



ORACLE®

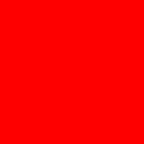


**ORACLE®**



## **Scalability and Performance made easy with Oracle Database 11g and Oracle Optimized Warehouses**

Hermann Baer  
Director, Product Management



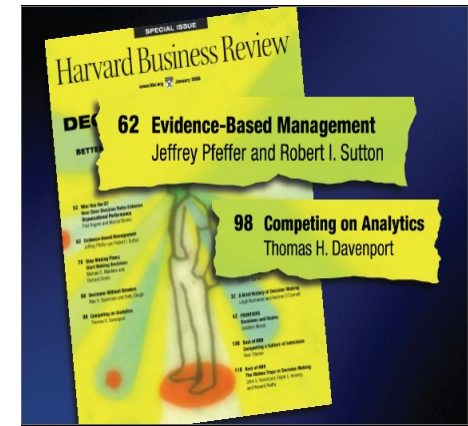
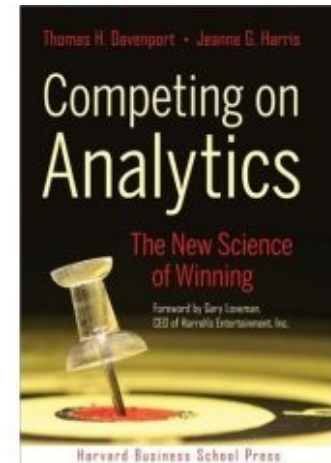
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

# Agenda

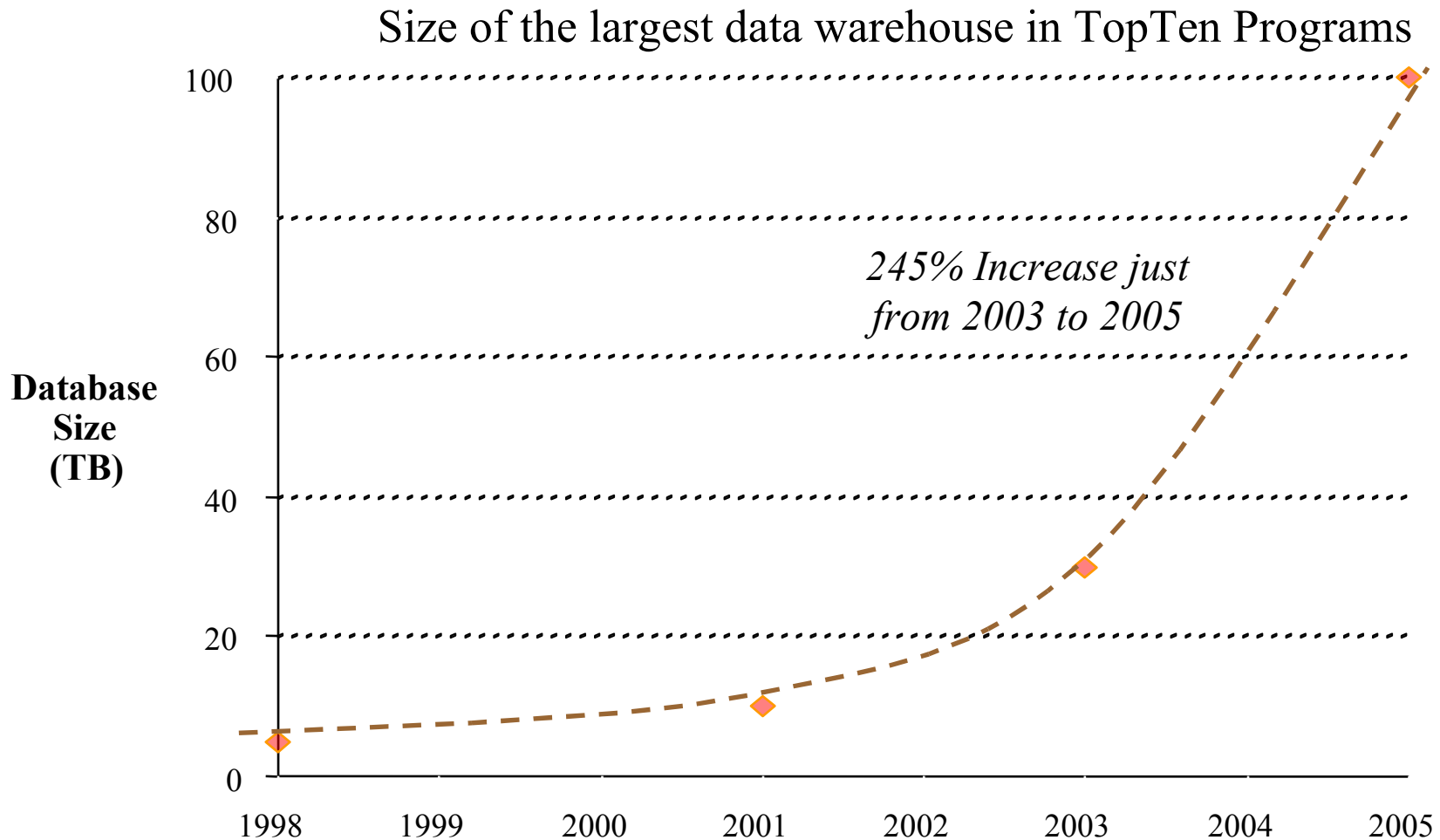
- Data warehouse market overview
- Oracle Data Warehouse strategy
- Oracle Optimized Warehouse Initiative

# Data Warehousing Business Drivers

- How did sales do this quarter and what is the year over year growth?
- How much revenue is driven by my current marketing campaign?
- How does this year's employee turnover compare to the last 5 years'?
- How has company profitability changed over the past 3 years?
- What are my most profitable products by region by year?

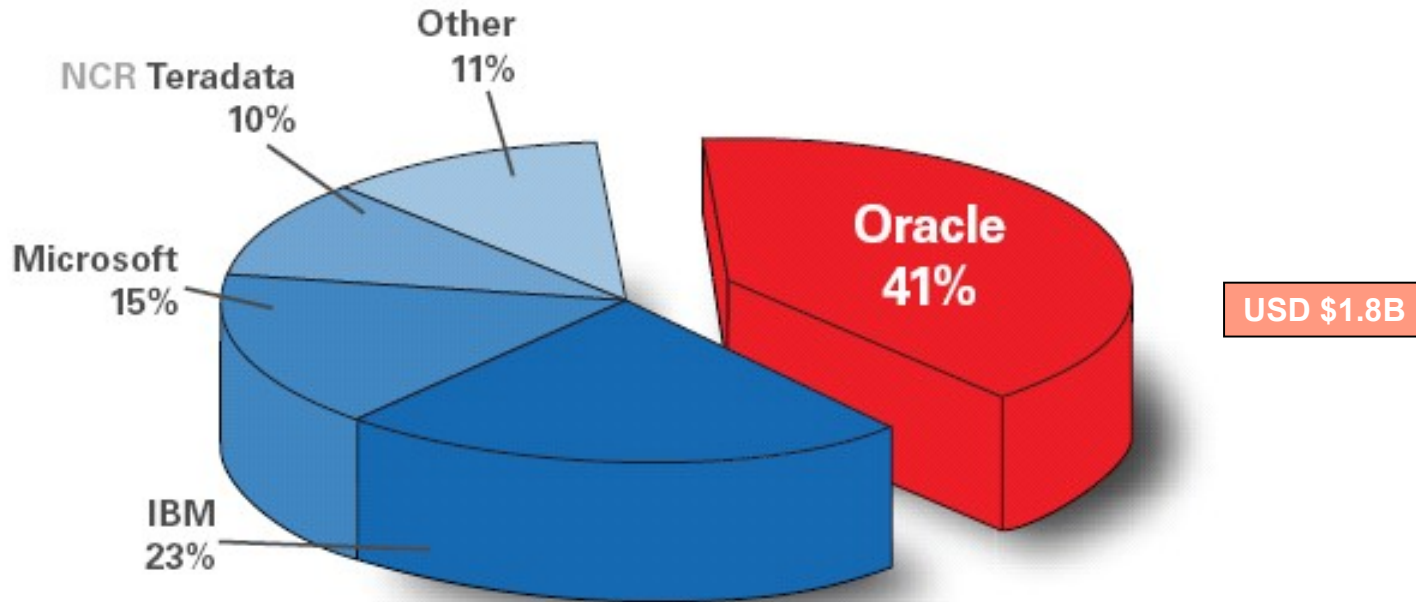


# Data Volumes Growing Exponentially



Source: 2005 TopTen Program, November 2005 © Winter Corporation, Waltham, MA, USA

