Source Explorer



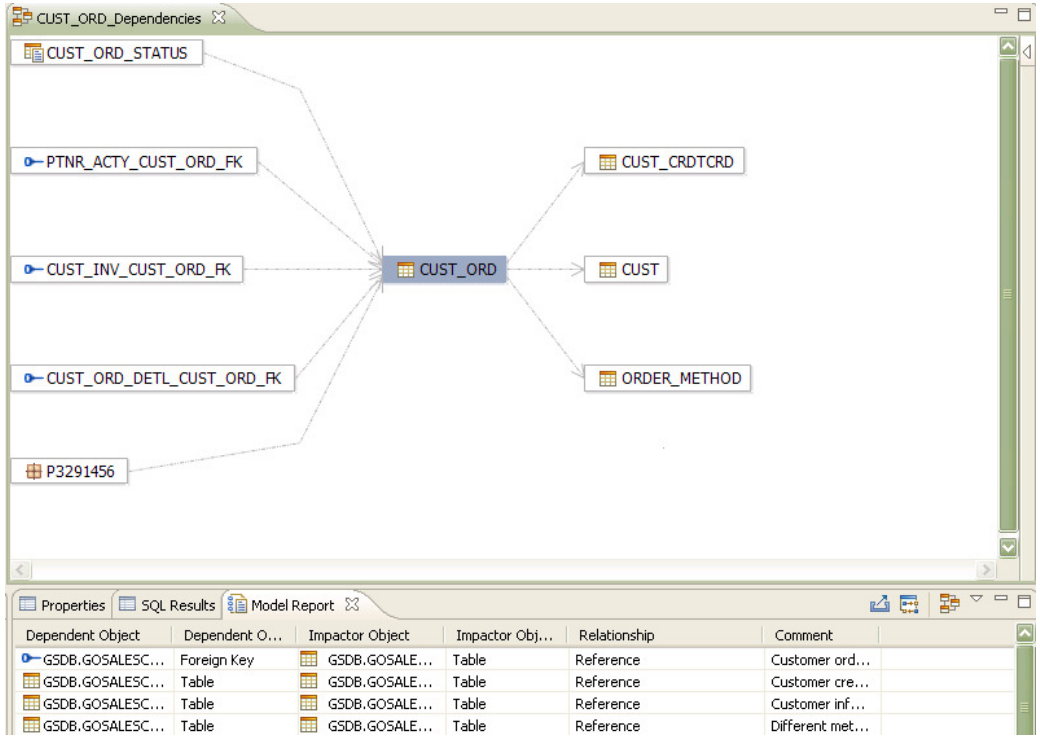
© 2009 IBM Corporation

Impact Analysis

Select the object you want to do impact analysis on

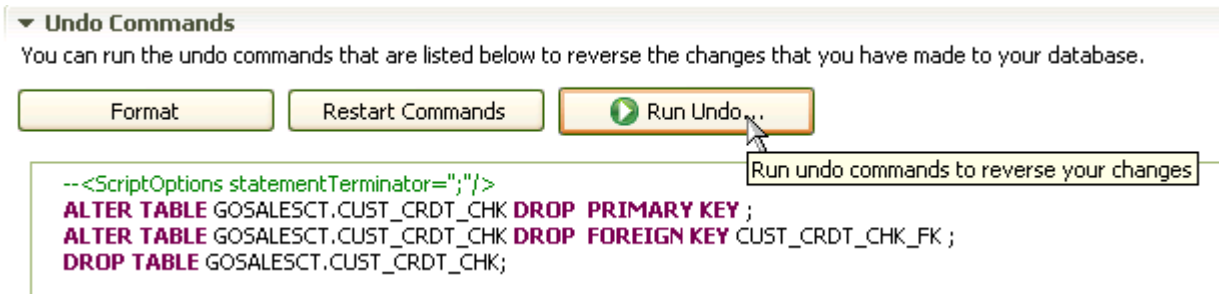


Impact and dependency information can be viewed both in tabular and visual formats

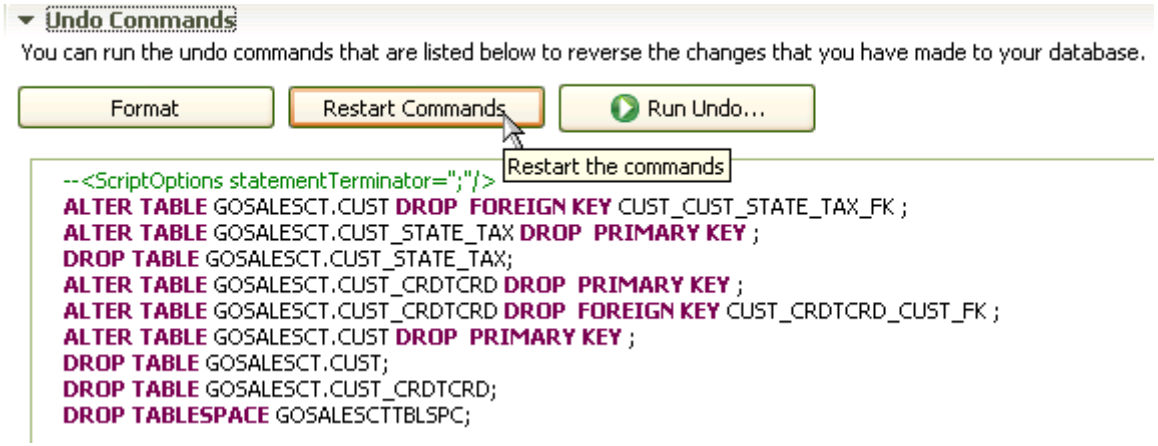


Auto-Generated Undo Commands / Restart From Failure

Undo Commands



Restart From Failure



© 2009 IBM Corporation

IBM Data Studio Administration Console

IBM Data Studio Administration Console provides a health and availability dashboard used to prevent downtime or performance issues before the business is impacted.

- Avoid outages and response time issues
  - Connect to and monitor multiple databases across different platforms from a single console
  - Configure thresholds for warnings and alerts for key performance indicators.
  - Collect and retain alert history for 72 hours
  - Use expert recommendations for problem solving
  - View the system and database parameters and statistics associated with a warning or alert

© 2009 IBM Corporation

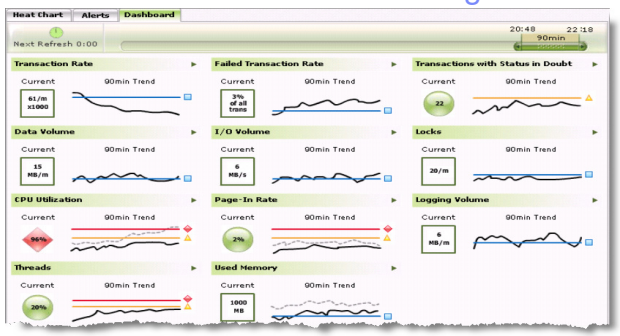
Administration Console
Health Monitoring

Problem Determination

Heatchart – Overall Health Status

Where are the most important hotspots that need my attention?

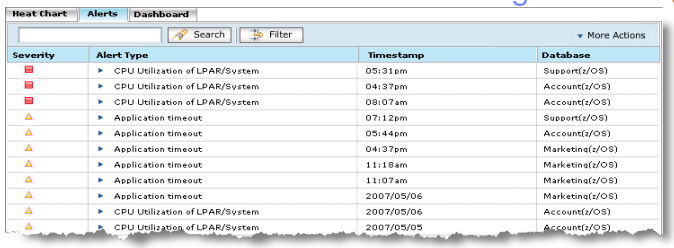
Dashboard – Adhoc Investigation



Something doesn't seem quite right. I wonder what's happening?

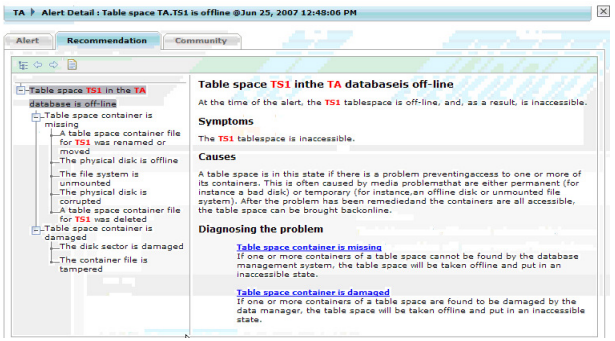
Administrator

Alert List – Historical Investigation



What happened when I was out for lunch? ... Away for weekend?

Recommendations – Root Cause Analysis



Guide me to the root cause and help me fix it properly; I need to know all the revelant info to make the best decision.

