

# IBM System z Technology Summit



zEnterprise – The First System Of Systems

**Data Serving And Business Analytics On A  
Single Platform**

**Victor Leith**

27<sup>th</sup> Jan 2011



# Data Plays A Key Role In Smarter Planet Solutions

**1 Trillion** connected intelligent devices

**4 Billion** mobile phone subscribers worldwide

**2 Billion** people on the web



Smart retail



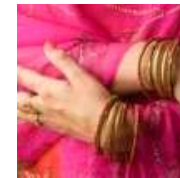
Smart supply chains



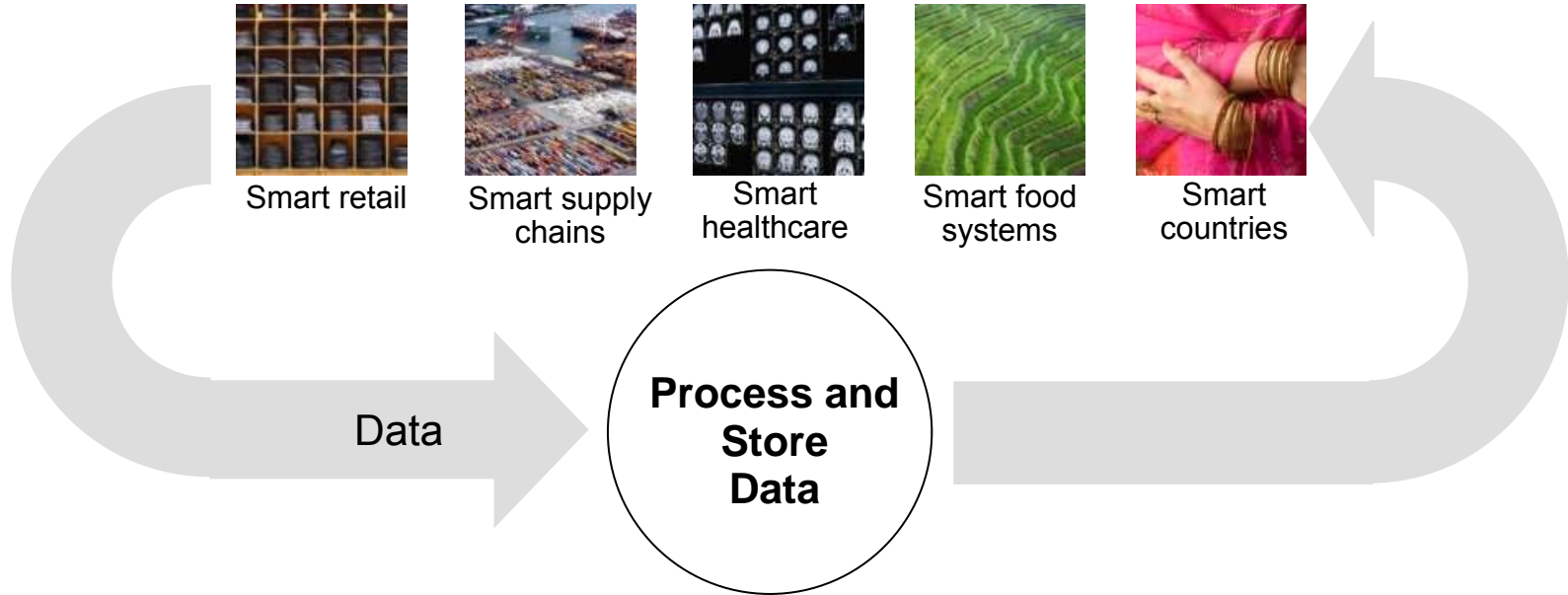
Smart healthcare



Smart food systems



Smart countries



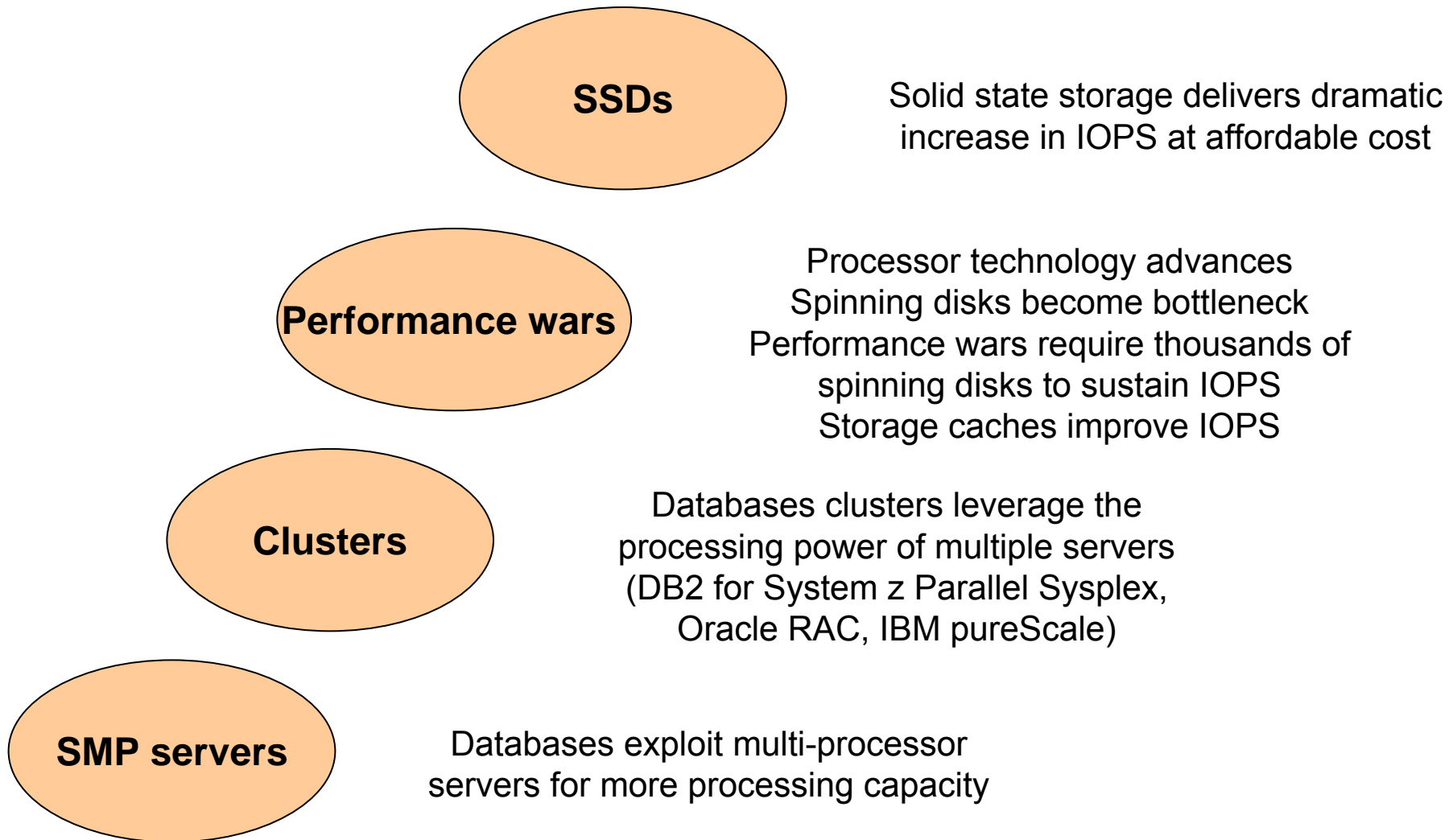
**15 petabytes** of new information generated daily  
**64 billion** credit card transactions per year (up 35% YTY)  
**30 billion** embedded RFID tags by 2010

# Having A Reliable And Scalable Database Is Vital

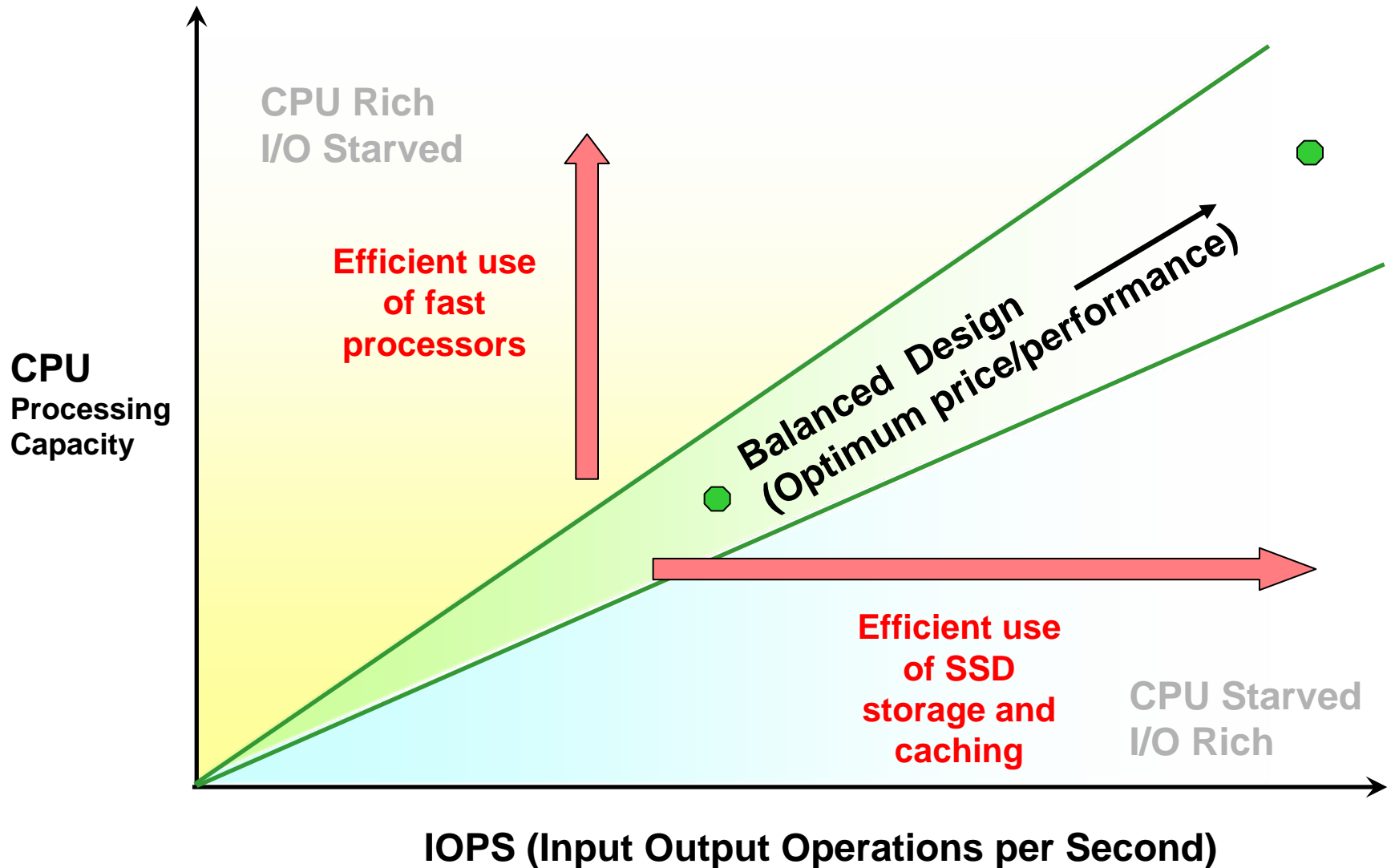
- Databases help you store, organize, and access information in an efficient manner
- A good database should:
  - ▶ Scale as your data processing needs grow
  - ▶ Be highly available to ensure access to information
  - ▶ Protect the integrity of the data stored
  - ▶ Support new data types and access methods
  - ▶ Protect the security of the data
  - ▶ Compression data to save disk space
  - ▶ Be cost competitive

**Let us explain why DB2 on z/OS is the best choice**

# A History Of Database Workload Optimization

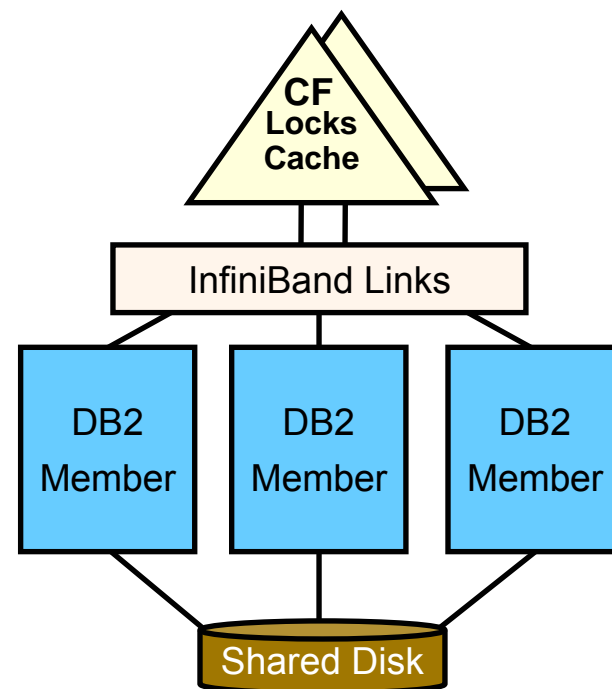


# Dramatic Increase In Database Performance Requires Two Tricks



# DB2 For z/OS Supports Parallel Sysplex Shared Data Clustering

- Shared data across nodes
- Hardware-based centralized lock and cache management
  - ▶ Provided by Coupling Facility
  - ▶ Supports near linear scalability
- Recovery from a node failure without a freeze
- Supports rolling upgrades with up to two different releases in a data sharing group



## Now Available DB2 10 for z/OS:

Up to 20% faster performance  
Hash access for faster OLTP  
Automatic snapshots of changing data  
Improvements in DB2 QMF and Tools suite

10x more concurrent users  
More online administration  
SQL and pureXML enhancements

# DB2 For z/OS Grows Database Capabilities Significantly Beyond SMP Solutions

- Lets you **Add Capacity** as you need it
  - ▶ Each z196 is capable of executing 50 billion instructions per second and you can cluster up to 32 z196 systems
  - ▶ New members automatically process new transactions
  - ▶ Load balancing is automatic
- Provides **Continuous Availability**
  - ▶ Available during unplanned outages
    - Redistribute workload to surviving members automatically
    - In-flight transactions automatically rolled back in as little as 15 seconds
  - ▶ Available during planned outages
    - Remaining members handle workload as a member goes offline
    - Cluster operates continuously during rolling software updates

# DB2 For z/OS Demonstrates Near Linear Scalability For Complex Banking Workload

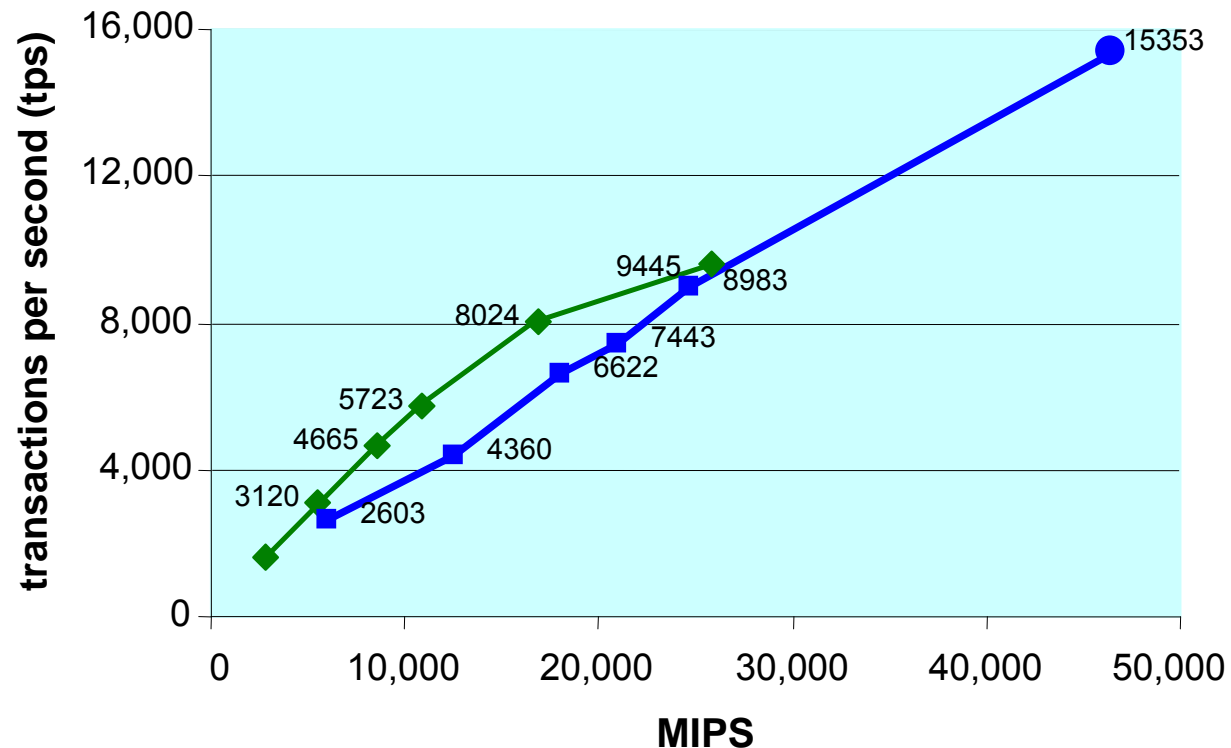
## System z and BaNCS Online Banking Benchmarks

### ■ Kookmin Bank

- ▶ IBM System z9 and DB2
- ▶ TCS BaNCS
- ▶ 15,353 Transactions/second
- ▶ 50 Million Accounts
- ▶ IBM benchmark for customer
- ▶ DB2 V9, CICS 3.1, z/OS V1.8

### ■ Bank of China<sup>1</sup>

- ▶ IBM System z9 and DB2
- ▶ TCS BaNCS
- ▶ 9,445<sup>2</sup> Transactions/second
- ▶ 380 Million Accounts
- ▶ IBM benchmark for customer



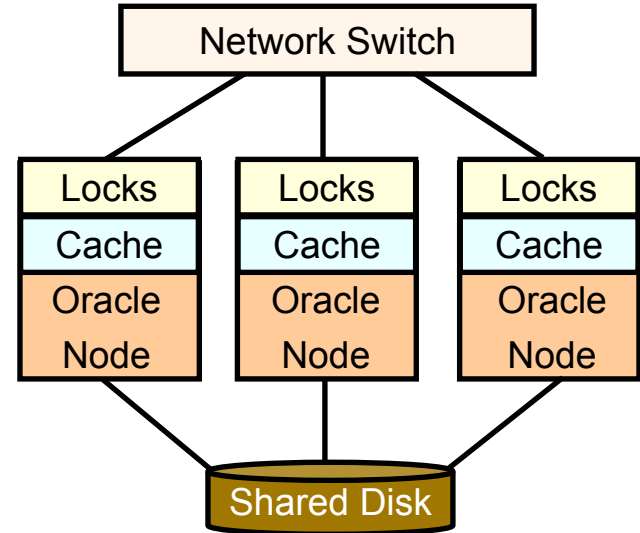
<sup>1</sup> Source: <http://www.enterprisenetworksandservers.com/monthly/art.php?2976> and *InfoSizing FNS BaNCS Scalability on IBM System z – Report Date: September 20, 2006*

<sup>2</sup> Standard benchmark configuration reached 8,024 tps, a modified prototype reached 9,445 tps



# Oracle Real Application Cluster (RAC) Supports A Distributed Network Clustering Architecture

- Shared disk across nodes
- Software-based distributed lock and cache management
- Poor scalability due to increasing interconnect traffic as cluster grows
- Upon node failure clusters can freeze during lock remastering process
- Database software has to be of same release on all nodes
- No mirroring



# There Are Expensive Consequences For Oracle RAC's Distributed Design

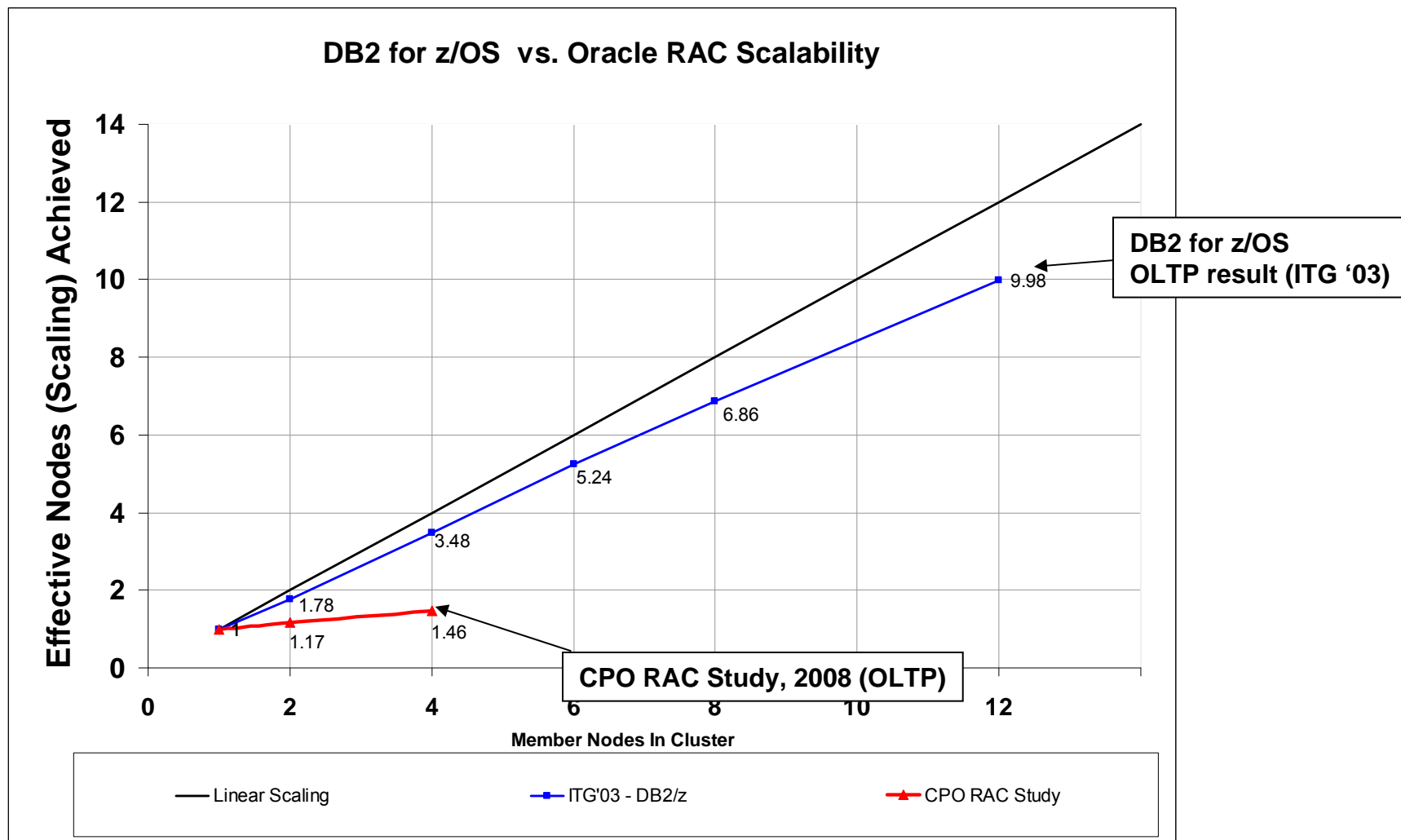
- Cluster members constantly share lock and cache data
  - ▶ In a 4-member cluster, obtaining a write lock could require 6 separate network communications
  - ▶ This is referred to as interconnect traffic
  - ▶ Interconnect traffic grows as the cluster grows and queries disperse across the cluster

***Per-member performance declines as the cluster grows.***

- During an unplanned outage Oracle RAC “freezes” I/O and lock requests
  - ▶ While re-mastering data blocks to surviving members
  - ▶ While locking pages that require recovery

***Oracle RAC does not offer continuous availability.***

# DB2 For z/OS Scales Near Linearly Whereas Oracle RAC Exhibits Poor Scalability

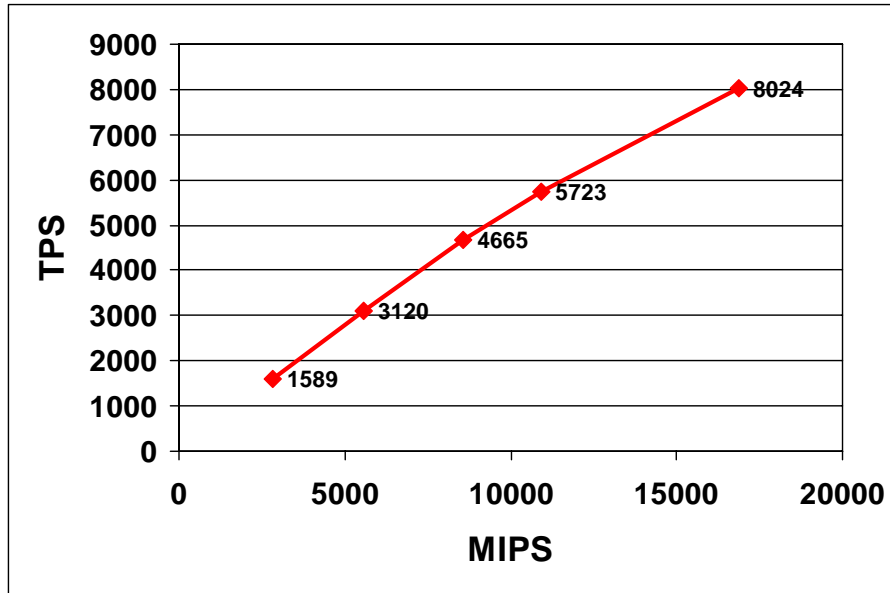


# Compare Techniques To Provide Sufficient I/O Bandwidth

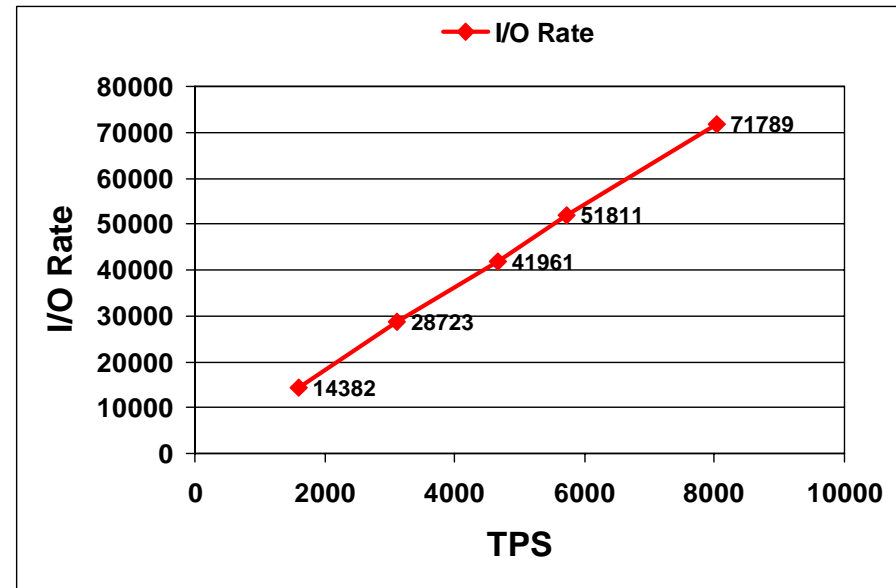
- DB2 on System z
  - ▶ I/O sub-system hardware is dedicated for I/O processing
    - Up to 14 SAP processors
    - Up to 336 channel processors
  - ▶ Typically attached to a DS8000 class storage sub-system with disk caches and large I/O bandwidth
- Oracle RAC on distributed server
  - ▶ No dedicated I/O sub-system
    - I/O operations executed by general purpose processors
  - ▶ Typically attached to mid tier storage

# Bank Of China Benchmark Demonstrates Linear Scalability With System z Across Both Resources

## Complex banking transactions (BaNCS – CICS/DB2 z/OS workload)



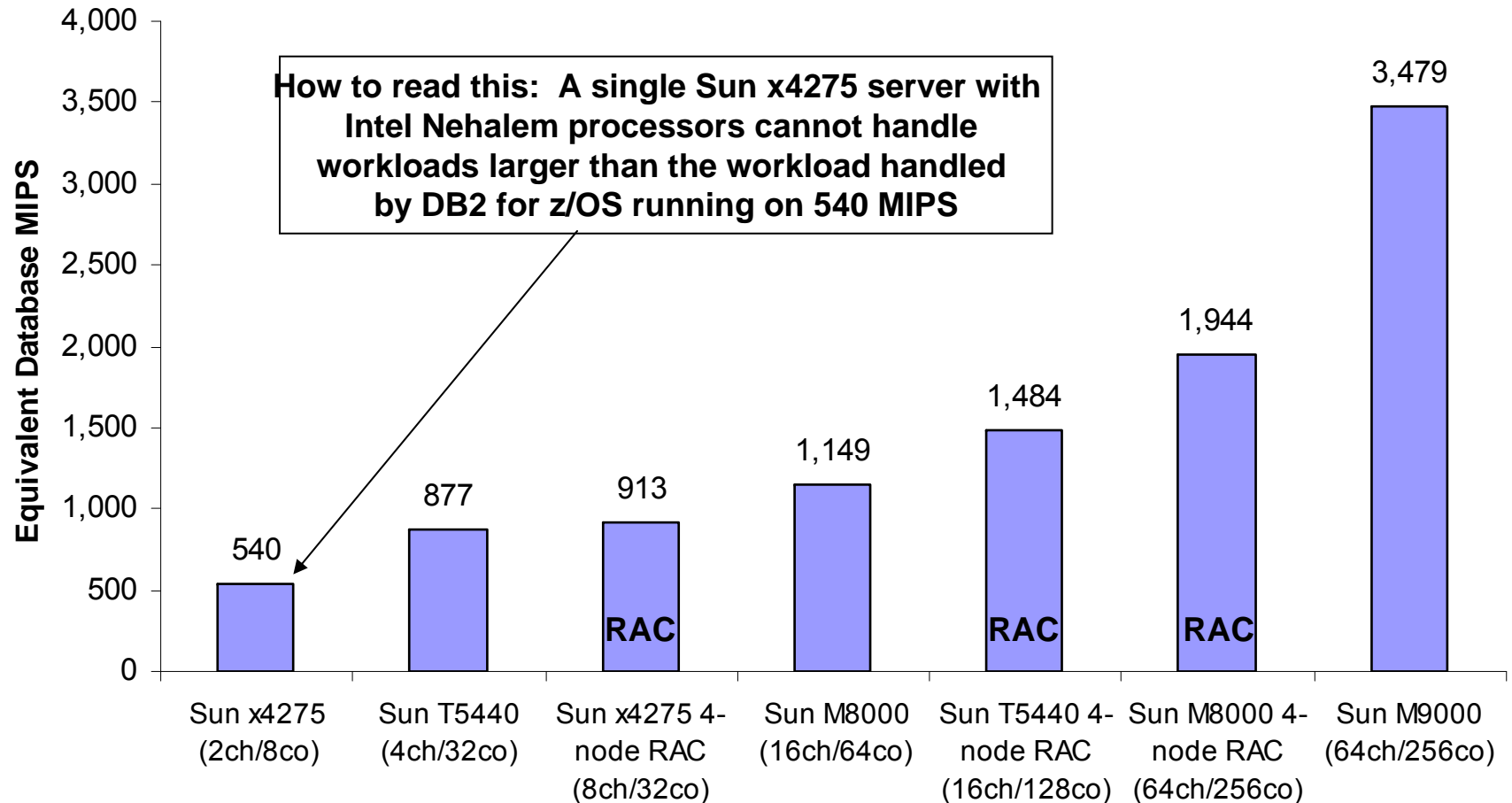
**Add MIPS as needed to increase transaction rate**



**I/O rate scales to sustain transaction rate**

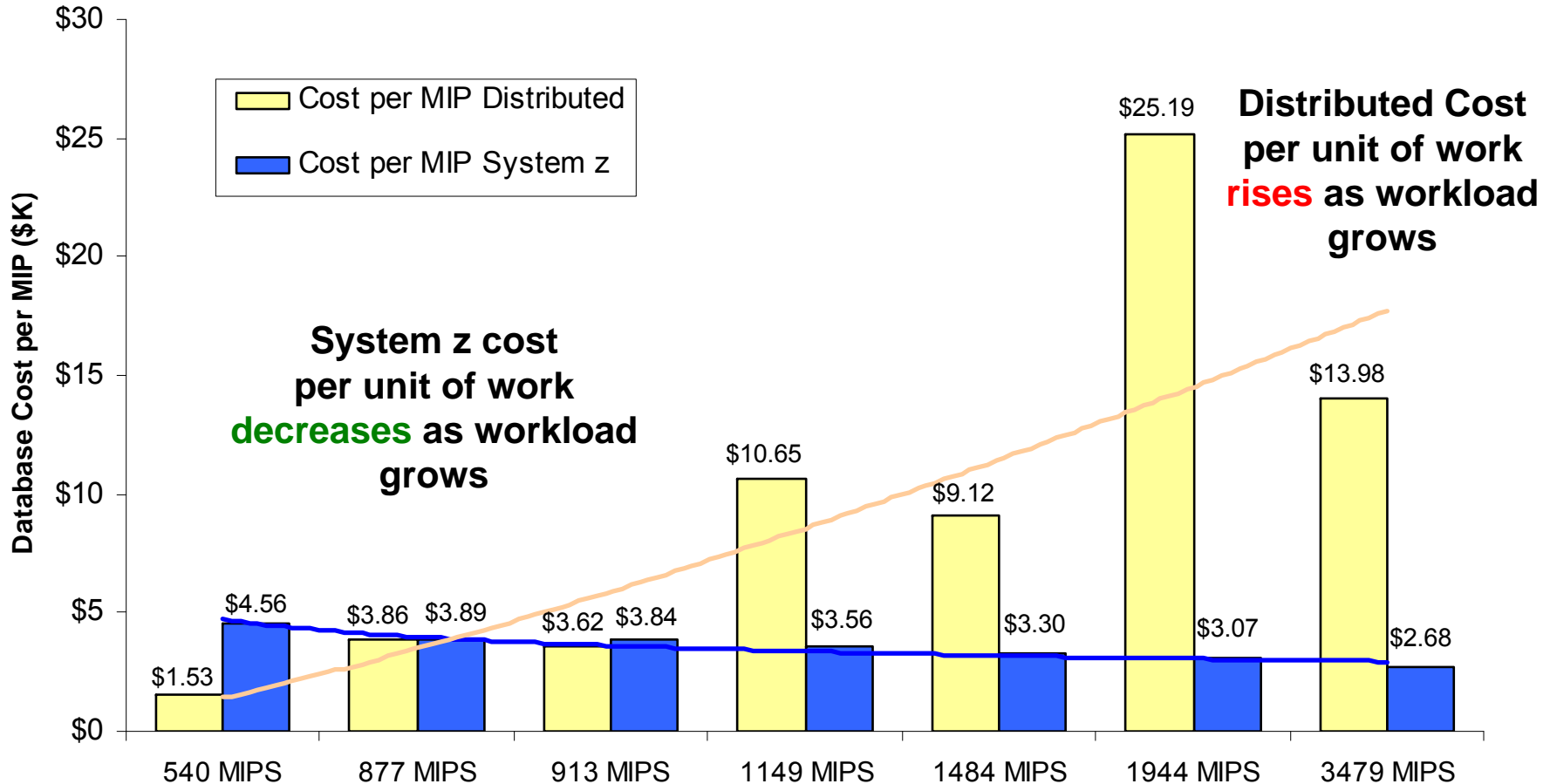
# What Equivalent DB2 For z/OS Performance Can You Get Out Of Oracle Running On Sun Servers?

Maximum Database Workload (in MIPS)



# Large Databases Are Best Fit On DB2 For z/OS

## Cost per Database MIP



# Result: Deploy SAP Database On DB2 For z/OS At 25% Of The Distributed Cost

*Add Production zEnterprise and Create 2 LPARs for SAP Database (DB2) and SAP Application (z/Linux)*



13,830 MIPS  
DB + App  
workload

*Add DR zEnterprise*



3 year  
TCA  
\$1.2M

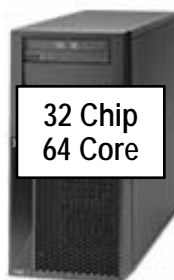
Production zEnterprise:  
1 zIIP 890 MIPS DB2  
2 GP 2,260 MIPS DB2 & Tools  
12 IFL 10,680 MIPS SAP App  
131 GB memory, 3.4TB DASD

Backup zEnterprise:  
1 zIIP, 2 GP, 12 IFL (CBU)  
131 GB memory, 3.4TB DASD

“Medium” Solution Edition for SAP promotional pricing

*Or add HP Integrity Superdome s2k 9150 N Server for Production*

Active RAC node



32 Chip  
64 Core

211,761  
Performance Units

*And add 1 server for Disaster Recovery, Dev & QA*

Failover RAC node



32 Chip  
64 Core

211,761  
Performance Units

3 year  
TCA  
\$4.7M



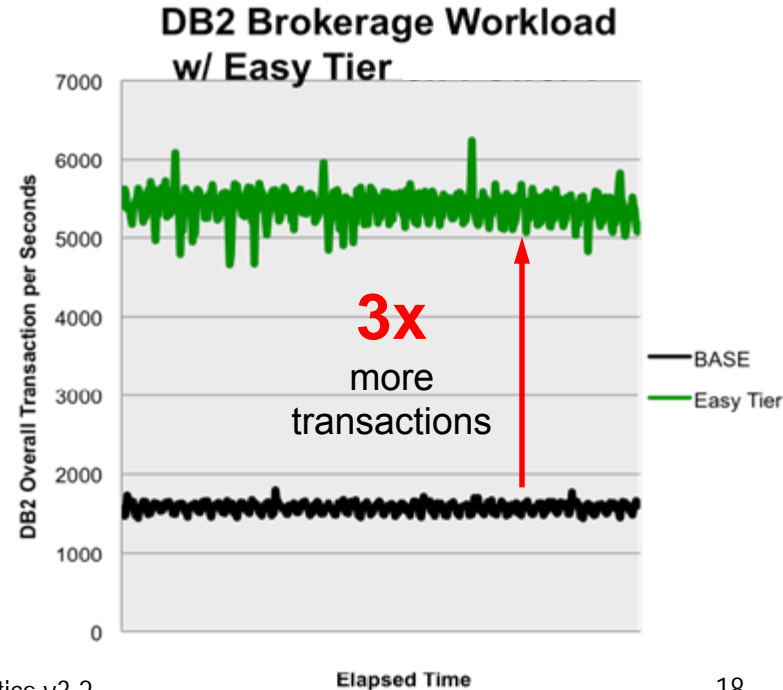
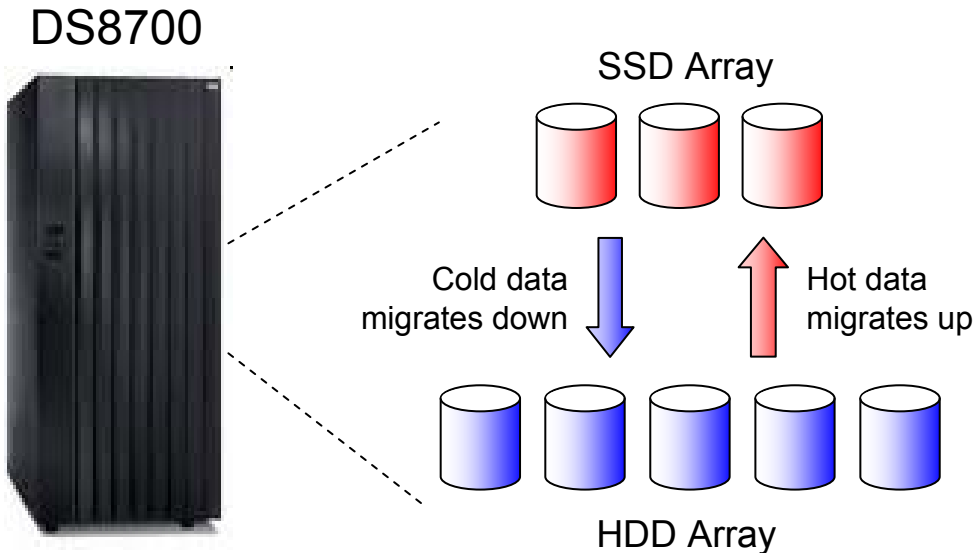
# Solid State Disk Drives Are Here To Revolutionize Storage

- Response times is around 0.8 milliseconds in contrast to 6 milliseconds for a typical hard disk drive
  - ▶ 5-10x improvement in throughput and queries
  - ▶ SSD drives can sustain I/O rates from 6,000 to 20,000 compared to spinning disk rates of 250 - 300
  - ▶ Cost per TB is 10x cost of spinning disks
  - ▶ Reduce the “batch window”
- Semiconductor storage available in DS8700 storage sub-systems
  - ▶ Random access solid state storage – no moving parts
- Benefits
  - ▶ 75% reduction in space
  - ▶ 80+% reduction in power and cooling
  - ▶ Reduce RAM requirements



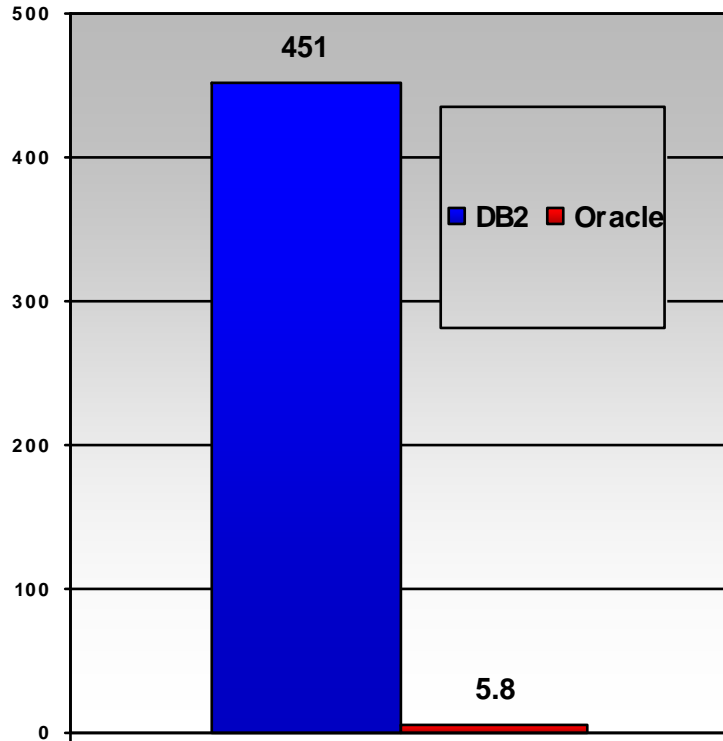
# DS8700 Easy Tier Capability Automatically Migrates Frequently Accessed Data To SSD

- **Automated** hotspot detection and migration of data between SSD and HDD
  - ▶ Transparent to applications, **no code changes** required
- Easy Tier maximizes SSD performance gains while minimizing costs
  - ▶ Increase performance by up to **300%**
  - ▶ Relocating just 5% of data to SSDs can reduce response time by 78%
- **No charge feature** – