



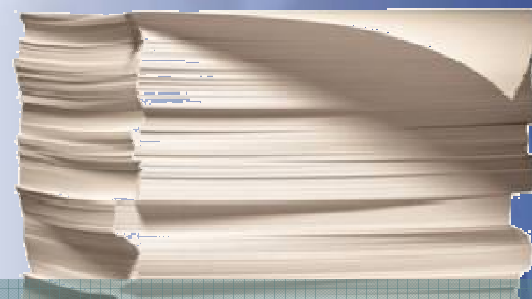
IBM Information Management

# DB2 Warehouse 9.5

*What's New in the Platform for  
High Performance Analytics*

DB2 Chat with the Lab

November 7, 2007



Pat Bates, [jpbates@us.ibm.com](mailto:jpbates@us.ibm.com)  
Product Manager, DB2 Warehouse Analytics

# Agenda

- DB2 Warehouse 9.5 Introduction
- No Copy Analytics
  - Text Analytics
  - Mining and Cubing Services
- Embedded Analytics Delivery
- High Performance Analytics
  - Workload Management
  - Performance Monitoring and Analysis



# Dynamic Warehousing

## *The Foundation for Effective Embedded Analytics*

**Information On Demand**  
to Optimize Real-Time  
Processes



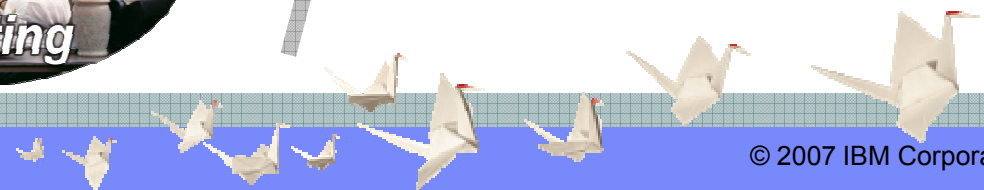
*Dynamic  
Warehousing*

**OLAP & Data Mining**  
to Understand Why and  
Recommend Future Action



*Traditional Data  
Warehousing*

**Query & Reporting**  
to Understand  
What Happened



# Dynamic Warehousing

## *The Foundation for Effective Embedded Analytics*

**Information On Demand**  
to Optimize Real-Time  
Processes

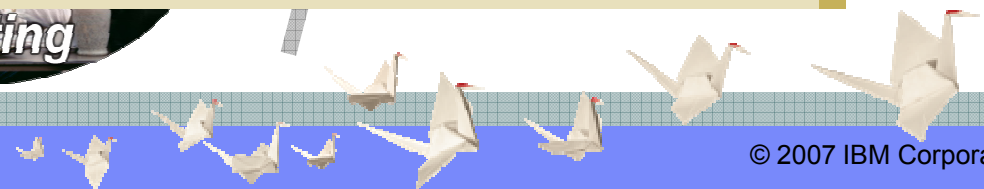


**Dynamic  
Warehousing**

### **Dynamic Warehousing Requires:**

1. *Real-time access – in context*
2. *Analytics – as part of a business process*
3. *Unstructured information – extracted knowledge*
4. *Extended infrastructure – tightly integrated*

Reporting



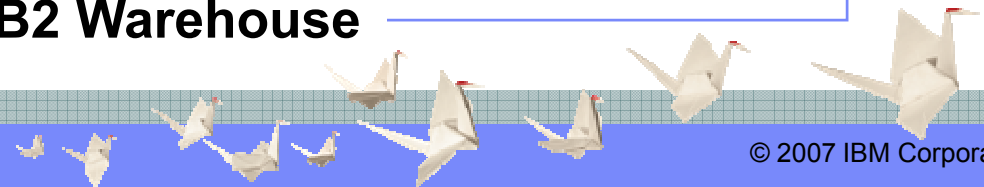
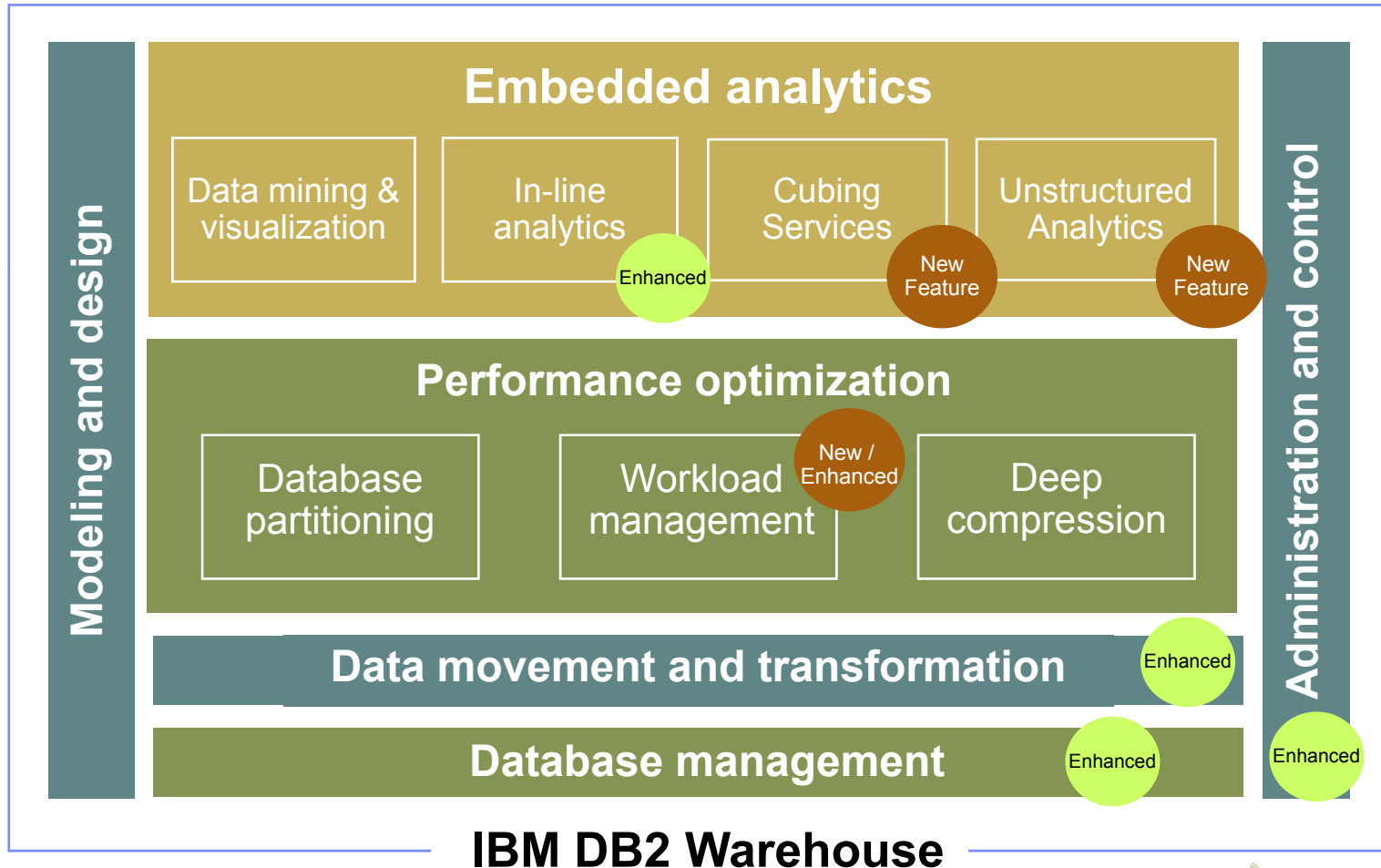
# Introducing DB2 Warehouse 9.5

**Insight without Boundaries**  
Reach Farther. Look Deeper. Act Faster.



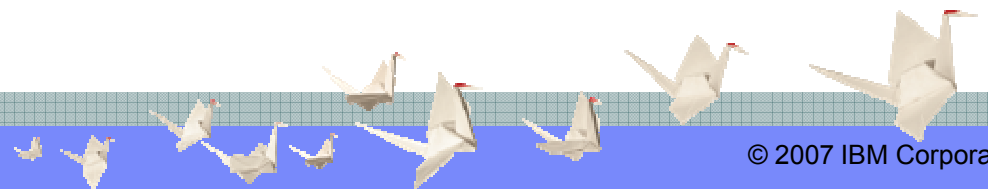
# IBM DB2 Warehouse 9.5

*A complete, integrated platform for End-to-End Analytics*



# DB2 Warehouse 9.5 Feature Highlights

- Multidimensional Analytics
  - Cubing Services - Open interface for cubing over DB2 data
- Text Analytics
  - Integration of unstructured data into analytics
- Extreme Workload Management
  - DB2 Warehouse Tooling for Workload Management
- In-Line Analytics
  - Blox Builder -- Easy to Build development tooling



# Reach farther. Look deeper. Act faster.



## Embedded unstructured analytics

Leverage previously untapped information to better understand customer and product issues



## Embedded OLAP cubing services

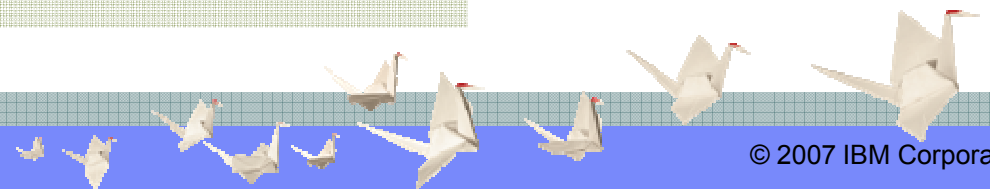
Enable deeper analysis of multiple business variables and dimensions to generate insight, while reducing cost



## Extreme workload management

Deliver real-time insight to more users and reduce costs, without compromising performance

## No copy analytics





# Unstructured Analytics

## The problem...

A **huge amount of** unstructured information (**text**)

- call center notes
- problem reports
- repair reports
- insurance claims
- email with customers
- product reviews

**cannot be used with existing business intelligence tools** to create insight and answer **forward-looking** business questions.