# Oracle #1 for Data Warehousing



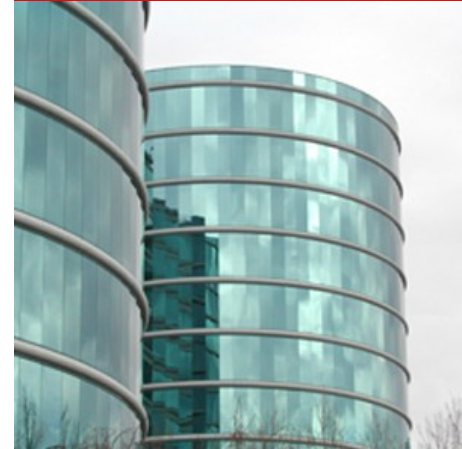
## Worldwide Data Warehouse Management Market Share, 2006

Total Market = USD \$4.4B, 12.4% growth YoY

Source: IDC, 2007 – Data Warehouse Platform Tools 2006 Vendor Shares

**Oracle**

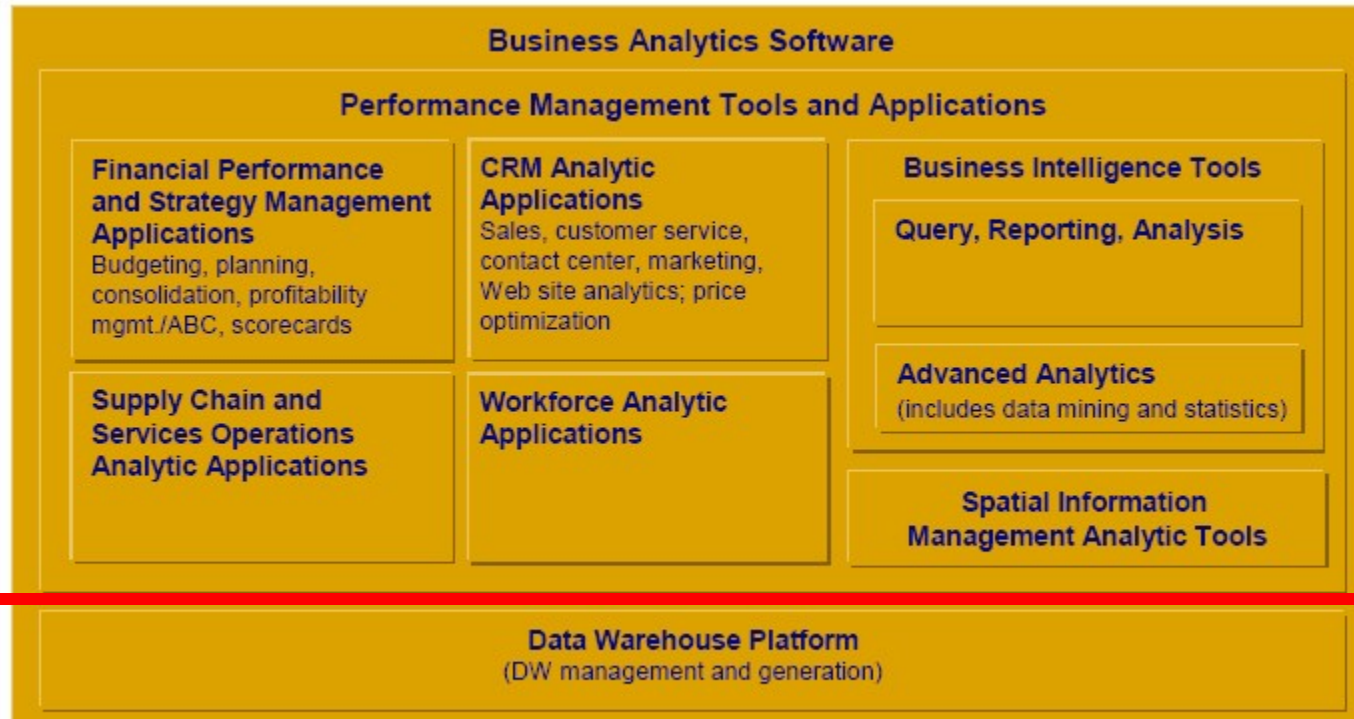
# Data Warehouse Strategy





# Why a Data Warehouse?

IDC's Business Analytics Taxonomy, 2007



Source: IDC, 2007

**Industry-Standard Best Practice:**  
**The Data Warehouse is the Foundation**  
**of Business Intelligence and Performance Management**

# Data Warehousing Strategy

- Best database for BI and data warehousing
  - No other database can compare on the breadth and sophistication of Oracle's database features
- Complete solution portfolio
  - Complete database platform including ELT and Analytics
  - Oracle BI and Performance Management solutions
  - Broadest array of third-party technologies and solutions
- On the right hardware infrastructure
  - Offer customers a choice of solutions tailored to meet needs

# Data Warehousing

## with Oracle Database 11g

- Embedded ELT and Data Quality
- Industry leading performance and scalability
- Comprehensive integrated analytics
- Reliable, easy to manage, low cost grids

# Data Warehousing

## with Oracle Database 11g

- **Embedded ELT and Data Quality**
- Industry leading performance and scalability
- Comprehensive integrated analytics
- Reliable, easy to manage, low cost grids

# Oracle Warehouse Builder

## Data integration

- Embedded into Oracle Database 11g installation
  - Feature of enterprise and standard editions
- Integrate data from disparate data sources
  - High performance data loading optimized for Oracle Database
  - Extreme scalability with support for RAC and Partitioning
  - Declarative, highly productive drag and drop design capabilities
- Optional ERP/CRM connectors
  - eBusiness Suite, Peoplesoft, Siebel
  - SAP

# Oracle Warehouse Builder

## Data quality

- Improve enterprise data quality
  - Name and Address cleansing
  - Fuzzy matching and merging
  - Data profiling and data auditing
  - Business rule support in profiling and ETL
- Provide data modeling capabilities for Oracle
  - Relational (tables, views, MVs, UDT, etc.)
  - Dimensional (slowly changing dimensions, OLAP support, etc.)
- Enterprise metadata management
  - End-to-end attribute level lineage and impact analysis
  - Extensible repository to hold any metadata

# Oracle Warehouse Builder

## Packaging

### Enterprise ETL Option

- Performance
- Productivity
- Reusability
- Metadata Management

### Data Quality Option

- Data Profiling
- Anomaly Detection
- Business Rules
- Audit

### ERP/CRM Connectors

- Oracle EBS
- PeopleSoft
- Siebel
- SAP

## Core Features

(No extra cost with database SE/SE1/EE)

# Data Warehousing

## with Oracle Database 11g

- Embedded ELT and Data Quality
- **Industry leading performance and scalability**
- Comprehensive integrated analytics
- Reliable, easy to manage, low cost grids



# New in Oracle Database 11g

- Advanced compression
- Online upgrades and patching
- Real Application Testing
- Active Data Guard
- Secure Files
- Total Recall
- Real Application Clusters Performance Management
- OLAP-based Materialized Views
- And more...

