

# System z Premier Executive Event



## Introducing IBM Smart Analytics Optimizer – Enabling a new class of high speed business intelligence and analytical capabilities

**Dan Wardman**  
Vice President  
Information Management  
Mainframe Software  
IBM Software Group



## The Resurgence of Data Warehousing and Business Intelligence on System z

→ **IBM has invested hundreds of millions of dollars to bring new state of the art capabilities and solutions to System z in support of customers' warehouse and BI requirements**

- A 2007 study by IDUG found that nearly 50% of IDUG respondents are already using DB2 for z/OS for data warehousing. 78% indicated a desire for more capabilities in warehousing, query and reporting.
- *Analysts Agree! IBM's Data Warehousing & BI breadth on System z is a game changer in the market.*  
*Donald Feinberg, Gartner*

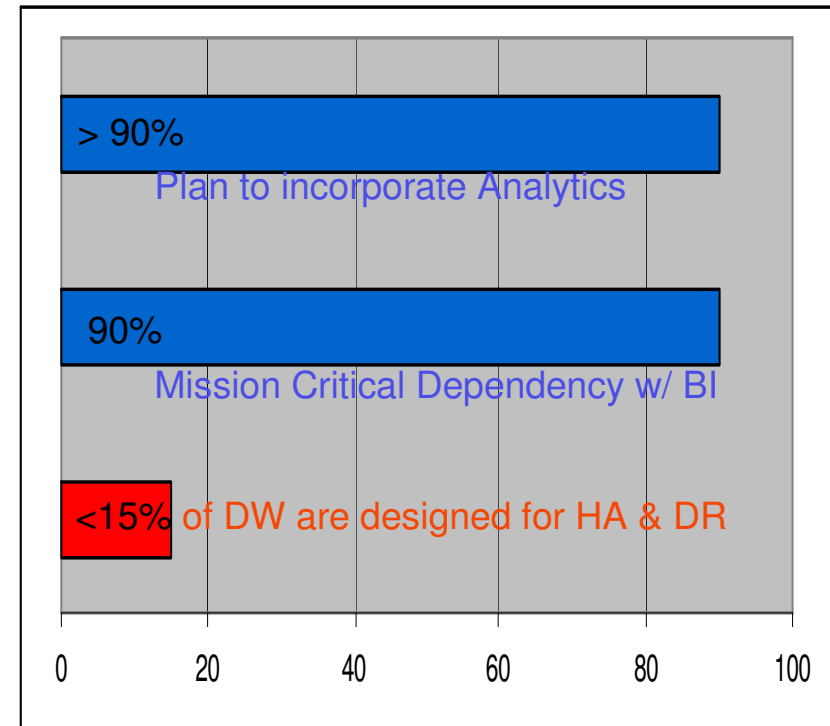


**Gartner.**

# Mission Critical Workloads Require Highest QoS

- **More than 90%** of Global 2000 companies plan to incorporate analytics into multiple operational applications that access the data warehouse by 2010, but fewer than 15% of data warehouses have been designed to provide high availability, failover, disaster recovery and the remaining components of mission-critical systems.
- By the end of 2009, **90% of Global 2000** companies will have implemented some type of mission-critical dependency between the warehouse and at least one revenue supporting or cost-controlling operational application — up from less than 25% in 2007.
- **Fewer than 15% of data warehouses** in 2007 have been designed to provide high availability, failover, disaster recovery and the remaining components of mission-critical systems.

Global 2000 customers



- Sounds like a good match for System z value proposition
  - Proven reliability and continuous availability capabilities
  - Exploiting synergistic effects of proximity to the operational data

<sup>1</sup> Operational Analytics and the Emerging Mission-Critical Data Warehouse, 14 May 2007

# IBM zEnterprise System – Best in Class Systems and Software Technologies



## Unified management for a smarter system: **zEnterprise Unified Resource Manager**

- Part of the IBM System Director family, provides platform, hardware and workload management
- Unifies management of resources, extending IBM System z® qualities of service across the infrastructure

The world's fastest and most scalable system:  
**IBM zEnterprise™ 196 (z196)**

- Ideal for large scale data and transaction serving and mission critical applications
- Most efficient platform for Large-scale Linux® consolidation
- Leveraging a large portfolio of z/OS® and Linux on System z applications
- Capable of massive scale up, over 50 Billion Instructions per Second (BIPS)

Scale out to a trillion instructions per second:  
**IBM zEnterprise BladeCenter® Extension (zBX)**

- Selected IBM POWER7™ blades and IBM System x® Blades<sup>1</sup> for tens of thousands of AIX® and Linux applications
- High performance optimizers and appliances to accelerate time to insight and reduce cost
- Dedicated high performance private network



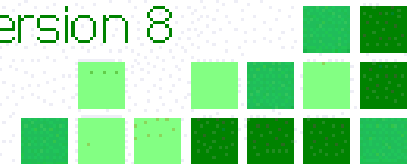
<sup>1</sup> All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.

## Warehouse/BI Highlights of DB2 for z/OS V8

- Materialized Query Tables
  - **Improved performance for query**
- Multi-row insert and fetch
  - **Improved speed of warehouse ETL and query**
- Online Schema evolution
  - **Improves availability and efficiency**
- Longer Table and Column names
  - **Increased compatibility with ETL and BI tools**
- Increased SQL vocabulary
  - **Increased compatibility with ETL and BI tools**
- 64 bit support
  - **Expanded size capability**

DB2 UDB for z/OS

Version 8



## Warehouse/BI Highlights of DB2 9 for z/OS

- Index over Expression
  - **Improved performance on ETL and BI tools**
- Index Compression
  - **Improved use of space (up to 50%)**
- Not Logged Tablespace
  - **More efficient for temporary tables such as Staging tables**
- Universal Table Space – partition by growth
  - **Easier to manage growth**
  - **Best of segmented tables and partitioning**
- Utility Improvements
  - **More online utility operations, reduced CPU**
- Overall DB2 9 for z/OS– reduced CPU



## Top 10 in DB2 10 for z/OS

1. CPU reductions for transactions, queries, & batch
2. Ten times more users by avoiding memory constraints
3. More concurrency for catalog, utilities, and SQL
4. More online changes for data definition, utilities and subsystems
5. Improved security with more granularity
6. Temporal or versioned data
7. SQL enhancements improve portability
8. pureXML performance and usability enhancements
9. Hash, index include columns, access path stability, skip migration, ... Pick your favorite!
10. Productivity improved for database & systems administrators, and application programmers



# OLAP Portfolio from IBM

## *Different methods for different uses*

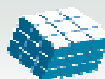
MODELING

### Specialized OLAP Modeling Tools

**Transformer**

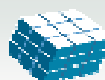
**Turbo Integrator**

**Data Architect**



**IBM Cognos PowerCube**

Optimized for broad, general purpose BI usage



**IBM Cognos TM1**

Optimized for write back and high volatility applications



**IBM IW Cubing Services**

Optimized for very large datasets with very large dimensions

OLAP

- Large read only user communities
  - Moderate 'focused area' data sets
  - Pre-aggregated static fast results
- 
- Off-line portable/ partitioned storage
  - Rapid startup for advanced business user self service modeling
  - Automatic time series analysis & trending
  - Point in time data
  - Personal and corporate data sources

- Read/write user communities
  - 64 bit in-memory fast performance
  - Larger highly dynamic data sets
- 
- Dynamic changing dimensions and hierarchy, what-if scenarios, data contribution
  - Common dimensions shared across multi-cube models
  - Budgeting and planning
  - Personal and corporate data sources

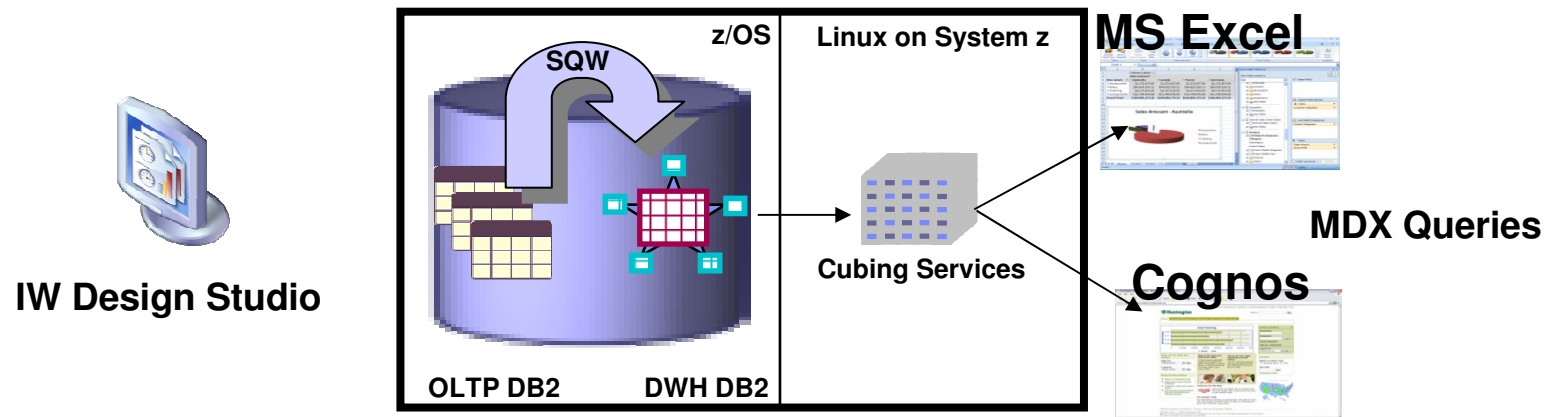
- Large user communities
  - Large data sets
  - Planned performance/optimization facilities
- 
- Optimized ROLAP
  - Zero Latency
  - High volume concurrency
  - Largest Data Volumes
  - Centralized IT management of information



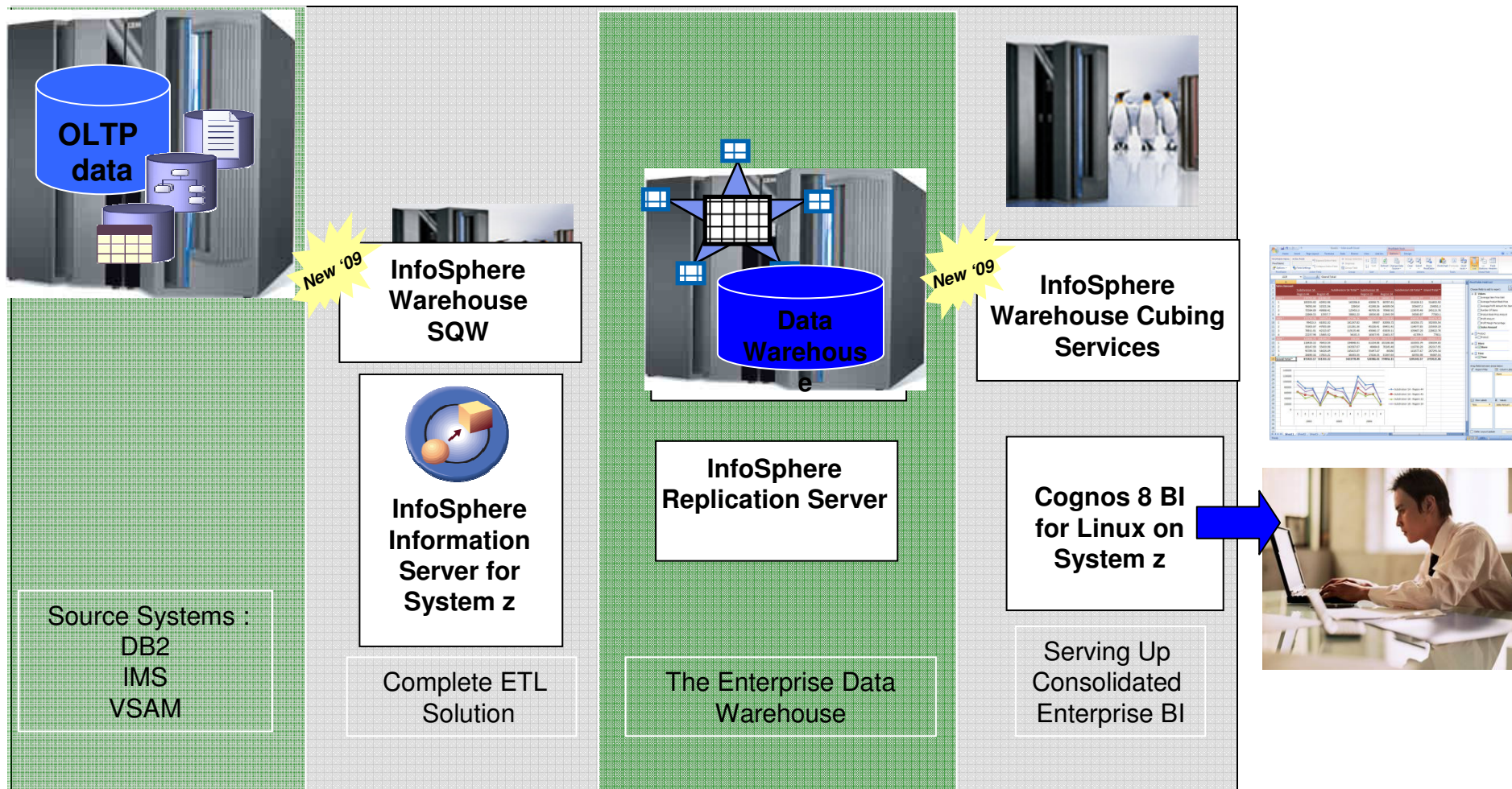
# InfoSphere Warehouse on System z

**Adds core data warehouse and analytics capability to DB2 for z/OS:**

- Advanced physical database **modeling and design**
- In-database data movement and manipulation capabilities of **SQL Warehouse Tool (SQW)**
- Optimize multidimensional reporting and analysis of data with **Cubing Services**

