



Solving Complex Data Warehousing Problems with Advanced Tooling of IBM InfoSphere Warehouse Advanced Editions

Jon Lind
IBM Product Manager, InfoSphere Warehouse software
February 23, 2012





DB2 Tech Talk series host: Serge Rielau

SQL Architect,
DB2 for Linux, UNIX, and Windows at IBM

<http://www.sqltips4db2.com/>



Today's presenter: Jon Lind

Product Manager,
InfoSphere Warehouse at IBM

<http://www.ibm.com/software/data/infosphere/data-warehousing/>
<http://www.facebook.com/pages/IBM-InfoSphere-Warehouse>



Special thank you: Rav Ahuja

Senior Program Manager, IBM Cloud and Big Data



IBM Data Warehouse & Analytics Solutions

Simplicity, Flexibility, Choice

IBM Netezza



True Appliance

IBM Smart Analytics System



Flexible Integrated System

IBM InfoSphere Warehouse



Custom Solution

Warehouse Accelerators

*Information Management Portfolio
(Information Server, MDM, Streams, etc)*

Simplicity

The right mix of simplicity and flexibility

Flexibility

IBM InfoSphere Warehouse

Solve problems – Create new opportunities



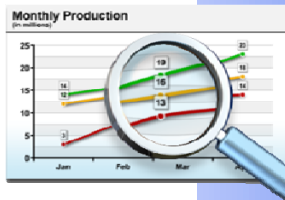
FLEXIBILITY

To Adapt to Your Business Needs



PERFORMANCE

For Business Agility



SOPHISTICATED ANALYTICS

For Smarter Decision-Making



ISW Departmental Editions

Departmental Edition

- **Core server**
 - DB2 9.7 Enterprise Server Edition
 - Data Partitioning Feature
 - Workload Manager Feature
 - Storage Optimization Feature
- **ELT and Administration**
 - Design Studio
 - Replication Server
 - Data Studio
 - Performance Manager
- **Analytics**
 - Cubing
 - Intelligent Miner
 - Text Analytics
 - Cognos 10 (5 licenses)

Advanced Departmental Edition Adds

- ***Warehouse packs***
 - *Customer Insight*
 - *Market and Campaign Insight*
 - *Supply Chain Insight*



ISW Enterprise Editions

Enterprise Edition

- **Core server**
 - DB2 9.7 Enterprise Server Edition
 - Data Partitioning Feature
 - Workload Manager Feature
 - Storage Optimization Feature
 - **Label Based Access Control**
- **Design, ELT, and Administration**
 - Design Studio
 - Replication Server
 - **Federation Server**
 - Data Studio
 - Performance Manager **Extended Edition**
 - **Query Patroller**
- **Analytics**
 - Cubing
 - Intelligent Miner
 - Text Analytics
 - **Mashup Center**
 - Cognos 10 (5 licenses)

Advanced Enterprise Edition Adds

- ***Warehouse packs***
 - *Customer Insight*
 - *Market and Campaign Insight*
 - *Supply Chain Insight*
- ***Design, ELT, and Administration***
 - *InfoSphere Data Architect (10 licenses)*
 - *High Performance Unload*
 - *Merge Backup*
 - *Recovery Expert*
 - *Query Workload Tuner*

Terabyte Pricing Makes Good Business Sense

- **Pay for what you analyze**
 - Data is the currency of the Information Age
 - The more data people can analyze the more value they receive
 - Data warehouses are a gold mine for business decision makers
- **Flexibility**
 - “Pay as you Grow”
 - Deploy on the system that best meets your business needs
- **Enterprise license planning made easy**
 - Companies know how much data they have and the growth rates
 - Companies can easily plan their warehouse license requirements
- **Lower cost of Growth**
 - Customers can provision hardware platforms for end game capacity
 - Simple lower cost of growth by simply Increasing license entitlement

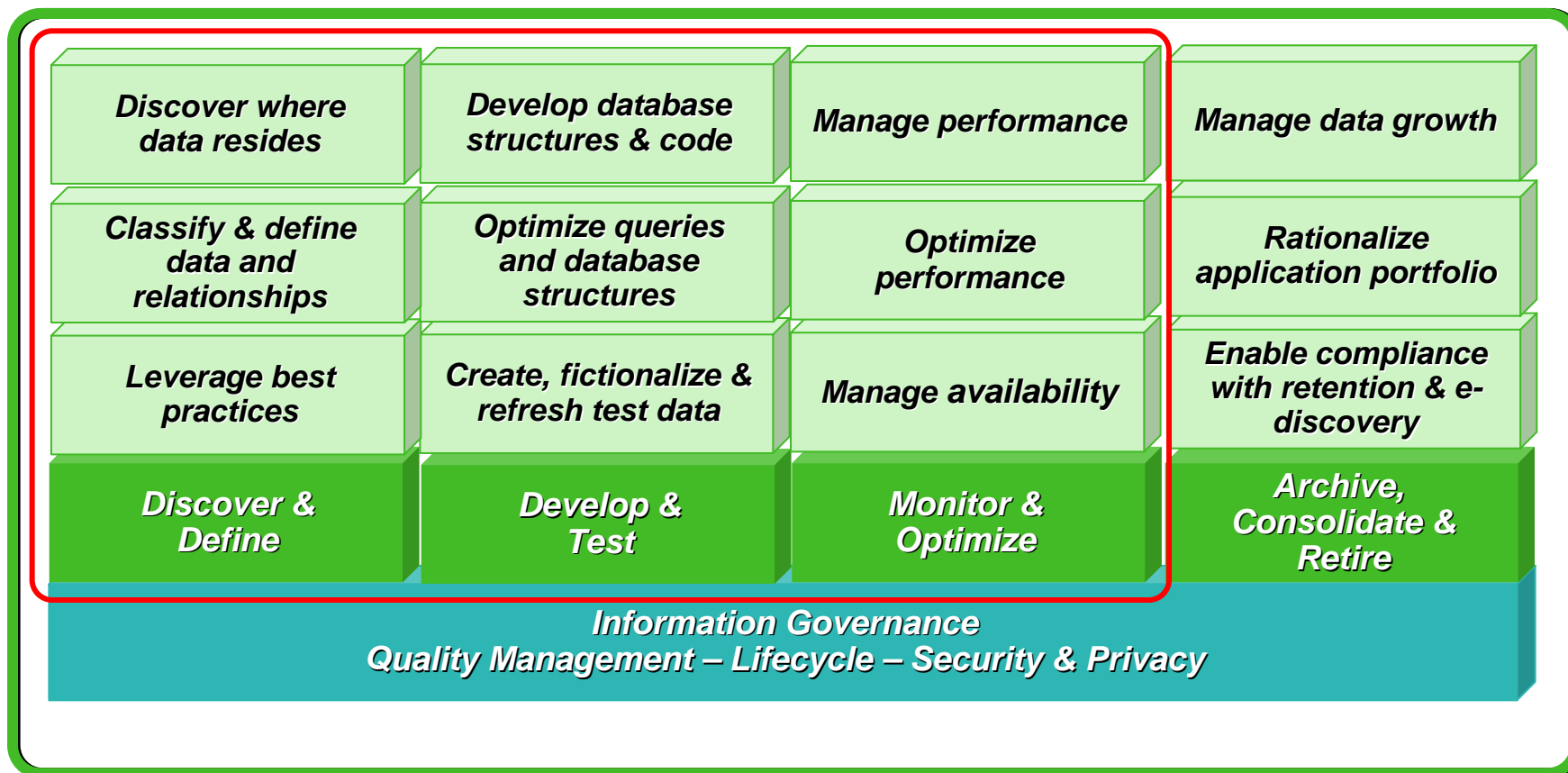
What's New in Advanced Editions?

- **Accelerate value delivery**
 - Warehouse packs with industry models and Cognos reports
 - Markets and Campaign Insight
 - Customer Insight
 - Supply Chain Insight
- **Improve governance**
 - InfoSphere Data Architect
- **Improve performance**
 - InfoSphere Optim Query Workload Tuner
- **Improve availability**
 - InfoSphere Optim High Performance Unload
 - DB2 Merge Backup
 - DB2 Recovery Expert

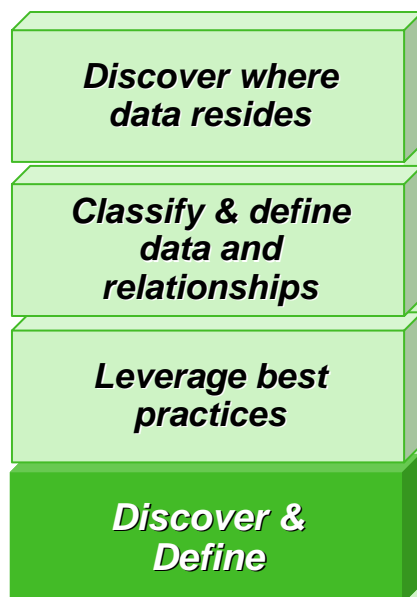


InfoSphere Optim Offerings

Streamlining the Management of Data Across its Lifecycle



InfoSphere Warehouse Solutions to Discover and Define



- **Data Architect**

- Design and manage enterprise models
- Enforce model conformance to enterprise standards
- Leverage industry data models for best practices