© 2009 IBM Corporation

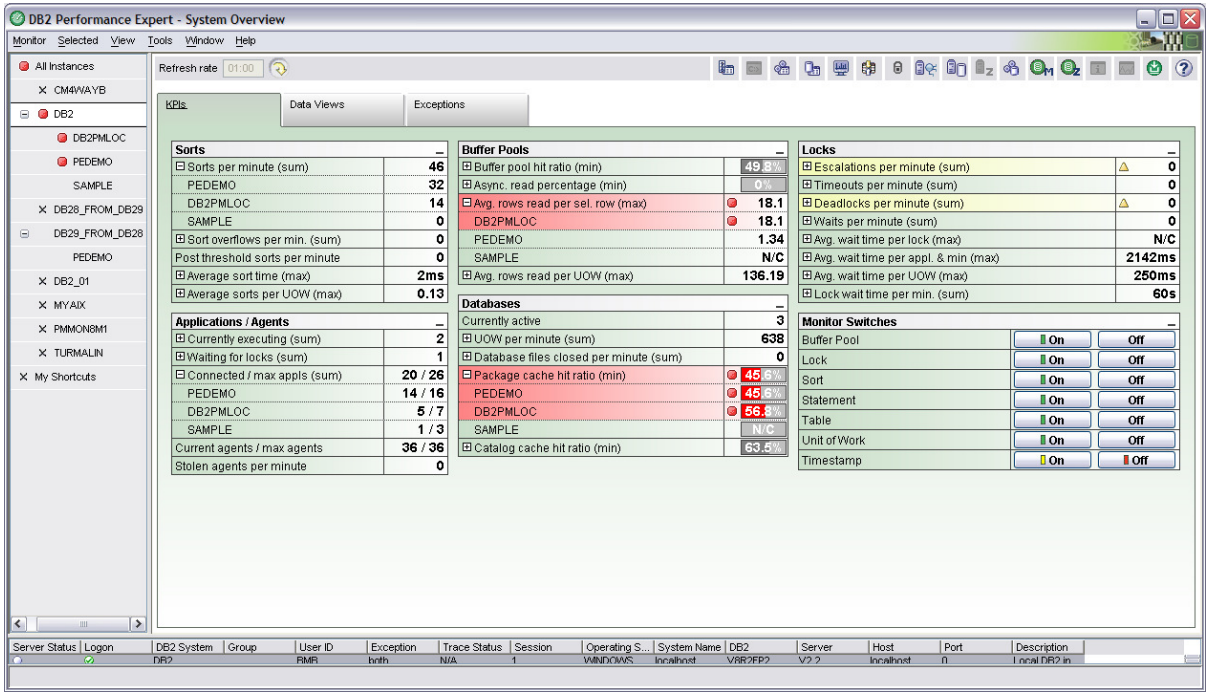
IBM Data Studio Performance Expert

Reduce downtime and quickly resolve performance issues with this comprehensive DB2 monitor.

Provides real-time and history data

DB2 Application monitoring

DB2 Engine monitoring



© 2009 IBM Corporation

Performance Expert Advanced Features

Provides real-time and history data

DB2 Application monitoring

- Monitor resource consumption, SQL Activity, transactions
- Identify locking conflicts, heavy hitter applications and SQL statements
- Force applications

DB2 engine monitoring

- Provide metrics on the instance, database, table spaces, buffer pools, tables
- Provide DB and DBM configuration
- Determine potential and existing bottlenecks

Integrated OS monitoring

Performance mgmt of many partitions

- Single partition, aggregated and customized multi partition views
- Problem partition and skew detection

Proactive performance problems alerts

SQL Heavy Hitter Identification

SQL Tracing & Analysis

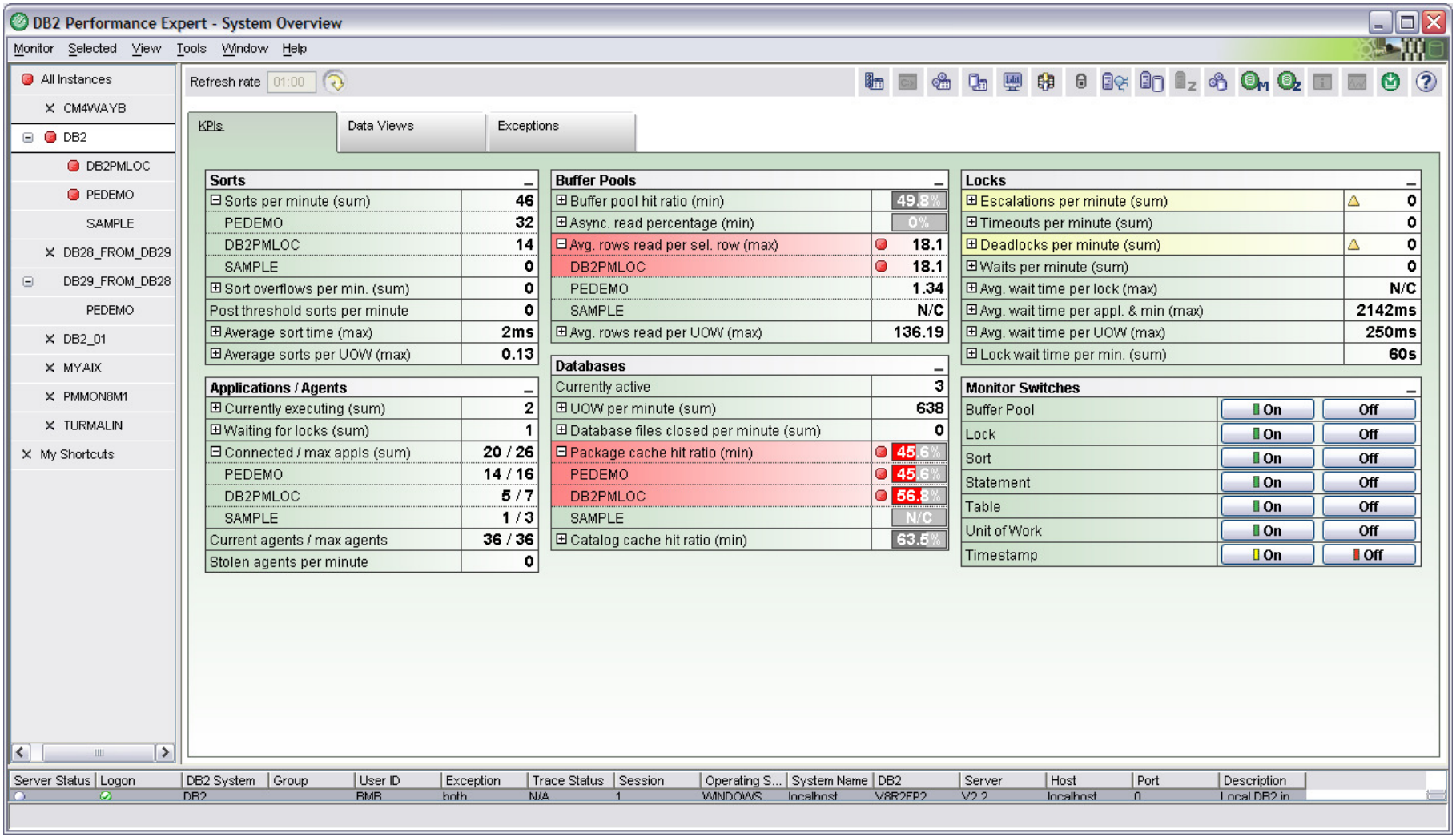
Performance History and Warehouse

- Post-morten analysis, trend analysis

WLM Monitoring (statistics and activities)