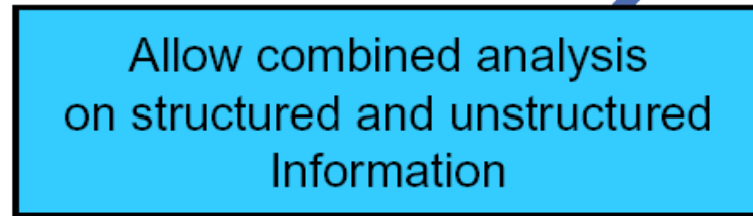
## ...the solution

**Transform unstructured information into structure** that can be **analyzed** in the warehouse **together with existing structured information using existing tool**

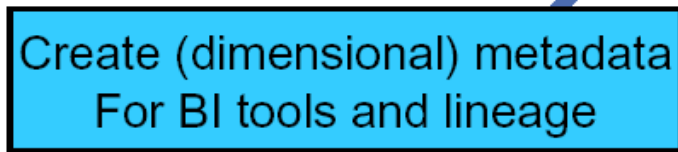
# Unstructured Analytics in DB2 Warehouse 9.5

Solution approach

Insight

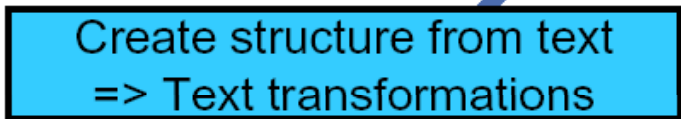


Ranking (top n)  
OLAP  
Data Mining



Cubing Services  
metadata extensions

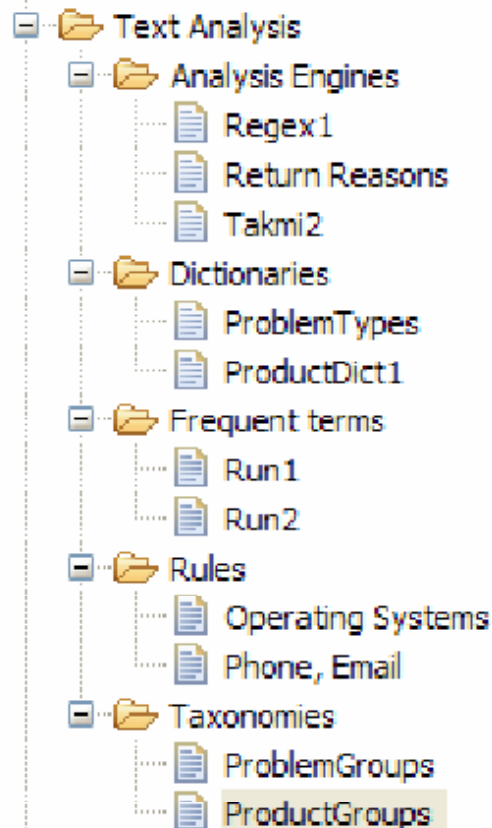
„ELT for Text“



UIMA runtime  
DWE configurable annotators  
IBM Research annotators  
Partner annotators

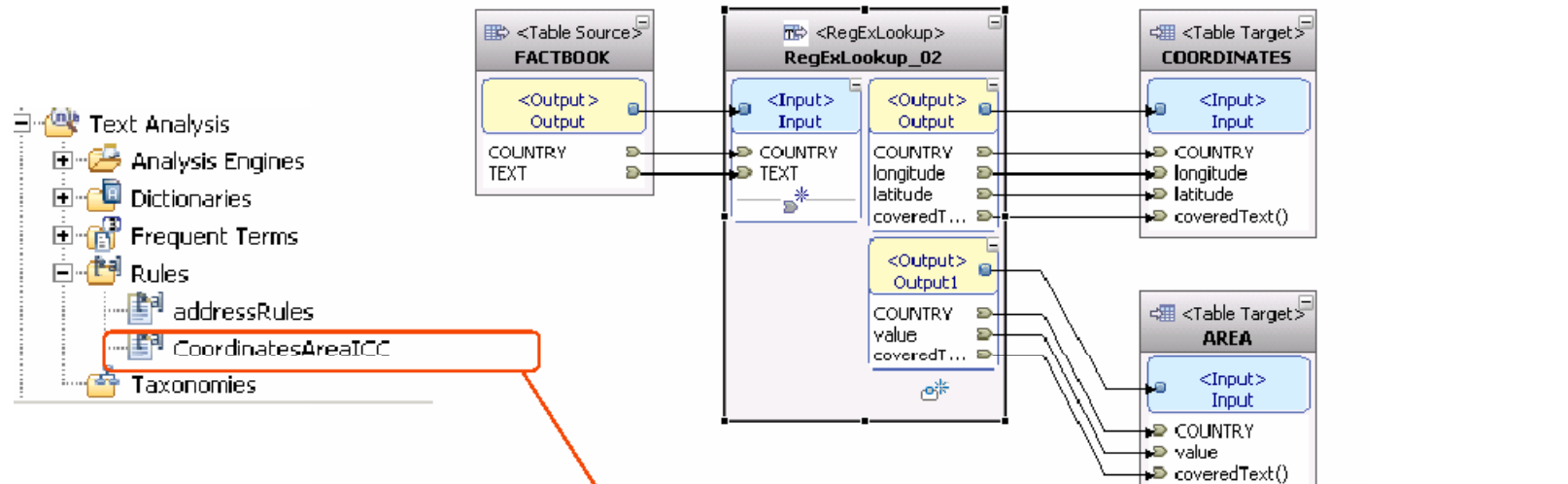


# DB2 Warehouse Design Studio text analysis overview



- Data understanding
  - ▶ View text columns in database, text statistics
  - ▶ frequent terms analysis with NLP in 30 languages
- Dictionary editor
  - ▶ Create named entity dictionary from frequent terms or import from table
- Rule editor
  - ▶ Create regular expression rule set
- Taxonomy editor
  - ▶ Create hierarchical classification for concepts extracted from text
  - ▶ Create dimension table from taxonomy for multidimensional analysis
- Analysis Engines
  - ▶ Import pre-configured analysis engine (from IBM research, business partners, Omnifind solutions)
- Flow editor with text specific operators
  - ▶ Create transformation flow that reads text columns in source tables, applies a dictionary lookup or rule lookup and writes the results to target table
  - ▶ TextAnalyzer, Dictionary Lookup, ItemAggregator and Regular Expression Lookup operators

# Miningflow with regular expression lookup operator



- Text Analysis
- Analysis Engines
- Dictionaries
- Frequent Terms
- Rules
  - addressRules
  - CoordinatesAreaICC**
- Taxonomies

Properties window for 'RegExLookup\_02'.