Standardize and collaborate

- **Warehouse Packs**

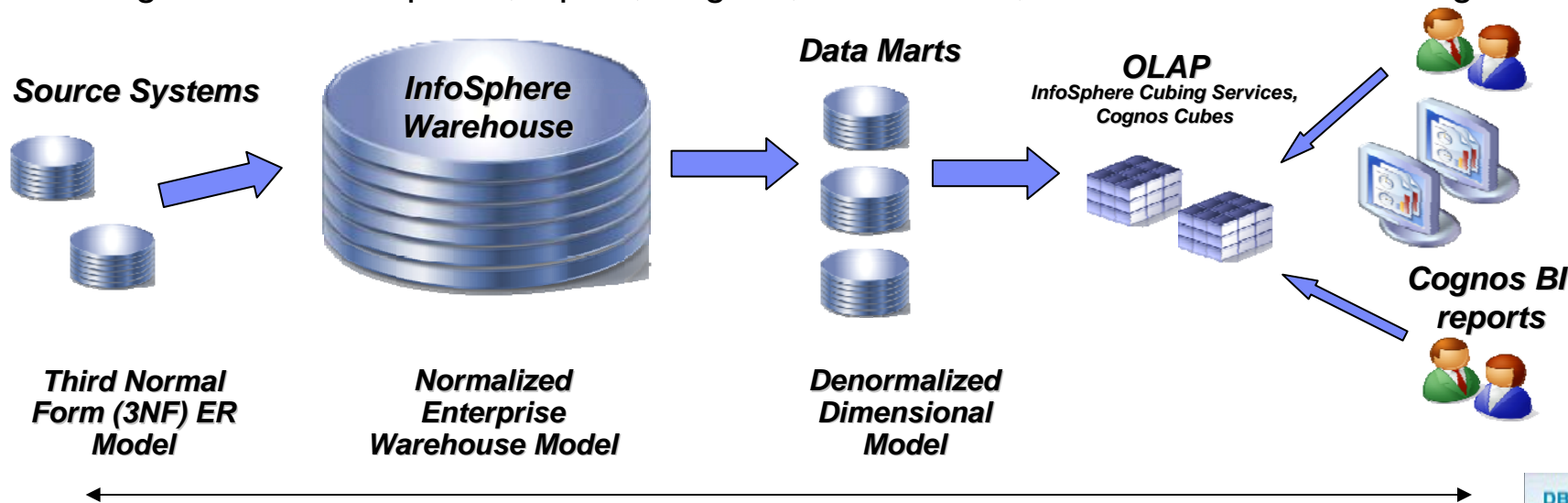
- Data models
- Business reports

Accelerate projects

Managing Your Information Models

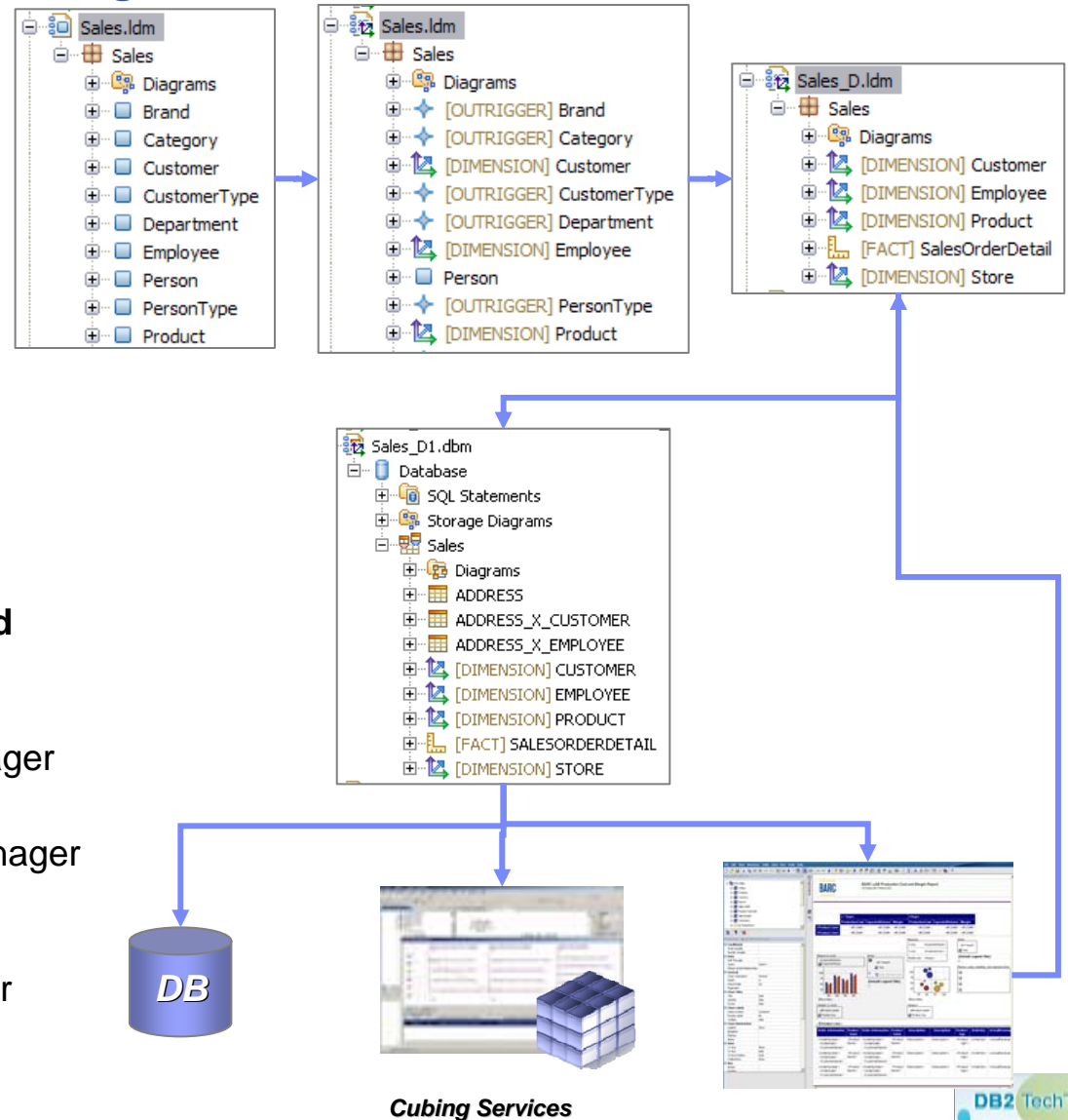
InfoSphere Data Architect

- **Full service data modelling**
 - Create logical, physical, and dimensional models
- **Validate models against enterprise standards**
 - Standardize names, terms, data types, transformations, etc
- **Leverage existing assets**
 - Forward engineer, reverse engineer, and import from source or other
- **Collaborate across teams**
 - Integrate with InfoSphere, Optim, Cognos, Data Studio, and Rational and offerings



Multi-Dimensional Modeling

- **Automatic Discovery**
 - Facts and measures
 - Dimensions and outriggers
- **Dimensional Notation**
- **Automatic De-normalization**
 - Star schema, snowflake, starflake
- **Lossless integration with Cognos and Cubing Services**
 - Transform dimensional-physical data model to a Cognos Framework Manager model or Cubing Services model
 - Transform a Cognos Framework Manager dimensional model to a dimensional-logical data model
 - Support Cognos Framework Manager model data source “re-targeting”



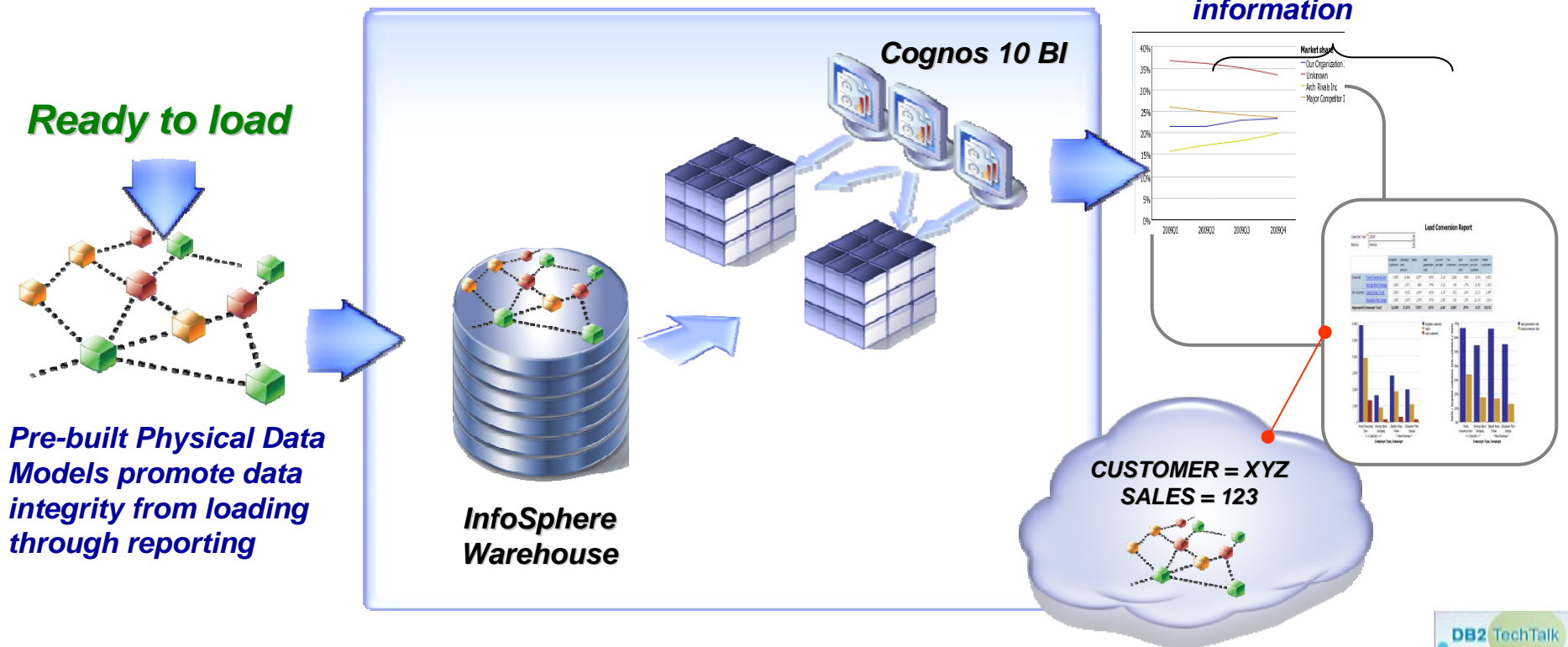
Business Models – InfoSphere Warehouse Model Packs

Solution acceleration brings immediate value – the work is done!