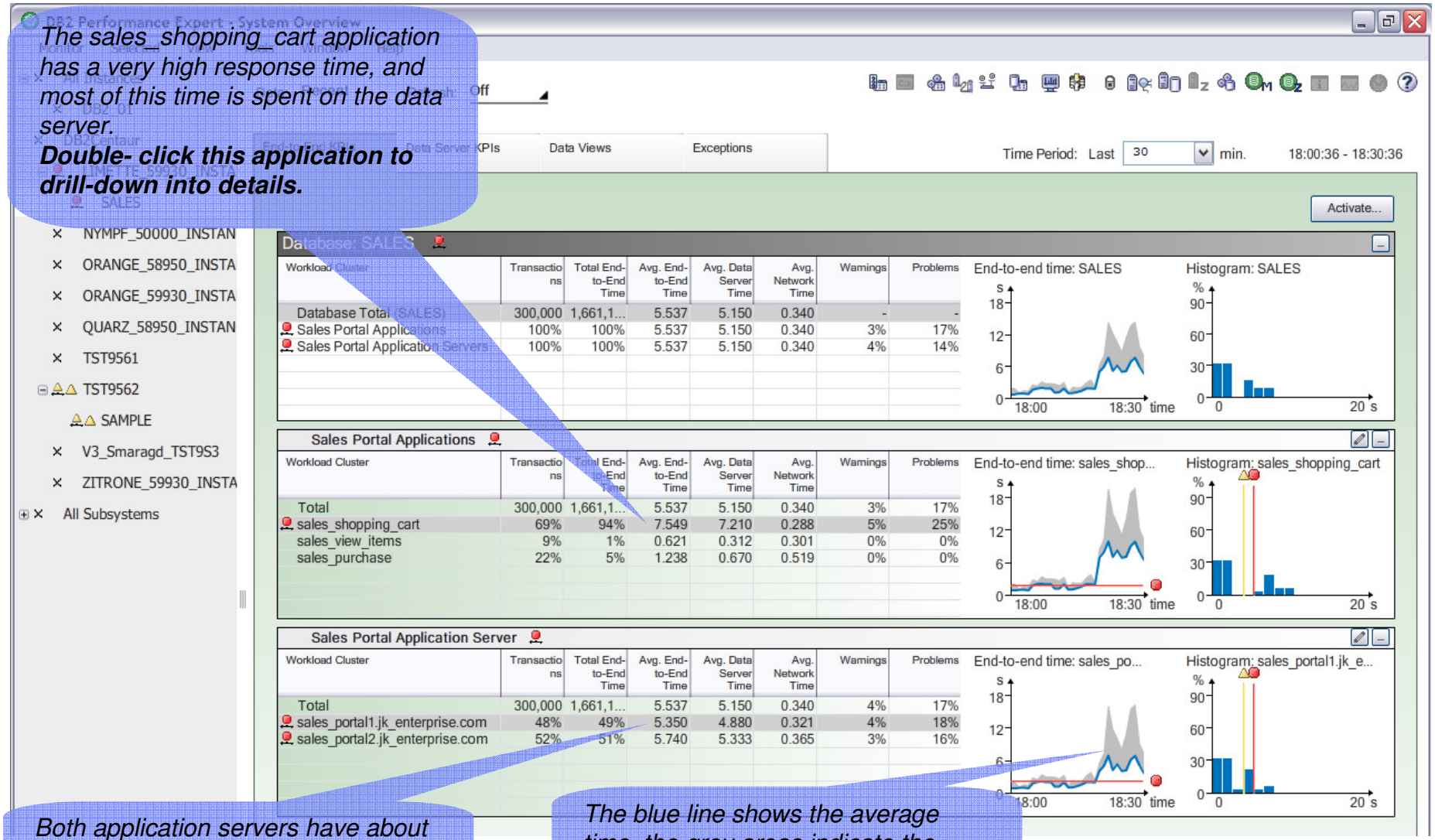
© 2009 IBM Corporation

Performance Expert - System Overview



© 2009 IBM Corporation

Diagnose Application High Response Time



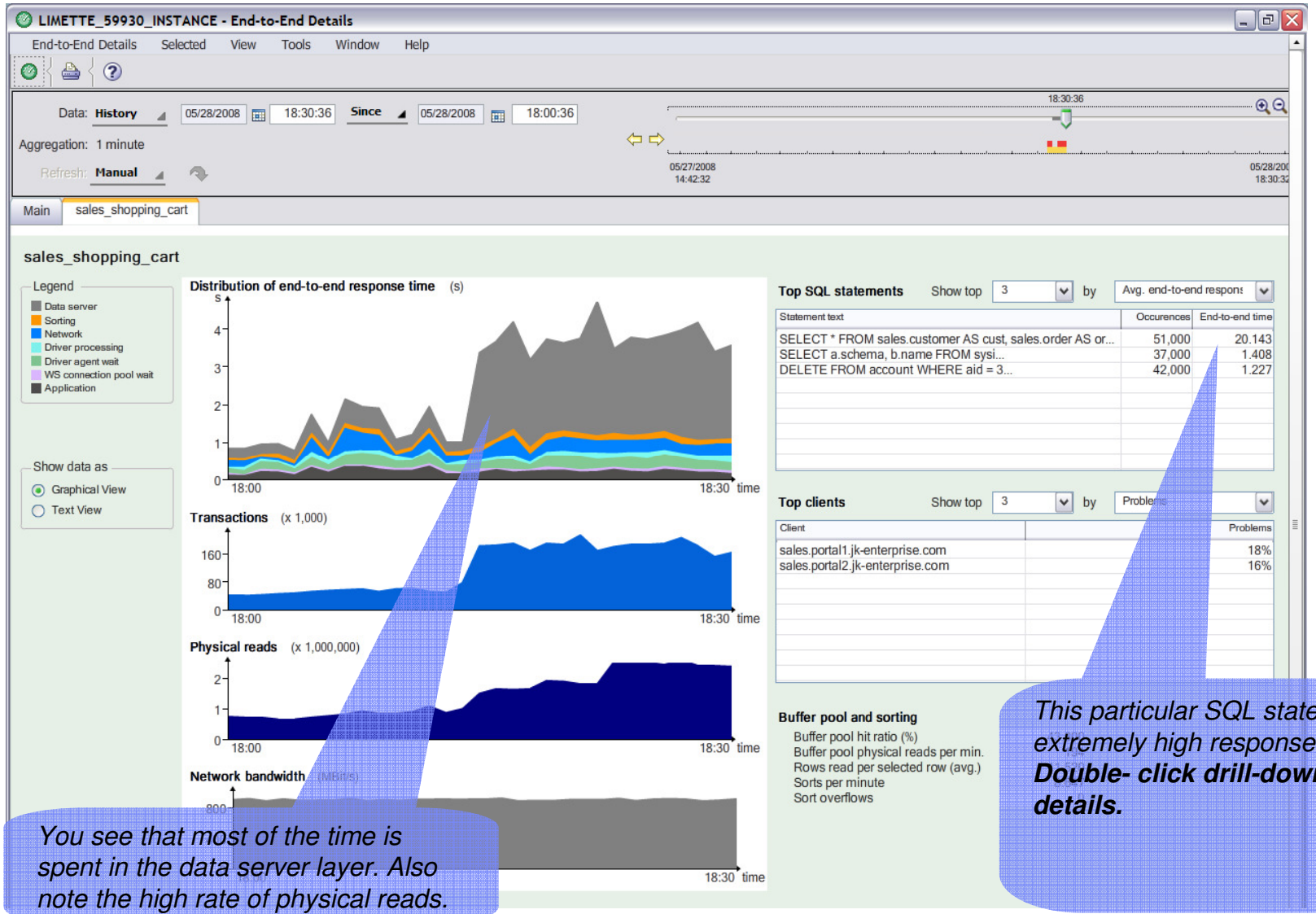
The sales_shopping_cart application has a very high response time, and most of this time is spent on the data server.
Double- click this application to drill-down into details.

Both application servers have about the same response time.

The blue line shows the average time, the grey areas indicate the maximium value.

© 2009 IBM Corporation

Diagnose Time Spent

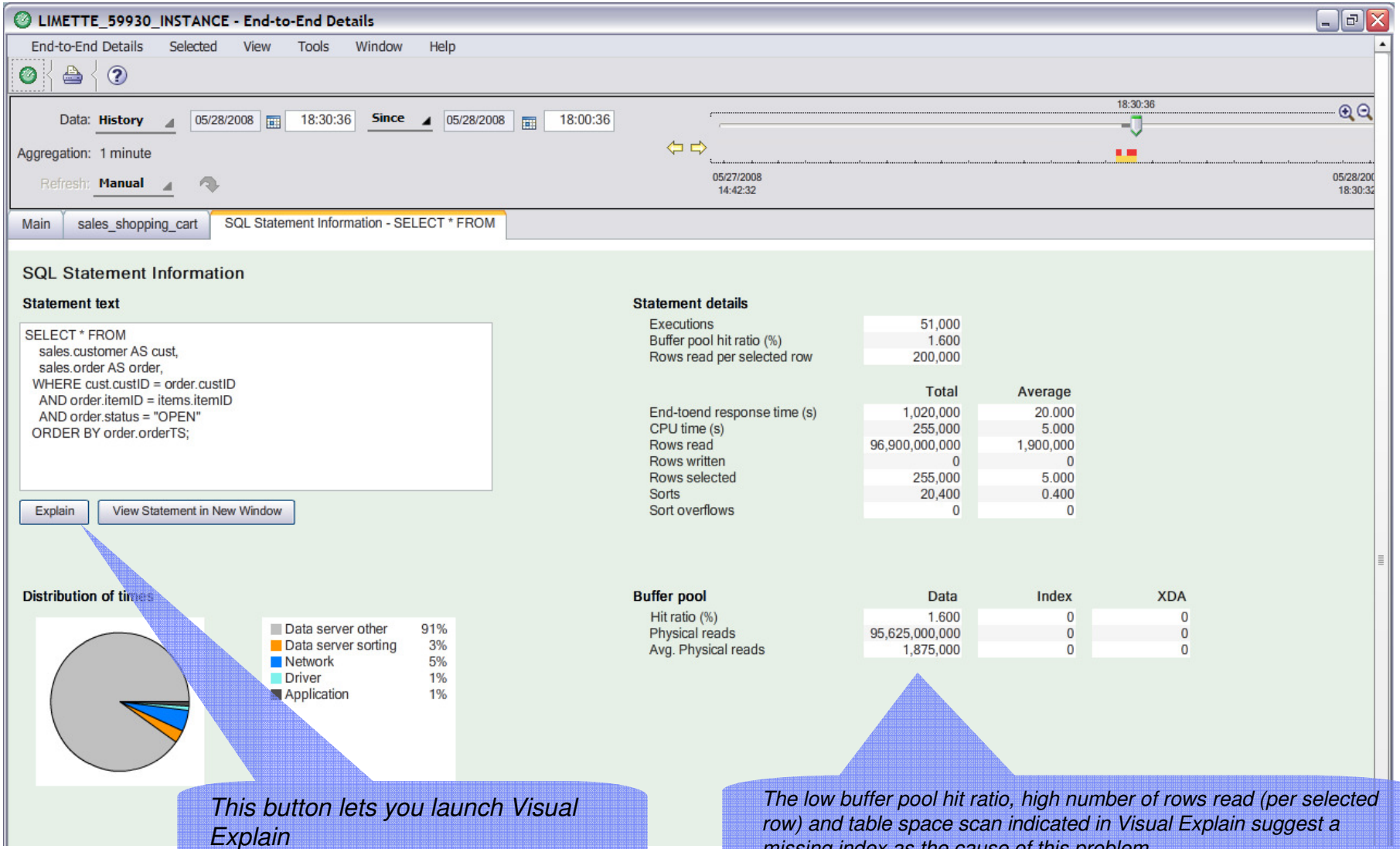


You see that most of the time is spent in the data server layer. Also note the high rate of physical reads.