General: RegExLookup\_02

Settings: Analysis Results

Analysis Results: Output: Output1

Output columns: Rules File: **CoordinatesAreaICC**

Runtime Options: Annotation type: .uima.workbench.demo.coordinates

Typedescription: Geographic location of the country (latitud)

Available Columns: com.ibm.uima.workbench.demo.coordinates(40)

Result Columns:

Expression	Column Name	Data Type
COORDINATES."lon...	longitude	VARCHAR
COORDINATES."latit...	latitude	VARCHAR
COORDINATES."cov...	coveredText()	VARCHAR



# Data Mining in the DB2 Warehouse 9.5

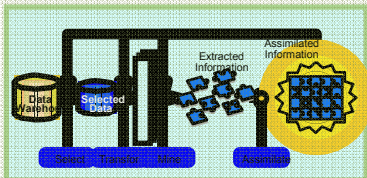


**Business Analyst**

**DWE**



**Partner**



**Statistician & Data Mining Workbench**

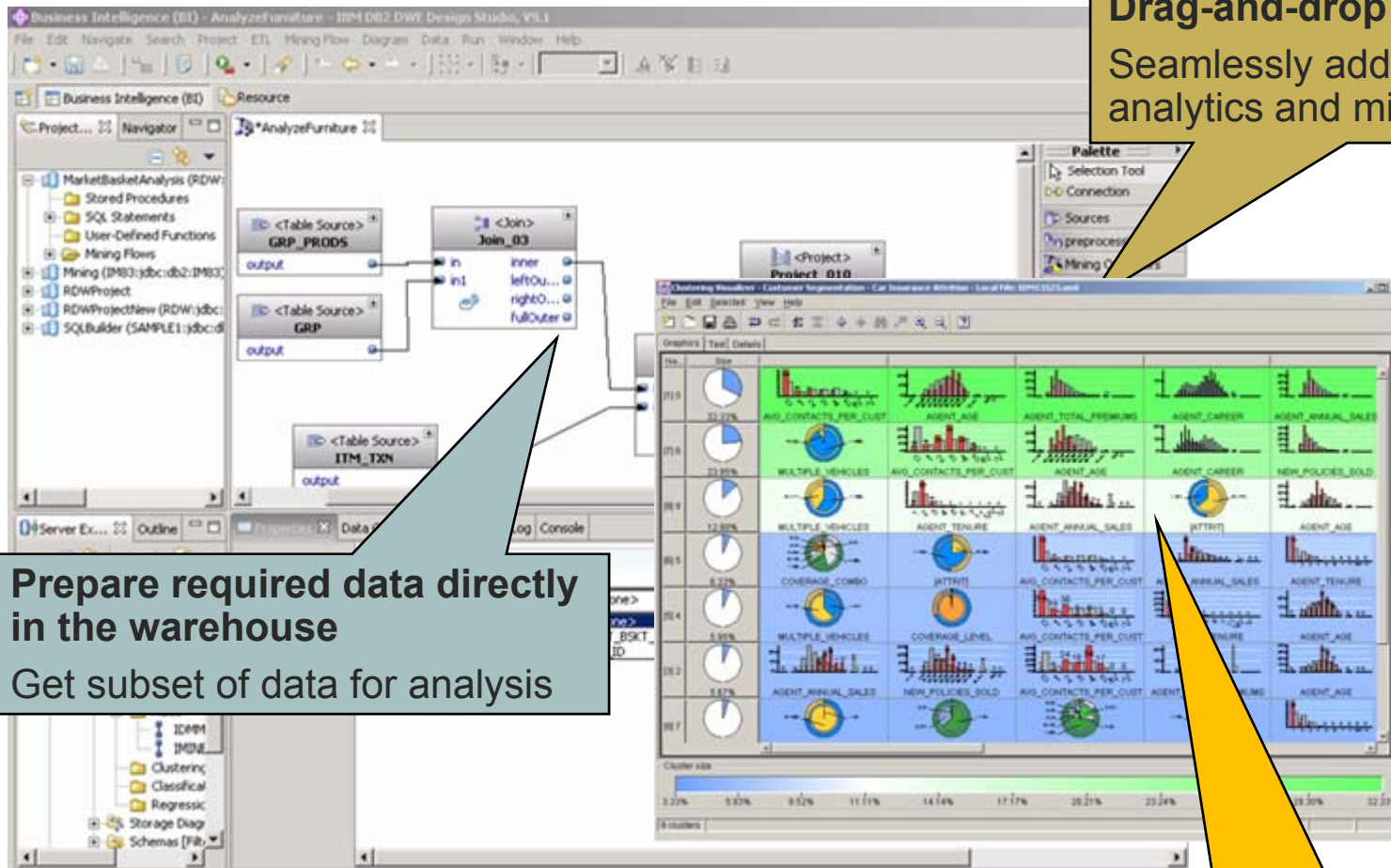
“Easy Mining” algorithms

- **Associations**
  - Which item affinities (“rules”) are in my data?
- **Sequences**
  - Which sequential patterns are in my data?
- **Clustering**
  - Which interesting groups are in my data?
- **Classification**
  - How to predict categorical values in my data?
- **Prediction**
  - How to predict numerical values in my data?

**Score data directly in DB2, scalable and real time**



# Embedded Mining with Integrated Tools



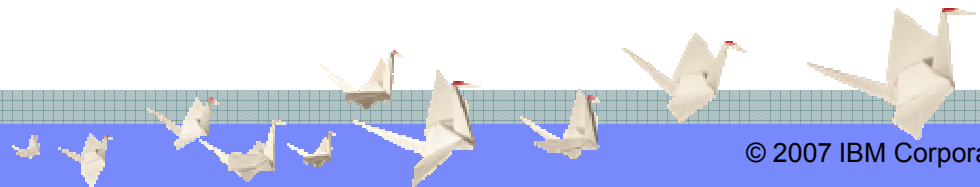
**Drag-and-drop interface**  
Seamlessly add specific analytics and mining operations

**Prepare required data directly in the warehouse**  
Get subset of data for analysis

**Visualize, Test, Debug and Deploy**

# Open Standards, Interoperable, Pervasive

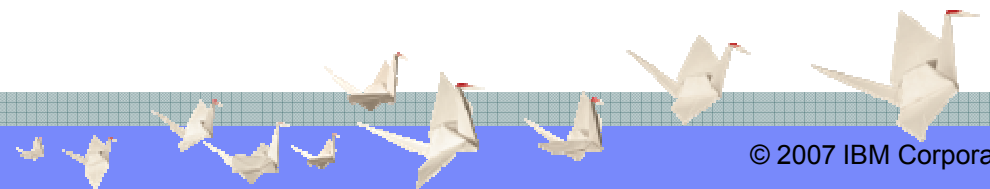
- Modeling and Scoring API: SQL/MM open standard
  - ISO/IEC 13249-6: <http://www.iso.org>
- Created models: PMML open standard (XML)
  - <http://www.dmg.org>
- Every artifact created by DWE 9.1 Data Mining (mining task definitions, models) is stored in XML format in a DB2 table in schema IDMMX. Can be easily accessed and modified from outside DWE 9.1
- Can score PMML models from other vendors (SAS, SPSS, Microstrategy, ...)
- Can invoke mining operations from SQL, Web Services (SOA) or Alphablox.
  - Maximum embeddability!



## What is Cubing Services?

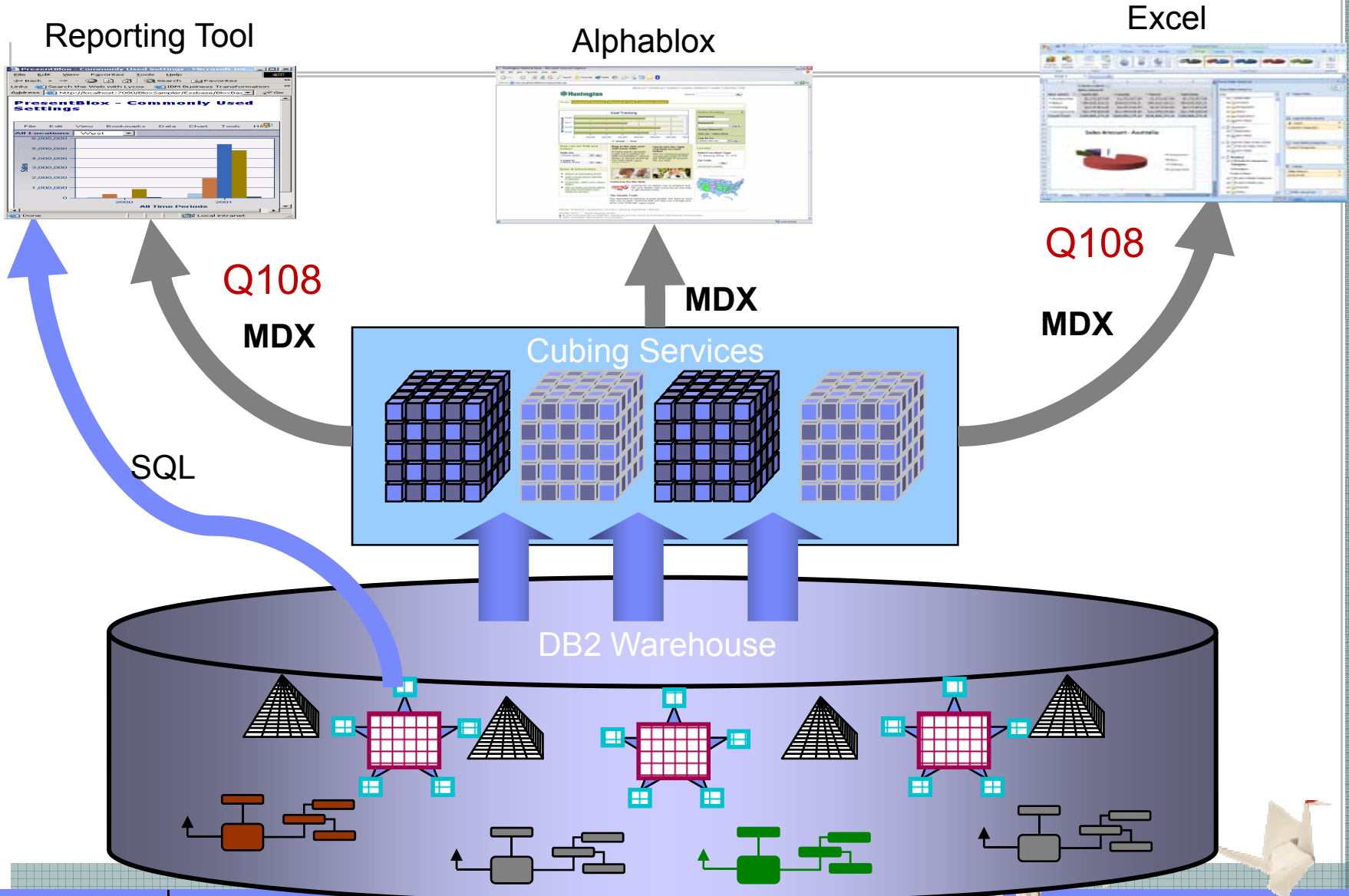
### **Warehouse-based multidimensional (OLAP) analytics built in to the DB2 Warehouse platform**