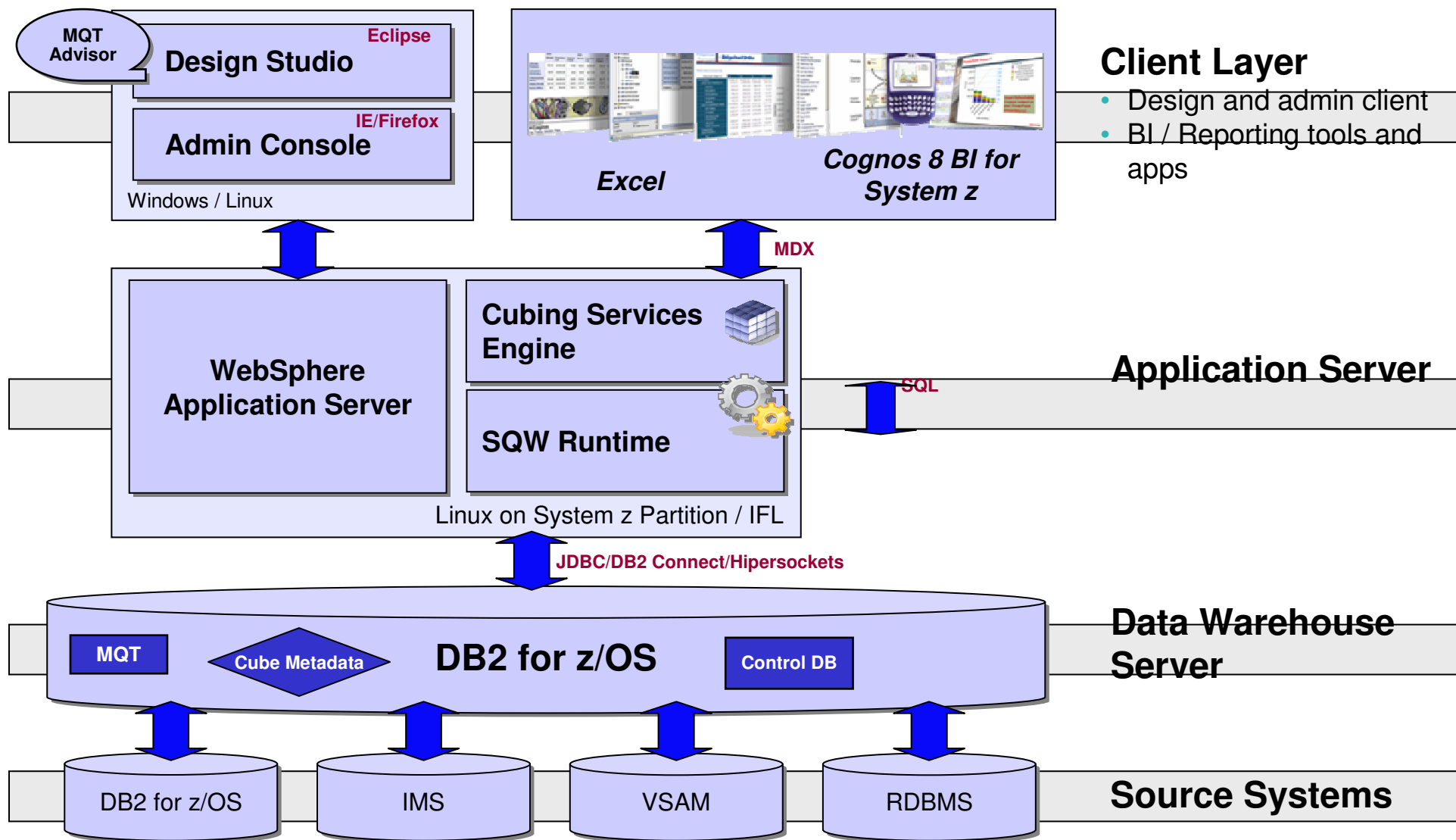


# The Data Warehouse and BI Solution on System z



***Combining the Reliability and Availability of DB2 for z/OS with Cost Effective Applications running on Linux for System z***

# InfoSphere Warehouse on System z - Architecture

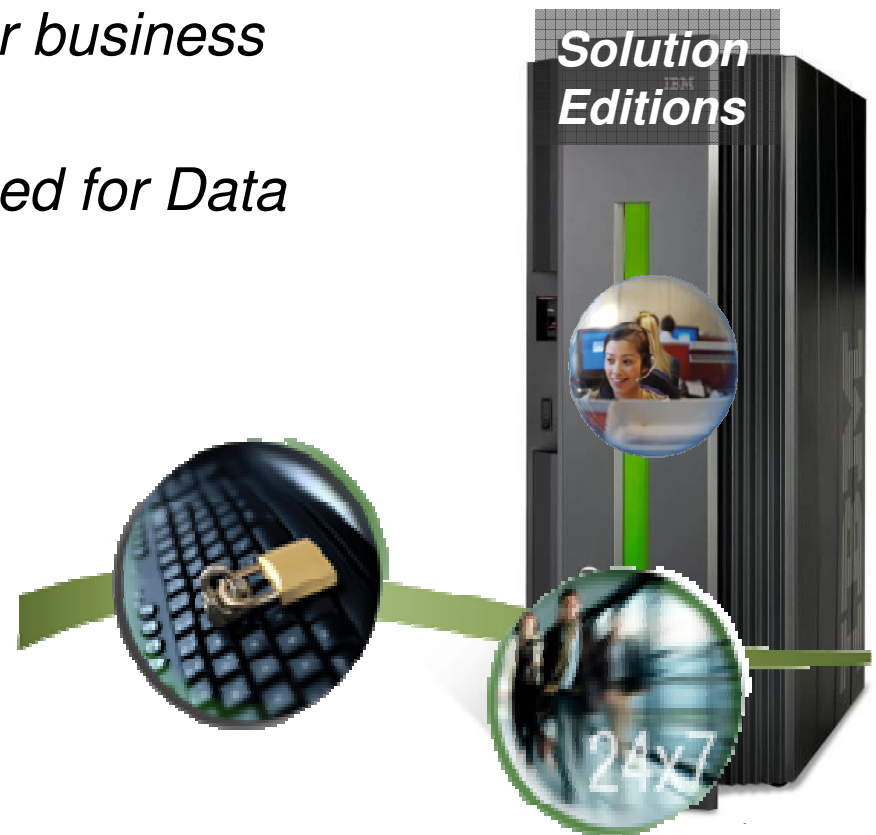


# *IBM Smart Analytics System 9600*



## Solution Edition for Data Warehousing – *Legendary System z quality, priced aggressively to fit within a shrinking budget*

- **Complete Offering:**
  - *Solution – based offering*
  - *Customized hardware to meet your business requirements*
  - *Pre-determined software stack tuned for Data Warehousing workloads*
  - *Maintenance for the hardware*
  
- *For a highly available, energy efficient*
- *Data Warehousing infrastructure designed for growth*
  
- **Created to simplify your IT decision!**



# IBM Smart Analytics System Offerings

## Deliver Analytics Value Across Platforms

*Faster Time to Value, Faster Business Results*

*Meeting clients where  
their information is...*

**5600**

**New**

*Based System x*



*Consistent value,  
Right sized to your needs*

**7600**

**2009**

*Based Power System*

**9600**

**New**

*Based on System z*

# IBM Smart Analytics System 9600

## What is it?

*The IBM Smart Analytics System 9600 is an integrated solution of hardware, software and services that enables customers to rapidly deploy cost effective game changing analytics across their business.*



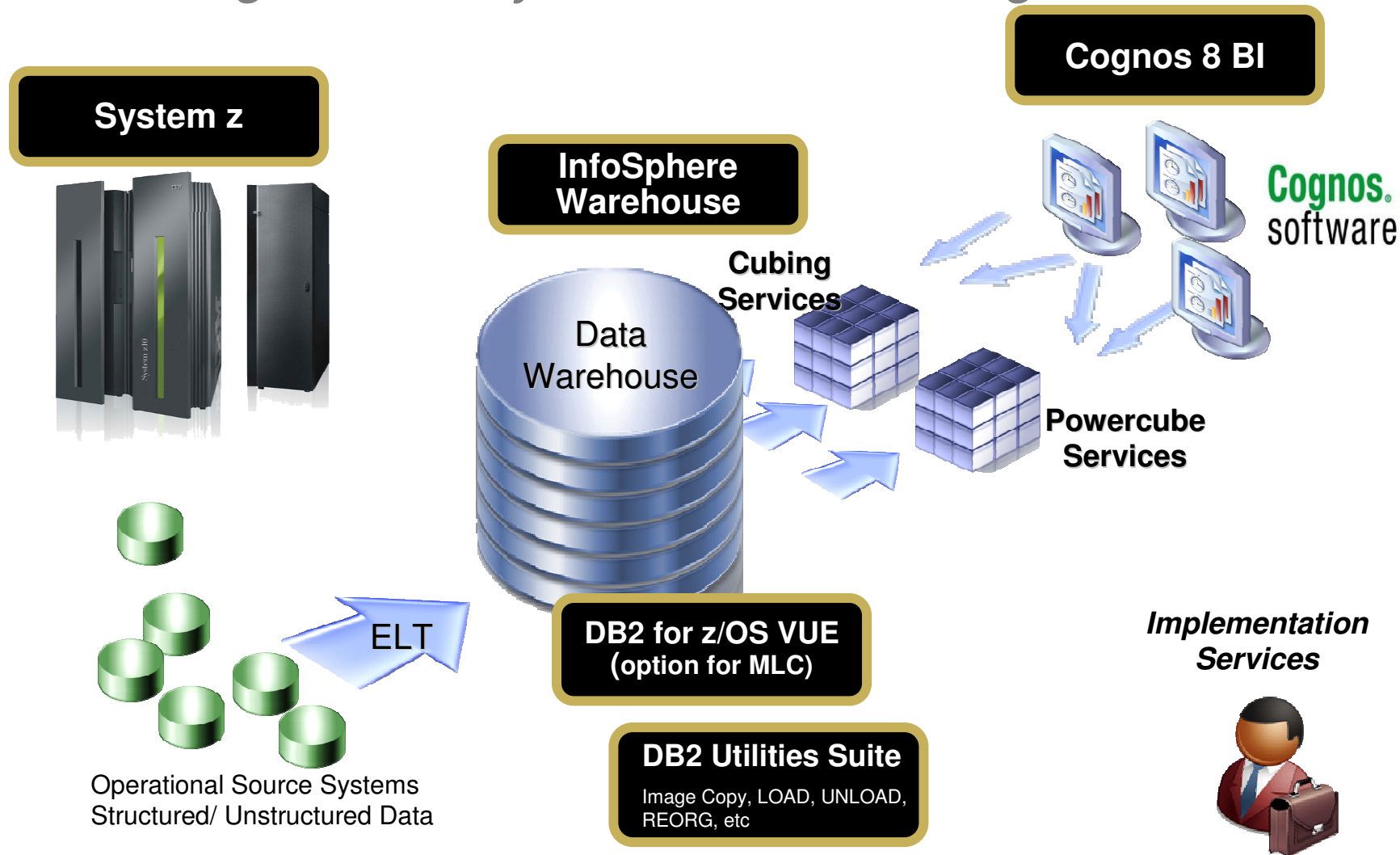
## How is it different?