This particular SQL statement has an extremely high response time. ***Double- click drill-down into details.***

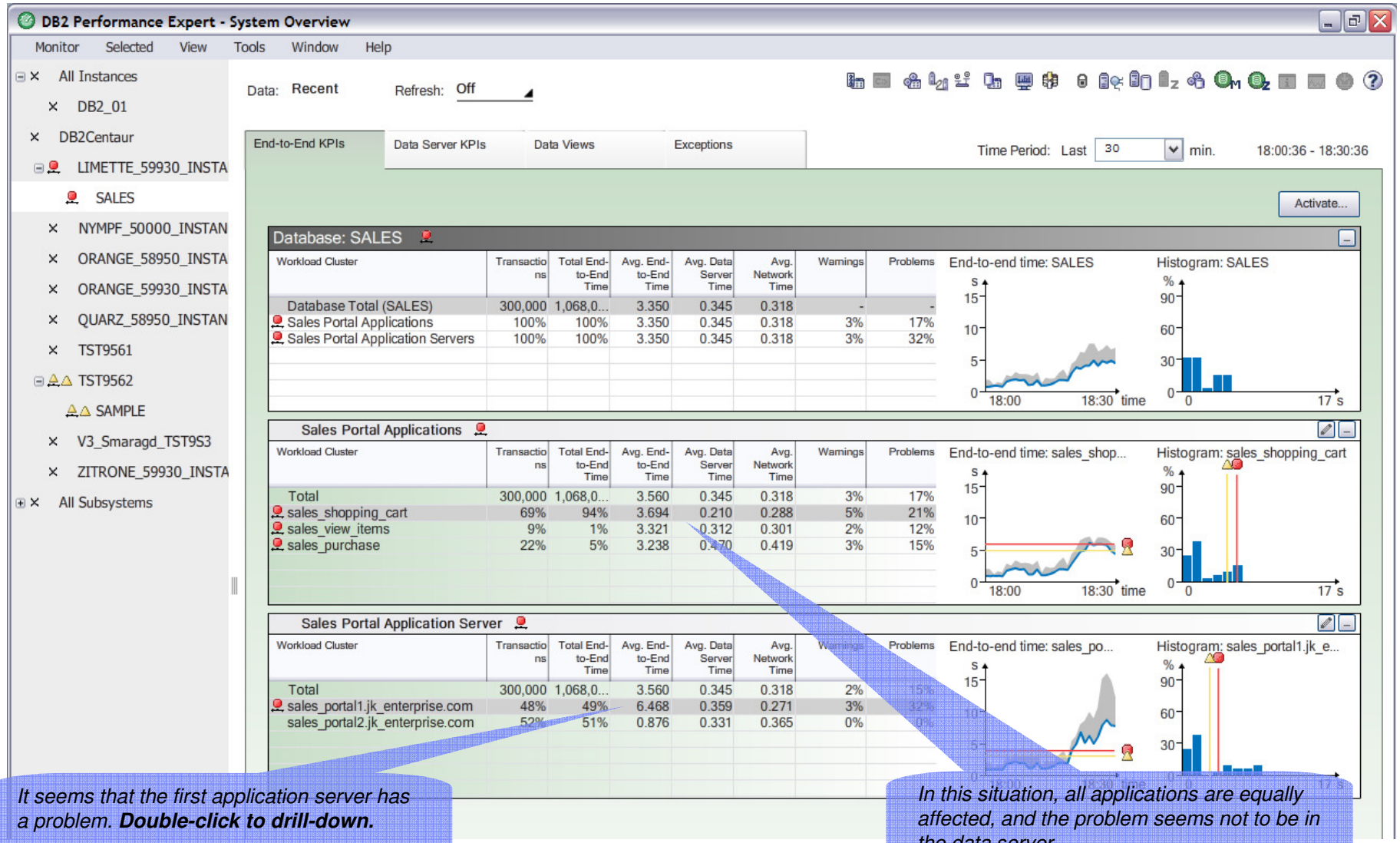
© 2009 IBM Corporation

Diagnose SQL Statement



© 2009 IBM Corporation

Diagnose Overall Response Time Problem

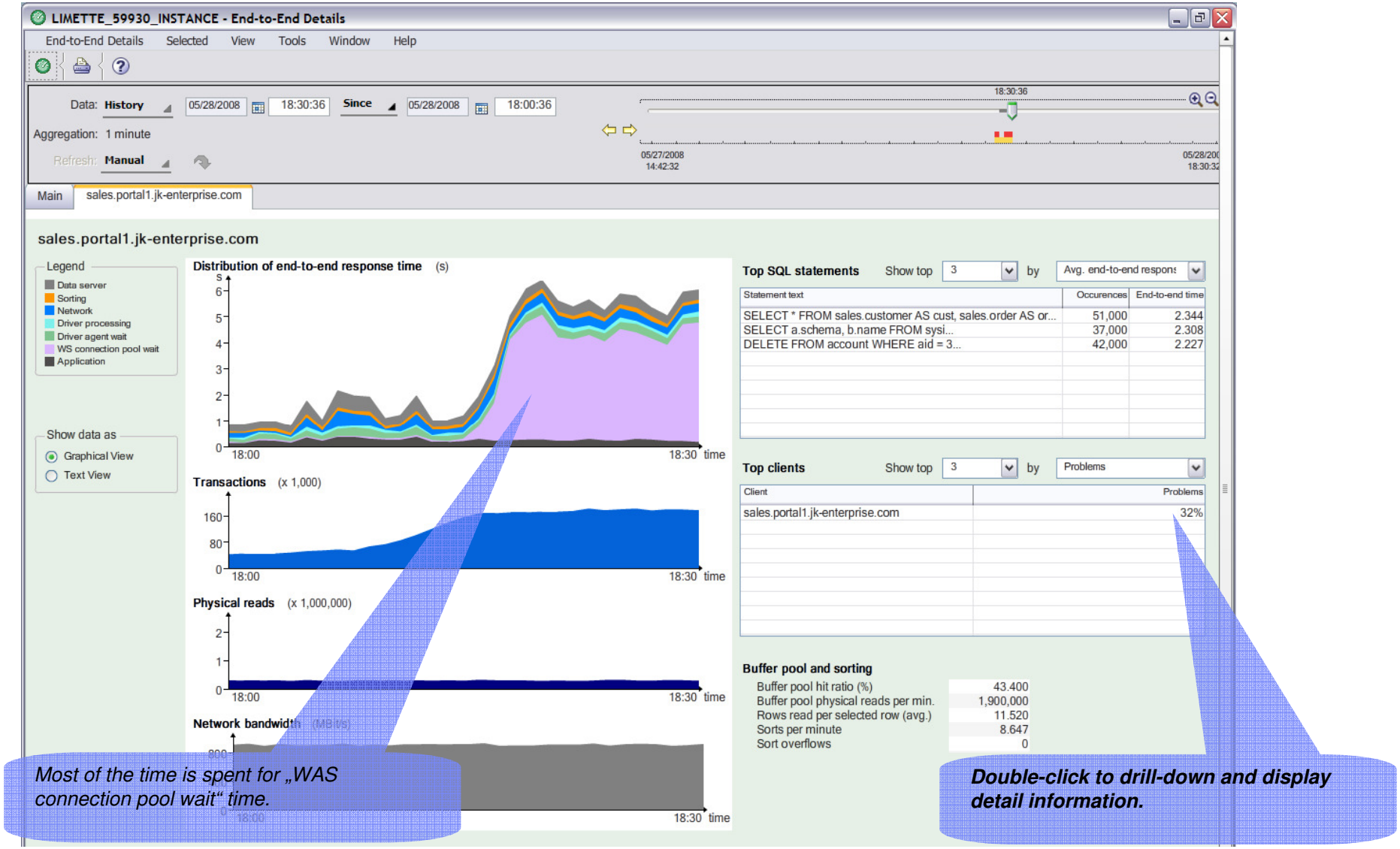


It seems that the first application server has a problem. ***Double-click to drill-down.***

In this situation, all applications are equally affected, and the problem seems not to be in the data server.