Leverage 'pre-built' data models targeted at specific business issues

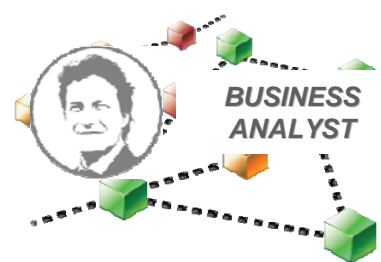
- **Customer Insight**
- **Market and Campaign Insight**
- **Supply Chain Insight**

Pre-built reports illustrate IMMEDIATE representation of your information

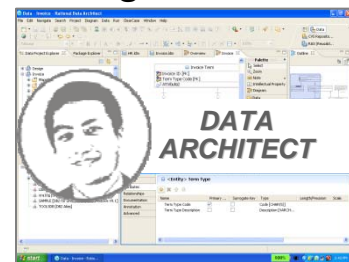


Managing Database Objects to Align Business and IT

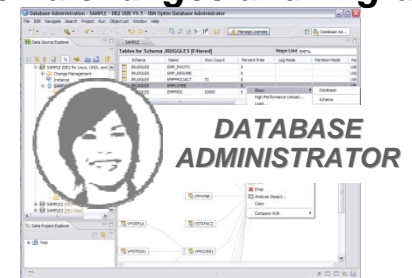
Leverage best practices



Design, extend, and govern data



Compare, generate, and deploy schema changes and migrate



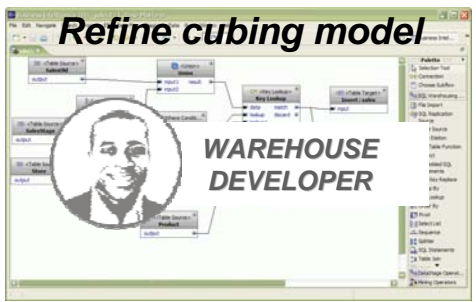
Warehouse pack models

InfoSphere Data Architect

Data Studio

Define ETL transformations

Refine cubing model



Design Studio

Develop business reports

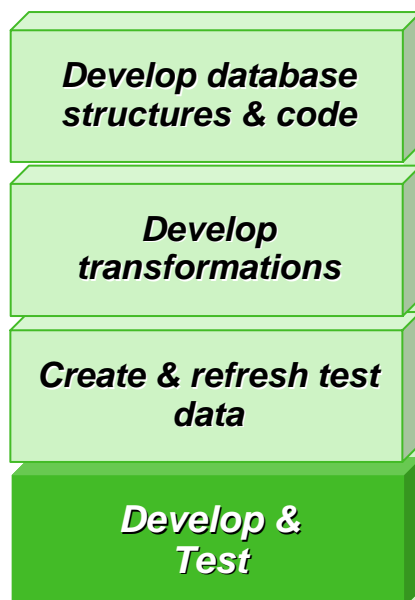


Cognos



InfoSphere Warehouse Solutions to Develop and Test

*Accelerate development
Save time, reduce errors*



- **Data Studio**
 - Manage databases and database changes
 - Develop database routines and access methods
 - Develop SQL
 - Copy and edit data

Optimize query performance

- **Design Studio**
 - Develop in-database transformations
 - Leverage cubing services

Install Data Architect, Data Studio, and Design Studio into the same Eclipse workspace

Strategic replacement for Control Center

Data Studio 3.1

- Object navigation, filtering and management
- Complex change management
- Configuration and maintenance assistants
- Task scheduling
- SQL and routine development
 - Query analysis
 - Oracle and Sybase Skin compatibility
- Data browsing and editing
- Shell sharing with Design Studio

The screenshot shows the IBM Data Studio 3.1 interface. The Administration Explorer on the left displays a tree view of databases and objects. The main pane shows a filtered table list with columns: Database, Schema, Name, Location, Row Count, and Tablespace. A context menu is open over a table, showing options like Show, Load, Unload, Edit, etc. The bottom pane shows the properties for the selected table, CBMT011_PROD_GRP.

Server -> Database: Callout pointing to the Administration Explorer tree.

Filtering: Callout pointing to the 'Name Like' filter box in the Tables [Filtered] pane.

Contextual launch: Callout pointing to the context menu over the table list.

Object properties: Callout pointing to the Properties pane for the selected table.

Database	Schema	Name	Loca...	Row Count	Tablespace
FORDDDB	ADMF001	CBMT010_REF_PART	STLEC1		CBMT010R
FORDDDB	ADMF001	CBMT011_PROD_GRP			
FORDDDB	ADMF001	CBMT012_PR_META			
FORDDDB	ADMF001	CBMT014_STRUC_I			

Connection: STLEC1
Showing 4 of 299 items

Table: CBMT011_PROD_GRP

General

Name: CBMT011_PROD_GRP

Label:

Schema: ADMF001

Audit: NONE

Data capture: NONE

Edit procedure:

Encoding: EBCDIC

Accelerate Development

Query editing

- Template-enabled
- Content assist with connected database
- Platform specific syntactic and semantic validation
- SQL formatting
- Control on commit and error processing and execution environment
- Access path visualization
- Statistics advice
- Performance visibility and validation
- Deployment management

Routine development

- Template enabled
- SQL/PL, PL/SQL, Sybase ANTS, Java
- Routine debugging
- Deployment management

Java development

- Query editor features
- Facilitate collaboration with SQL Outline
- Identify SQL hot spots
- Simplify problem isolation to source code

SQL	Run Id	Connection Profile	Annotation	Data Server Time	Network Time	Rows Read	Rows Returned	Rows Returned Ratio	CPU Time	Total Sorts	T	Physical Page I/Os	Logical Page I/Os
select * from sales whe...	sp_621719694...	v97fp4		19	1	41.0	19.0	2.16	0.0	0	-	2.0	2.0
select * from sales	sp_922533502...	v97fp4	second run	41	1	41.0	41.0	1.00	0.0	0	-	2.0	2.0
select * from HUNGPLE...	sp_541254062...	v97fp4	3rd run	9	1	70.0	2.0	35.00	0.0	0	-	1.0	2.0
select * from HUNGPLE...	sp_524356021...	v97fp4		35	1	70.0	0.0	-	0.0	0	-	2.0	2.0

Save Time and Reduce Errors

Complex change management

- Multiple objects at a time
- Automatic dependency handling
- Compare objects side by side
- Data and privilege migration
- Undo scripts

Maintenance

- Configuration and utilities task assistants

Data management

- Browsing, editing, copying, loading, unloading

Job Management

- Scheduling jobs, view applications, view history

The screenshot displays the IBM DB2 Information Management software interface. The top window, titled 'SAMPLE', shows a 'Default Change Plan' with a table list. The table list has columns: Schema, Name, Percen, Row Count, Log Mode, Primary Key, Partition M, and Regular Tables. The tables listed are: ARLANF.ACT, ARLANF.CATALOG, ARLANF.CL_SCHED, ARLANF.CUSTOMER, ARLANF.DEPARTMENT, ARLANF.EMPLOYEE, ARLANF.EMPMDC, ARLANF.EMPPROJACT, ARLANF.EMP_PHOTO, and ARLANF.EMP_RESUME. A callout bubble points to the 'Tables' button in the toolbar, labeled 'Implicit change plan creation'. Another callout bubble points to the 'Change plan navigation' button, labeled 'Change plan navigation'.

The bottom window, titled 'Properties', shows the 'Table ACT' properties. The 'Columns' tab is active, displaying a table with columns: Name, Primary Key, Data Type, Length, Scale, Not Null, Generate, and Default Value. The columns listed are: ACTNO (Primary Key, SMALLINT, Not Null), ACTKWD (CHAR, Length 6, Not Null), ACTDESC (VARCHAR, Length 20, Not Null), and NEWCOLUMN (CHAR, Length 5). A callout bubble points to the 'NEWCOLUMN' row, labeled 'Change in place on properties tab'.