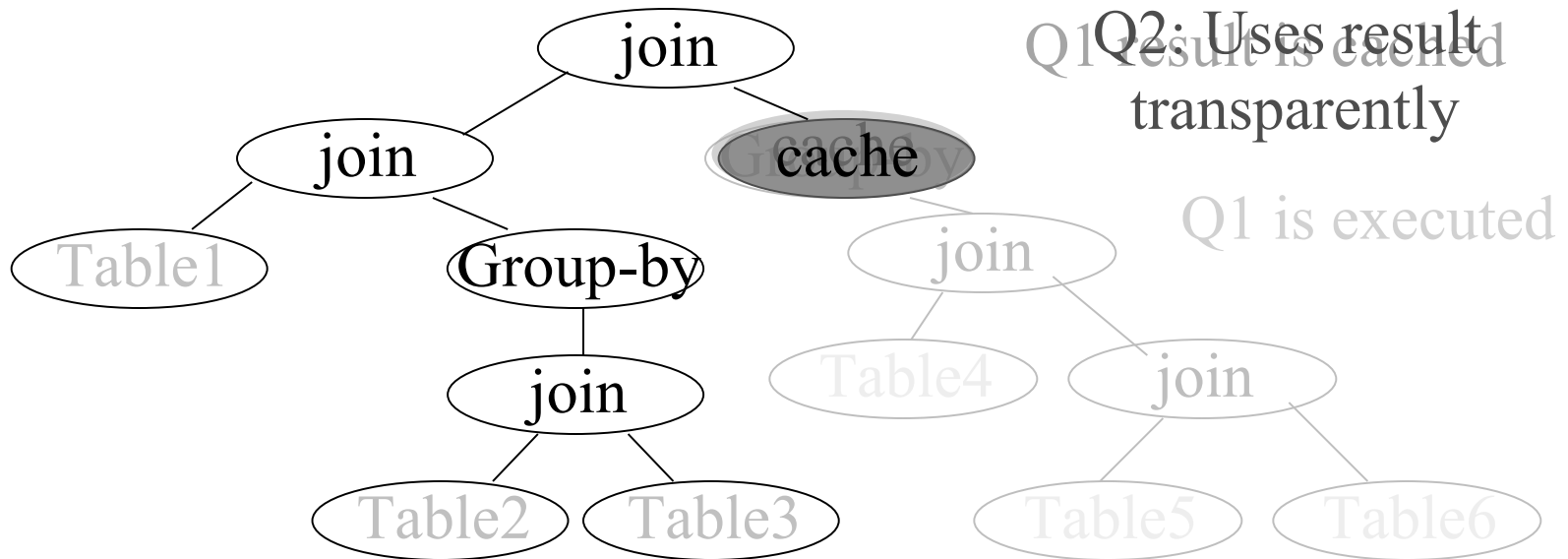
# New in Oracle Database 11g

## for business intelligence and data warehousing

- VLDB
  - Composite Range-Range
  - Composite List-Range
  - Composite List-List
  - Composite List-Hash
  - REF Partitioning
  - Virtual Column Partitioning
  - Compression enhancements
- Performance
  - Query Result Cache
- Data loading
  - Change data capture enhancements
  - Materialized view refresh enhancements
- Manageability
  - Partition Advisor
  - Interval Partitioning
  - SQL Plan Management
  - Automatic SQL Tuning with Self-Learning Capabilities
  - Enhanced Optimizer Statistics Maintenance
  - Multi-Column Optimizer Statistics
  - ASM Fast Resync, Fast VLDB Startup and other enhancements
- SQL
  - SQL Pivot and Unpivot
  - Continuous Query Notification
- OLAP
  - Continued database integration
    - Cube metadata in the Data Dictionary
    - Materialized view refresh and SQL rewrite
    - Fine-grained data security on cubes
  - Simplified application development
    - Fully declarative cube calculations
    - Cost-Based Aggregation
    - Simpler calculation definitions
- Data Mining
  - Simplified development and deployment of models
    - Supermodels: data preparation combined with mining model
    - Additional packaged predictive analytics
    - Integration in database dictionary
  - New algorithms: “General Linear Models”
    - Encapsulates several widely used analytic methods
    - Multivariate linear regression; logistic regression

# Database Result Cache

- Automatically caches results of queries
  - Cache is shared across statements and sessions on server
  - Significant speed up for read-only / read-mostly data
  - Full consistency and proper semantics
  - Cache refreshed when any underlying table updated

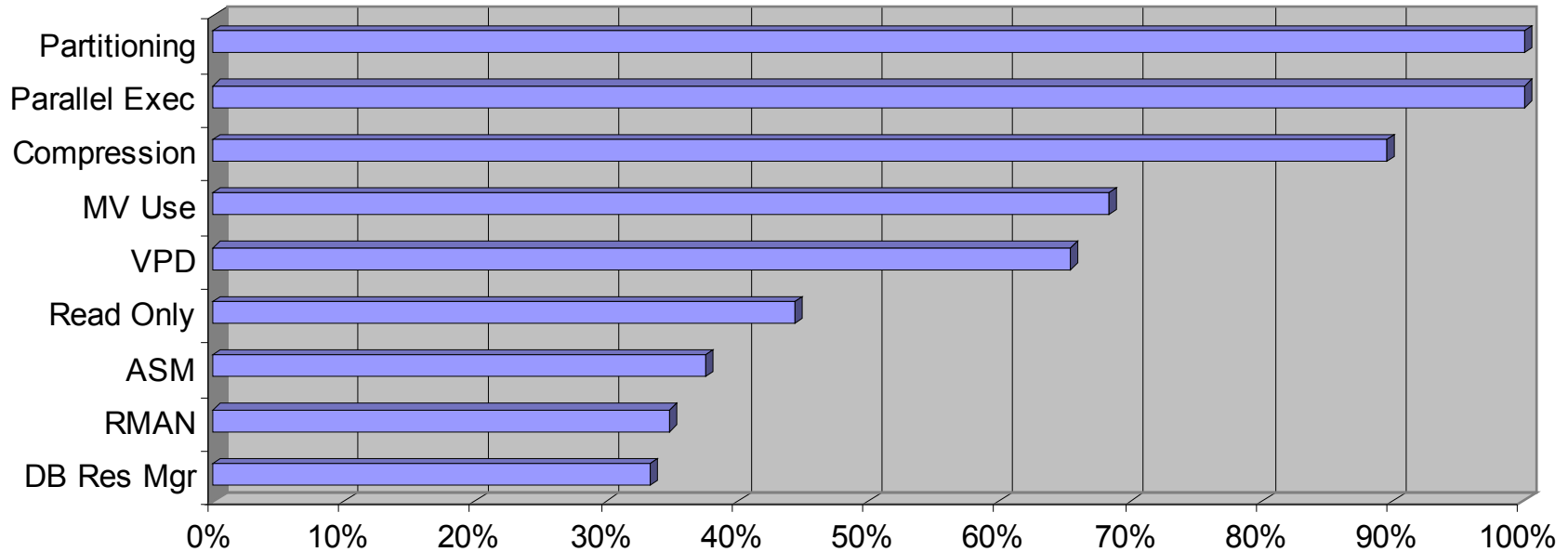


# Query Result Cache Opportunity

- Retail customer data (~50 GB)
  - Concurrent users submitting queries randomly
    - Executive dashboard with 12 heavy analytical queries
  - Cache results only at in-line view level
  - 12 queries run in random, different order – 4 queries cached
- Measure average, total response time for all users

# Users	No cache	Cache	Improvement
2	186 s	141 s	24%
4	267 s	201 s	25%
8	447 s	334 s	25%

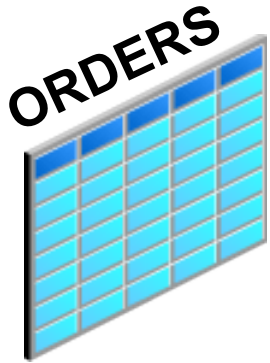
# Large-Scale Data Warehouse Features



Partitioning, parallelism, and compression are the foundation for large-scale data warehousing