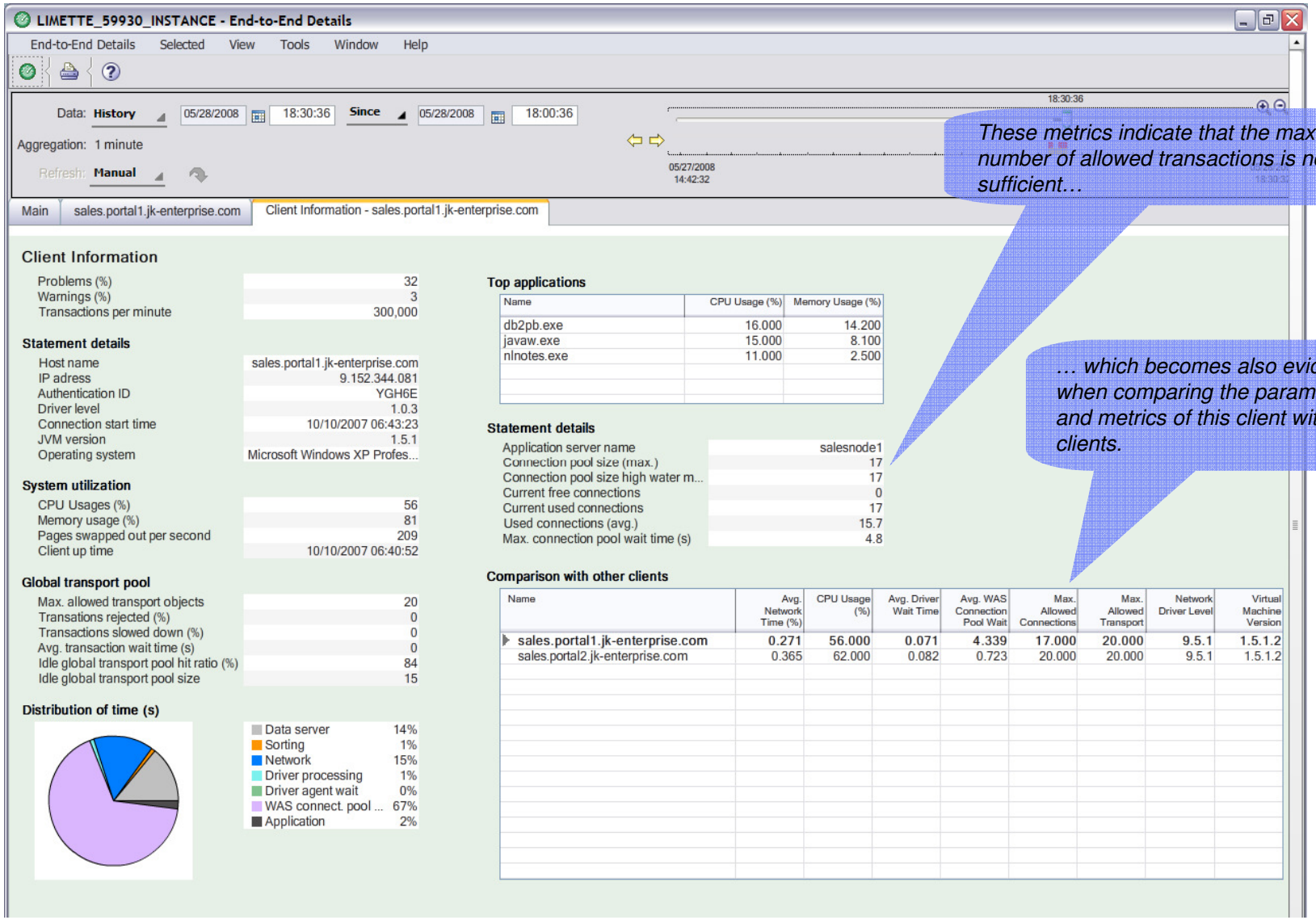
© 2009 IBM Corporation

Diagnose Application Server



© 2009 IBM Corporation

Diagnose Connection Pool



© 2009 IBM Corporation

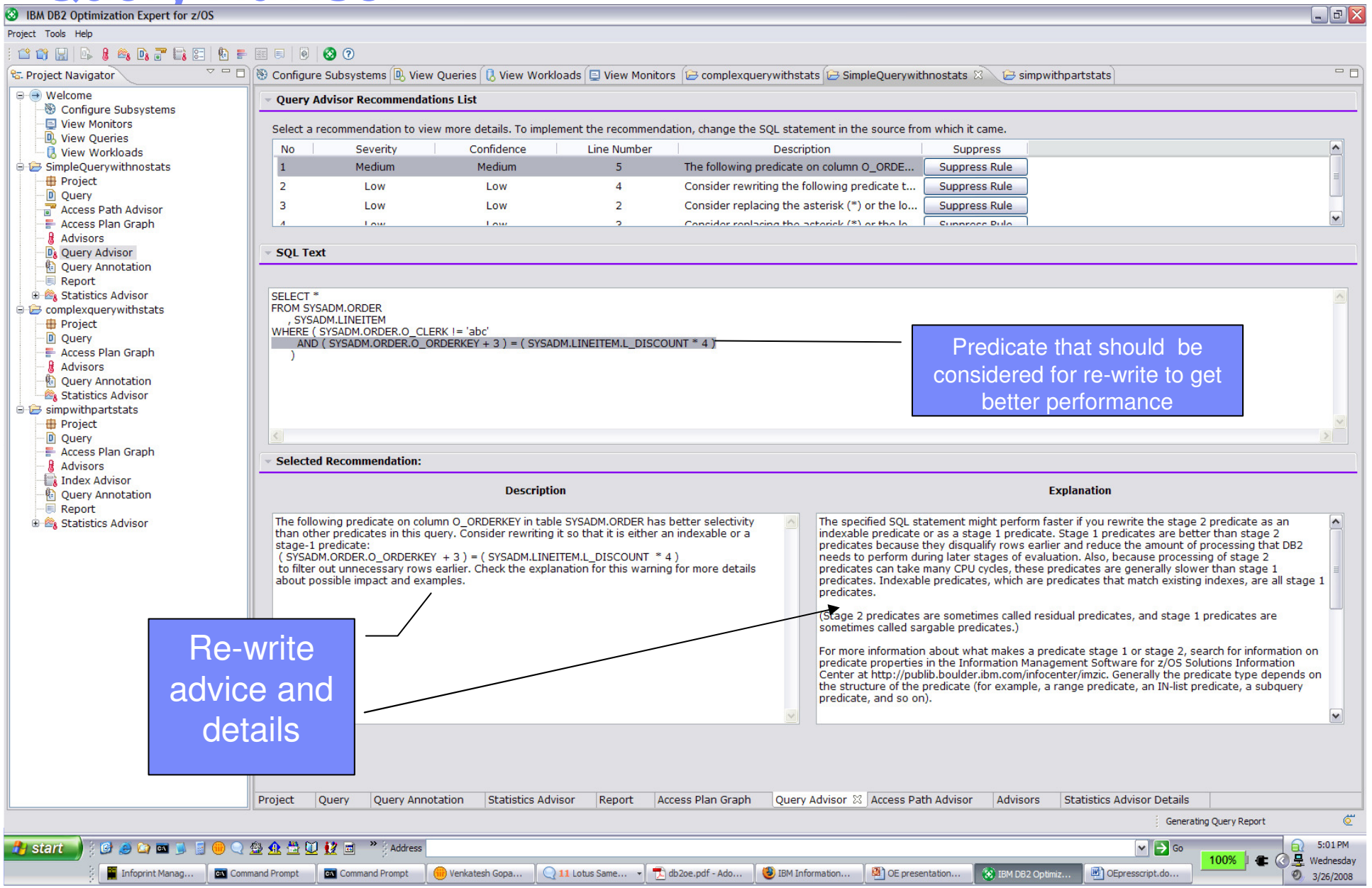
Data Studio Optimization Expert

IBM DB2 Optimization Expert tunes your SQL in context with your workload.

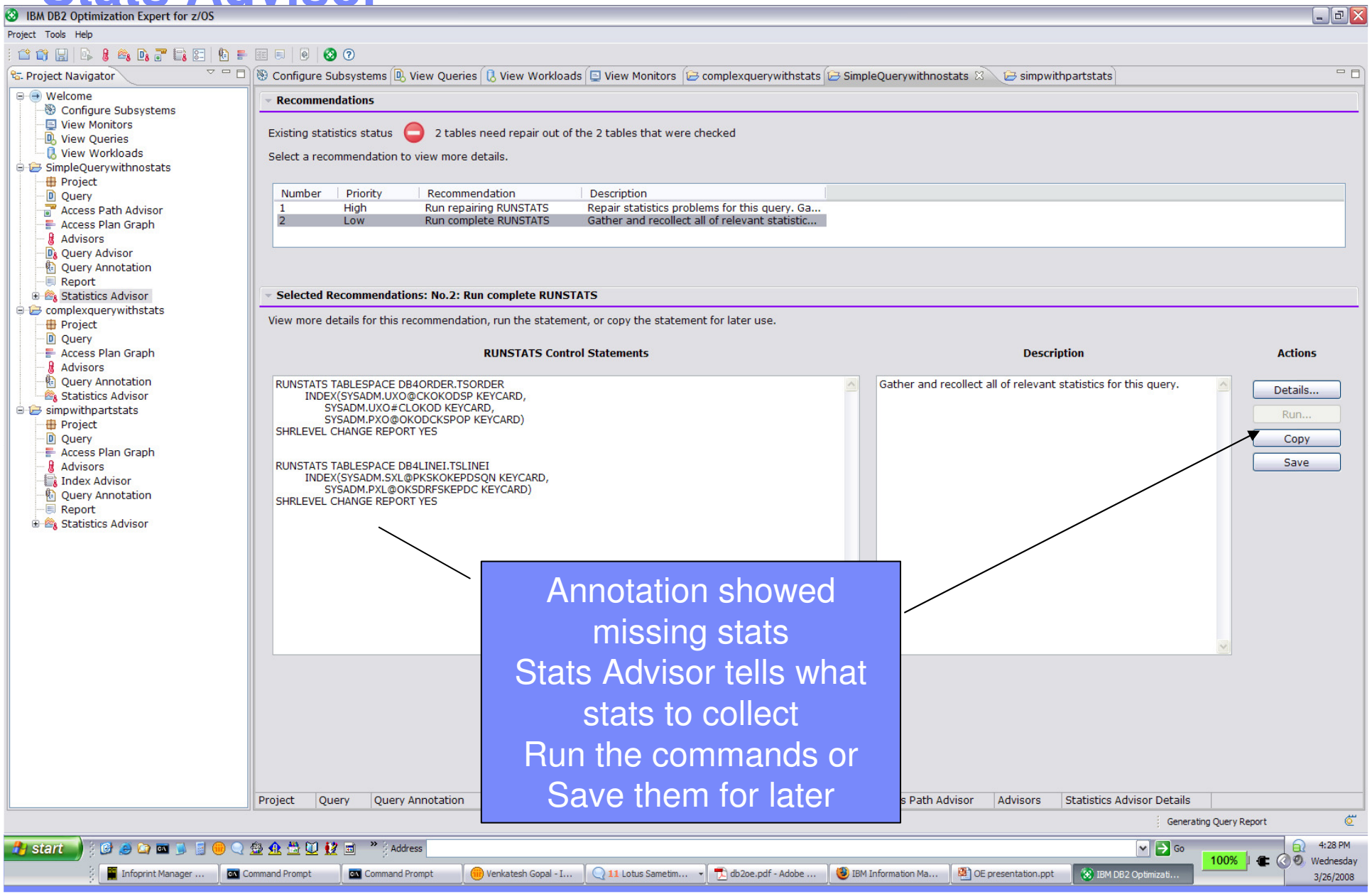
- **DB2 Optimization Expert offers a comprehensive set of expert advisors that can help improve system performance and lower total cost of ownership.**
  - Locates performance problems by being able to view query activity
  - Provides expert tuning advice to improve the performance of an SQL query
  - DB2 Optimization Expert provides profile-based monitoring and tuning for SQL statements that run as part of a workload on DB2 9 for z/OS or DB2 for z/OS Version 8 subsystems.
  - The latest release of DB2 Optimization includes support for DB2 data sharing environments and an improved user experience.