For data movement & transformation (ELT)

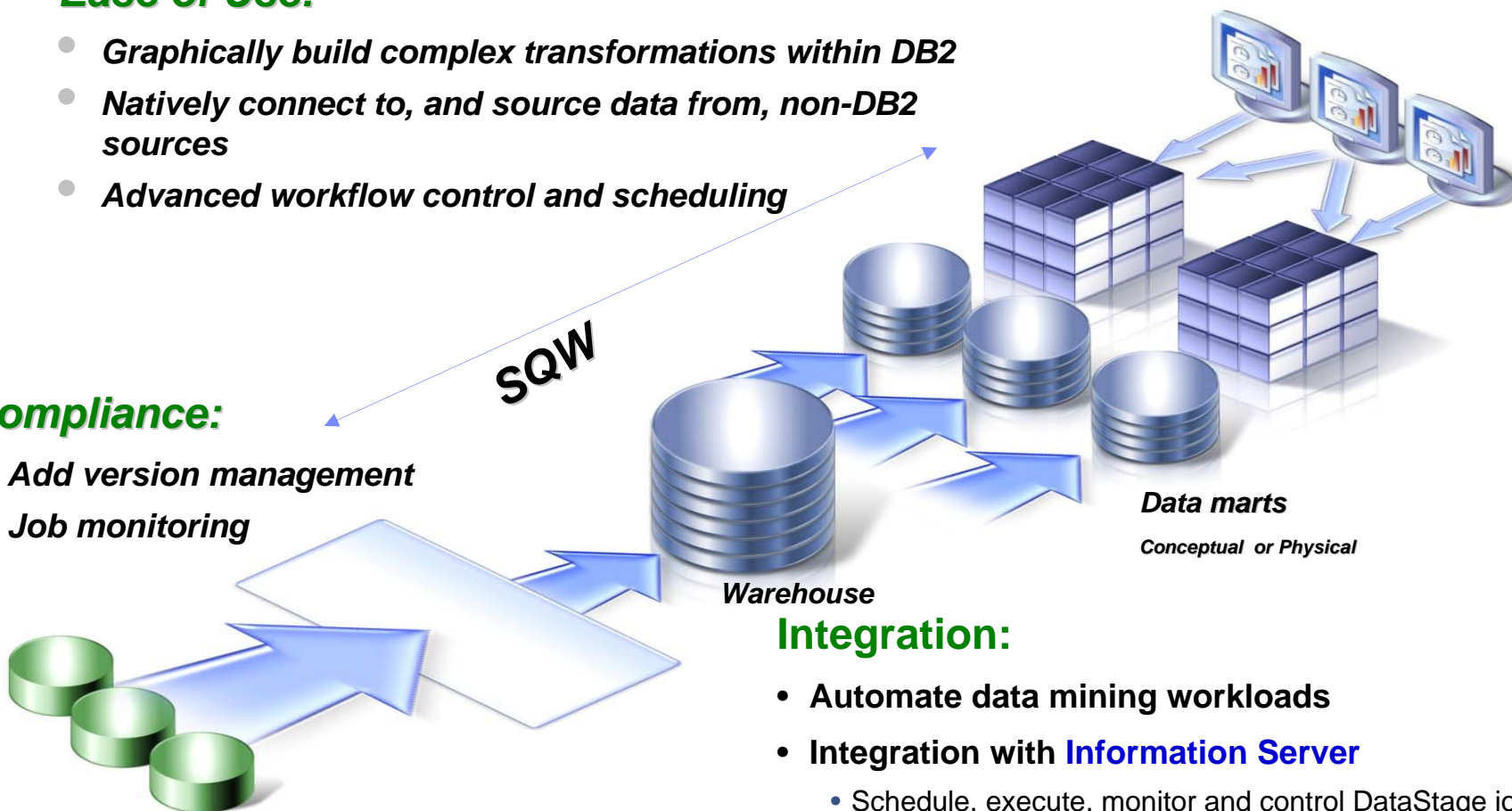
Design Studio includes the SQL Warehousing Tool

Ease of Use:

- Graphically build complex transformations within DB2
- Natively connect to, and source data from, non-DB2 sources
- Advanced workflow control and scheduling

Compliance:

- Add version management
- Job monitoring

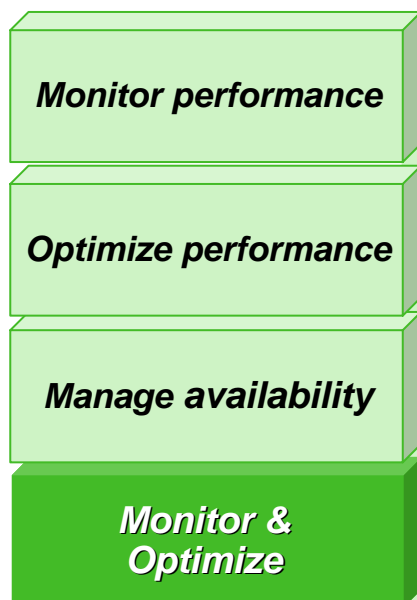


Integration:

- Automate data mining workloads
- Integration with **Information Server**
 - Schedule, execute, monitor and control DataStage jobs
 - Integrate DataStage work streams

InfoSphere Optim Solutions to Monitor and Optimize

*Meet service level objectives
Improve service
Reduce costs*



- **Performance Manager**
 - Monitor database performance and health
 - Align resource use with business priorities
 - Get expert advice on queries, indexes, statistics, and access plans for entire workloads
 - Optimize and stabilize performance of Java applications
- **Query Workload Tuner**
 - Get expert advice on queries, indexes, statistics, and access plans
 - Leverage monitoring data to select high cost queries
- **DB2 Advanced Recovery Solutions**
 - Minimize impact of database backups
 - Minimize recovery time
- **Replication Server**
 - Disaster recovery
 - Active - active

Optimize query performance

Optimize recovery time

Support high availability scenarios

Monitoring and Optimizing Performance

Get early warning of potential problems

Diagnose database problems with resource specific

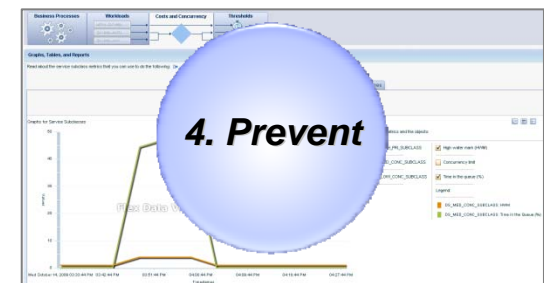
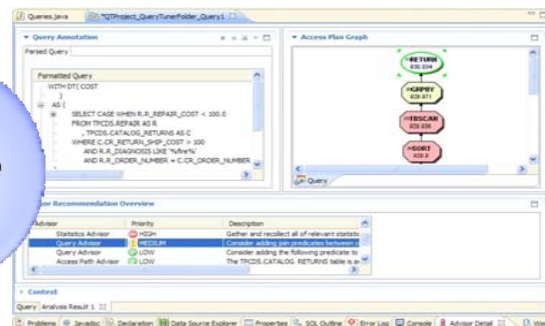
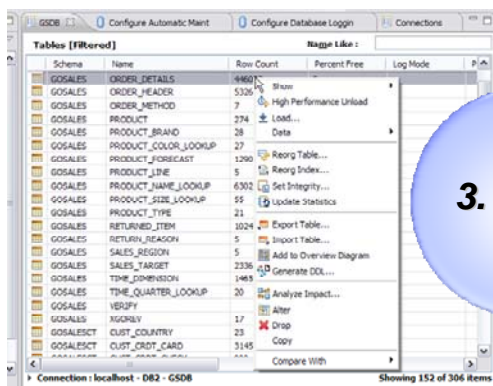