- Integrates Alphablox cubing technology, Cube Views and DB2 optimization technology
- Delivers mainstream OLAP (MDX) Function / Robust Data Cubes
- Supports premier OLAP (MDX) client tools through industry standard OLAP API(s)
- Improves Time-to-Value, Ease of Use / Deployment, lower TCO

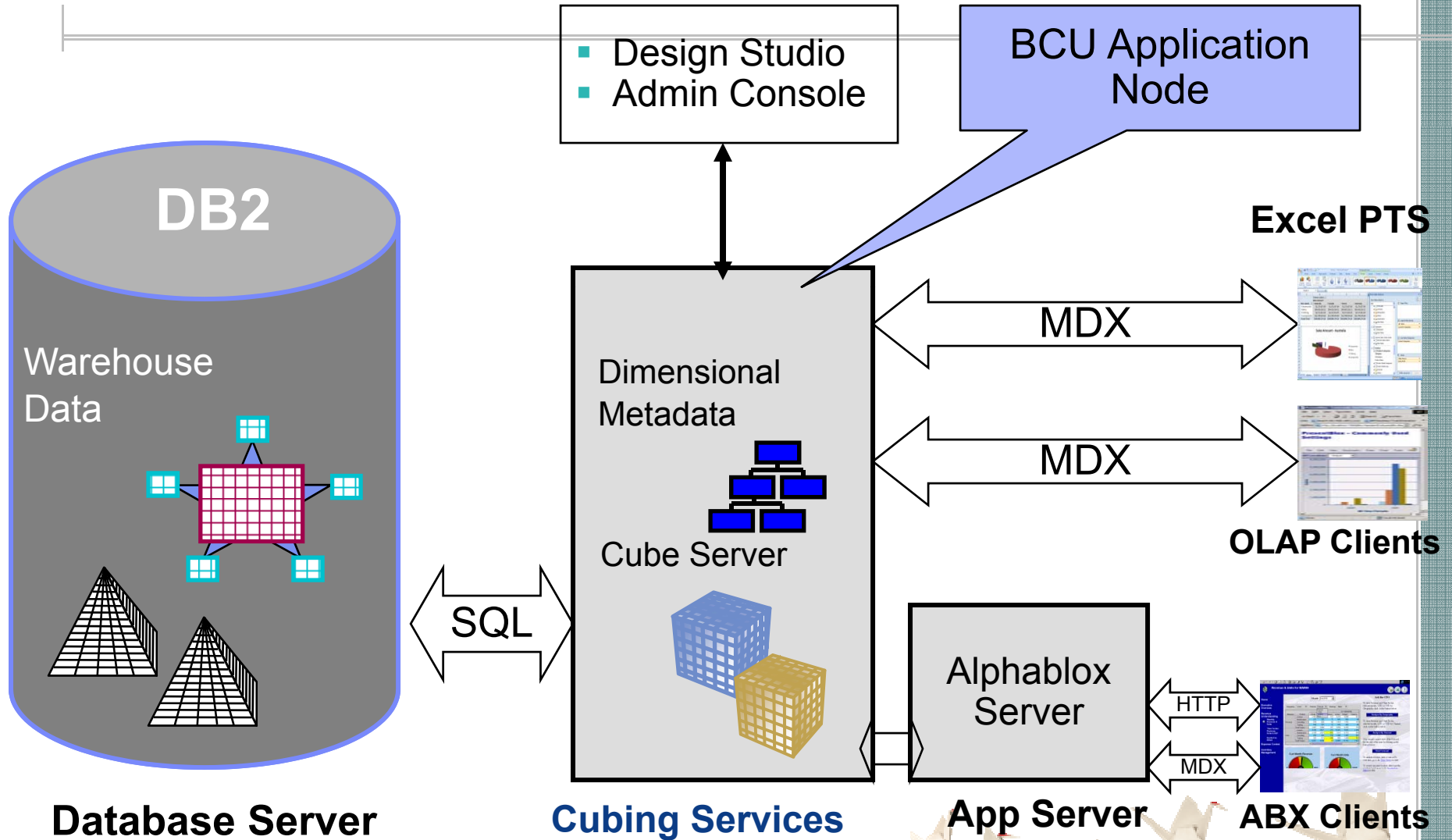




# Cubing Services in DB2W 9.5: OLAP Analytics and Open Access



# DB2 Warehouse 9.5 Cubing Services Architecture

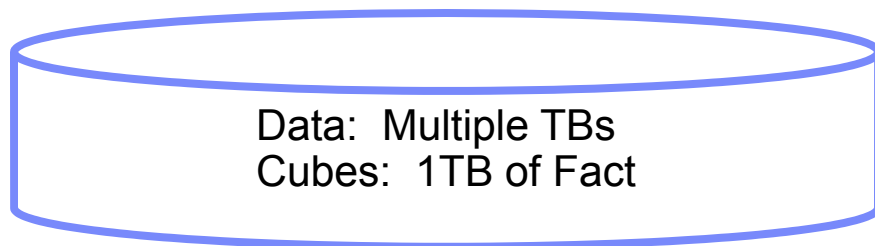


# OLAP Coverage in DB2 Warehouse – 9.5

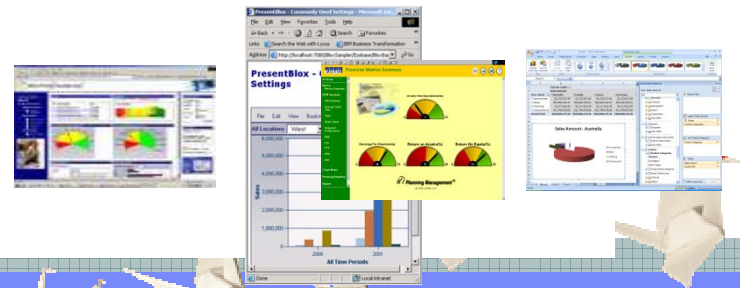
75% of Applications		5-10%	10-15%
<b>Dashboards</b>	<b>Mainstream R/O</b>	<b>Advanced Financials</b>	<b>Planning / Budgeting</b>

- Functional Target: Mainstream read-only analytics and dashboards
  - Multidimensional calcs, aggregates and time series intelligence
  - Cross industry applications: Retail, Telecommunications, Financial Services, Gov't, etc.