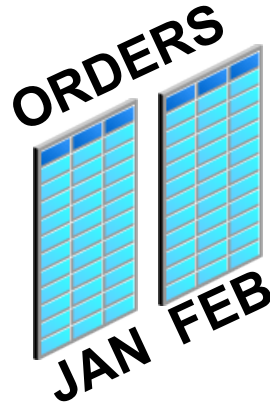
# Oracle Partitioning

## Transparent to applications



**Large Table**

**Difficult to Manage**

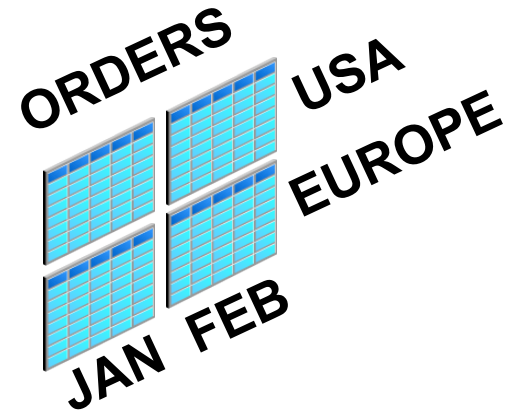


**Partition**

**Divide and Conquer**

**Easier to Manage**

**Improve Performance**



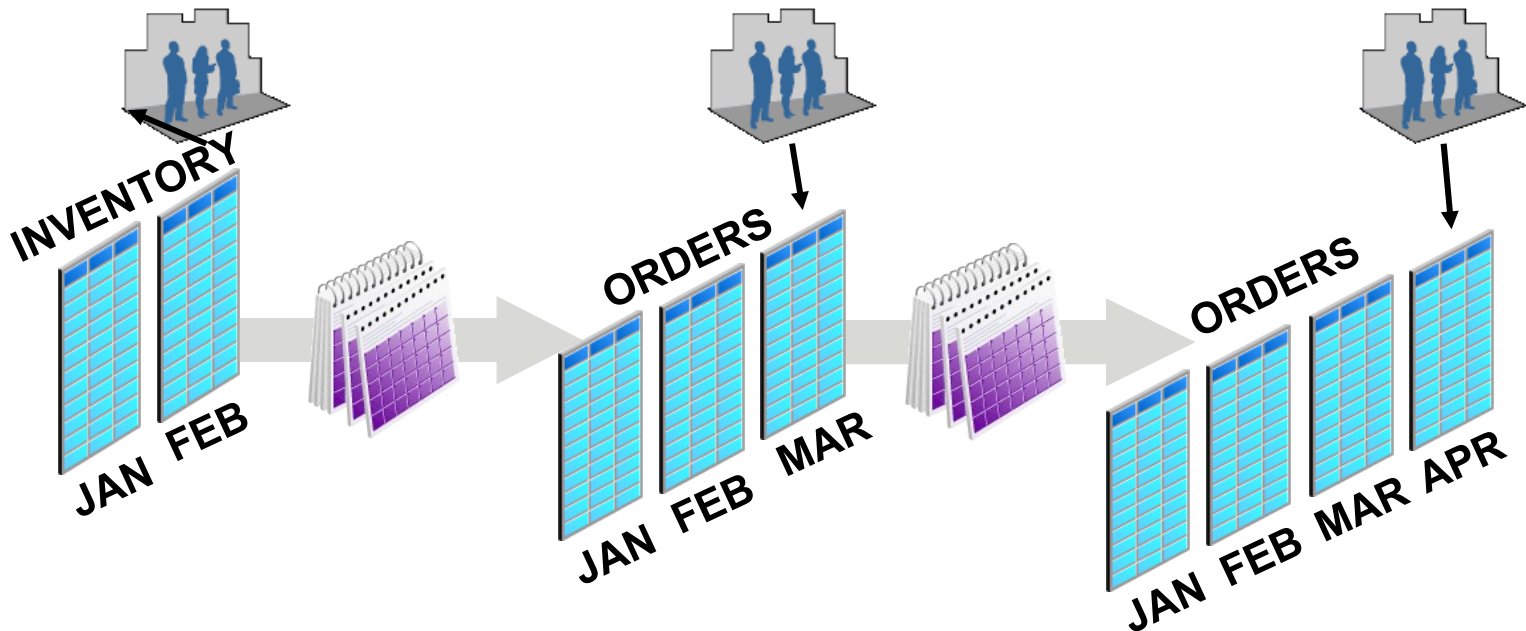
**Composite Partition**

**Better Performance**

**More flexibility to match  
business needs**

# Interval Partitioning

- Partitions are created automatically as data arrives



# Composite Partitioning

- Two-dimensional partitioning schemes
- Extensions in Oracle Database 11g

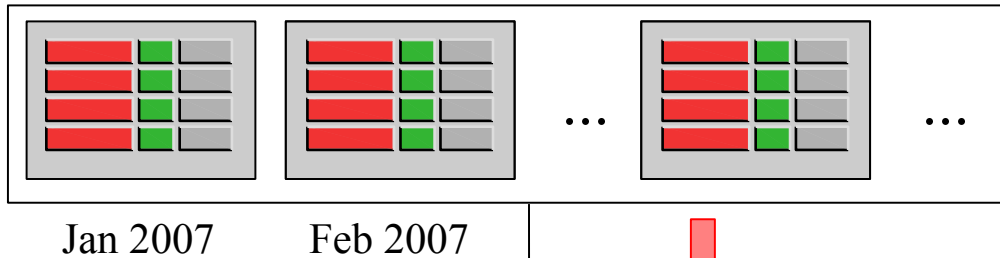
	Range	List	Hash
Range	11g	9i	8i
List	11g	11g	11g

- e.g. List-range:
  - Partition by country, then by week
  - Partition by line-of-business, then by week



# Before REF Partitioning

Table ORDERS



- RANGE(**order\_date**)
- Primary key **order\_id**

- Redundant storage of order\_date
- Redundant maintenance

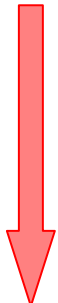
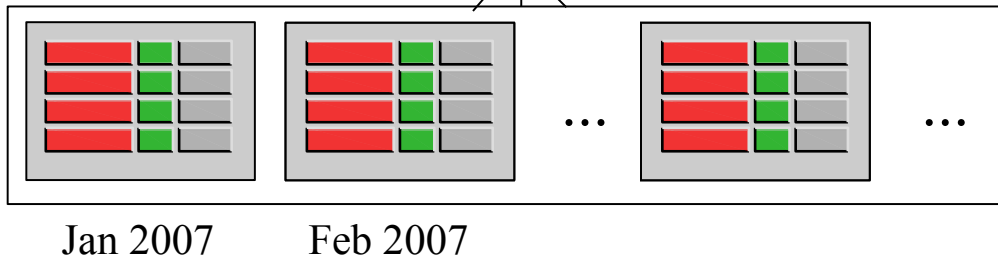


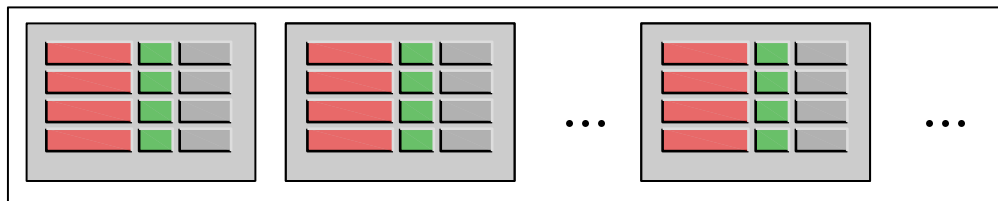
Table LINEITEMS



- RANGE(**order\_date**)
- Foreign key **order\_id**

# With REF Partitioning

Table ORDERS



Jan 2007

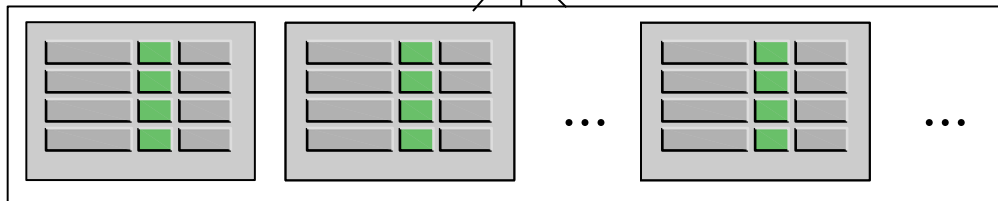
Feb 2007

- RANGE(**order\_date**)
- Primary key **order\_id**

## PARTITION BY REFERENCE

- Partitioning key inherited through PK-FK relationship

Table LINEITEMS



Jan 2007

Feb 2007

- RANGE(**order\_date**)
- Foreign key **order\_id**

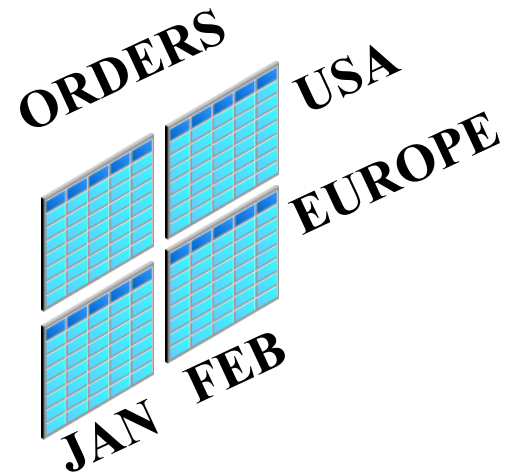
# Virtual Column Partitioning

## ORDERS

ORDER_ID	ORDER_DATE	CUSTOMER_ID...
9834- <b>US</b> -14	12-JAN-2007	65920
8300- <b>EU</b> -97	14-FEB-2007	39654
3886- <b>EU</b> -02	16-JAN-2007	4529
2566- <b>US</b> -94	19-JAN-2007	15327
3699- <b>US</b> -63	02-FEB-2007	18733

REGION	AS (SUBSTR (ORDER_ID, 6, 2))
US	
EU	
EU	
US	
US	

- REGION requires **no storage**
- Partition by ORDER\_DATE, REGION



# Oracle Partitioning

## What customers are saying



“... We are able to do a lot of partition elimination for customer queries.”

--Mike Ogden, Senior Technical Architect, AT&T.



“... You can't run very large databases effectively without partitioning.”

--Douglas Miller, Director of IT Operations, R L Polk.



“... Partitioning is part of the secret sauce that helps us scale our database environment.”

--Logan McLead, IT Straregy, Dell.

# Advanced Compression



## Compress Large Application Tables

- Transaction processing, data warehousing

## Compress All Data Types

- Structured and unstructured data types

## Typical Compression of 2-4X

- Cascade storage savings throughout data center

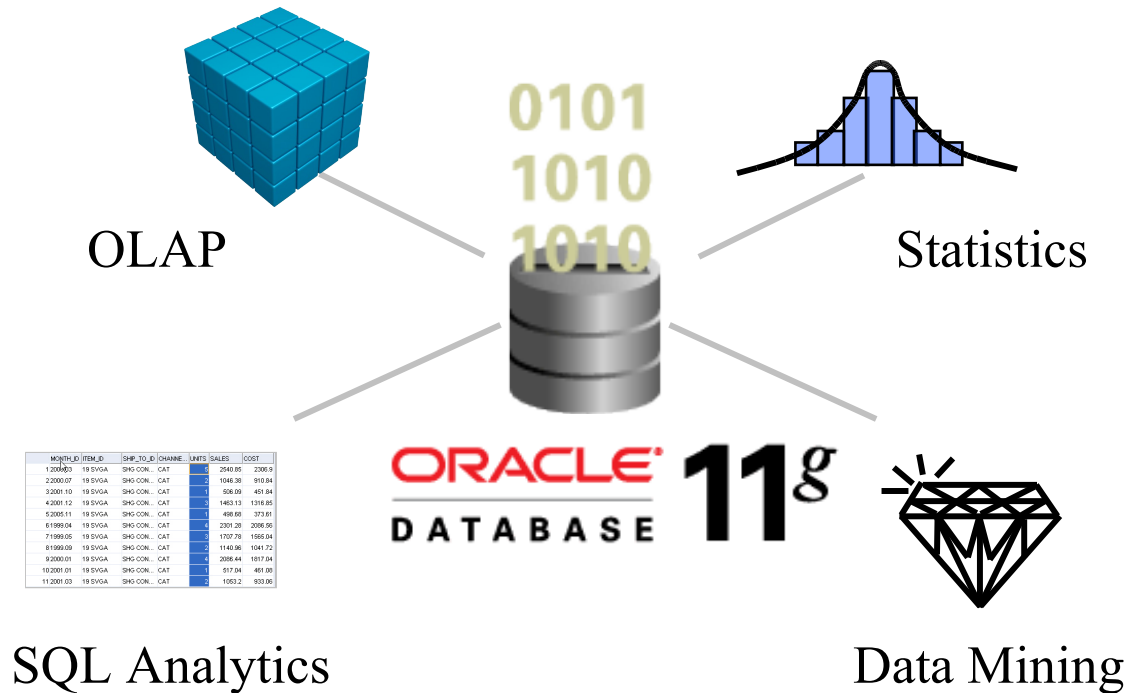
Up To  
**4X**  
Compression

# Data Warehousing

## with Oracle Database 11g

- Embedded ELT and Data Quality
- Industry leading performance and scalability
- **Comprehensive integrated analytics**
- Reliable, easy to manage, low cost grids

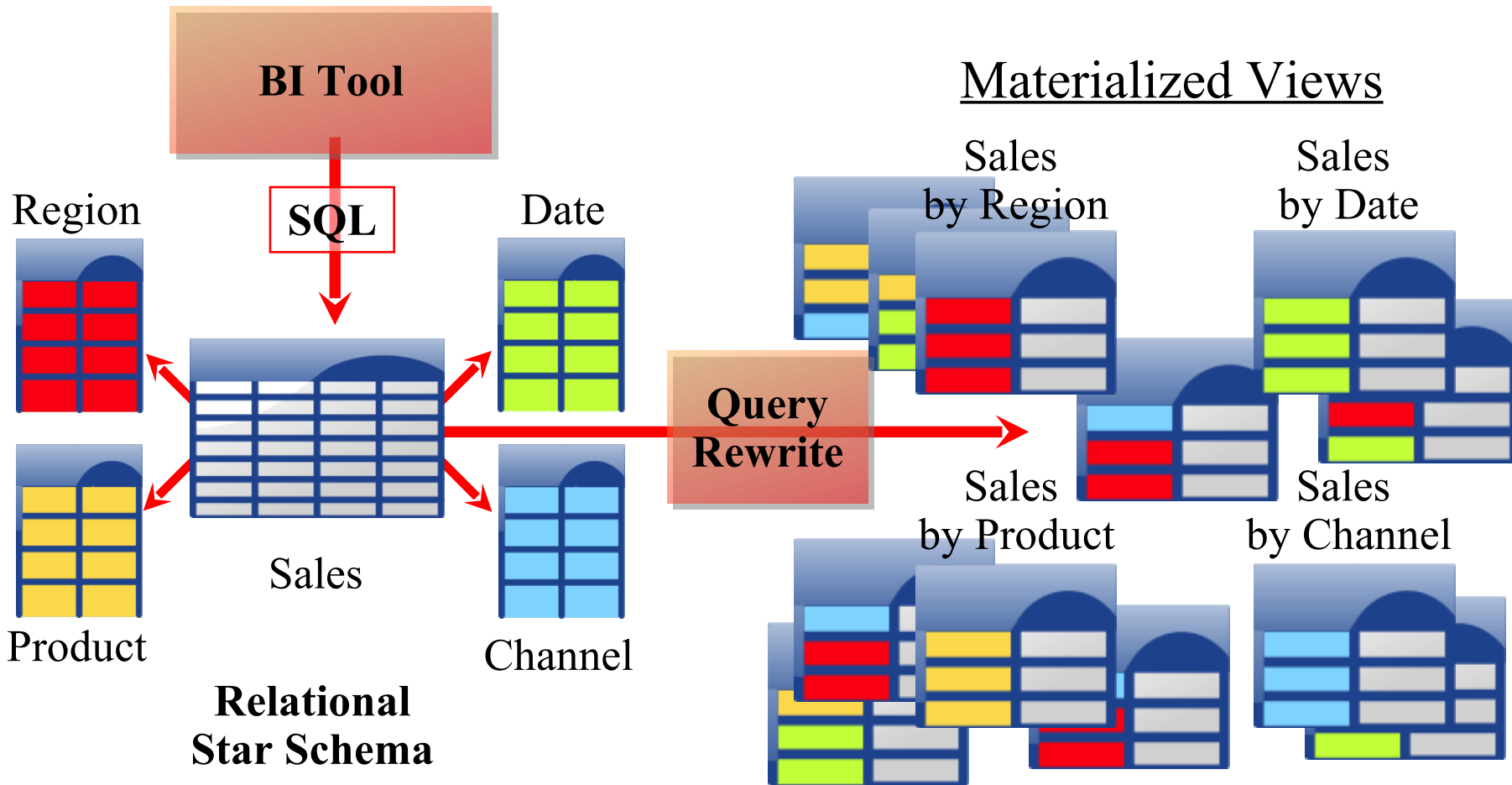
# Integrated Analytics



- Bring the analytics to the data
- Leverage core database infrastructure

# Business Intelligence Analysis

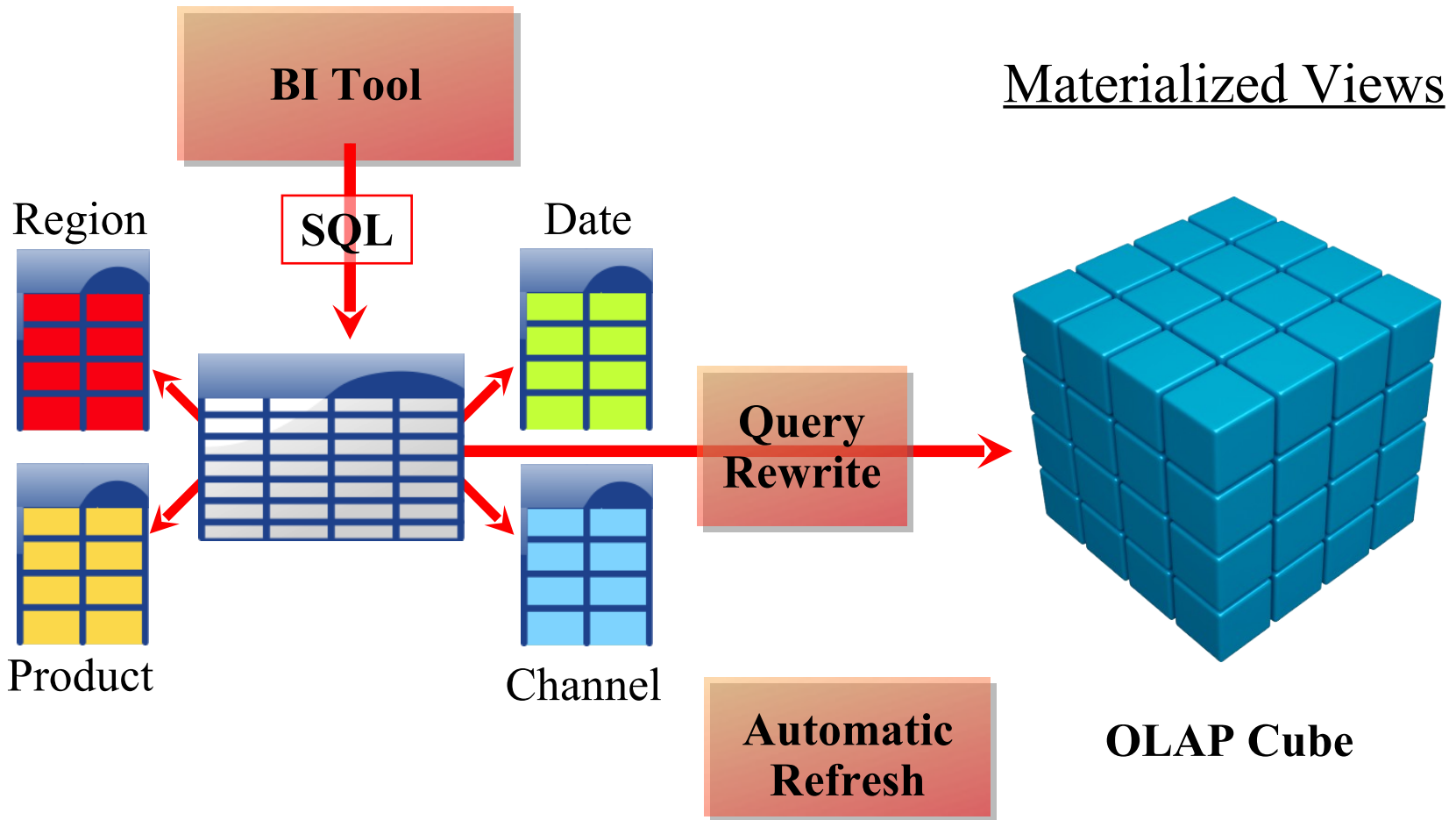
## Typical Architecture Today





# New in Oracle Database 11g

## Cube-Organized Materialized Views



# Data Warehousing

## with Oracle Database 11g

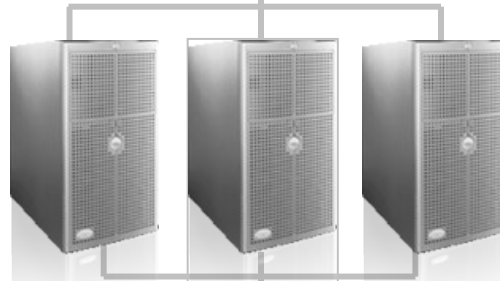
- Embedded ELT and Data Quality
- Industry leading performance and scalability
- Comprehensive integrated analytics
- **Reliable, easy to manage, low cost grids**

# Data Warehousing Grid

**Fusion Middleware**



**Real Application Clusters**



**Automatic Storage Management**



**Grid Control**



# Data Warehousing Grid

**Transient**

**Fusion Middleware**



**Real Application Clusters**



**Persistent**

**Automatic Storage Management**



**Grid Control**



**Oracle**

# Optimized Warehouse Initiative



# Why Optimized Warehouses?

- Today our customers can buy any combination of hardware components to build a data warehouse.
- Some percentage of customers inevitably end up with poorly configured data warehouses
- The solution: Oracle Optimized Warehouse Initiative
  - Phase I: **Optimized Warehouse Reference Configurations**
    - Best practices and reference configurations
  - Phase II: **Oracle Optimized Warehouse**
    - Pre-packaged, pre-configured, pre-installed, calibrated & validated systems

# Balanced Configurations Drive Optimized Performance

## An Unbalanced Configuration

**100%**  
Possible  
Efficiency

Database CPUs Memory Actuators LUNs Disks Raid

**< 50%**  
Achieved  
Efficiency

## A Balanced Configuration

**100%**  
Possible  
Efficiency

Database CPUs Memory Actuators LUNs Disks Raid

**100%**  
Achieved  
Efficiency

# Goals for Oracle Data Warehouse Solutions



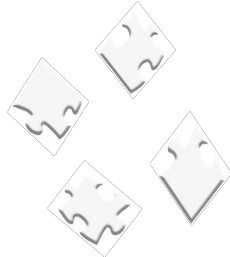
- Provide superior system performance
- Provide a superior customer experience
- Offer flexibility and choice of solutions