© 2009 IBM Corporation

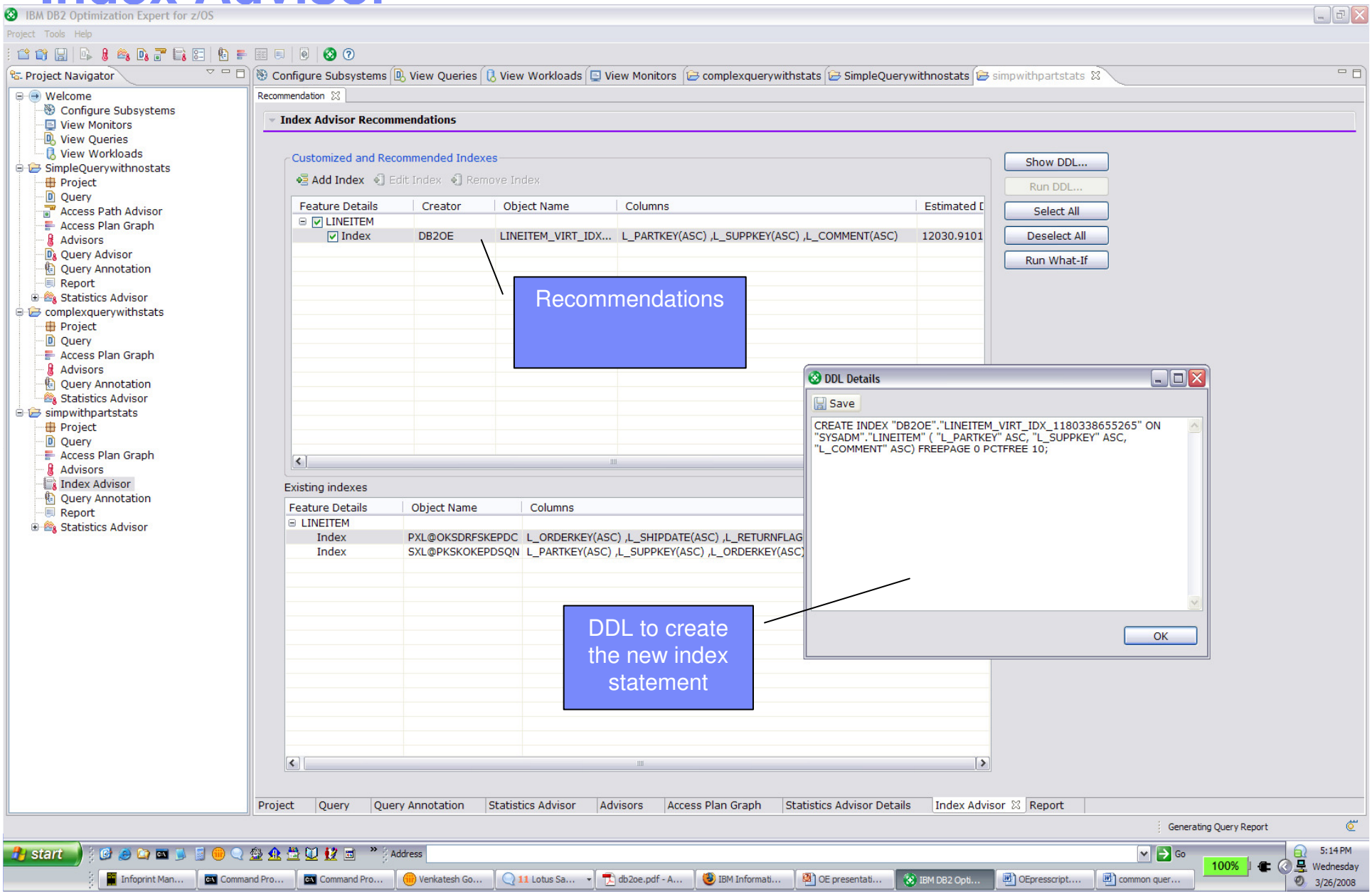
Query Advisor



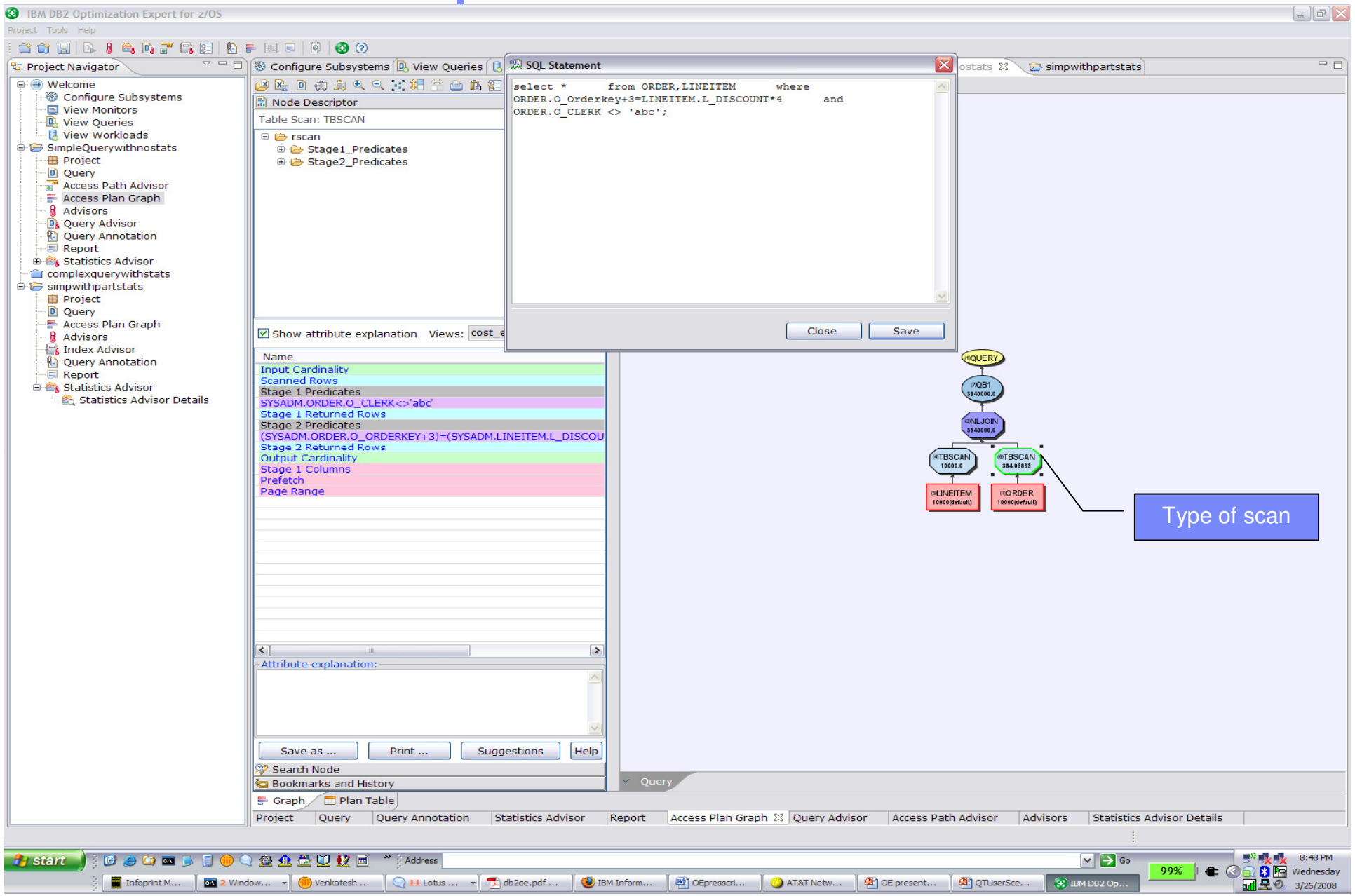
Stats Advisor



Index Advisor

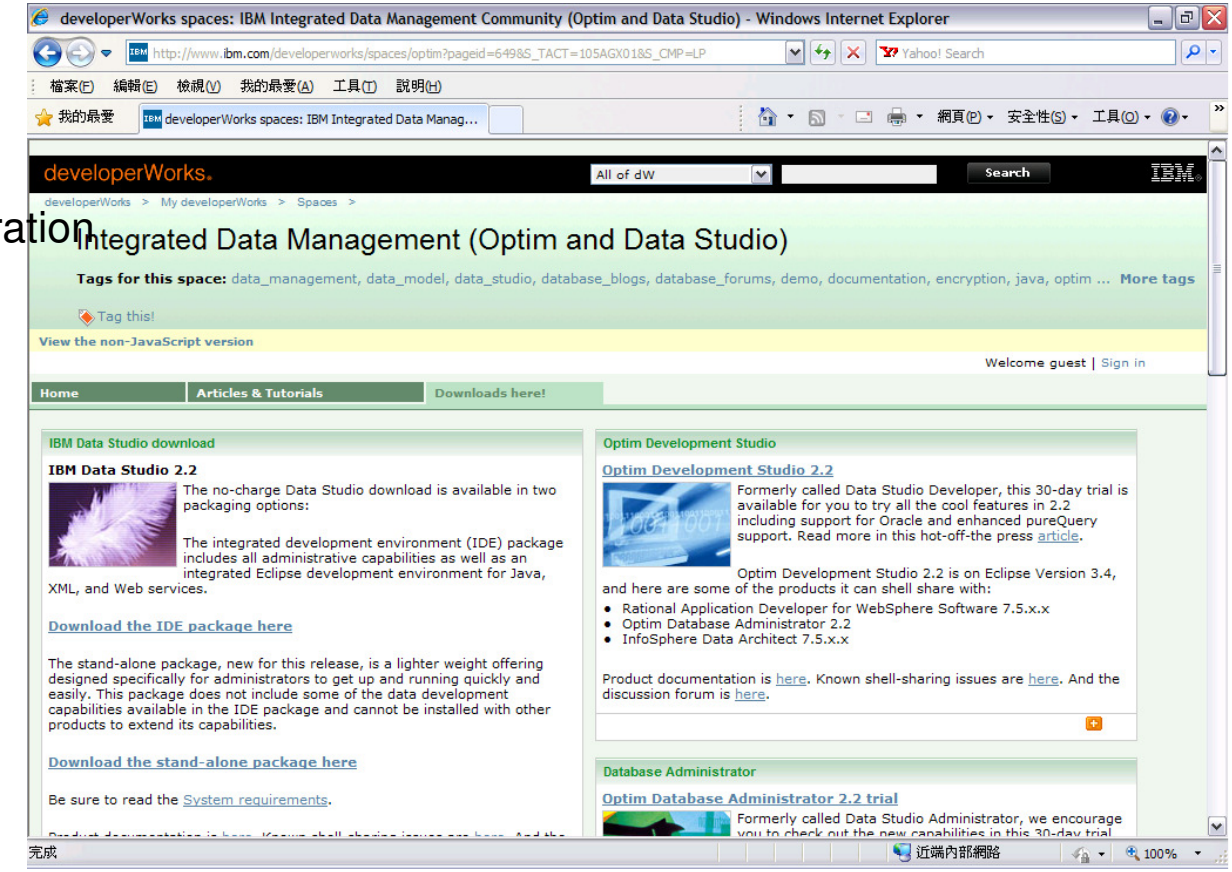


Access Plan Graph



IBM Data Studio

- **Getting started**
  - Downloads
    - IBM Data Studio
    - IBM Data Studio Administration Console
  - Community
    - developerWorks
      - forums
      - zone and space
      - articles
      - tutorials
  - Documentation
  - Support



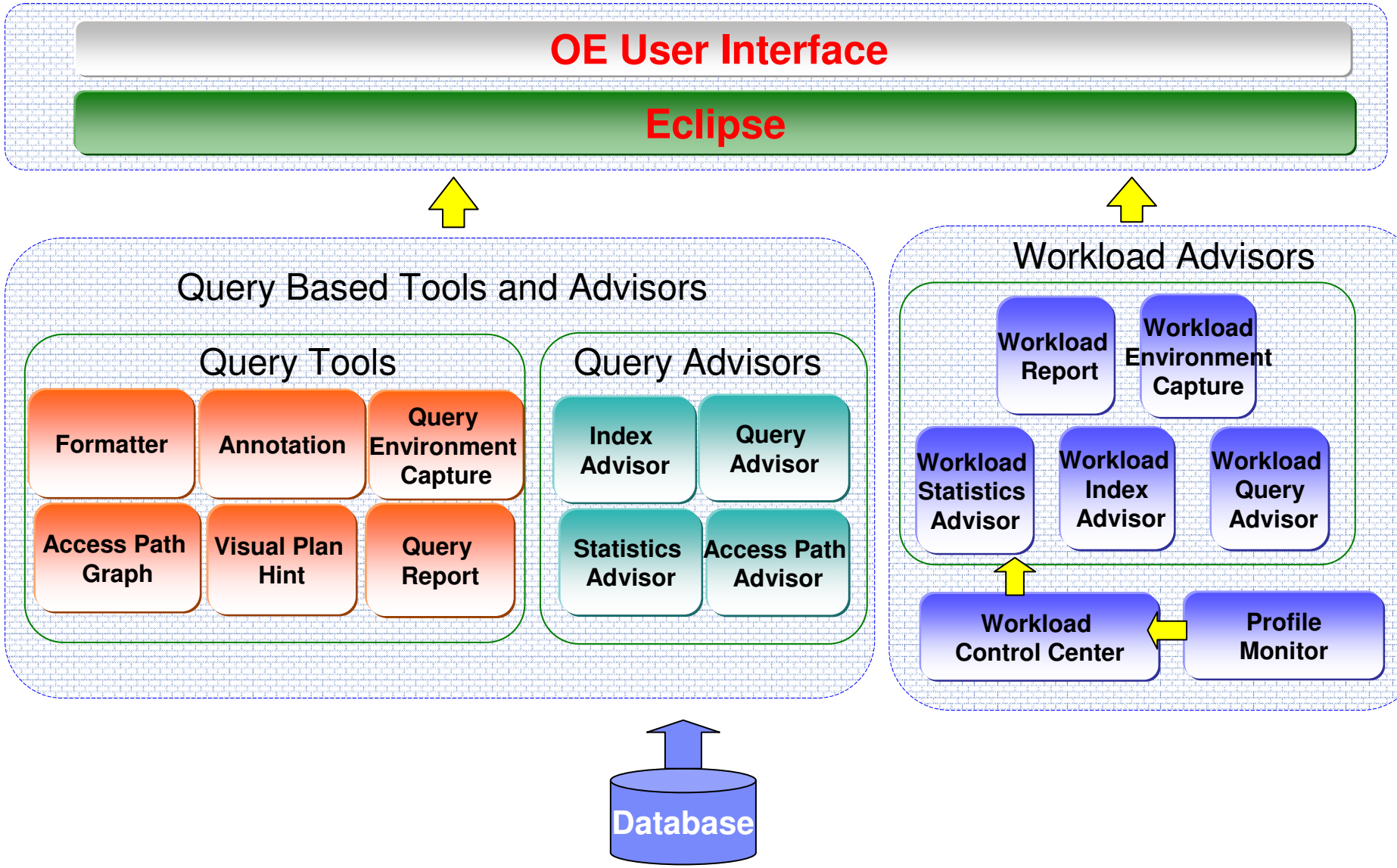
Get IBM Data Studio

http://www.ibm.com/developerworks/spaces/optim?pageid=649&S_TACT=105AGX01&S_CMP=LP