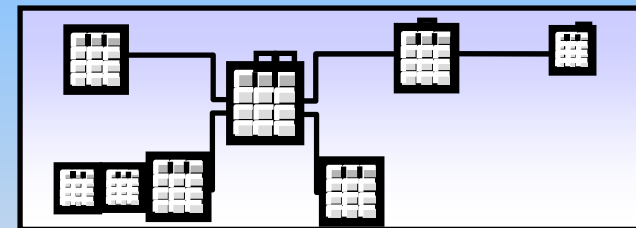
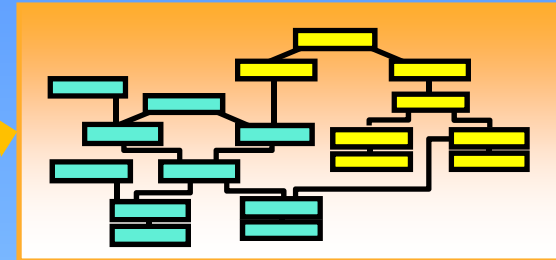
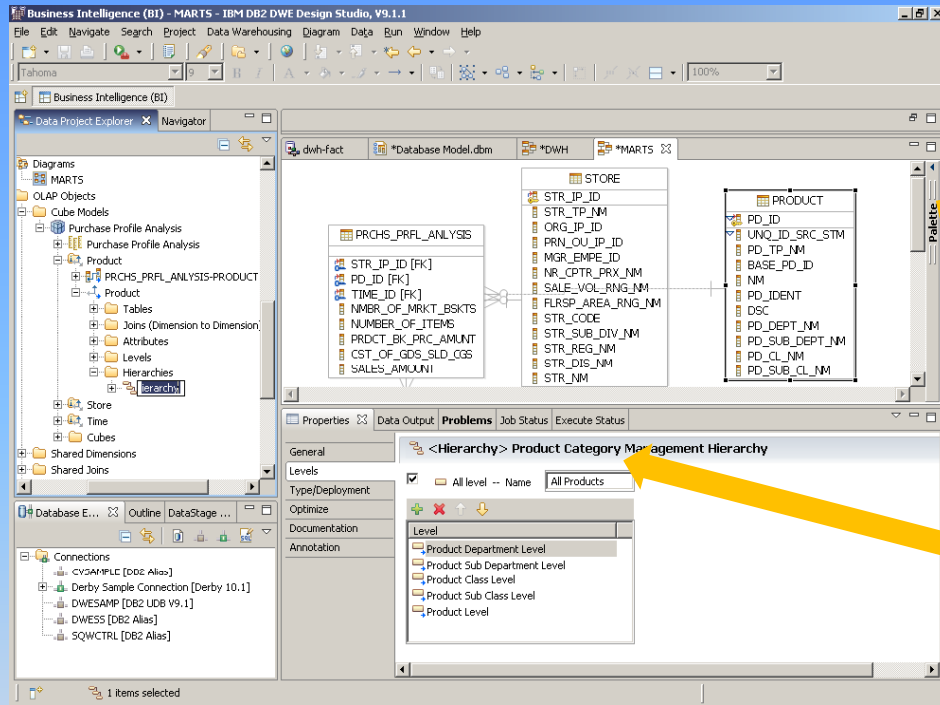
- Data Volumes and Latency
  - Cube up to 1TB of Fact data
  - Up to intra-day update frequency



- Client Support
  - Alphablox
  - Q1 2008: Excel, BOBJ, others
    - ODBO support



# Cube Modeling in Design Studio

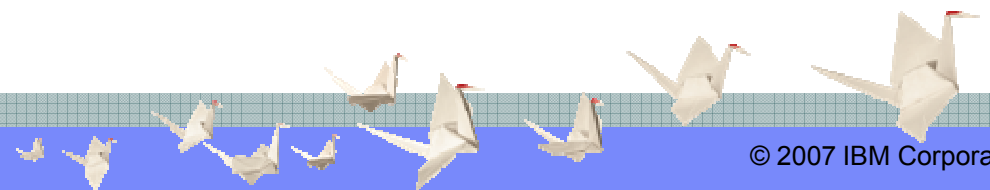


- Multidimensional Business Model
  - ▶ Maps logical business concepts to physical tables

- Deploy to the Cube Server
  - Available to Applications
  - Can be optimized

# DB2W Integrated Analytics Value Propositions

- **Reduced data latency**
  - Fewer data copies
  - Analytics IN the warehouse
  
- **Reduced cost, single vendor solution**
  - Reduced software licenses, hardware, maintenance
  - Common design and administration tooling
  - Reduction of piecemeal and outboard analytics engines
  
- **Embeddable and customizable delivery**
  - See the following.....



# Alphablox in DB2W 9.5 Delivery Vehicle for Embedded Analytics

- Platform for Customized Analytic Applications and Inline Analytics
- Pre-built components (Blox) for analytic functionality
- Allows you to create **customized** analytic components that are **embedded** into existing business processes and web applications

The screenshot displays an IBM WebSphere Portal interface. The main content area is titled 'Welcome Category Manager!' and features several analytics dashboards. One dashboard, 'Category KPI's', shows 'Net Sales' at \$13.71M (Actual) and \$13.40M (Target), and 'Billing Gross' at \$6.54M (Actual) and \$6.02M (Target). Another dashboard, 'Markdowns % Sales Variance by Division', includes a map of Alberta. The portal also has a 'People Finder' and 'Alerts' section on the left. Overlaid on the bottom right is a 'Huntington' bank website, which includes a 'Goal Tracking' bar chart, a login form, and various service links.



# Alphablox New BloxBuilder in Design Studio in 9.5

The screenshot shows the BloxBuilder Design Studio interface. On the left is a 'Palette' with various components like Selection Tool, Connection, Blox Components, ChartBlox, DataBlox, DataLayoutBlox, GridBlox, PageBlox, and PresentBlox. The main workspace contains a visual model with components: <Button> Button\_05, <DataBlox> DataBlox\_01, <Message Dialog> Message\_Dialog\_04, and <PresentBlox> PresentBlox\_02. Lines connect these components, indicating relationships. Below the workspace, a browser window displays a 'Credit Card Analysis' report for card 4123-4409-1187-3284. The report includes a summary table, two bar charts showing balance and activity, and a detailed activity table.

**Model / Layout**

**Generate**

**Deploy**

**Credit Card Analysis - Microsoft Internet Explorer**

Address: http://localhost:8080/ibm/

>Hello Admin  
17 October 2007

**Credit Card Analysis**

Analyst for Card 4123-4409-1187-3284 For the 6 months ending 200708

Current Balance	Credit Limit	Lifetime Value	Cardholder Name	Credit Score	Member Since
4495.6	5000	211.99	James Kingston	765	Aug 21, 1992
			Jamie Kingston	700	Aug 21, 1992

Activity	Fees	Payment	Purchases	Return	Transfer	Net	Month End Balance
200703	-1584.55	595.69	-23.84	5000.00	3977.41	3977.41	3977.41
200704	-572.02	1677.42	-394.64	710.76	4888.23	4888.23	4888.23
200705	32.50	-500.00	1842.90	-200.35	1275.05	5963.28	5963.28
200706	33.40	-2957.83	1309.05	-80.84	-1295.57	4667.71	4667.71
200707		-1252.25	1462.91	-60.33	140.32	4808.03	4808.03
200708		-1462.58	1200.39	-99.97	-41.27	-338.43	4499.60
<b>Total</b>	<b>65.90</b>	<b>-7889.24</b>	<b>8190.98</b>	<b>-858.77</b>	<b>4958.73</b>	<b>4499.60</b>	

Month	Sold	Posted	Reference	Activity	Type	Amount
200708	Jul 31, 2007	Jul 31, 2007	FDWTRCT	CALIFORNIA SUNSHINE - JACKSON NJ	Purchases	93.65
200708	Jul 31, 2007	Jul 31, 2007	TDHELPV	RADIOBROCK CORRO1151616 FREEHOLD NJ	Purchases	21.39
200708	Jul 31, 2007	Jul 31, 2007	G339SR30	STRIDE RITE #6071 JACKSON NJ	Purchases	46.98
200708	Aug 1, 2007	Aug 1, 2007	TC4PBR30	STRIDE RITE #6071 JACKSON NJ	Return	-46.98
200708	Aug 1, 2007	Aug 1, 2007	G6D8MPCY	RADIOBROCK CORRO1151616 FREEHOLD NJ	Purchases	21.39
200708	Aug 1, 2007	Aug 1, 2007	*WORGSHS	SKETCHERS-USA #0123 JACKSON NJ	Purchases	42.45
200708	Aug 1, 2007	Aug 1, 2007	SD4PBR30	STRIDE RITE #6071 JACKSON NJ US 000	Purchases	19.99
200708	Aug 1, 2007	Aug 1, 2007	WYRPNZ7	JACKSON STORE 107 JACKSON NJ	Purchases	68.42
200708	Aug 2, 2007	Aug 2, 2007	SLRRRLD	CENTRAL JERSEY POOLS FREEHOLD NJ	Purchases	35.24
200708	Aug 3, 2007	Aug 3, 2007	WR1052V3	MCLAUGHLINS AUTO SERVICE MORGANVILLE NJ	Purchases	36.34

Book1 - Microsoft Excel | PivotChart Tools: Design, Layout, Format, Analyze

Chart 1

Column Labels	Australia	Canada	France	Germany
Row Labels				
Accessories	\$1,272,057.89	\$1,272,057.89	\$1,272,057.89	\$1,272,057.89
Bikes	\$94,620,526.21	\$94,620,526.21	\$94,620,526.21	\$94,620,526.21
Clothing	\$2,117,613.45	\$2,117,613.45	\$2,117,613.45	\$2,117,613.45
Components	\$11,799,076.66	\$11,799,076.66	\$11,799,076.66	\$11,799,076.66
<b>Grand Total</b>	<b>\$109,809,274.20</b>	<b>\$109,809,274.20</b>	<b>\$109,809,274.20</b>	<b>\$109,809,274.20</b>

**Sales Amount - Australia**

Legend:

- Accessories
- Bikes
- Clothing
- Components

**PivotTable Field List**

Show fields related to:

- Employees
- Contacts
- Demographic
- History
- Organization
- More fields
- Geography
  - Geography
  - More fields
- Internet Sales Order Details
  - Internet Sales Orders
  - More fields
- Product
  - Product Categories**
    - Category
    - Subcategory
    - Product Name
  - Product Model Categories
  - Product Model Lines
- Financial
- History

Report Filter

Legend Fields (Series)

- Σ Values
- Customer Geography

Axis Fields (Categories)

- Product Categories

Σ Values

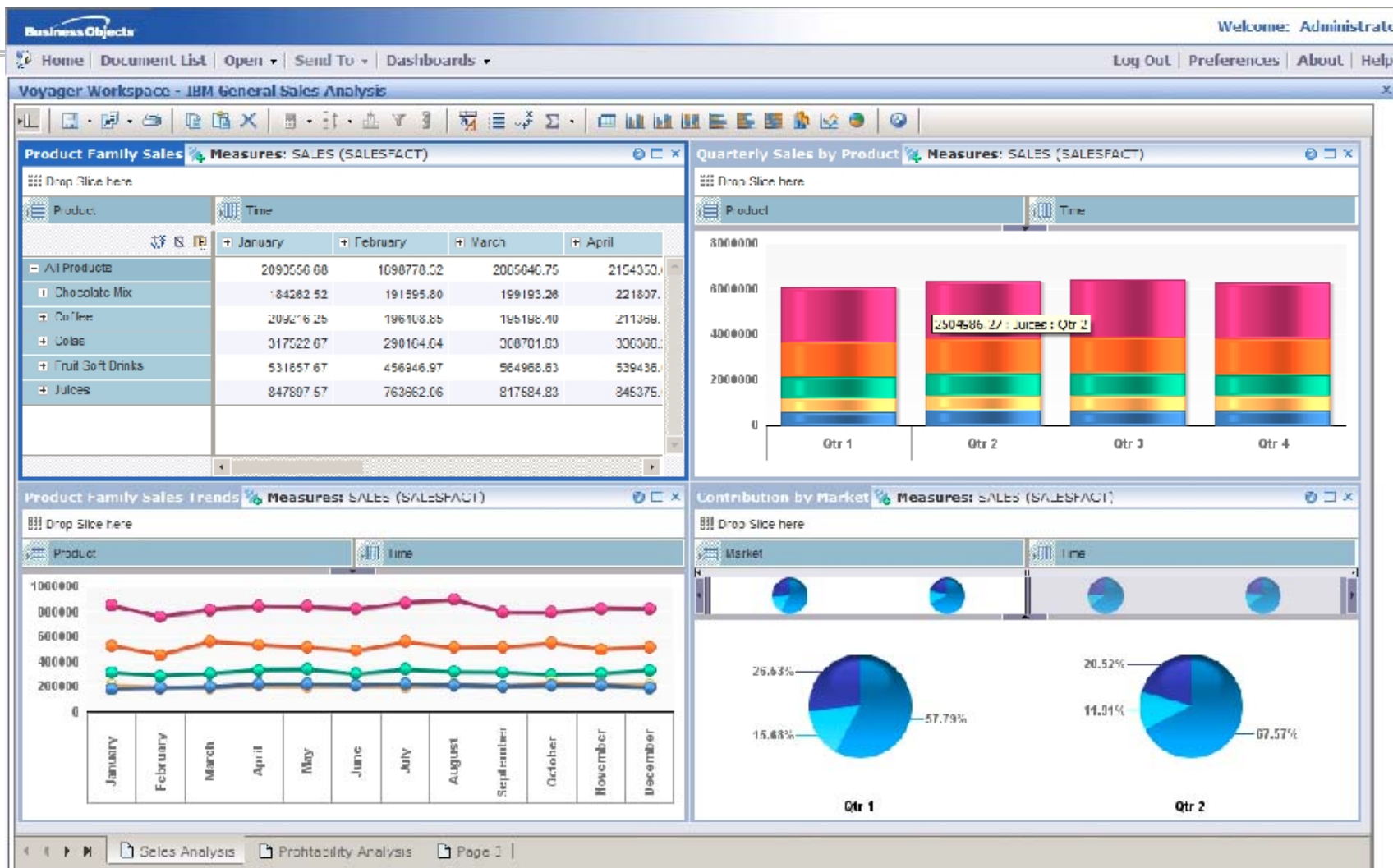
- Sales Amount
- Gross Profit

Defer Layout Up... Update

Sheet1 | Sheet2 | Sheet3 | Ready | 100%



# Business Objects Voyager Support in 2008



# Act faster with extreme workload management

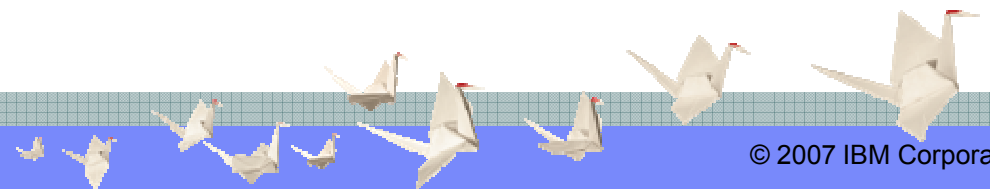


## Deliver right-time insight by...

- Setting business policy based priorities for different users and applications, with appropriately dedicated resources
- Delivering actionable insights to a broader set of users operational business processes and decision making



***A bank delivers individualized offers at the point of contact, based on customer specific history, to optimize promotions and increase revenue per customer***



# DB2 WLM Objectives

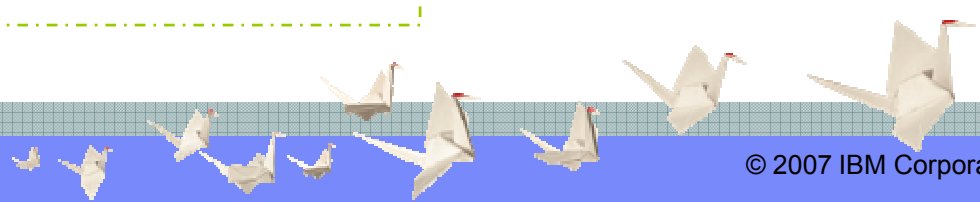
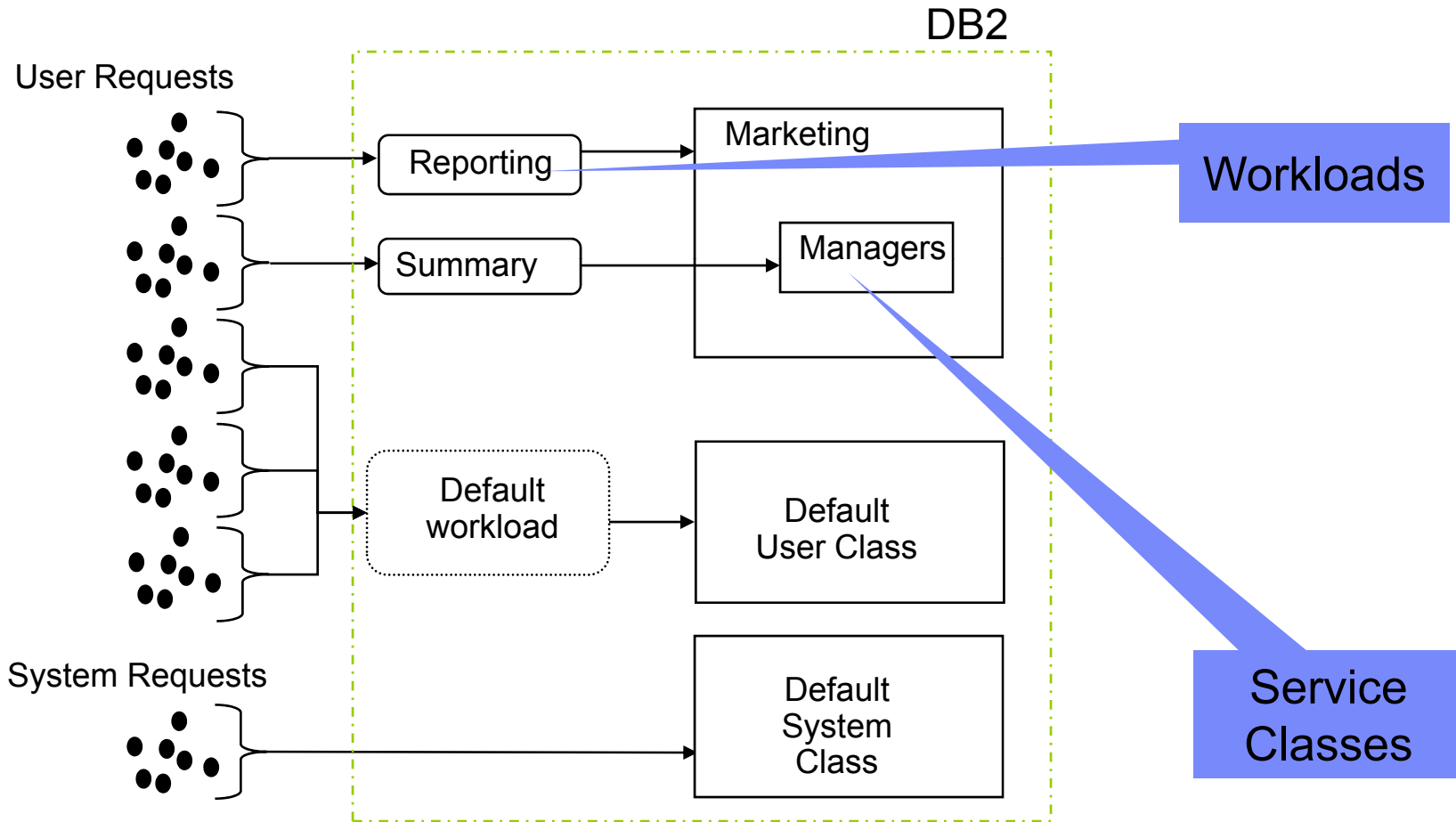
- A stable, predictable execution environment
- A light-weight, granular way to monitor active work
- Better resource management
  - Be able to explicitly allocate resources amongst work
  - Be able to limit excessive, unexpected resource consumption
- Better request management
  - Be able to manage work based on its business priority
  - Be able to track performance of work
- End-to-end workload management solutions