InfoSphere Optim Performance Manager Extended Edition

Resolve database problems

Resolve query problems

Prevent problems



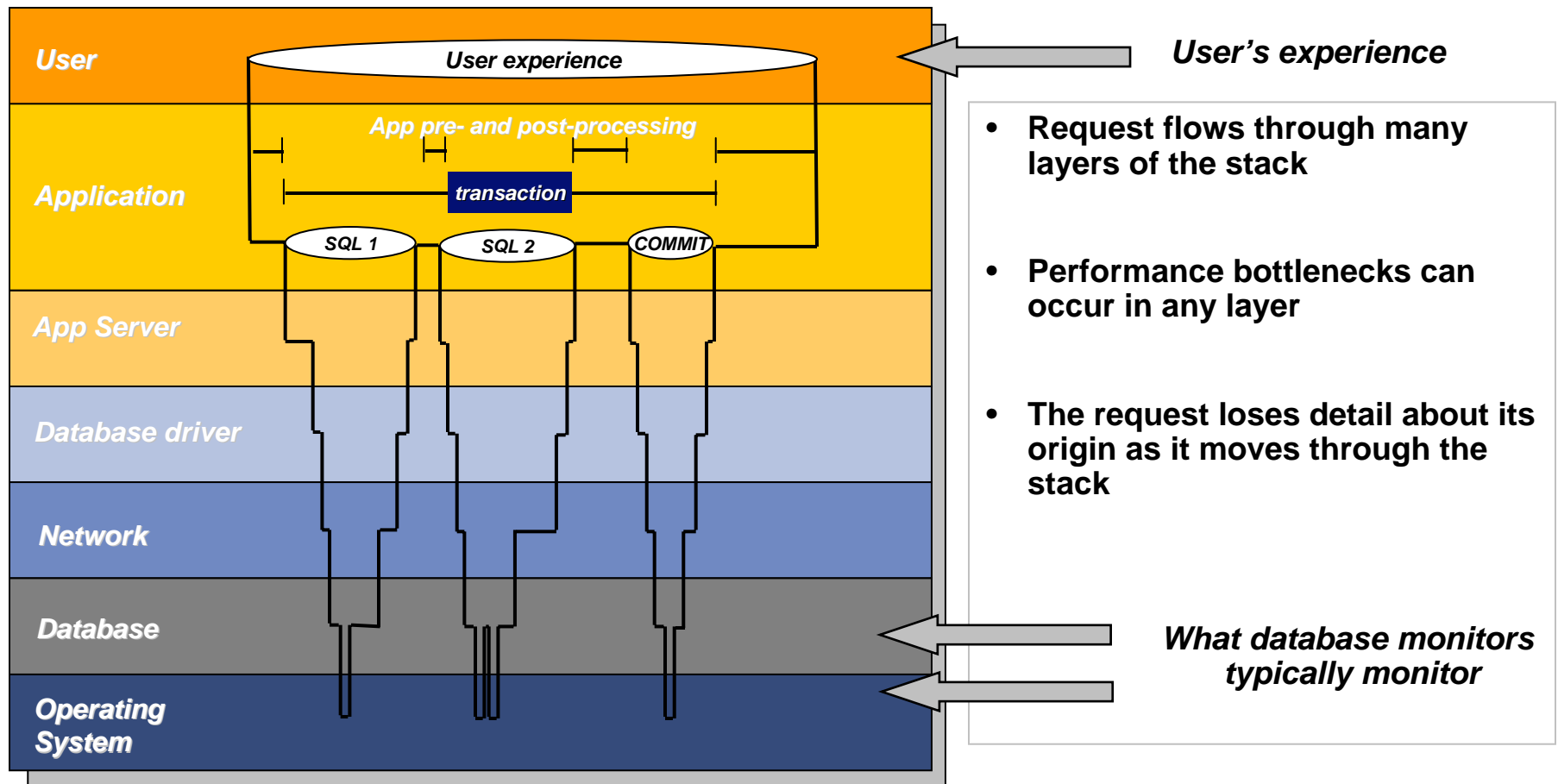
Data Studio

InfoSphere Optim Query Tuner

DB2 Workload Manager



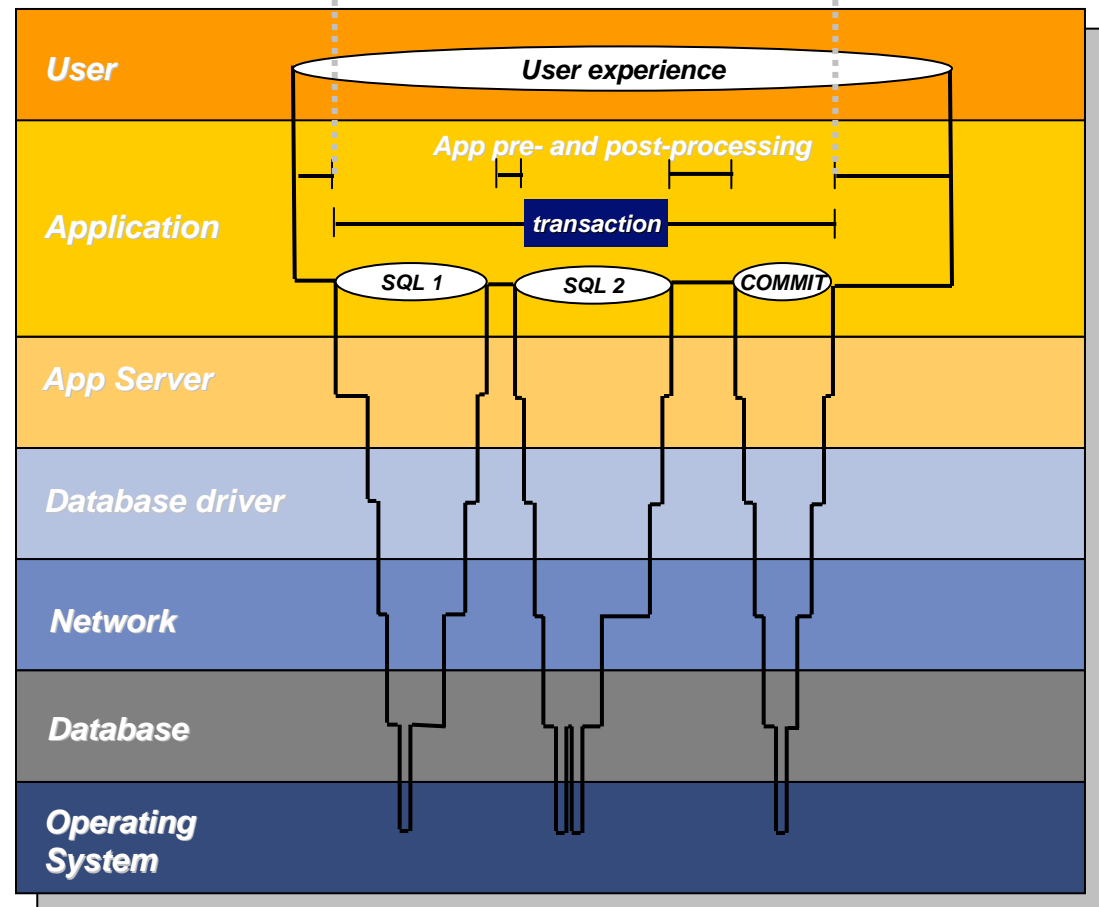
Database Monitor Limitations



InfoSphere Optim Performance Manager Extended Insight

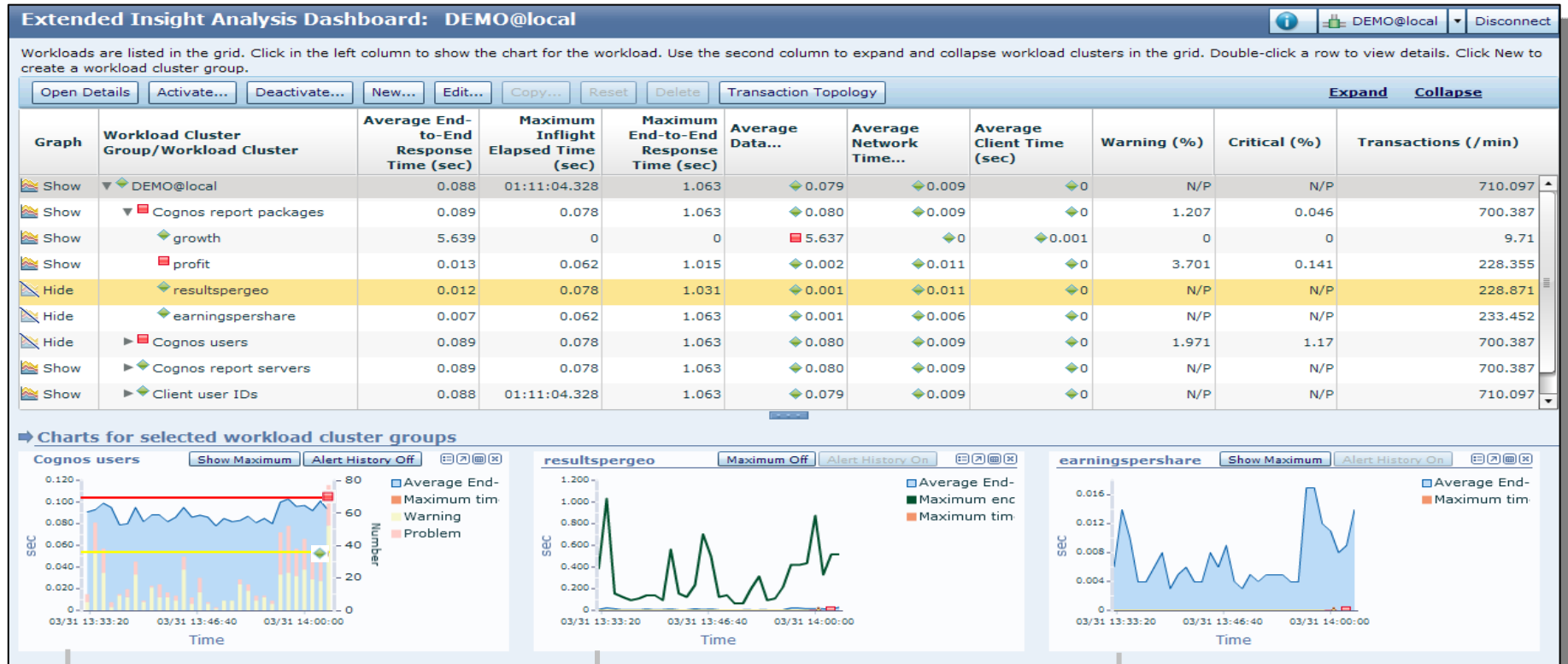
Extended Insight measurement

- **Manage response time SLAs**
 - By workload
- **Pinpoint the problem**
 - Problem workload
 - Problem layer
 - App server, driver, network, database, operating system
 - Problem timeframe
 - Problem SQL
- **Down to database component**
 - Breakdown database time by transaction or statement*
 - Compilation, workload manager queue, lock wait, I/O, sort, routine, transaction end,...
 - * with DB2 9.7



Transaction Response Time Monitoring

By Workload



Response time objective violations

Maximum transaction response time

Average transaction response time



Analyze Workload Response Time

Shows where workload is spending time

Provides detail on queries, clients, and partitions

Extended Insight Analysis Dashboard: DEMO@local

Locate the source of performance problems, determine how those problems affect different parts of the workload, and analyze the performance of individual SQL statements, clients, and partitions.