- **Secure, Available Business Analytics**
  - Rapidly delivers analytic information to decision makers at the time of decision.
  - New environment for the availability, reliability and scalability necessary to stay aligned with the operational systems
- **Simplified administration**
  - Appliance-like delivery
  - Faster deployment at lower cost.
  - Leverages customers existing disaster recovery, and backup processes
- **Proven Operational Characteristics**
  - Extends the qualities of service of System z.
  - Reduces risk through extending System z manageability and security across the entire system.
- **High Value Operational BI**
  - Cost effective way to drive daily operational decisions

***Delivering business results in days, not months***

# IBM Smart Analytics System 9600

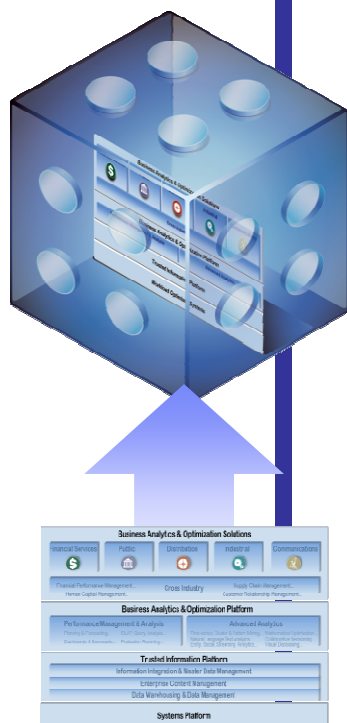
## *High Value Dynamic Warehousing*





# IBM Smart Analytics System 9600 Software

Deeply Optimized by IBM Experts... Flexible Growth...



## Powerful Data Warehouse and BI Software

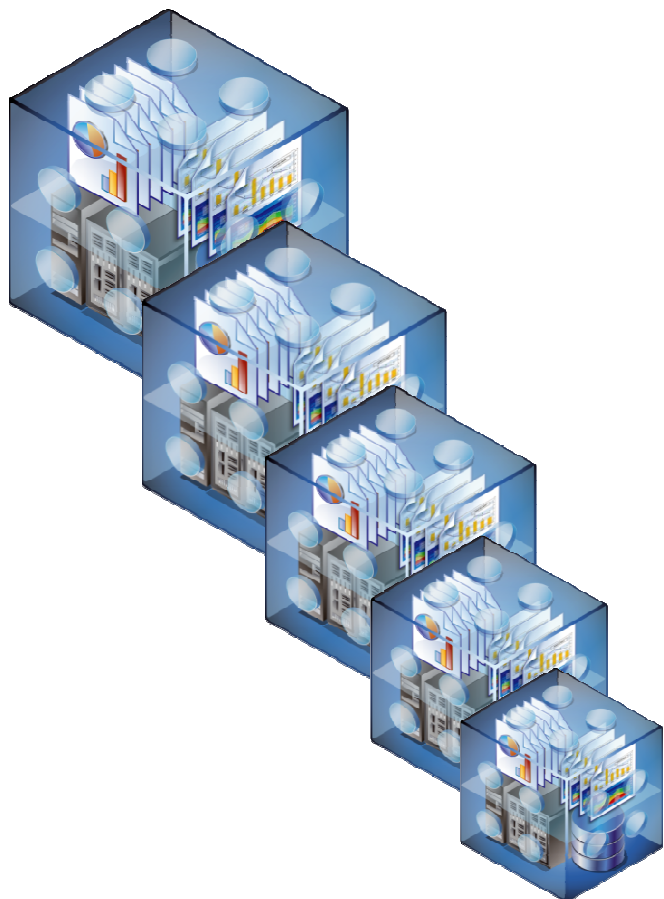
- DB2 for z/OS Value Unit Edition (primary) V9 - *Option for MLC*
- DB2 Utilities Suite V9
- DB2 Connect
- InfoSphere Warehouse on System z V9.5.2
- IBM Cognos 8.4 BI for Linux on System z  
IBM Cognos® 8 BI reporting, IBM Cognos® 8 BI analysis, IBM Cognos® 8 BI dashboard, Cognos 8 Go! Mobile, Cognos 8 Go! Search, Cognos 8 Go! Office
- z/OS Operating System Stack V1.11
- z/VM 6.1

## Optional Value Priced Add-ons

- Tivoli OMEGAMON for DB2 Performance Expert
- Tivoli Directory Server
- InfoSphere Master Data Management Server
- InfoSphere Information Server
- InfoSphere Replication Server
- Q-Rep, CDC and Event Publisher eligible
- InfoSphere Federation Server plus Classic Federation on System z
- SPSS
- Tivoli ITCAM, ITUAM
- Cognos Now! For Linux on System z
- Cognos Blueprints for Healthcare, Banking and others...
- BI User on-boarding application (as proposed for Smart Analytics Cloud)

# IBM Smart Analytics System

## ***Pre-Set Configurations Sized to Meet Business Requirements***



### Data Capacity Sizing

Database Size	
Amt Usable Disk*	Amt User Data*
244TB	100TB
122TB	50TB
61TB	25TB
27TB	12TB
10TB	4TB

\*Usable = Disk for compressed table data, indexes, work files, MQTs etc.

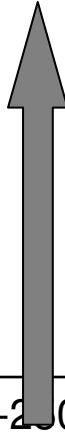
\*User Data = Raw, uncompressed, user data

# IBM Smart Analytics System

## ***Pre-Set Configurations Sized to Meet Business Requirements***



### BI Capacity Sizing

<i>Cognos Workload</i>	
Named Users	Max Concurrent Users
5000 - 10,000	100
	50
	25

## Choosing A Deployment Style



### ***Standalone Delivery***

Server delivered in  
“appliance-like”  
fashion

### ***Virtual Upgrade Delivery***

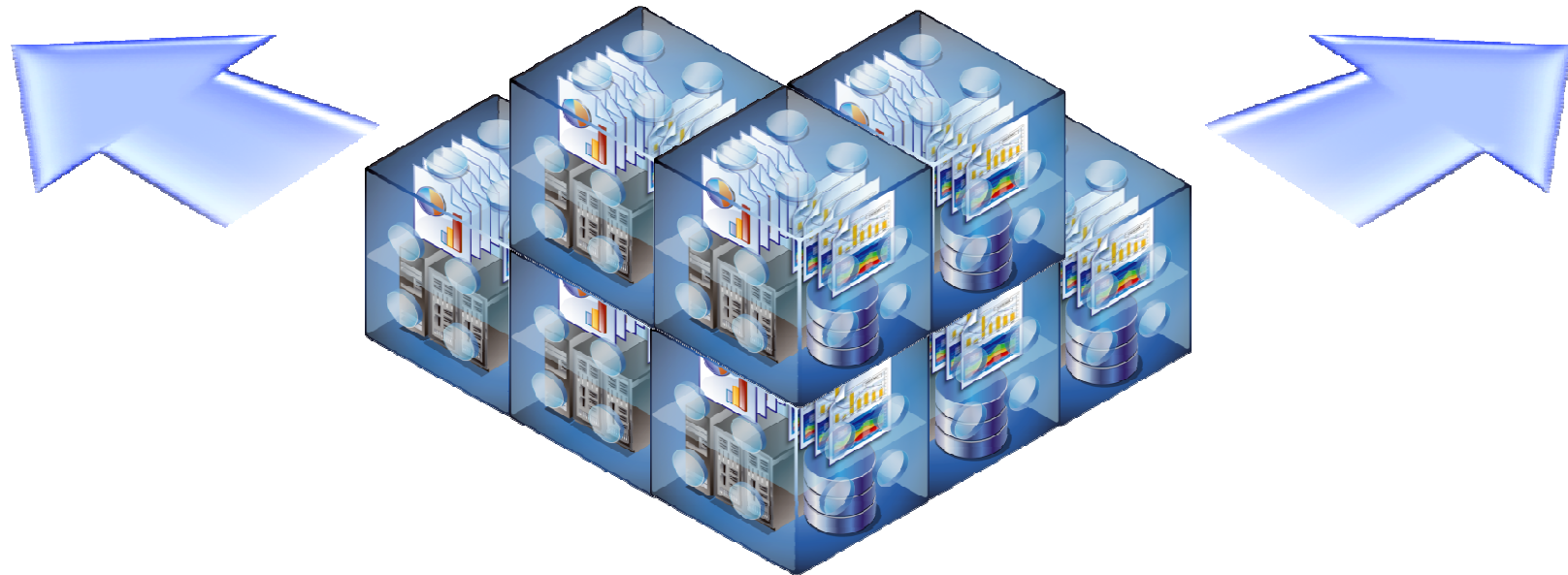
Capacity delivered as  
a “bolt-on” to  
existing server(s) or  
as new server(s)  
depending on  
customer needs in a  
“virtual appliance-  
like” fashion

# Scales for Your Unique Needs

Deployed in scalable increments to match each type of need

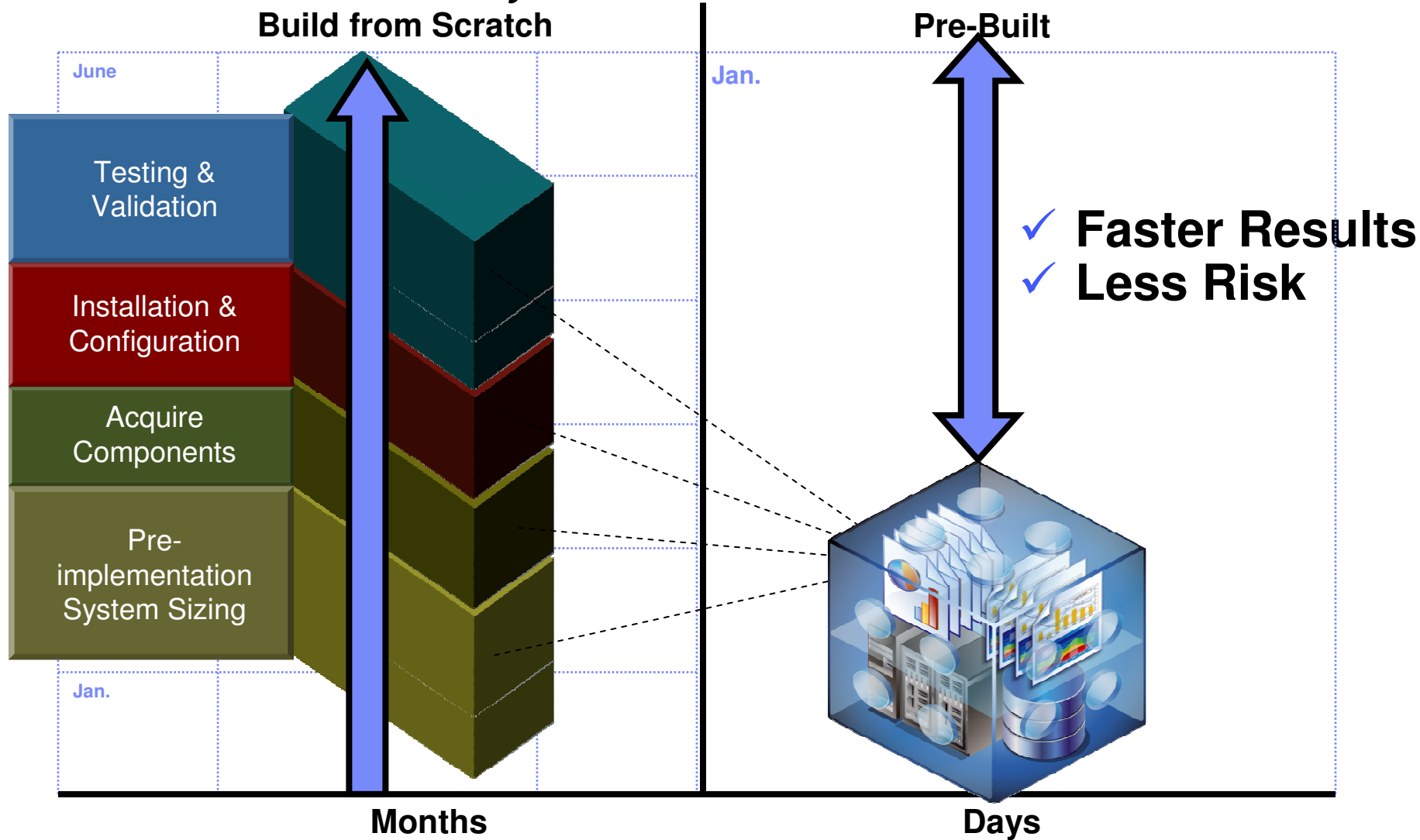
***Rapidly & easily add  
new capacity as  
information & new  
analytic capabilities grow***