## Key Concepts in WLM

- A workload categorizes work to be controlled
- All work runs in a service class
  - Associate DB2 service classes with AIX WLM for strict control of CPU resources
- Use thresholds to
  - Enforce limits
  - Control concurrency
  - Specify fine grained monitoring
- Use work action sets for sophisticated controls

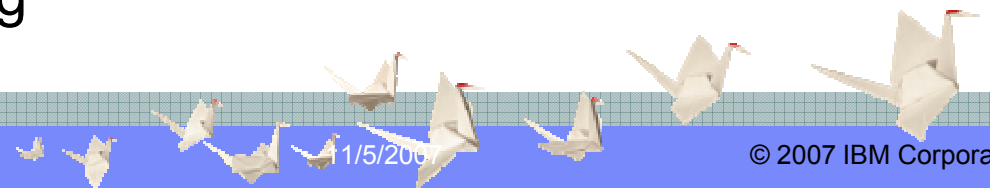


# Workload Management Design



## DB2W Design Studio Simplifies WLM configuration

- Reverse engineering
  - Visualize your WLM configuration
- Guided configuration
  - Encapsulates best practice solution templates
- Robust execution
  - Provides error recovery when deploying
- Delta execution
  - Hide complexity of executing WLM DDL
  - Implement scheduling



# DB2 Warehouse Design Studio for WLM Design

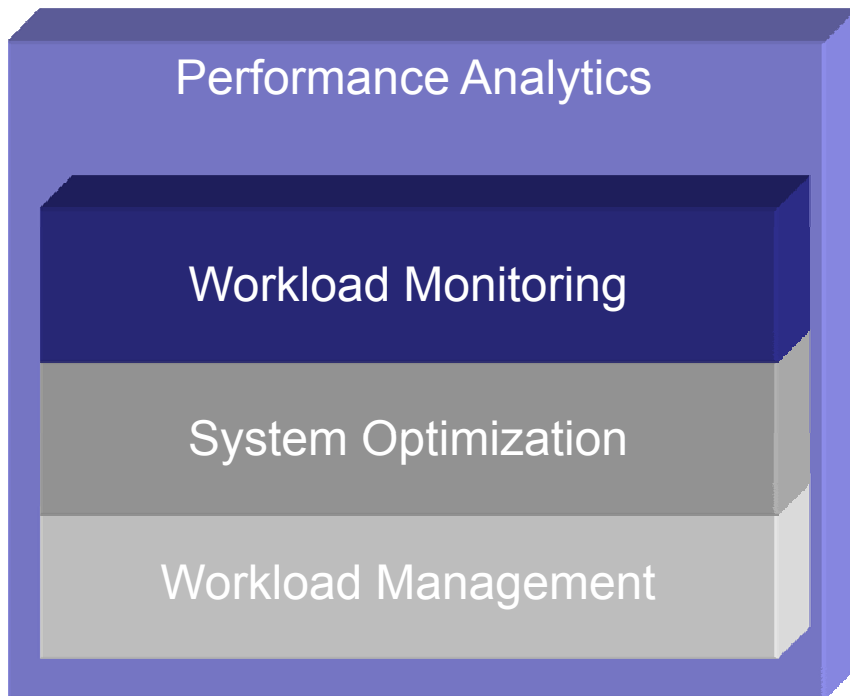
The screenshot displays several overlapping windows from the DB2 Warehouse Design Studio for WLM Design:

- Control and Share System Resources:** A window with tabs for General, Work Identities, Superclasses, and Create Relationships. It contains a table for mapping work identities to superclasses.
- Enforce Limits for Activities:** A window with tabs for General, Superclasses, Work Identities, Work Types, and Create Limits. It is used to define resource usage limits for activities.
- Create Work Type:** A dialog box for defining a new work type, including fields for work type set, name, and type.

Work Identity	Superclass
MARK_WL	MARK_SUPER
SALES_WL	SALES_SUPER

# DB2 Warehouse Performance Management Suite

## *Instrumenting Your Warehouse*



***Monitor. Analyze. Optimize.***

*Software solution to monitor, analyze and optimize the complete lifecycle of enterprise BI and Data Warehouse deployments.*



# DB2 Warehouse Performance Management Suite

## Components

### Performance Analytics

#### Workload Monitoring

Performance Management Feature (Appfluent)

#### *DB2 Warehouse Performance Monitoring Feature*

*Query workload analysis that correlates user activity, data usage and query performance metrics - without disrupting or adding overhead to production systems.*

### System Monitoring

DB2 Performance Expert

### Workload Management

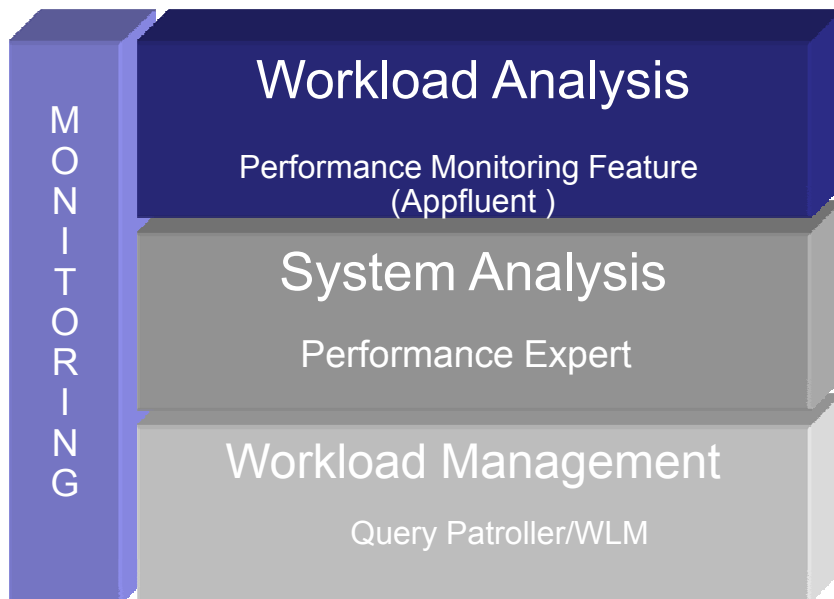
DB2 QP/WLM

#### *DB2 Warehouse Performance Optimization Feature*

*Database and operating system level analysis for system-level optimization and regulation/control of resource consumption.*

# DB2 Warehouse Performance Management Suite

## Monitoring



### Performance Monitoring Feature

Monitor 100% of query workload and get detailed historical query usage and workload performance analysis - without adding overhead on production systems.

### Performance Expert

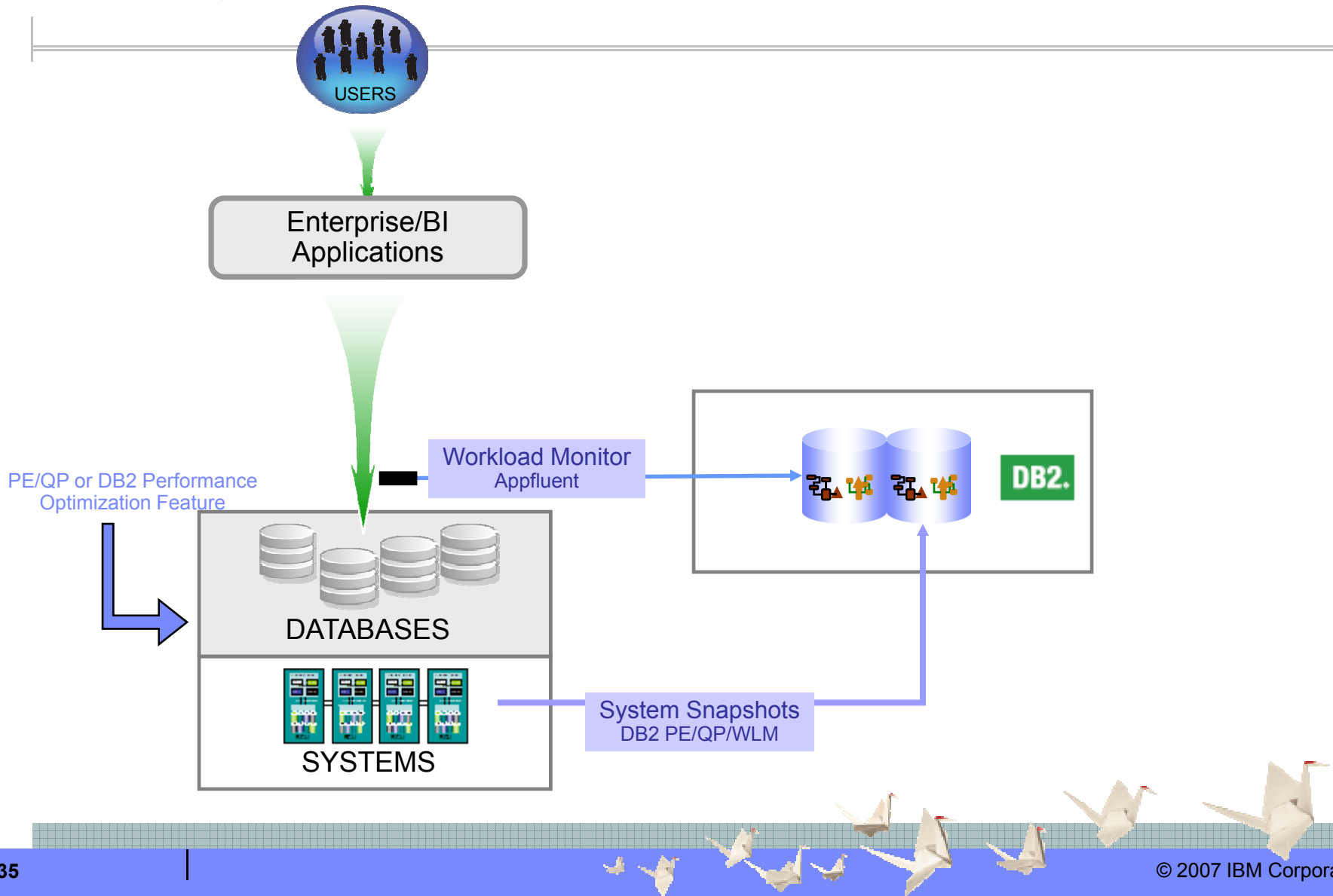
Use information discovered via workload analysis to collect and examine relevant system snapshot information for tuning the servers (I/O, Memory, Buffer Pools etc.) - minimizing PE snapshot overhead.

### QP/WLM

Set optimal policies for workload control based on workload and system information provided by Appfluent/PE to get the maximum performance improvements in the database.

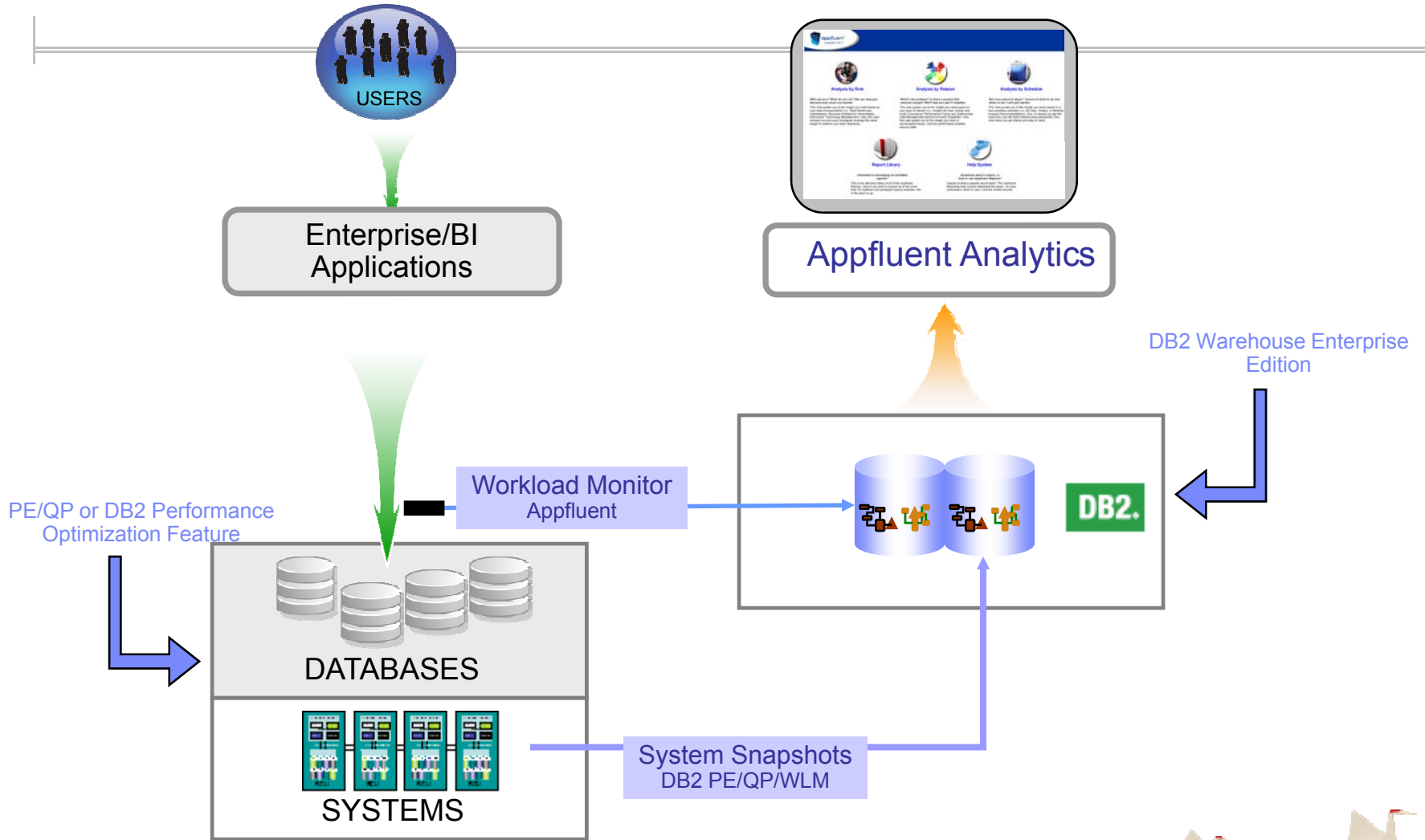
# DB2 Warehouse Performance Management Suite

## Monitoring



# DB2 Warehouse Performance Management Suite

## Performance Analytics



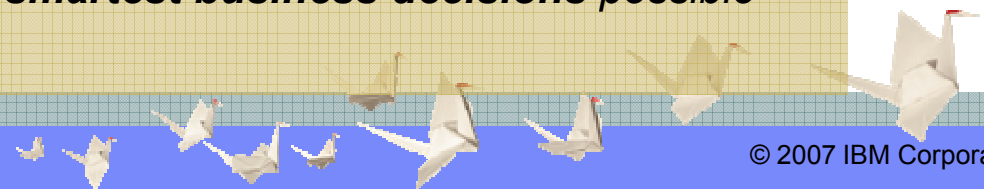
## Changing Businesses with Information

# Insight without Boundaries

Reach Farther. Look Deeper. Act Faster.



*Successful companies execute with **maximum efficiency and effectiveness** and make the **smartest business decisions possible***



---

# Thank You & Q&A