Response Time Details: profit

Graph | Grid

Selected layer: Average End-to-End Response Time | Show Maximum

SQL Statements | Clients | Partitions/Members

Show highest 10 by Average Data Server Time (sec)

Statement Text	Statement Executions	Average Data Server Time (sec)	Average End-to-End Response Time (sec)
INSERT INTO WE_ZXZ...	140	0.001	0.049
INSERT INTO WE_ZXZ...	139	0.002	0.020
INSERT INTO WE_ZXZ...	150	0.002	0.020

Detail Area for SQL Statements

Statement information: INSERT INTO WE_ZXZRFJP1G.WE_1G0_85894 (value) values(100)

Statement Performance:

- Number of Executions: 140
- Average end-to-end elapsed time: 0.049 sec
- Average client time: 0 sec
- Average driver time: 0 sec
- Average network time: 0.048 sec
- Average data server time: 0.001 sec

Statement Time Distribution (%): Client time, Driver time, Data server time

Launch Optim Query Tuner

Provides detail on selected area

InfoSphere Optim Query Workload Tuner

Expert-Enabled Query Tuning Out of the Box

- **Identify query candidates from numerous sources**
 - DB2 catalog
 - Package cache
 - Development Studio hot spots
 - Optim Performance Manager
- **Facilitate analysis**
 - Query formatting
 - Query annotation
 - Access path visualization, annotation, and warnings
- **Get expert tuning advice**
 - Improve query design
 - Improve statistics quality
 - Improve database design

The screenshot displays the 'Query Tuner Workflow Assistant' interface. The left sidebar shows a workflow with steps: 1. Status, 2. Capture, 3. Manage, 4. Invoke, 5. Review, and 6. Compare. The main area shows 'Analysis Results' for 'Query 3', including 'Analysis Result 1' with a tuning report. Below this, there are options to 'Open Single-Query Recommendation', 'Open Formatted Query', 'Open Access Plan Graph', 'Open Access Plan Explorer', 'Open Summary Report', and 'Compare Access Plan Graphs'. The right pane is titled 'Review Single-Query Advisor Recommendations' and shows a table of recommendations for 'Analysis Result 1'.

Recommendation	Number	Priority	Description
Statistics	1	HIGH	Gather and recollect a
Query revision	2	LOW	Provide a predicate of
Query revision	3	LOW	Provide a predicate on
Access path	4	LOW	Avoid reading all index
Index	5	LOW	Index recommendatio

Analyze single queries or query workloads

Improve Query Design

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

Recommendation	Number	Priority	Description
Shown			
Query revision	1	LOW	Provide a predicate on column SR_RETURNED_DATE_SK.
Access path	2	LOW	Avoid table space scan (Operator ID = 6) on table TPCDS.STORE_RETURNS.
Access path	3	LOW	Avoid table space scan (Operator ID = 7) on table TPCDS.DATE_DIM.
Index	4	LOW	Index recommendations found.
Statistics	5	MAINTENANCE	Gather and recollect all of relevant statistics for this query.

Filter recommendations by severity

Summary - 1 Initial Analysis **Query Revision**

Recommendation 1: Provide a predicate on column SR_RETURNED_DATE_SK.

```

SQL Text
SELECT TPCDS.STORE_RETURNS.SR_CUSTOMER_SK AS CTR_CUSTOMER_SK
, TPCDS.STORE_RETURNS.SR_STORE_SK AS CTR_STORE_SK
, SUM(TPCDS.STORE_RETURNS.SR_RETURN_AMT) AS CTR_TOTAL_RETURN
FROM TPCDS.STORE_RETURNS
, TPCDS.DATE_DIM
WHERE TPCDS.DATE_DIM.D_YEAR = 2000
AND TPCDS.STORE_RETURNS.SR_RETURNED_DATE_SK = TPCDS.DATE_DIM.D_DATE_SK
GROUP BY TPCDS.STORE_RETURNS.SR_CUSTOMER_SK,
TPCDS.STORE_RETURNS.SR_STORE_SK
    
```

Highlight 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
 - Use local predicates for joining tables or single table.

Recommendation Details

Provide a predicate on column SR_RETURNED_DATE_SK.

Consider adding the following predicate to column SR_RETURNED_DATE_SK in table STORE_RETURNS:
 SR_RETURNED_DATE_SK IS NOT NULL
 to filter the table earlier and unlock more possible join sequences.

Explanation

If a predicate filters out NULL columns and data type mismatches earlier in processing, performance can be improved.

Improve Statistics Quality and Collection

Statements Summary **Statistics**

The following table lists all of the tables that are referenced in the query workload. Highlight a row to see the statistics

Automatic statistics collection: is enabled

6 Tables have statistics recommendations 0 Tables has all the relevant statistics

Tables that have recommendations

Table	Type	Status	SQL Reference	Cardinality	Cumulative Total Query Cost	Collection Timestamp
TPCD.PART	TABLE	PR...		2	200000	52240840956837888
TPCD.ORDERS	TABLE	PR...		2	1500000	52240840956837888
TPCD.PARTSUPP	TABLE	PR...		2	800000	52240840956837888
TPCD.NATION	TABLE	PR...		3	25	52240840956838800
TPCD.LINEITEM	TABLE	PR...		5	6001215	52240840959557688
TPCD.SUPPLIER	TABLE	PR...				2011-06-03 15:15:07

Detail info for selected row in the summary table

- Provides advice on
 - Missing statistics
 - Conflicting statistics
 - Out-of-date statistics
 - Statistical views
- Simplify use
 - Provides Runstats control statements
 - Run immediately or save

Table Details

TPCD.LINEITEM has missing conflicting table statistics

Columns Frequency Values

Column	Cardinality	LowZKey	HighZKey	Frequency Count	Sum of Freque...	Quantiles
L_ORDERKEY	0			0	0	
L_DISCOUNT	0			0	0	
L_TAX	0			0	0	
L_PARTKEY	0			0	0	
L_SUPPKEY	0			0	0	
L_EXTENDEDPRICE	0			0	0	
L_SHIPDATE	0			0	0	

*WARNING: The following columns have statistics information available in the catalog but are not referred in the workload. The statistics for these dropped if 'non-consolidated' RUNSTATS commands are run. The 'consolidated' RUNSTATS may contain all/subset of these columns.

Column	Cardinality	LowZKey	HighZKey	Frequency Count	Sum of Freque...	Quantiles
L_SHIPINSTRUCT	4	'DELIVER IN PE...	'NONE	0	0	
L_COMMENT	6001215	'0 6L2BQA0Ph...	'zzzzQ3n0C1R...	0	0	
L_RETURNFLAG	3			0	0	
L_LINENUMBER	-1			0	0	
L_SHIPMODE	7	'FOB	'SHIP	0	0	

View details statistics and status before deciding on collecting recommended statistics

RUNSTATS Commands for Selected Recommendations

Select Command Type and Final Selection

Select consolidated or non-consolidated commands for each relevant recommendation

Consolidated commands for all RUNSTATS recommendations
 New RUNSTATS recommendations are merged with RUNSTATS stored on the server

Non-consolidated commands for all RUNSTATS recommendations
 Select command type for each individual RUNSTATS recommendation

Recommendations

Tables	Command Type
TPCD.PART	Consolidated
TPCD.ORDERS	Consolidated
TPCD.PARTSUPP	Consolidated
TPCD.NATION	Consolidated
TPCD.LINEITEM	Consolidated
TPCD.SUPPLIER	Consolidated

Generates RUNSTATS control statements

- Better query performance
- Reduced CPU consumption
- Improved maintenance window throughput

Workload Statistics Advisor and Statistical Views

- **Challenge**
 - Data skew and join predicate filter [mis]calculation
 - WHERE salary+bonus > ...
 - Overloaded dimension in a star join
 - Time dimension may contain 20 years of data but the transaction table keeps only the last 5 quarters of data. Uniform distribution is the wrong assumption after the join.
- **Solution: Generate statistical views**

Estimated RUNSTATS execution time: 656,012.569580 ms
 Recommended new statistical views: 3
 Recommendations for existing statistical views: 0
 Existing statistical views without recommendations: 0

Recommendations:

	Owner Name	Owner Name	State	Tables in Definition	Affected Statements	Current Estimated Execu...	Estimated Time to R...	Statist
<input checked="" type="checkbox"/>	cleung	STATSVIEW201...	PROPOSED	4	3	3,463,315.187500	647,308.250000	NO
<input checked="" type="checkbox"/>	cleung	STATSVIEW201...	PROPOSED	2	1	497,285.312500	7,888.009277	NO
<input checked="" type="checkbox"/>	cleung	STATSVIEW201...	PROPOSED	2	3	3,053,967.533585	816.310303	NO

Recommendation Details:

Statement Text	Number of Executions	Current Time to Execute Once (ms)	Current Estimated Cumulative Execution Time (ms)
selectnation,o_year,su...	1	497,285.312500	497,285.312500
selectnation,o_year,su...	1	2,556,499.750000	2,556,499.750000
selectn_name as nation...	1	182.471085	182.471085

Indexing Advice to Improve Query Efficiency

- **Improve query efficiency**
 - Indexing foreign keys in queries that do not have indexes defined
 - Support for index only access
- **Simplify use**
 - Consolidate indexes and provide a single recommendation
 - Enables what-if analysis
 - Shows estimated improvement
- **Test before deployment**
 - Utilize virtual index capabilities built into the DB2 engine

The screenshot displays the 'Workload Index Advisor' interface. At the top, it shows performance metrics: Estimated performance improvement: 68.52 %, Disk space required (DASD space): 673.08 MB, and Total IUD Cost: 0.0 MS. A callout bubble highlights the 'Estimated performance improvement' value. Below this, a 'Recommendation' table lists several indexes to be created for tables like CUSTOMER, ORDER, and LINEITEM. A 'DDL for Selected Indexes' window is open at the bottom, showing the SQL commands for creating these virtual indexes.