***Rapidly & easily add  
new users as business  
needs expand***



***Building blocks for each growth  
increment***

# Faster Time to Value *Results in Days Versus Months*





# *IBM Smart Analytics Optimizer*

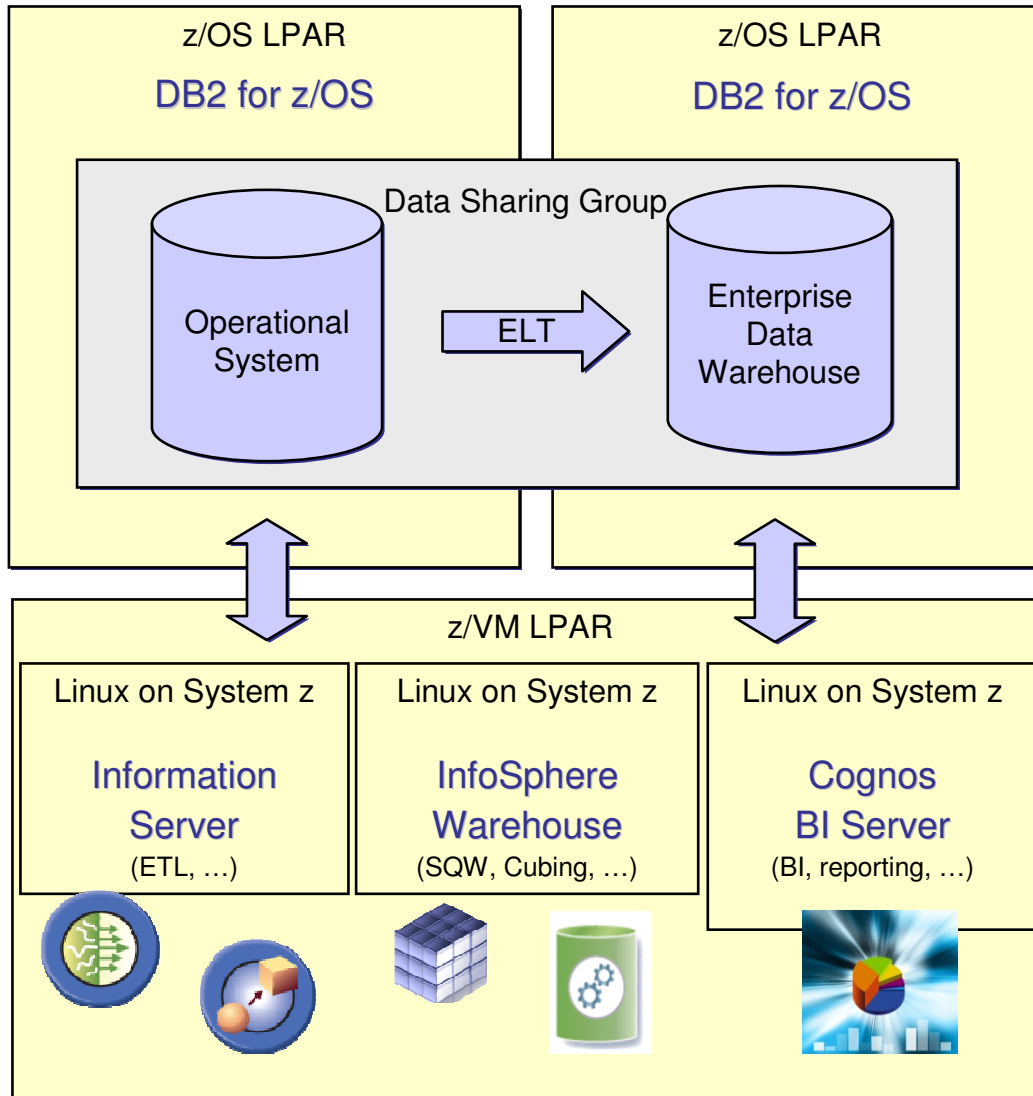


# Business Challenges and Technology Trends

- **Changing business requirements**
  - BI/DW becoming mission critical and requires OLTP-like QoS
    - reliability, continuous availability, security, mixed workload management, ...
    - orders of magnitude faster execution of complex, ad hoc queries
    - predictable query performance
  - Shift towards dynamic DW and operational BI
    - Combining OLTP and OLAP workloads
- **Traditional performance tuning tools of the trade such as indexing, prebuilt aggregates and MQTs struggling to keep the pace**
  - Require top DBA expertise and sophisticated tools
  - Even then not good enough due to ad-hoc, unpredictable nature of the workload
- **Technology trends**
  - Very large number of processor sockets and cores
  - Massive amounts of real memory
  - Specialized physical data designs: row-store vs. column-store



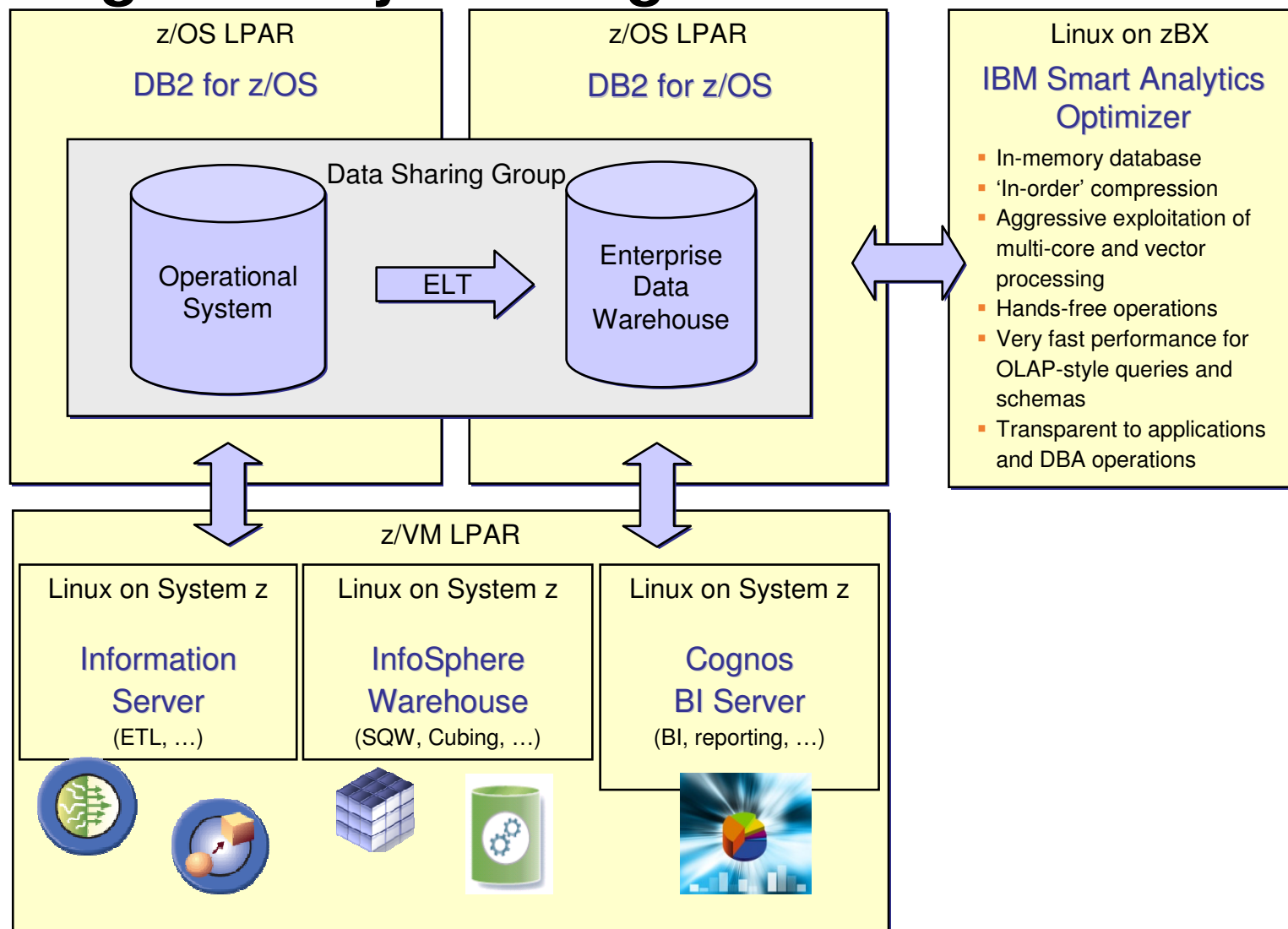
# Ultimate Consolidation Opportunity



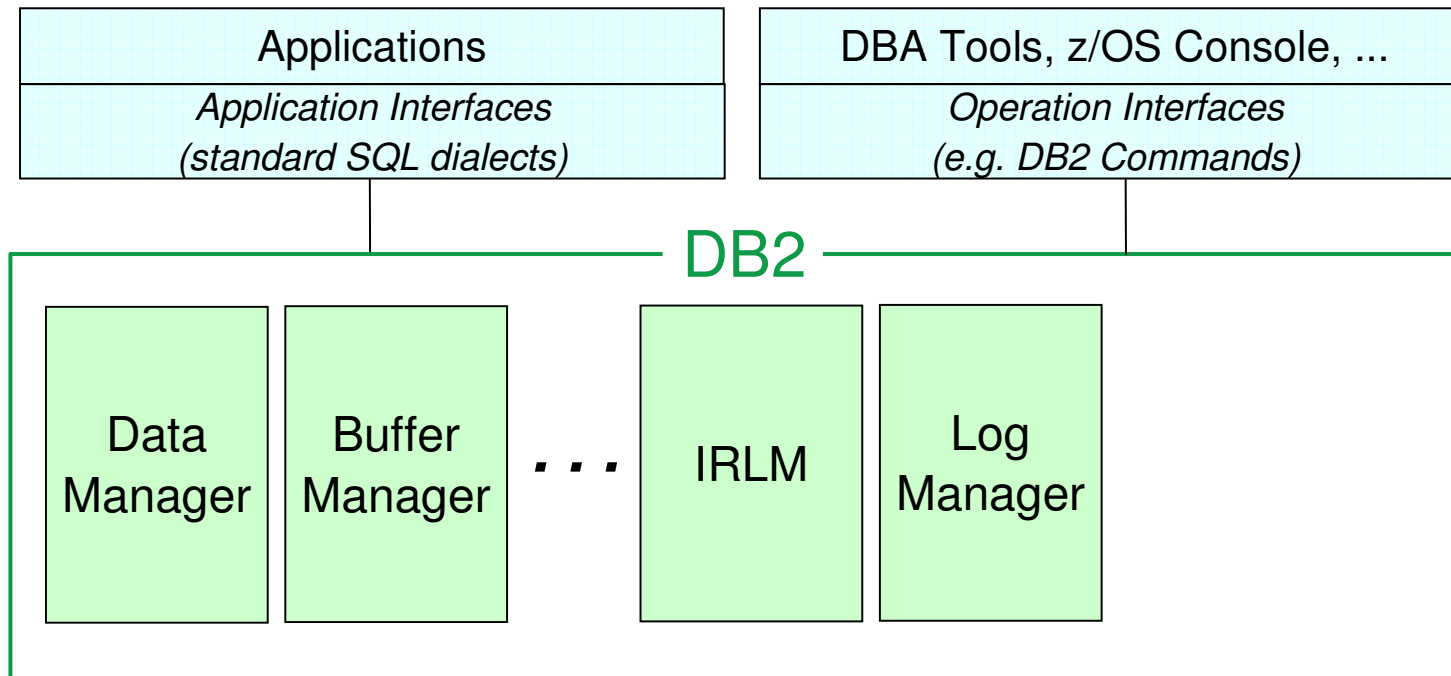
- Consolidation of mission-critical data on System z
- Leveraging existing environment, high availability, backup and governance procedures as well as skills
- Efficient data movement within a data sharing group (no network)
- Performance and TCO improvements through cubing services (data marts) and DB2 enhancements
- Complex transformations and data quality are driven from Linux on System z with Information Server