# Full Range of DW Options

## Custom Solutions



**Flexibility for the most demanding data warehouse**

## Reference Configurations



**Documented best-practice configurations for data warehousing**

## Optimized Warehouse



**Scalable systems pre-installed and pre-configured: create-table ready**

Pre-configured, Pre-installed, Validated

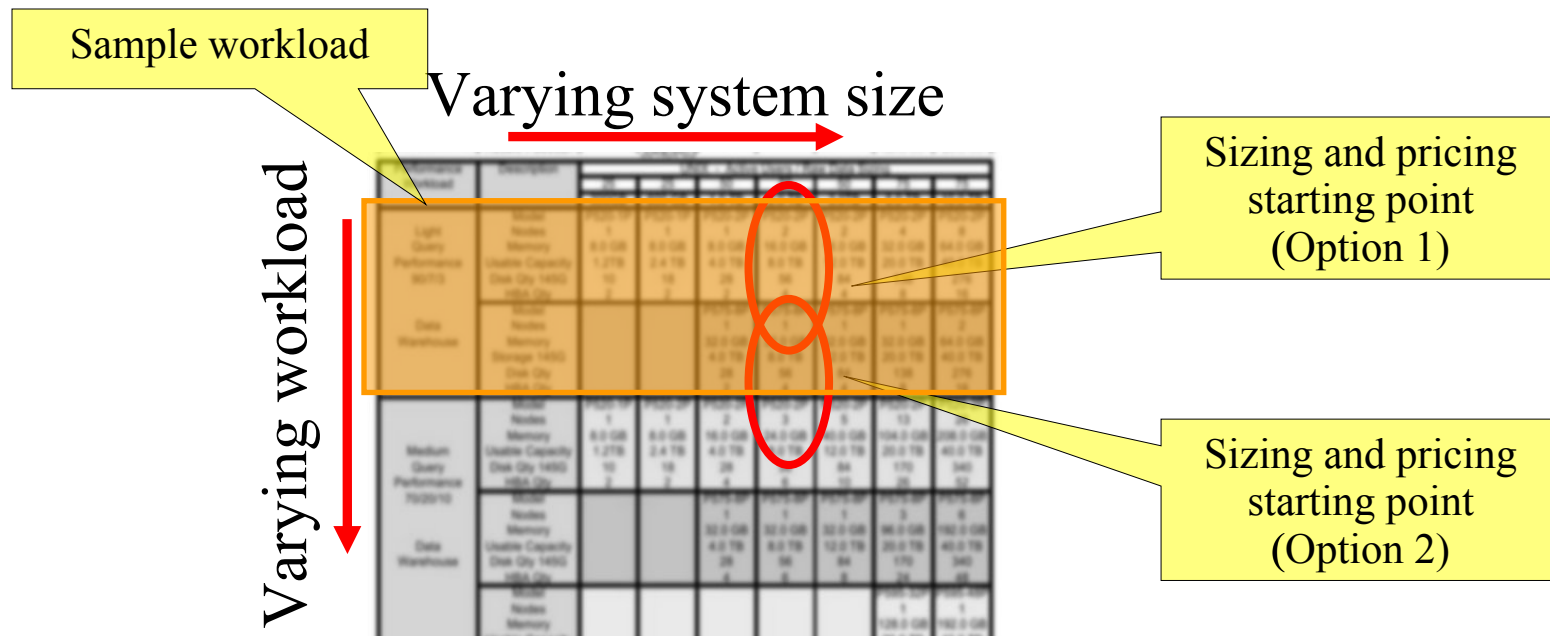
Flexibility

# Oracle Optimized Warehouse Initiative

## Sample Reference Configuration



- Spreadsheet-like guidance
- Choice of workload and data volume
- Choice of hardware and operating system

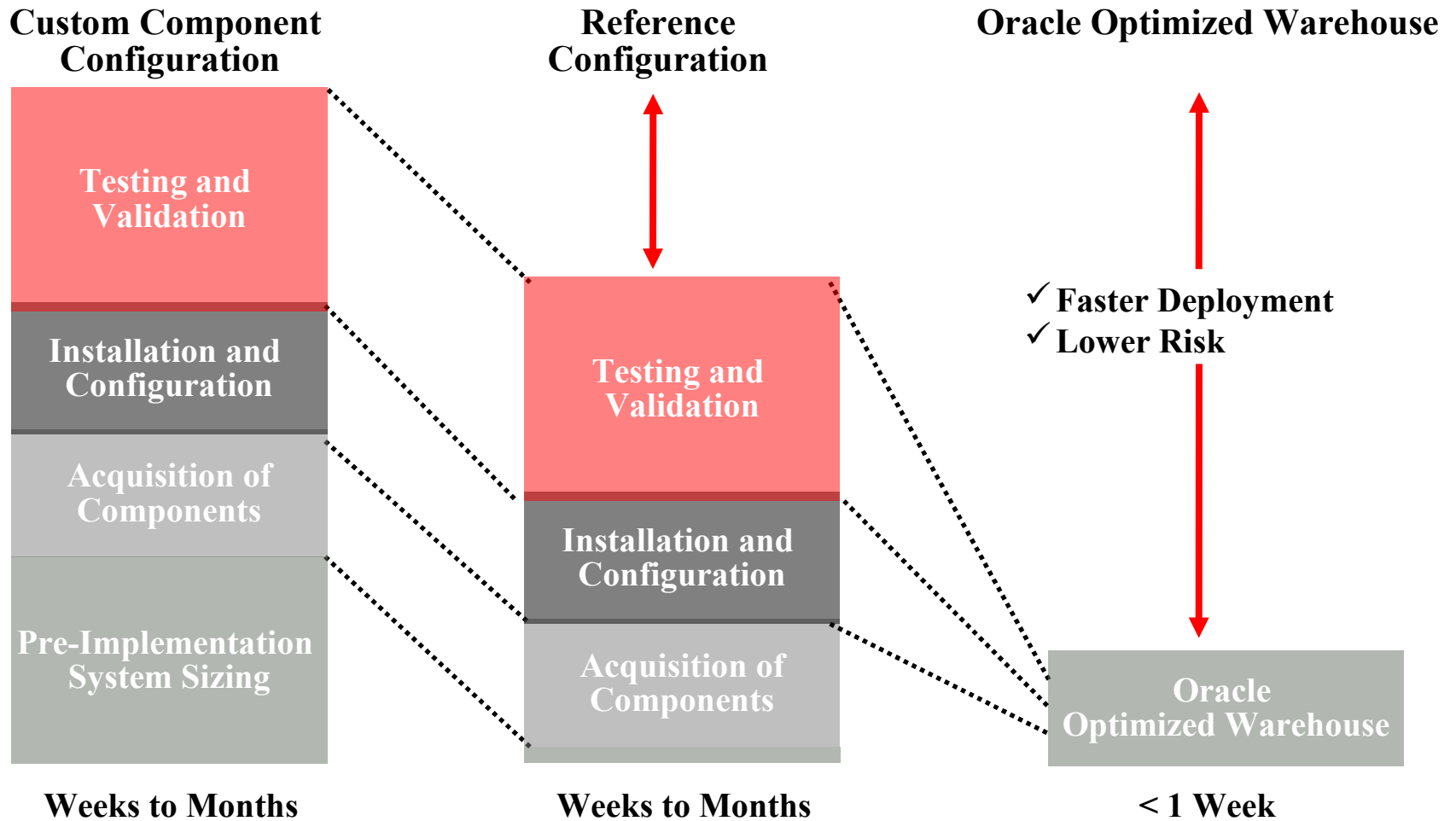


# Definition: Oracle Optimized Warehouse



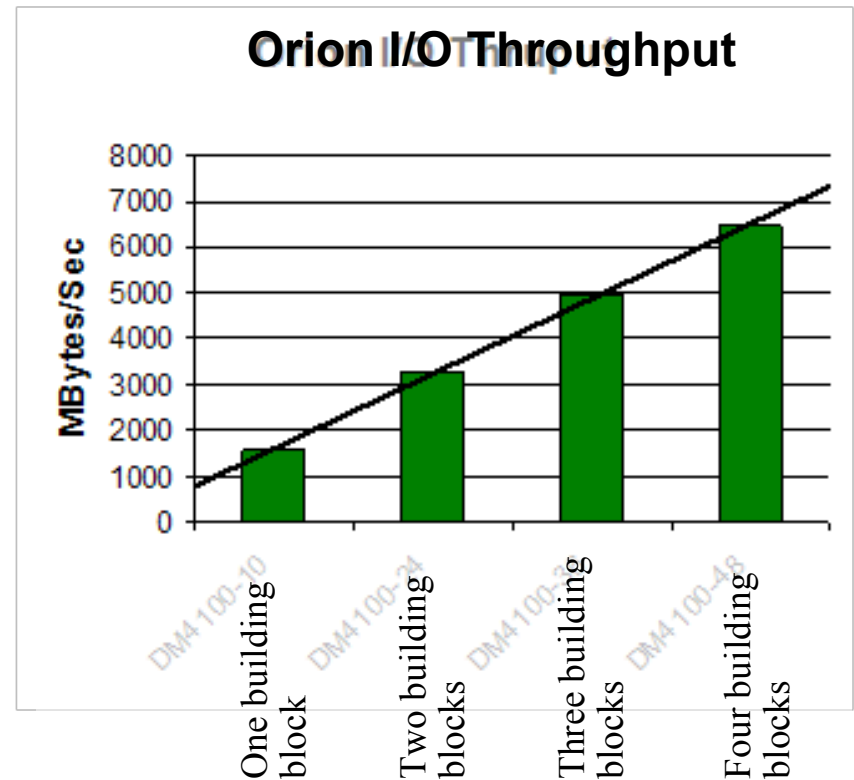
- Database, servers, storage
- *Pre-installed, pre-configured*
  - Oracle Database Enterprise Edition
  - Oracle Real Application Clusters
  - Oracle Partitioning
- *Validated* performance
- Sold as a single product
- Supported as a single product