Feature Details	Action	Object N...	New Index Columns	Old Inde
[-] <input checked="" type="checkbox"/> Create Indexes				
[-] <input checked="" type="checkbox"/> CUSTOMER				
<input checked="" type="checkbox"/> Index	Create	CUSTOME...	C_MKTSEGMENT(ASC), C_...	
<input checked="" type="checkbox"/> Index	Create	CUSTOME...	C_MKTSEGMENT(ASC), C_...	
[-] <input checked="" type="checkbox"/> 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...	
<input checked="" type="checkbox"/> Index	Create	ORDER_V...	O_ORDERSTATUS(ASC), O...	
[-] <input checked="" type="checkbox"/> LINEITEM				
<input checked="" type="checkbox"/> Index				
<input checked="" type="checkbox"/> Index				

```

CREATE INDEX DB2OE.LINEITEM_VIRT_IDX_1159535563777 ON SYSADM.LINEITEM (
L_SUPPKEY ASC, L_EXTENDEDPRI ASC, L_TAX ASC) FREEPAGE 0 PCTFREE 10;

CREATE INDEX DB2OE.ORDER_VIRT_IDX_1159535563657 ON SYSADM.ORDER ( O_CUSTKEY
ASC, O_ORDERDATE ASC, O_SHIPPRIORITY ASC, O_ORDERKEY ASC) FREEPAGE 0 PCTFREE
10;

CREATE INDEX DB2OE.CUSTOMER_VIRT_IDX_1159535564238 ON SYSADM.CUSTOMER (
C_MKTSEGMENT ASC, C_NATIONKEY ASC, C_ACCTBAL ASC) FREEPAGE 0 PCTFREE 10;

CREATE INDEX DB2OE.ORDER_VIRT_IDX_1159535554165 ON SYSADM.ORDER (
O_ORDERKEY ASC, O_ORDERDATE ASC) FREEPAGE 0 PCTFREE 10;

CREATE INDEX DB2OE.LINEITEM_VIRT_IDX_1159535563807 ON SYSADM.LINEITEM (
L_QUANTITY ASC, L_EXTENDEDPRI ASC, L_TAX ASC) FREEPAGE 0 PCTFREE 10;

CREATE INDEX DB2OE.PARTSUPP_VIRT_IDX_1159535562295 ON SYSADM.PARTSUPP (

```

IBM DB2 Advanced Recovery Solutions

Save time, reduce errors, and meet SLAs

DB2 Merge Backup

- *Improve speed of your backup & recovery processes*
- *Minimize application impact*

Backup

DB2 Recovery Expert

- *Recover faster with greater granularity while protecting your critical business data*
- *Eliminate data errors before they compound into costly business mistakes*
- *Track and report data changes in response to auditing requests*

Recover

Optim High Performance Unload

- *Extract large amounts of data quickly and with minimal impact on system productivity*
- *Perform full data and system migrations from one DB2 instance to another*

Unload

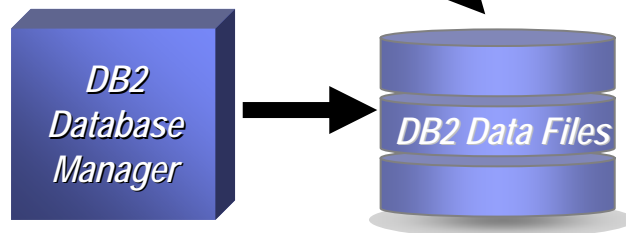


InfoSphere Optim High Performance Unload

Rapidly unload, extract, and repartition data across enterprise systems

```
SELECT * FROM MY_TABLE
```

Optim High Performance Unload



Total Data Unloaded:	2,905,294,900 rows
Complete database export time by usual methods	23 hours 54 minutes
Time Taken by DB2 HPU	1 hour 30 minutes

“That’s about 16 times faster without using any special features, just out of the box performance.”

Requirements

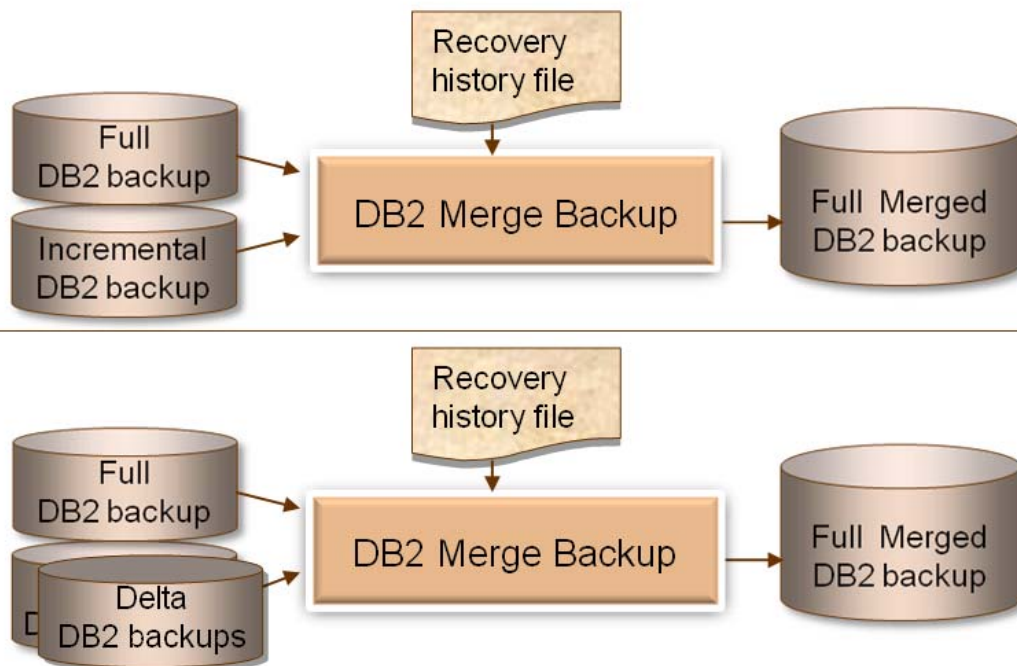
- Quickly move large volumes of data between different sources without affecting the production system
- Reduce storage costs
- Fit more work into shrinking batch windows

Benefits

- Meet SLA’s with high speed data unload capabilities
- Move data between systems while production systems are fully functioning
- Eliminate the need for intermediate storage
- Shorten data recovery time

DB2 Merge Backup

Accelerate DB2 backups for improved availability



Requirements

- *Meet SLAs for accuracy and responsiveness*
- *Perform backup/recovery with minimal impact*
- *Scaled approach for responding to unplanned service interruptions*
- *Ensure a complete, accurate DB2 backup*

Benefits

- *Perform backup and recovery faster*
- *Create incremental backup to minimize production impact*
- *Merge incremental with full backups to meet SLAs*

DB2 Merge Backup Scenario

Scenario

If you have the following backup schedule

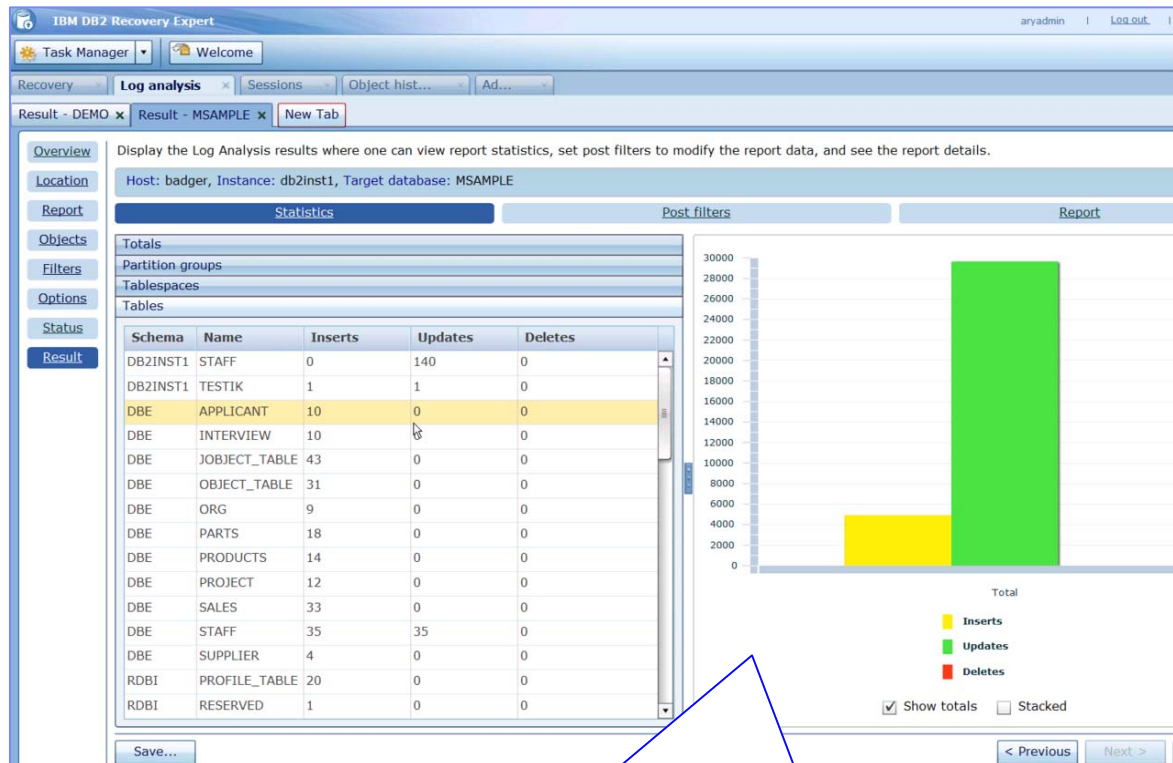
- Sunday – Full Backup
- Monday – Delta Backup, Create new full backup using DB2 Merge Backup
- Tuesday – Delta Backup, Create new full backup using DB2 Merge Backup
- Wednesday – Delta Backup, Create new full backup using DB2 Merge Backup
- Thursday – **BANG**. Table is accidentally dropped

- **Use DB2 Merge Backup and the delta backup to use in the recovery**
- **Without DB2 Merge Backup, you would need to restore from the full backup from Sunday and 3 delta backups!**

DB2 Recovery Expert

Reduce recovery window to increase availability

enterprise systems



Use Log Analysis to figure out who is doing what with your database

Requirements

- Recover lost data with minimal impact to the business
- Scaled approach for responding to accidental data loss or corruption
- Ensure a complete, accurate DB2 backup

Benefits

- Significantly reduce data recovery times
- Allows DBAs to quickly restore or correct erroneous data with less resource and minimal disruption

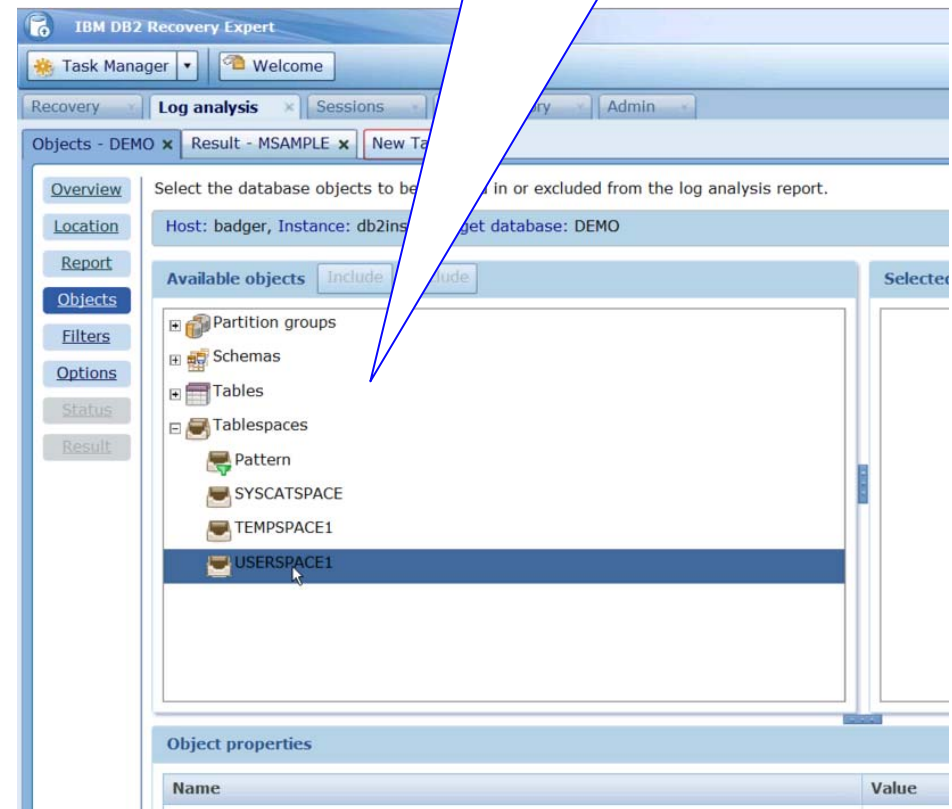
DB2 Recovery Expert

Scenarios

An accident occurs

- I ran the weekly batch job instead of the monthly job
- I dropped a production table instead of a test table
- I ran a delete with the wrong date parameters and ended up deleting too many rows

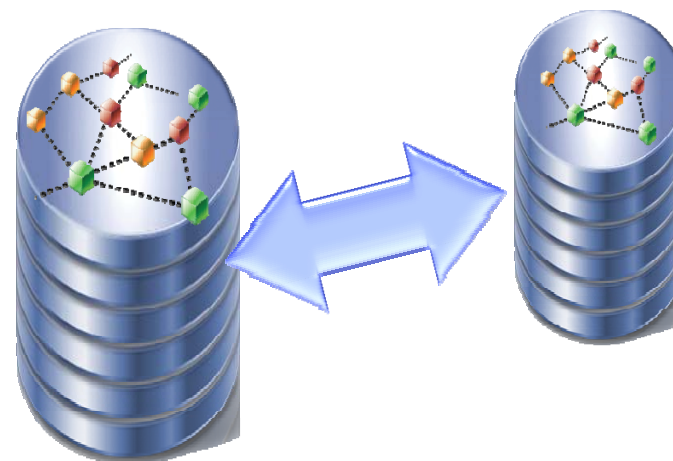
Web UI helps to easily navigate DB2 logs



High Availability

InfoSphere Replication Server

- **Bi Directional Synchronisation**
 - Ability to synchronize across sites in real-time
 - Synchronize Entire Warehouse or Subsets
 - Extendable to multiple sites
- **Seamless Customer Failover**
 - Automatic Failover of Clients to Secondary Sites
 - Seamless to Applications
- **Rolling Upgrades**
 - Ability to Upgrade Warehouse environments sequentially without loss of service
- **Active – Active**



IBM InfoSphere Warehouse

Solve problems – Create new opportunities



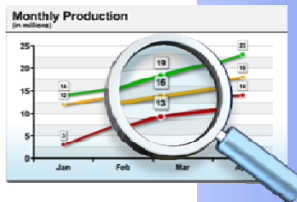
FLEXIBILITY

Everything to manage your data across its lifecycle



PERFORMANCE

DB2 9.7 with partitioning, compression, and workload management



SOPHISTICATED ANALYTICS

In database analytics and Cognos reporting

IBM Smart Analytics System Offerings Deliver Analytics Value Without Forcing Migration

Faster Time to Value, Faster Business Results

Meeting clients where their information is...

5710 **New**

Based on System x

- ...Cost-effective solutions for analytics and BI, reporting
- ...Compact, integrated single analytics solutions
- ...Available for mid-market



5600

Based on System x

- ...Designed for business analytics workloads
- ...Optional Solid State Disk – reducing data latency

7700

Based on POWER7 Servers

- ...Scaling to hundreds of terabytes of data
- ...Extract insights from untapped information

7710

New

Based on Power7 Server

- ...A single server warehousing and analytics solution
- ...Built on POWER7 based servers and designed for production data warehouses - sizes under 10 TB,
- ...For development and non-production use.

9600

Based on Based on System z

- ...Advanced query /workload management
- ...Database designed and optimized for system
- ...Disk controller, optimized to reduce data latency



Communities

- **On-line communities, User Groups, Technical Forums, Blogs, Social networks, and more**
 - Find the community that interests you...
 - **Information Management** ibm.com/software/data/community
 - **Business Analytics** ibm.com/software/analytics/community
 - **Enterprise Content Management** ibm.com/software/data/content-management/usernet.html
- **IBM Champions**
 - Recognizing individuals who have made the most outstanding contributions to Information Management, Business Analytics, and Enterprise Content Management communities
 - ibm.com/champion

More information

- **Analyst report from ITG**
 - [Bottom-line Advantages of IBM InfoSphere Warehouse](#)
- **IBM White Paper**
 - [IBM InfoSphere Warehouse 9.7: Deliver actionable business insights](#)

Next DB2 Tech Talk Coming in March:

Database Archiving: Best practices to improve performance of your DB2 LUW and z/OS databases

Registration coming soon to [DB2 Tech Talk page](#)

Questions



Thank You and Follow-up!

ibm.com/db2/labchats



Acknowledgements and Disclaimers:

Availability. *References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates.*

The workshops, sessions and materials have been prepared by IBM or the session speakers and reflect their own views. They are provided for informational purposes only, and are neither intended to, nor shall have the effect of being, legal or other guidance or advice to any participant. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided AS-IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other materials. Nothing contained in this presentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.

© Copyright IBM Corporation 2011. All rights reserved.

- *U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.*

IBM, the IBM logo, ibm.com, DB2, InfoSphere, and Optim are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml

Other company, product, or service names may be trademarks or service marks of others.