# IBM Smart Analytics Optimizer

## Adding Industry Leading Performance

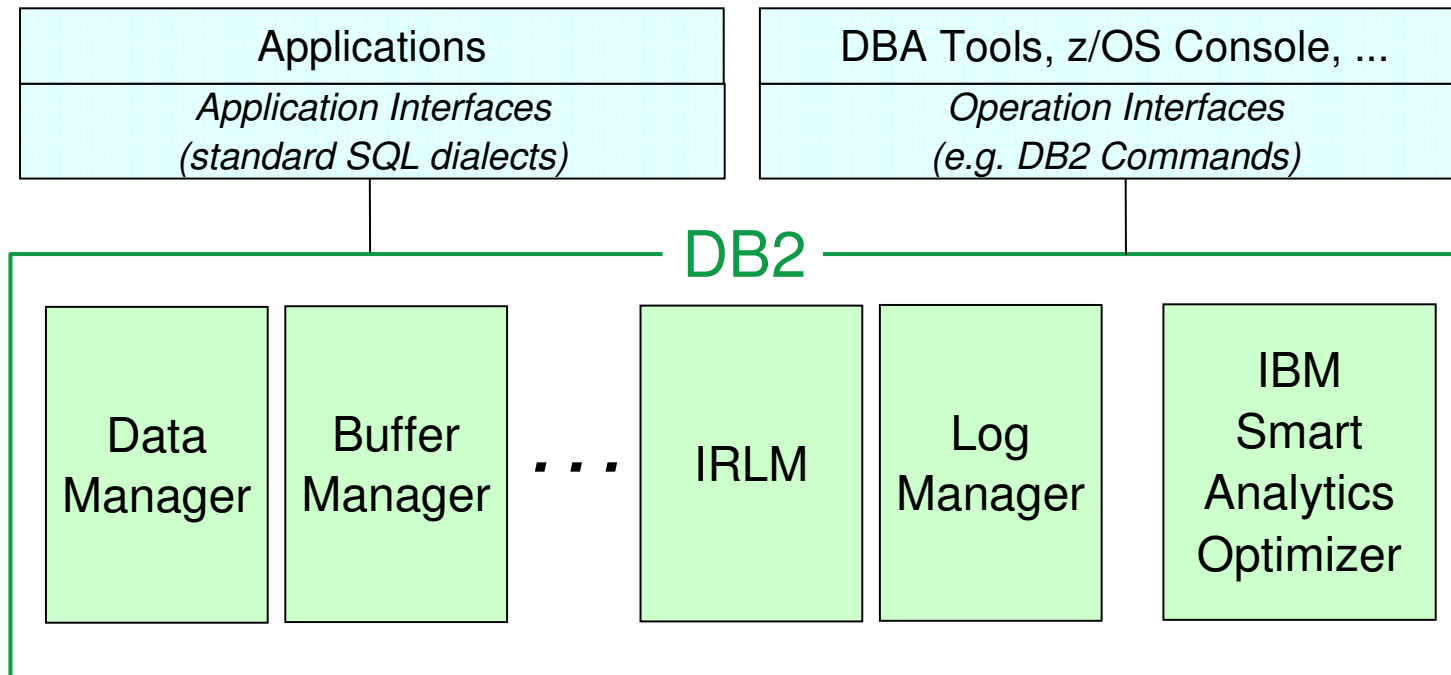


# Deep DB2 Integration within zHybrid Architecture

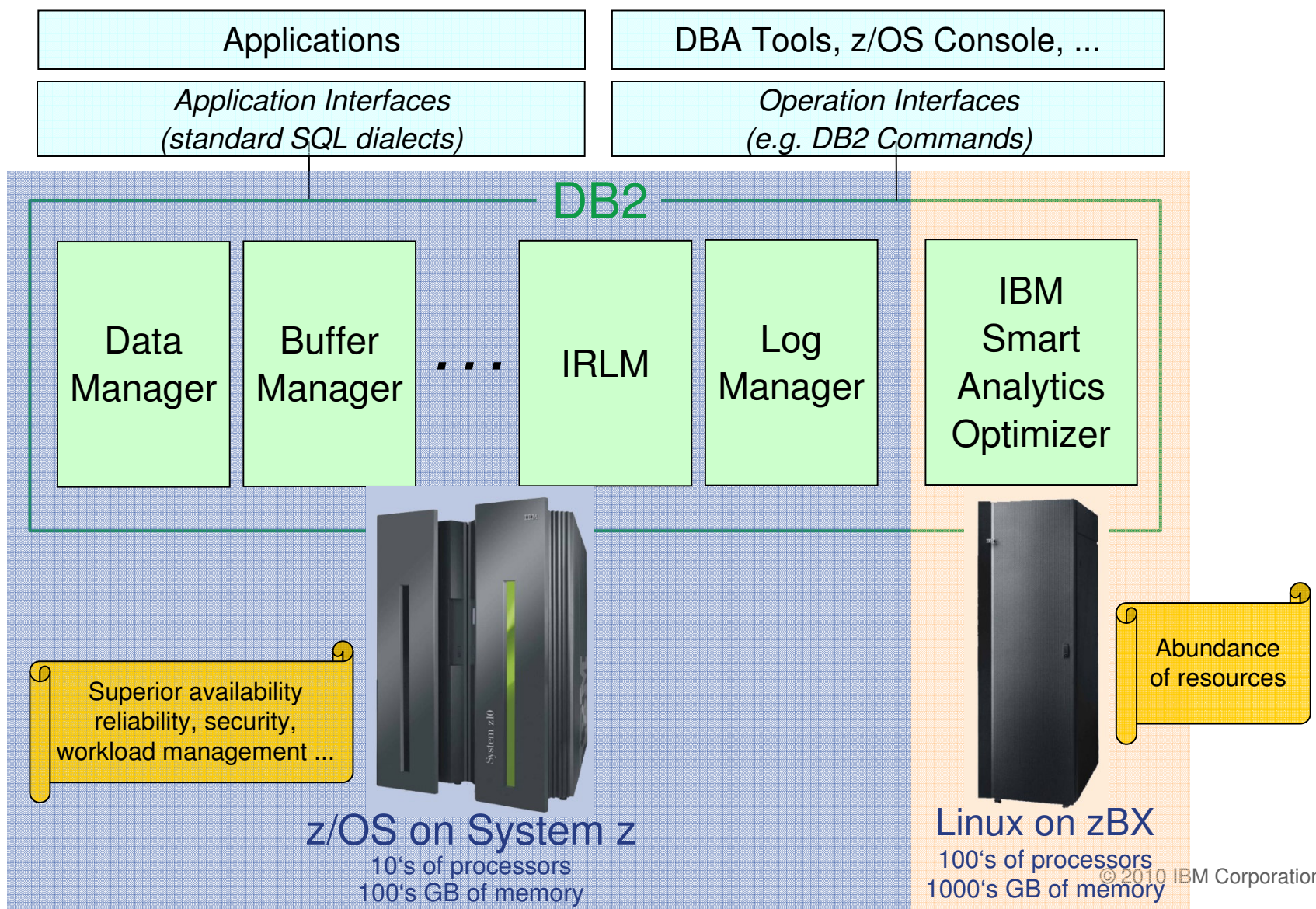


# Deep DB2 Integration within zHybrid Architecture

## IBM Smart Analytics Optimizer as a Virtual DB2 Component



# Deep DB2 Integration within zHybrid Architecture IBM Smart Analytics Optimizer as a Virtual DB2 Component



# IBM zEnterprise System – Best in Class Systems and Software Technologies

## *A system of systems that unifies IT for predictable service delivery*



Unified management for a smarter system:  
**zEnterprise Unified Resource Manager**

- Part of the IBM System Director family, provides platform, hardware and workload management
- Unifies management of resources, extending IBM System z® qualities of service across the infrastructure

The world's fastest and most scalable system:  
**IBM zEnterprise™ 196 (z196)**

- Ideal for large scale data and transaction serving and mission critical applications
- Most efficient platform for Large-scale Linux® consolidation
- Leveraging a large portfolio of z/OS® and Linux on System z applications
- Capable of massive scale up, over 50 Billion Instructions per Second (BIPS)



Scale out to a trillion instructions per second:  
**IBM zEnterprise BladeCenter® Extension (zBX)**

- Selected IBM POWER7™ blades and IBM System x® Blades<sup>1</sup> for tens of thousands of AIX® and Linux applications
- High performance optimizers (**IBM Smart Analytics Optimizer**) and appliances to accelerate time to insight and reduce cost
- Dedicated high performance private network

<sup>1</sup> All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.

## Enabling Technology – IBM Research Project BLINK

- **Various Compression Techniques**
  - Enables in-memory database
  - Order-preserving
  - Frequency partitioning
- **Register-store: a combination of row- and column-based stores**
- **Multi-core friendly scans**
  - Massive scale-out parallelism
  - Scans on compressed data
  - Vector processing
  - Evaluation of all predicates in parallel
- **Selective schema melting**

## OLTP vs. DW

- **In a typical transactional workload, you normally fetch and use all attributes of a tuple. If you for example have a CUSTOMERS table, you wouldn't fetch the STREETNAME w/o also fetching the house number or ZIP code.**
  - A transactional query is used to fetch few, very specific records of a relation.
  
- **In typical Data Warehouse workloads, you tend to fetch only a small subset of each record.**
  - The tables are usually very wide, having multiple measure columns.
  - Queries almost never touch all attributes of the tuples but only a small subset of the available attributes.
  - A query usually needs to evaluate/aggregate many tuples per relation.



# Row-Store – Optimal Choice for OLTP

- In traditional DBMS, we use a **Row – Store** approach where each row is stored contiguously and where multiple rows are stored sequentially in I/O optimized data structures.
- If only few attributes are required, the complete row needs to be fetched and uncompressed.
- Lots of the data is moved and decompressed w/o even being used.

COL 1	COL 2	COL 3	COL 4	COL 5	---	COL n	COL 1
COL 2	COL 3	COL 4	COL 5	---	COL n	COL 1	COL 2
COL 3	COL 4	COL 5	---	COL n	COL 1	COL 2	COL 3
COL 4	COL 5	---	COL n	COL 1	COL 2	COL 3	COL 4
COL 5	---	COL n	COL 1	COL 2	COL 3	COL 4	COL 5
---	COL n	COL 1	COL 2	COL 3	COL 4	COL 5	---

COL 1	COL 2	COL 3	COL 4	COL 5	---	COL n	COL 1
COL 2	COL 3	COL 4	COL 5	---	COL n	COL 1	COL 2
COL 3	COL 4	COL 5	---	COL n	COL 1	COL 2	COL 3
COL 4	COL 5	---	COL n	COL 1	COL 2	COL 3	COL 4
COL 5	---	COL n	COL 1	COL 2	COL 3	COL 4	COL 5
---	COL n	COL 1	COL 2	COL 3	COL 4	COL 5	---

Page X

Page Y

While a **Row – Store** is very efficient for transactional workloads, it is suboptimal for analytical workloads where only a subset of the attributes is needed!

# Column-Store: Optimized for Certain DW Workloads

- Query Engines, which are optimized for analytical queries, sometimes use a **Column – Store** approach.
- In a **Column – Store**, the data of a specific column is stored sequentially before the data of the next column begins.
- If attributes are not required for a specific query execution, they simply can be skipped, not causing any I/O or decompression overhead.

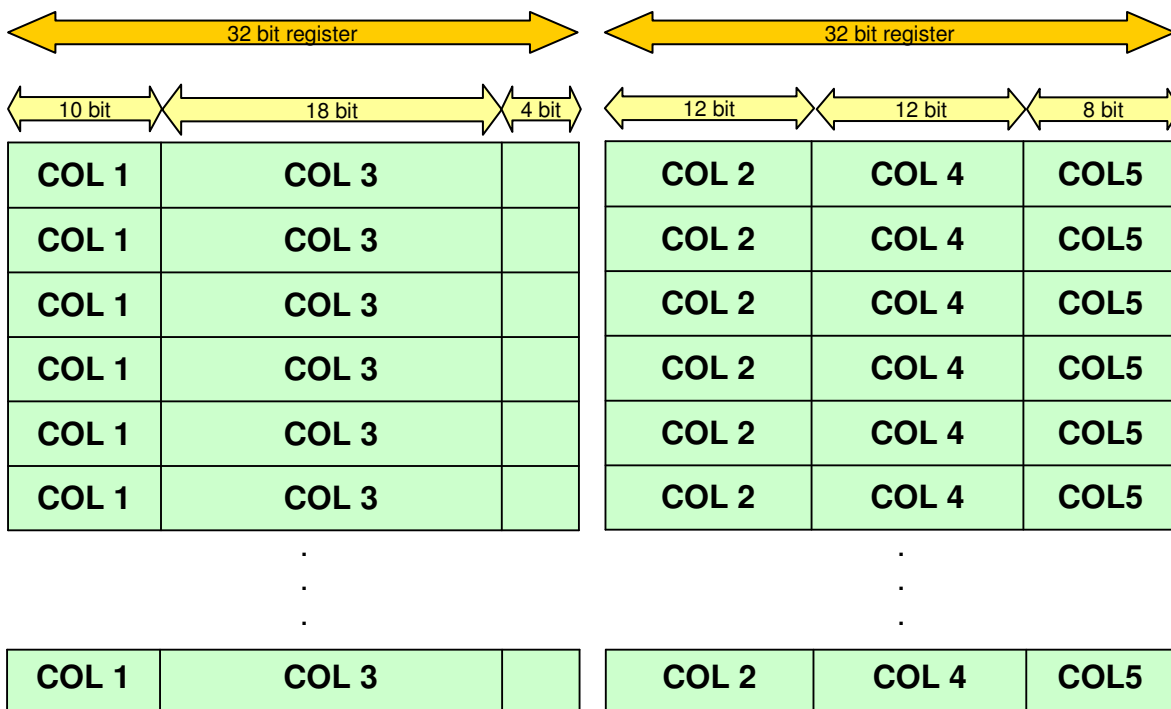
COL 1	COL 1	COL 1	COL 1	COL 1	COL 1
COL 1	COL 1	COL 1	COL 1	COL 2	COL 2
COL 2	COL 2	COL 2	COL 2	COL 2	COL 2
COL 2	COL 2	COL 3	COL 3	COL 3	COL 3
COL 3	COL 3	COL 3	COL 3	COL 3	COL 3
COL 4	COL 4	COL 4	COL 4	COL 4	COL 4
COL 4	COL 4	COL 4	COL 5	COL 5	COL 5
COL 5	COL 5	COL 5	COL 5	COL 5	COL 5
COL 5	COL 5	---	COL n	COL n	COL n
COL n	COL n	COL n	COL n	COL n	COL n
COL n					

In a **Column – Store**, the data is also compressed sequentially for a column. This is an optimized approach if you plan to perform a sequential scan over your data. Random access to specific attributes in this store is not performing well.

This is normally handled by limiting the number of tuples per column before the next column is stored. (The data is split into blocks.)

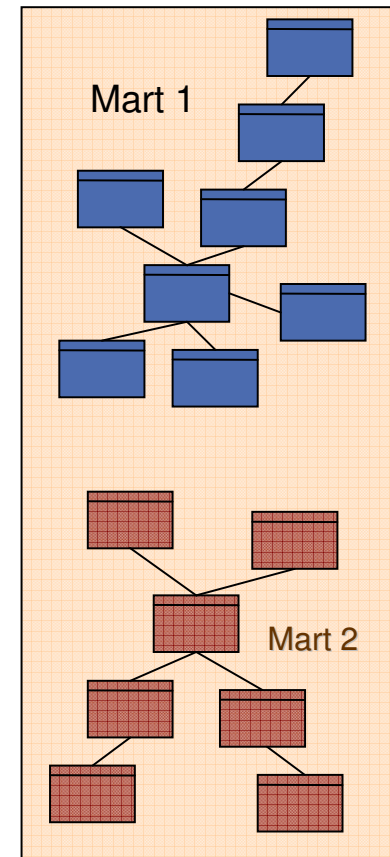
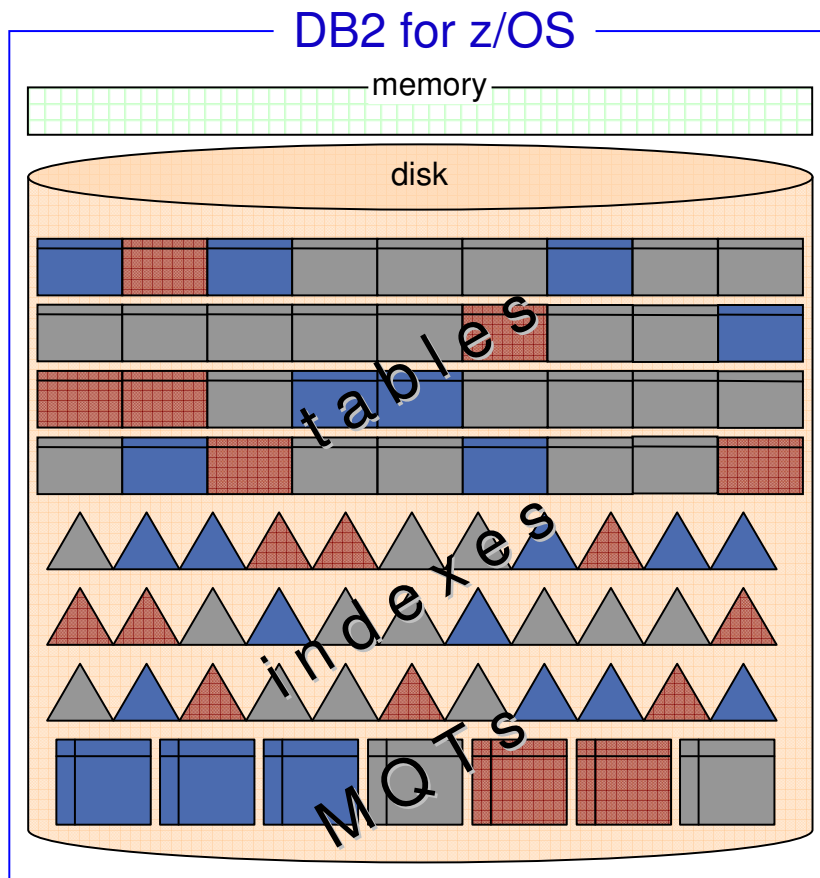
# Register-Store

- Within a **Register – Store**, several columns are grouped together.
- The sum of the width of the compressed columns doesn't exceed a register compatible width. This could for example be 32 or 64 bit for a 64 bit system. It doesn't matter how many columns are placed within the register – wide data element.
- It is beneficial to place commonly used columns within the same register – wide data element. But this requires dynamic knowledge about the executed workload (runtime statistics).
- Having multiple columns within the same register – wide data element prevents ANDing of different results.

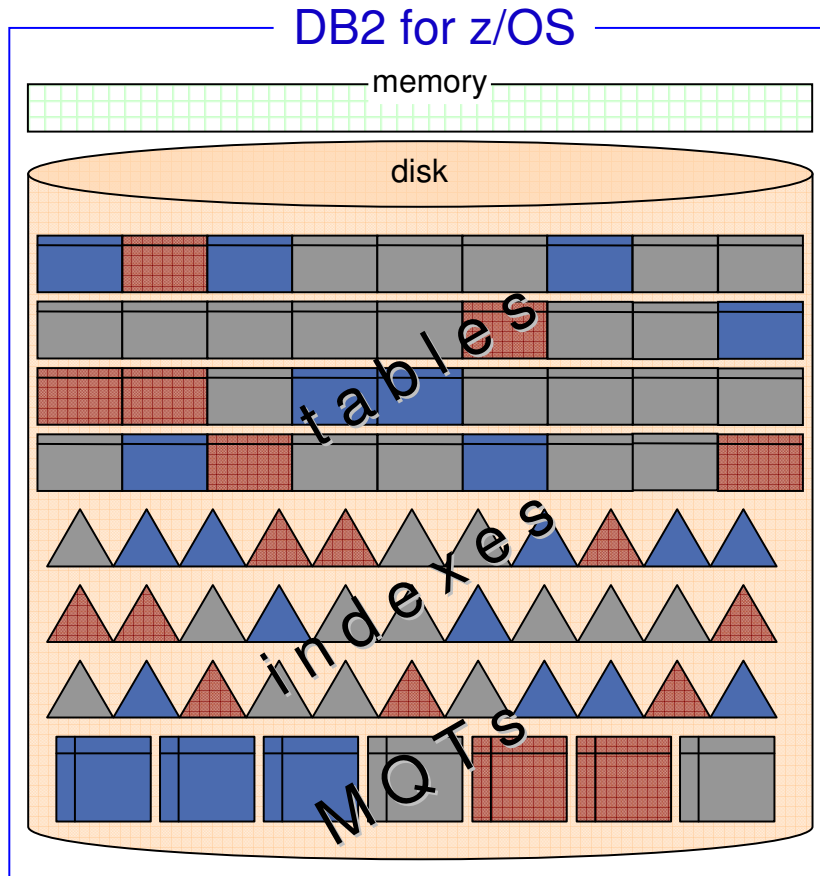


The **Register – Store** is an optimization of the Column – Store approach where we try to make the best use of existing hardware. Reshuffling small data elements at runtime into a register is time consuming and can be avoided. The **Register – Store** also delivers good vectorization capabilities.

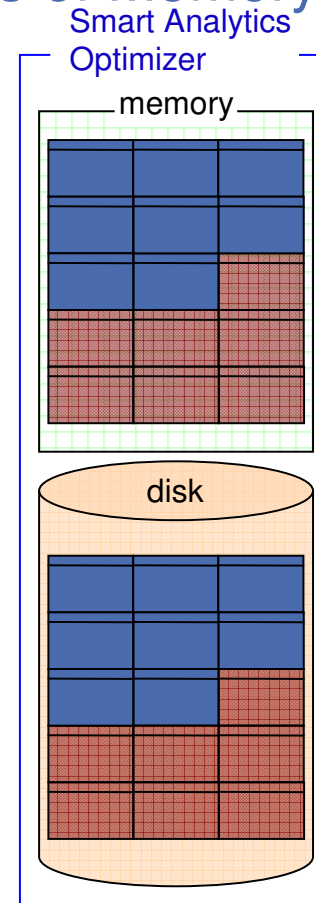
# Data View



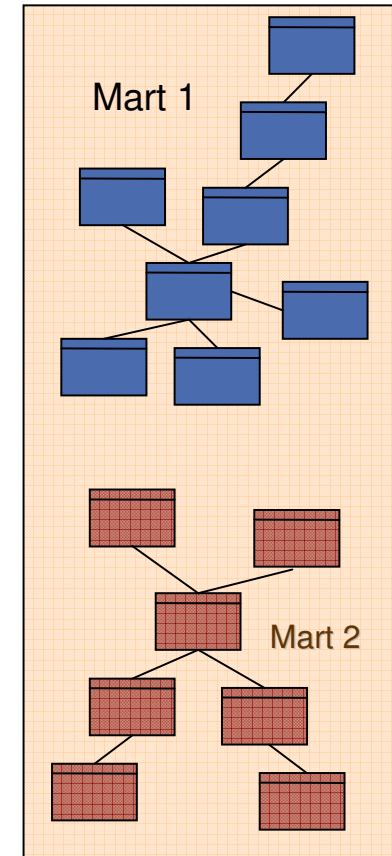
# Data view: Marts – Redundant Sets of Memory Resident Tables



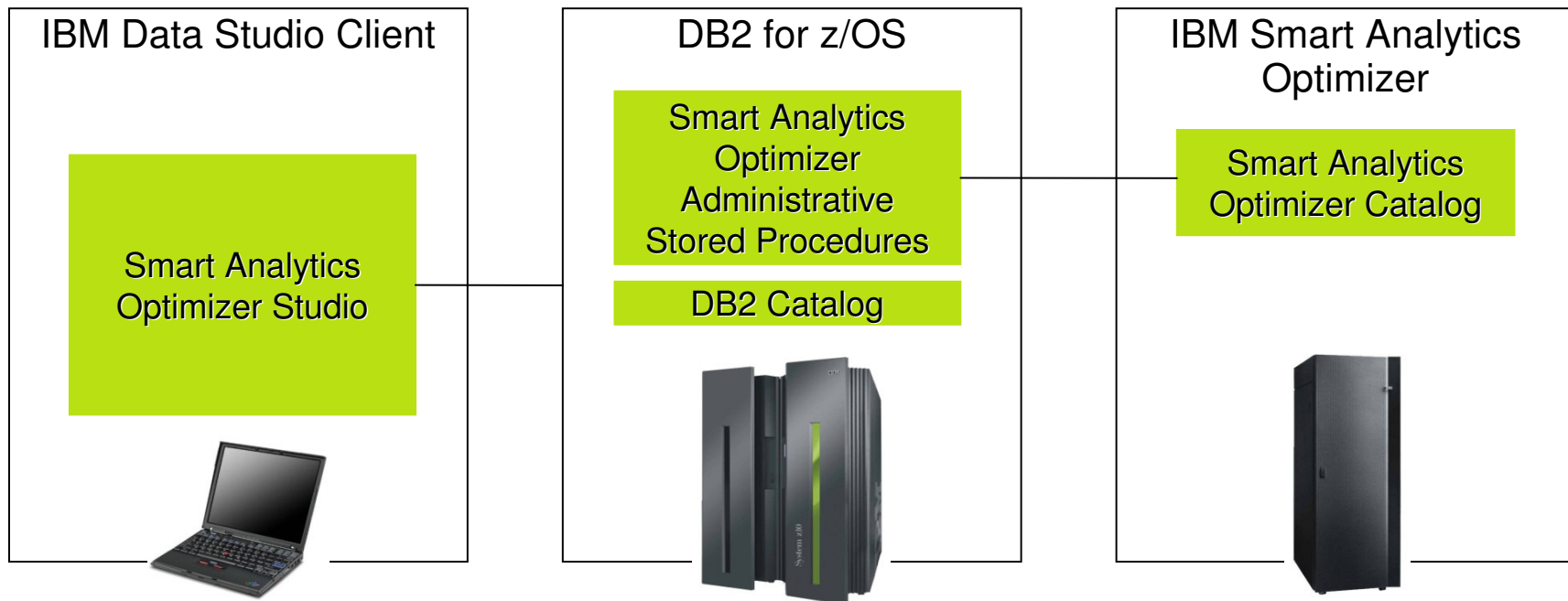
- DB2 continues to *own* and manage all data
- Access performance is influenced by traditional tuning mechanisms such as indexing, MQTs, aggregates, ...
- Typical usage: Enterprise Data Warehouse, large Data Marts



- Smart Analytics Optimizer contains fully memory resident, compressed *copies* of performance critical tables grouped into logically connected *mart*s, bound ideally by star schema constraints
- Similar but much broader than MQTs: no column projections, no row restrictions, no row aggregations
- Typical usage: Data Marts, MQTs consolidation and replacement

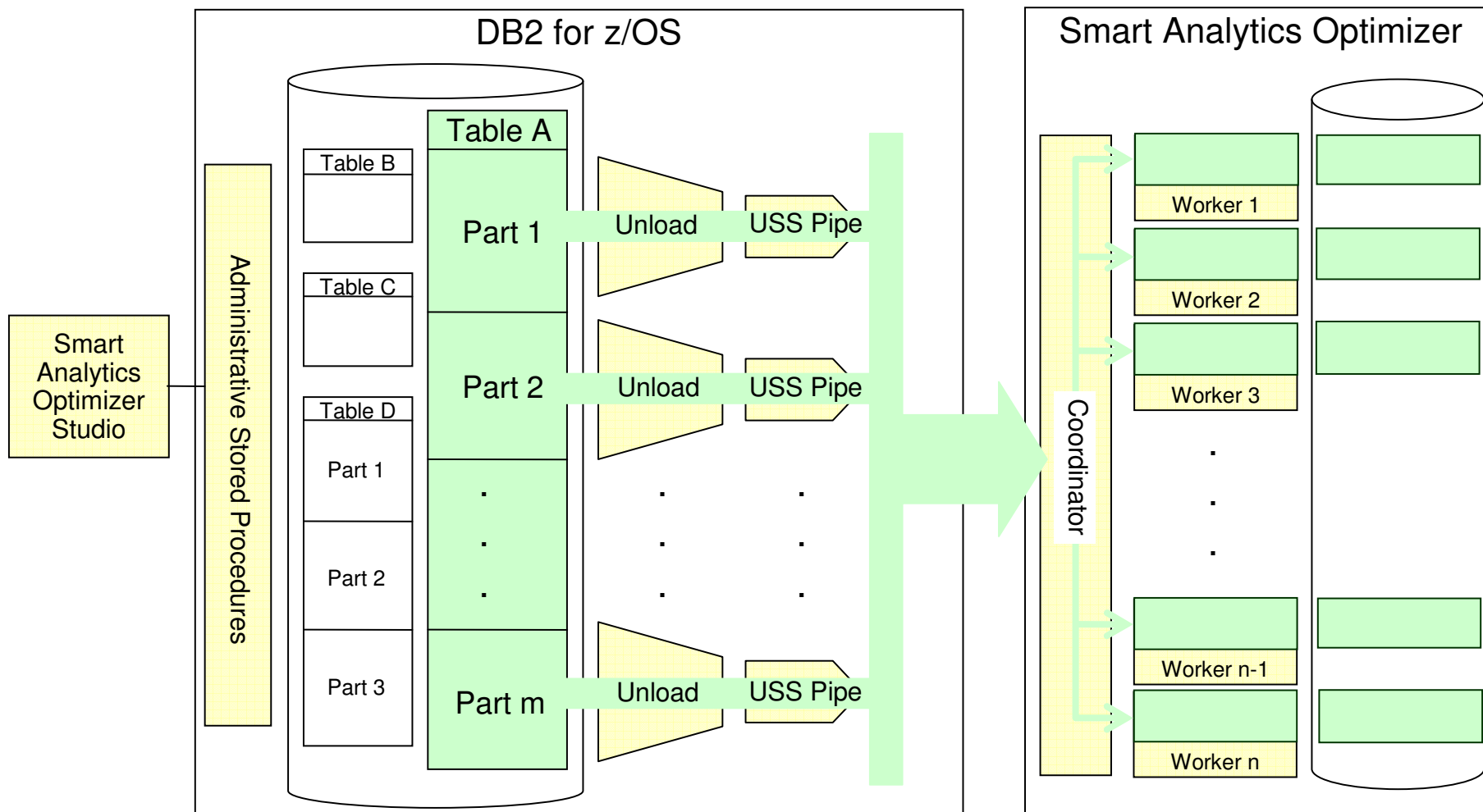


# Smart Analytics Optimizer Mart Definition and Deployment



- The marts need to be defined and deployed to Smart Analytics Optimizer before data is loaded and queries sent to it for processing.
  - Definition: identifying tables and relations that make up marts.
  - Deployment: making marts known to DB2, i.e. storing mart meta data in the DB2 and Smart Analytics Optimizer catalog.
- Smart Analytics Optimizer Studio guides you through the process of defining and deploying marts, as well as invoking other administrative tasks.
- Smart Analytics Optimizer Stored Procedures implement and execute various administrative operations such as mart deployment, load and update, and serve as the primary administrative interface to Smart Analytics Optimizer from the outside world including Smart Analytics Optimizer Studio.

# Smart Analytics Optimizer Mart Load

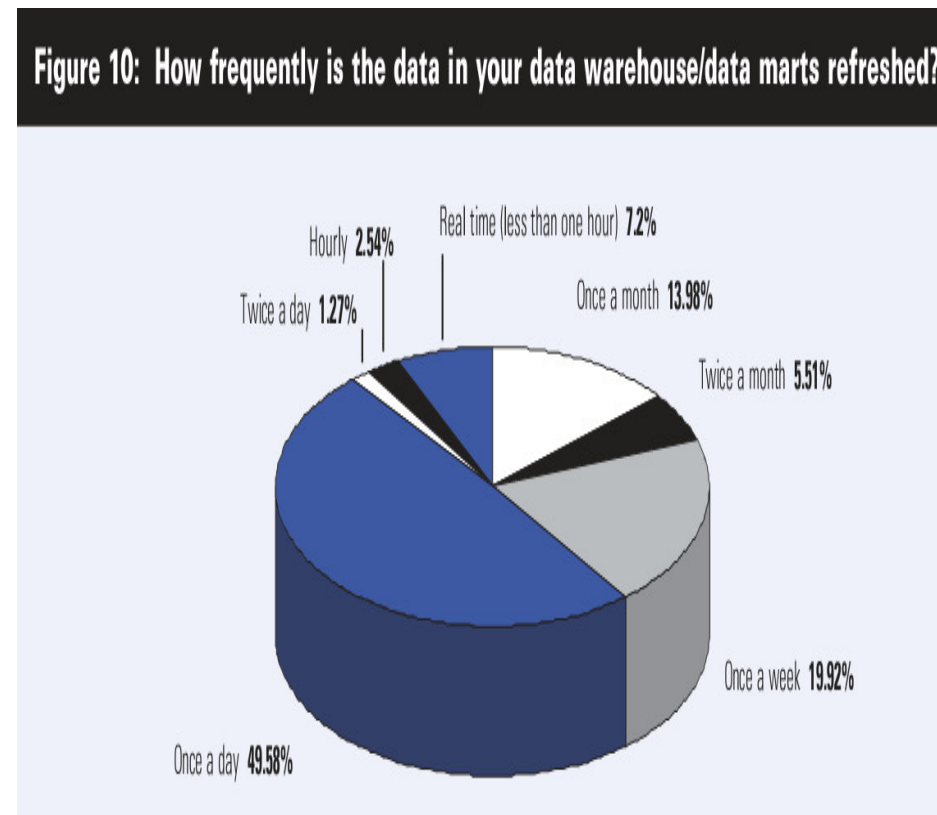


# Smart Analytics Optimizer Mart Update

- Typical DW update operations:
  - LOAD RESUME and REPLACE
  - ADD and ROTATE PARTITION
  - SQL INSERT, UPDATE, DELETE
  - Delete complete partition or table
  - TRUNCATE TABLE
  
- Smart Analytics Optimizer will over time phase-in support for all the typical operations in this order
  1. Full table reload
  2. Updated partition reload
  3. Individual row change
  
- The marts update is initiated and controlled through Smart Analytics Optimizer Studio
  
- Queries off-loaded to Smart Analytics Optimizer before the marts are refreshed can return different result set as compared to not being off-loaded
  - In case this is not acceptable use  
SET CURRENT REFRESH AGE = 0

IDUG study on DW update frequency

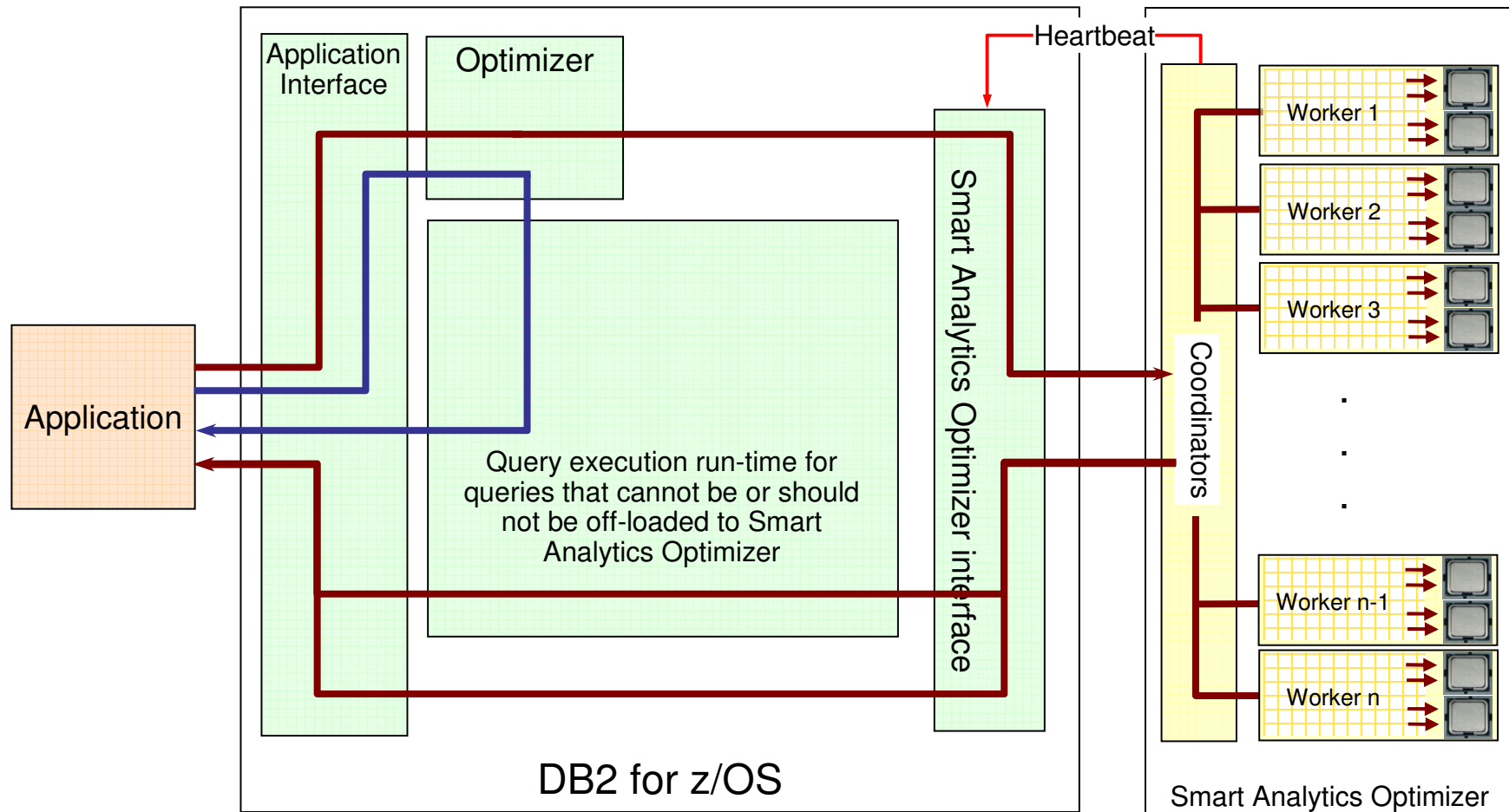
Figure 10: How frequently is the data in your data warehouse/data marts refreshed?



In 90% of cases DW is updated once a day or even less frequently



# Query Execution Process Flow



- Heartbeat (Smart Analytics Optimizer availability and performance indicators)
- Queries executed without Smart Analytics Optimizer
- Queries executed with Smart Analytics Optimizer

# Smart Analytics Optimizer Offering Options



z10 or  
zEnterprise 196

zBX: MTM 2458  
Model 001 or Model 002

DS5020

Offerings	XS 0.5TB	S 1TB	M 2TB	L 3TB	XL 4TB
<b>Enterprise Rack</b>	1	1	1	2	2
<b>BC-H Chassis, incl. redundant infrastructure (AMM, PDU, TOR 10Gb Ethernet switch for data, TOR 1Gb Ethernet for management, 4Gb F/C switches)</b>	1	1	2	3	4
<b>Blades</b>	7	14	28	42	56
<b>DS5020 with # of SATA disks (1 TB)</b>	16	16	16	32	32
<b>Storage Expansion Units</b>	0	0	0	1	1

- For z10, order Machine Type 2458 Model 001
- For zEnterprise 196, order Machine Type 2458 Model 002
- Model 001 is MES upgradable to Model 002, when z10 is upgraded to zEnterprise 196
- DSS5020 disk needs to be ordered separately from zBX, and housed in its own rack.

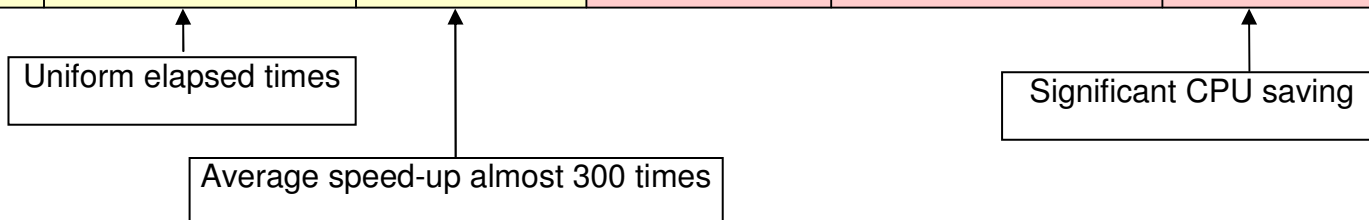
# Testing Results

- The problem queries provided by a customer
- Expert database tuning done on all the queries
  - Q1 – Q6 even after tuning run for too long and consume lots of resources
  - Q7 improved significantly – no Smart Analytics Optimizer offload is needed
- The table shows elapsed and CPU times measured in DB2 (without Smart Analytics Optimizer)

Query	Times measured in DB2 <b>without</b> Smart Analytics Optimizer			
	Total Elapsed	CP	zIIP	Total CPU Time
Q1	0:02:43	0:03:52	0:02:39	0:06:31
Q2	0:38:31	0:11:52	0:36:10	0:48:02
Q3	0:00:25	0:00:04	0:00:15	0:00:19
Q4	0:26:33	0:13:43	0:20:50	0:34:33
Q5	0:00:35	0:00:09	0:00:29	0:00:38
Q6	1:30:35	5:53:30	1:29:56	7:23:26
Q7	0:00:02	0:00:02	0:00:00	0:00:02

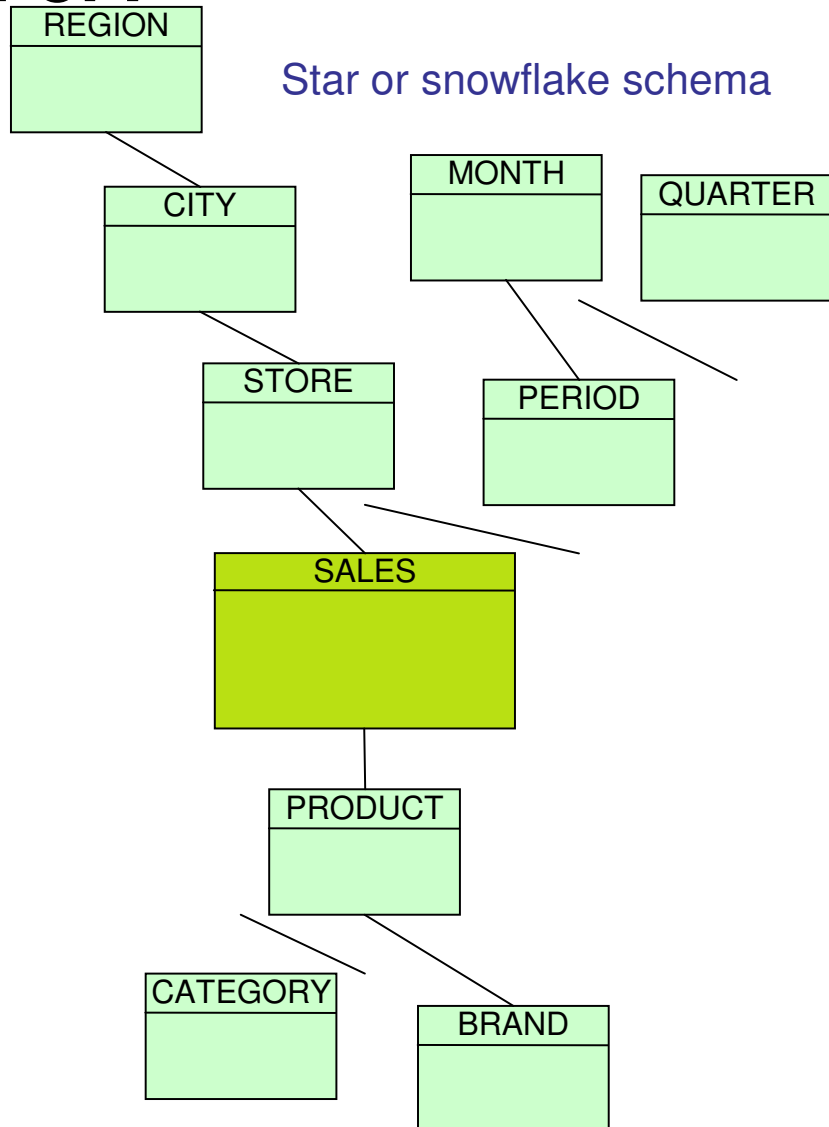
# Testing Results

Query	Query Elapsed Time			Query CPU Consumption on System z		
	DB2 only	DB2 with Smart Analytics Optimizer	Speed-up	DB2 only	DB2 with Smart Analytics Optimizer	Saving
Q1	0:02:43.0	0:00:03.4	48	0:06:31.0	0.004495	~100%
Q2	0:38:31.0	0:00:04.5	511	0:48:02.0	0.004713	~100%
Q3	0:00:25.0	0:00:02.2	12	0:00:19.0	0.099702	99.48%
Q4	0:26:33.0	0:00:07.8	206	0:34:33.0	0.005174	~100%
Q5	0:00:35.0	0:00:08.3	4	0:00:38.0	0.520915	98.63%
Q6	1:30:35.0	0:00:03.8	1424	7:23:26.0	0.003979	~100%
Q7	0:00:02.0	0:00:02.0	1	1.361983	1.361983	0.00%
<b>Total</b>	<b>2:39:24.0</b>	<b>0:00:32.0</b>	<b>298</b>	<b>8:53:31.0</b>	<b>2.000961</b>	<b>99.99%</b>



Note: Your mileage will vary. This particular query mix is suited exceptionally well for IBM Smart Analytics Optimizer

# What Is Smart Analytics Optimizer Ideally Suited For?



Complex, OLAP-style queries that typically:

- Need to scan large subset of data (unlike OLTP queries)
- Involve aggregation function such as COUNT, SUM, AVG.
- Look for trends, exceptions to assist in making actionable business decisions

```

SELECT PRODUCT_DEPARTMENT, REGION, SUM(REVENUE)
FROM FACT_SALES F
      INNER JOIN DIM_PRODUCT P ON F.FKP =
P.PK
      INNER JOIN DIM_REGION R ON F.FKR = R.PK
      LEFT OUTER JOIN DIM_TIME T ON F.FKT =
T.PK
WHERE T.YEAR = 2007
      AND P.TYPE = 'SOFTWARE'
      AND R.GEO = 'SOUTH'
GROUP BY PRODUCT_DEPARTMENT, REGION
  
```

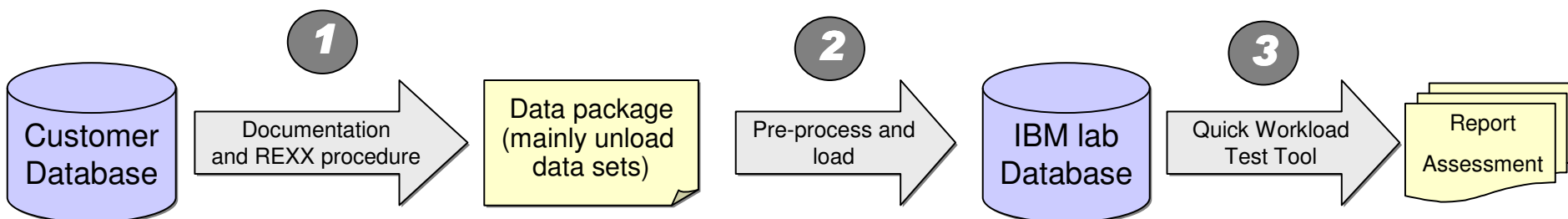
# Quick Workload Test

## Customer

- Collecting information from dynamic statement cache, supported by step-by-step instruction and REXX script (small effort for customer)
- Uploading compressed file (up to some MB) to IBM FTP server

## IBM / Center of Excellence

- Importing data into local database
- Quick analysis based on known Smart Analytics Optimizer capabilities



IBM Smart Analytics Optimizer -- V  
Center of Excellence, Datawarehouse on System z, IBM R

### Query Summary

	Total	With potential	Uncertain	W/o potential
Queries	23	11 (48%)	5 (22%)	7 (30%)
Query Blocks	23	11 (48%)	5 (22%)	7 (30%)
Elapsed Time	144801.47	106821.61 (74%)	8150.21 (6%)	29829.66 (21%)
CPU Time	21300.25	11420.12 (54%)	1453.14 (7%)	8426.98 (40%)

Queries	Count	Percentage
... no eligible blocks	7	30%
... with very large dim.	1	4%

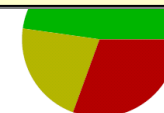
Leaf Query Blocks	Count	Percentage
... not read-only	0	0%
... with UDFs	0	0%
... with unsupported functions	6	26%
... with unsupported join types	0	0%
... with very large dimensions	1	4%
... with multiple refs to tables	0	0%
... with disjunction of joins	0	0%
... with unsupported subselects	0	0%
... with aggregations (info only)	23	100%

Start trace time: Apr 2, 2010 9:41 AM  
End explain time: Apr 2, 2010 4:37 PM  
Min stmt cached: Apr 2, 2010 9:44 AM  
Max stmt cached: Apr 2, 2010 4:37 PM

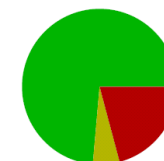
Disclaimer: Information provided in this document is for information purposes only and does not guarantee characteristics nor imply supported features of IBM products. V20100709

Report for a first assessment:

- Acceleration potential for
  - Queries
  - Estimated time
  - CP cost



Query blocks with acceleration potential  
Query blocks with uncertain potential  
Query blocks without acceleration potential



Elapsed time with acceleration potential  
Elapsed time with uncertain potential  
Elapsed time without acceleration potential

# Options for Workload Analysis

Stage	Purpose
Questionnaire	<ul style="list-style-type: none"> <li>▪ <b>Initial assessment based on size, query response time, update characteristics and customer pain points</b></li> </ul>
Quick Workload Test	<ul style="list-style-type: none"> <li>▪ <b>Assessment based on dynamic customer workload, runtime statistics, table sizes and SQL.</b></li> </ul>
Detailed Online Workload Analysis	<ul style="list-style-type: none"> <li>▪ <b>Assessment based on data mart definition for customer data model and offload capabilities in a real Smart Analytics Optimizer environment. Addresses all inhibitors for offload and data mart definition questions.</b></li> </ul>

## Resources:

- Instruction for the process at <https://w3.tap.ibm.com/w3ki08/display/isao/Process>
  - Includes capturing description and jobs/scripts to support the procedure
- Contact the Data Warehousing on System z Center of Excellence at [dwhz@de.ibm.com](mailto:dwhz@de.ibm.com)

# Sizing Process

1

'Rule of Thumb'  
based on offering

Smart Analytics Optimizer Offerings	XS .5TB	S 1TB	M 2TB	L 3TB	XL 4TB
Enterprise Rack	1	1	1	2	2
BC-H Chassis	1	1	2	3	4
Blades	7	14	28	42	56
DS5020	16	16	16	32	32
Storage Expansion Units	0	0	0	1	1

2

Information captured during  
workload assessment

## Sizing Information

Queries	Fact tables	Size fact tables (DASD)	Size fact tables (uncompressed)	Dim tables	Size dim tables (DASD)	Size dim tables (uncompressed)
Top 5	14	12.41 GB	42.92 GB	3	356.54 MB	540.14 MB
Top 10	18	50.05 GB	133.23 GB	4	356.62 MB	540.22 MB
Top 50	104	218.14 GB	573.11 GB	25	6.35 GB	15.05 GB
Elapsed time > 1 min	86	205.90 GB	540.44 GB	18	6.32 GB	14.91 GB
Elapsed time > 10 min	18	50.05 GB	133.23 GB	4	356.62 MB	540.22 MB

3

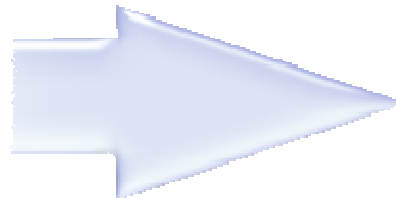
POC and/or other insights  
(e.g. data mart modelling with virtual accelerator)



# IBM Smart Analytics Optimizer

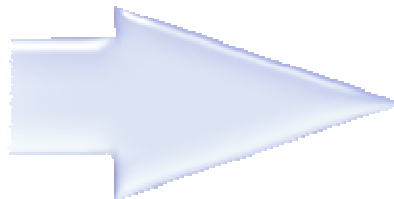
*Unlocking unprecedented value from enterprise data*

***Extreme Performance  
for Complex Business  
Analysis***



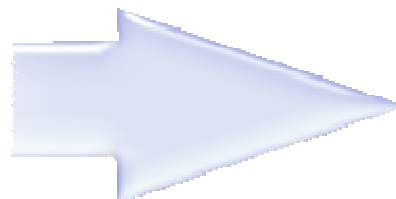
*Breakthrough technologies  
providing dramatic  
performance improvement*

***Database Performance  
Appliance***



*Quickly and simply deploy,  
hands-free operations, no  
query tuning, with application  
transparency*

***Proven operational  
characteristics***



*Inherits the availability,  
reliability, and security of  
System z*



Thank You