# Reduce Risk & Accelerate Deployment



# Calibration Testing

- Goal: Balanced hardware configurations
  - Verify raw platform performance
  - Generate data warehouse IO workload
    - Using Oracle's ORION utility
  - Linear scalability



# Validation Testing

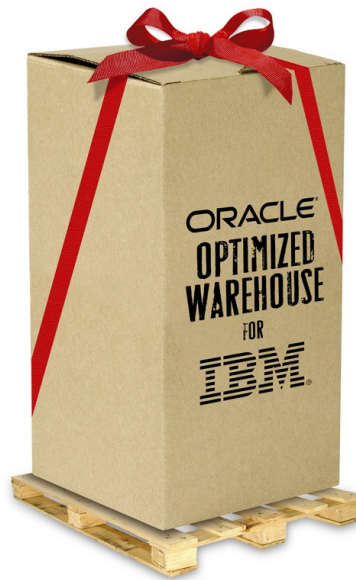
- Installation Verification:
  - ASM, Database, RAC
- Functional Testing:
  - Basic DBA operations
- Destructive Testing:
  - Test proper recovery from failures of nodes, HBA's, controllers, or disks
- Performance Testing:
  - Large-scale databases
  - Multi-user performance test
  - Single-node to multi-node scalability

The screenshot shows a Microsoft Excel spreadsheet titled "OIA\_verification\_test\_plan\_for\_panta.xls". The spreadsheet is organized into columns: A (Script Name), B (Test Description), C (Elapse time(sec)), D (Result (pass/Fail)), and E (Comments). The data is as follows:

Script Name	Test Description	Elapse time(sec)	Result (pass/Fail)	Comments
<b>Tablespace creation</b>				
cr_tbs_data.sql	Create tablespace ts_data	180	PASS	What should
build_out_tbs.sh	Add datafiles to tablespace ts_data	126	PASS	SET Timing is
build_out_temp.sh	Add datafiles to tablespace ts_temp	8	PASS	
<b>Create user</b>				
cr_user.sh	Create test user and allocate default tablespace as ts_data	21	PASS	
<b>Data Generation</b>				
make_data.sh	Generate data for the twelve tables	4513	PASS	
make_sales.sh	Extra data generation steps for two large tables	2287	PASS	
<b>Data Loading test</b>				
cr_ext_tables.sql	Create the external tables for the load.	100	PASS	How much ds
cr_tables_and_load.sql	Validate external tables are accessible on all nodes	300	PASS	Where is the
cr_tables_and_load.sql	Create, load, and analyze the twelve tables	8456	PASS	Speed of com
additional_load_sr.sql	Complete incremental load of the two largest tables	3194	PASS	ROW COUNT
<b>Multi-user tests</b>				
Single node test => keep adding users until the average response is greater than 300 seconds		<b>Avg Response Time</b>		
run_test.sh	1 user	78	PASS	Average resp
	2 users	93	PASS	Identify log file
	4 users	143	PASS	
	8 users	198	PASS	
	16 users	240	PASS	
	32 users	298	PASS	
	64 users	321	FAIL	
	Run the single node test on two nodes => should be able to do twice as many users	296	PASS	Keeping addi
	Run the single node test on three nodes => should be able to do three times as many users	299	PASS	
<b>Backup and Recovery</b>				
	Complete a full online backup of the database (60000000)			

# Oracle Optimized Warehouses

## Superior Customer Experience



# Oracle Optimized Warehouses

## Superior System Performance





# Optimized Warehouse Roadmap

- **Phase One: Bundling, Validation, and Optimization**
  - Deliver pre-packaged, pre-installed, and pre-configured
  - Validation and Calibration of balanced configurations
  - Built from off-the-shelf components
- **Phase Two: 2<sup>nd</sup> generation Optimized Warehouse**
  - Enhance Optimized Warehouse with significantly differentiated functionality
  - Integrated management tools
  - In-depth product-development efforts
- **Phase Three: Optimized Solution Warehouse**
  - Include built-in Analytics and Business Intelligence
  - Industry specific Business Intelligence schemas & applications
  - Enables business-level 'solution selling'

# The Oracle Difference

## Optimized Data Warehouse

- Reduced risk
  - Proven high availability & performance
  - Enterprise ready security
  - Single support call
- Reduced time to decisions
  - Ease of deployment
  - Proven support for mixed workloads
  - Open platform support for BI Tools, Applications & Analytics
- Flexibility
  - Scale through modular building blocks
  - Choice of platform and operating system
  - Leverage your current IT infrastructure

# Oracle Optimized Warehouse

## Benefits

### Optimized Warehouse Reference Configurations

- Consistent messages for the customer from Oracle and the hardware partner
  - Reference Configurations are joint recommendations
- Find a complete and well-balanced system quickly
  - Customers need a quick sizing starting point
- Mitigate deployment risk
  - Provide and reuse best practices
- Shorten the sales cycle

# Oracle Optimized Warehouse Benefits

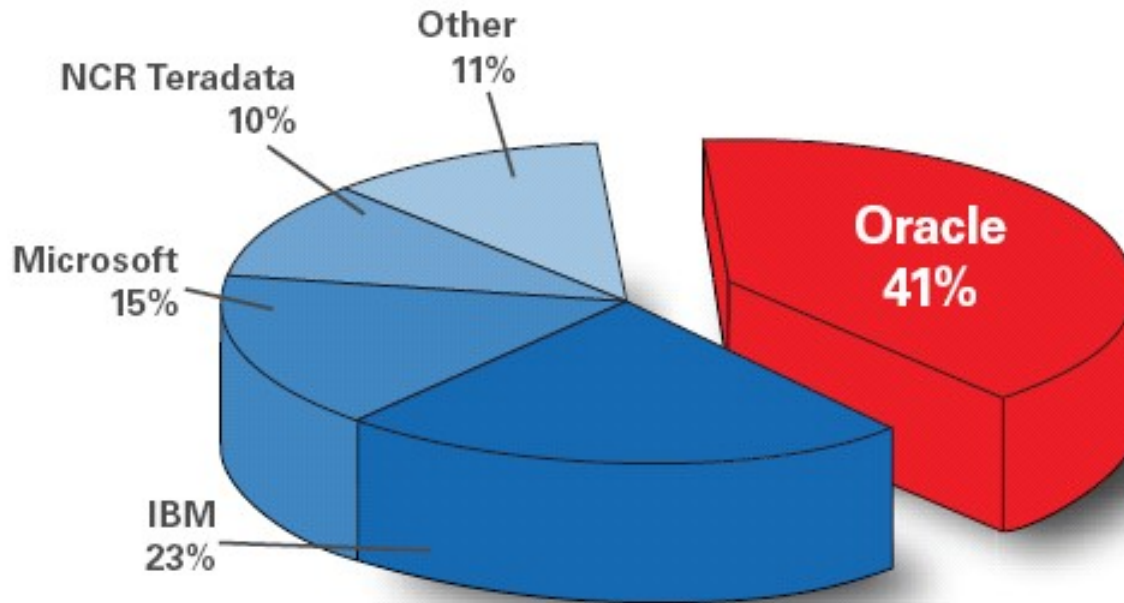
## Oracle Optimized Warehouse

- “Best of all worlds”
- Strengths of “Traditional” Appliances
  - High performance
  - Easy to buy
  - Simple to maintain
- Strengths of Oracle
  - Best-of-breed software
  - Widespread implementation expertise
  - Strong fit into existing IT infrastructure
  - Full-stack solutions

# Why Oracle Database 11g for Data Warehousing?

- Complete, integrated ETL and data-quality capabilities
- Consolidated platform optimized for grid computing
  - Eliminates the costs and inefficiencies of multiple data stores
  - Available in wide choice of configurations
- Unmatched analytic capabilities
  - Embedded OLAP, data-mining and statistics all accessible via SQL
- Next-generation BI with integrated OLTP and DSS
  - Build on world's most popular database
- Foundation of Oracle Business Intelligence solutions

# #1 for Data Warehousing



**Worldwide Data Warehouse Management  
Market Share, 2006**

# For More Information

<http://search.oracle.com>

oracle data warehousing



**or**

[www.oracle.com/bi](http://www.oracle.com/bi)



**ORACLE IS THE INFORMATION COMPANY**





ORACLE®