© 2009 IBM Corporation

© 2009 IBM Corporation

Optimization Expert Key Features at a Glance



© 2009 IBM Corporation

Optim Consolidation – Product Names

Existing Name	Rebranded Name
IBM Data Studio	**IBM Data Studio**
IBM InfoSphere Data Architect	**IBM InfoSphere Data Architect**
IBM Data Studio Developer	**IBM Optim Development Studio**
IBM Data Studio pureQuery Runtime	**IBM Optim pureQuery Runtime for Linux, UNIX, and Windows**
IBM Data Studio pureQuery Runtime for z/OS	**IBM Optim pureQuery Runtime for z/OS**
IBM Data Studio Administrator	**IBM Optim Database Administrator**
IBM DB2 Performance Expert	**IBM Optim Performance Manager for DB2 on Linux, UNIX, and Windows**
IBM DB2 Performance Expert Extended Insight Feature	**IBM Optim Performance Manager Extended Insight for DB2 on Linux, UNIX, and Windows**
IBM DB2 Optimization Expert	**IBM Optim Query Tuner for DB2 on z/OS** **IBM Optim Query Tuner for DB2 on Linux, UNIX, and Windows**
	IBM Optim Query Workload Tuner for DB2 on z/OS **IBM Optim Query Workload Tuner for DB2 on Linux, UNIX, and Windows**
IBM DB2 High Performance Unload	**IBM Optim High Performance Unload for DB2 on Linux, UNIX, and Windows**

© 2009 IBM Corporation

DB2 Developer Workbench vs. Data Studio

Developer Workbench

SQL Query Editor
- SQLJ Editor
- SQL Builder
- XQuery Builder
- SQL Routine Debugger
- Java Routine Debugger
- XML Editor
- XML Schema Editor
- Data Management
- Visual Explain
- Project Management

Data Studio is a full replacement of DB2 Developer Workbench plus much more

- ***DB2 for Linux, Unix, Windows v8.x, v9.1.x, v9.5***
- ***DB2 for z/OS v7, v8, v9***
- ***DB2 for i5/OS v5r2, v5r3, v5r4***
- ***Informix Dynamic Server (IDS) v9.x, v10.x, v11, v11.5***

Data Studio

- Integrated Query Editor – SQL + XQuery
- SQLJ Editor
- SQL Builder
- XQuery Builder
- SQL Routine Debugger
- Java Routine Debugger
- XML Editor
- XML Schema Editor
- Data Management
- Visual Explain
- Project Management

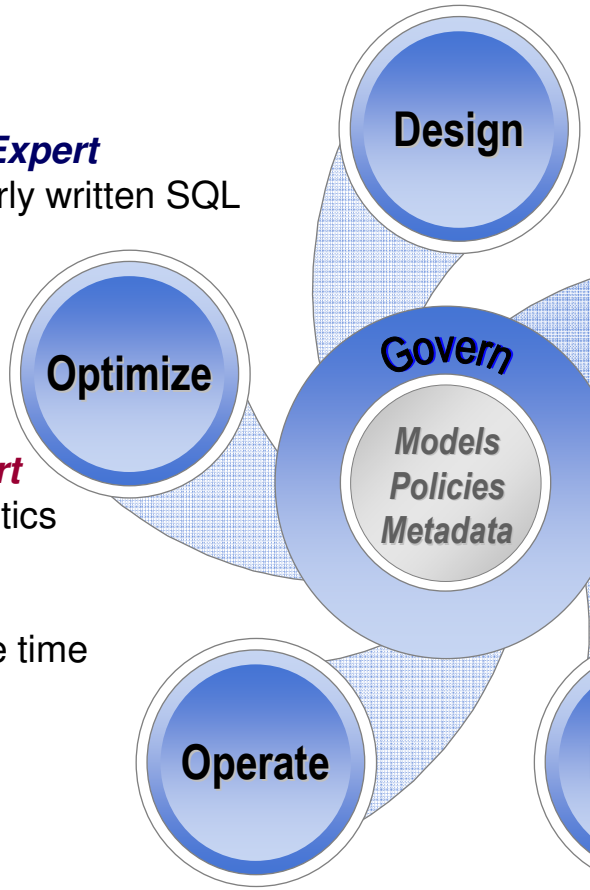
- ER Diagramming
- Data Distribution Viewer
- Object Management
- Privilege Management
- Browse & Update Statistics
- Security Access Control
- Connection Management integration with Kerberos and LDAP
- Data Web Services
- IDS Server Support
- Health Monitoring DB2 for LUW 9.5 and DB2 z/OS v9
- Instance and database commands
- Utility support
- Data management support
- View and force applications

Data Studio v2.1

© 2009 IBM Corporation

Data Studio Offerings

Managing the value of your data throughout its lifetime



InfoSphere Data Architect
Enterprise data modeling and integration design tool.

DB2 Optimization Expert
Detect and tune poorly written SQL

Data Studio Developer
Integrated SQL Development Environment

DB2 Performance Expert
DB2 monitor and diagnostics solution

DB2 Extended Insight
DB2 end-to-end response time monitor

**** Integrated in Nov. 09***

Data Studio pureQuery Runtime
A high-performance Java data access platform

High Performance Unload
Unloads online and offline data quickly for extract and recovery

Data Studio Administrator
Simplify and automate data and object management functions

© 2009 IBM Corporation

IBM Data Studio關鍵價值及其整合元件

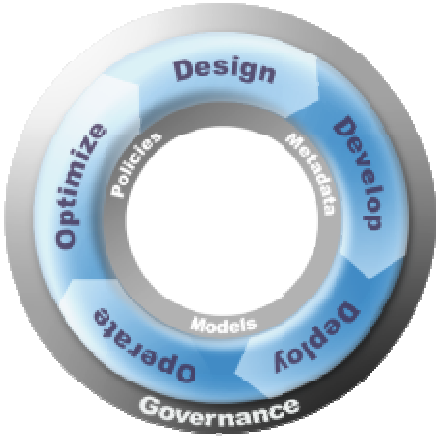
提高生產力
Java, XML, SOA 開發
作業自動化
團隊協同合作

提高服務品質
效能最佳化
增強的安全性
提升品質

提升效率
有效利用CPU性能
降低錯誤率
問題隔離

IBM InfoSphere Data Architect
A collaborative data design tool to understand information assets and their relationships, model data and integration designs, and enforce enterprise standards for data quality and consistency

IBM Optim Database Administrator
An administration environment to reduce application outages by automating and simplifying complex DB2 structural changes



IBM Optim Development Studio
An integrated development environment for rapidly creating and testing database and pureQuery applications and services.

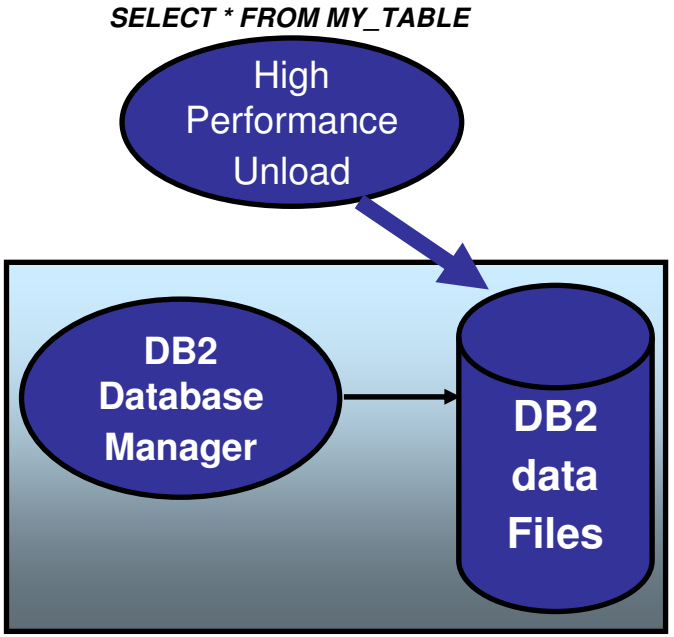
IBM Optim pureQuery Runtime
A high-performance Java data access platform to improve performance, security, and manageability of Java connections to databases.

© 2009 IBM Corporation

DB2 High Performance Unload

DB2 High Performance Unload helps meet service agreements and eases application upgrades by extracting data quickly and efficiently.

- **Fast unload capability**
  - Often 4 to 6 times faster than export
  - Parallel processing for higher speeds
- **Flexibility to meet all use cases**
  - Full Partitioned DB2 support
  - Subset of SELECT syntax used to filter columns and/or rows
  - Many automatic data type conversions & formats available
  - Offline or Online Backups
- **Repartitioning feature**
  - Built-in “DB2 Splitter”
    - Unload and split in a single operation



Data Studio v2.1

© 2009 IBM Corporation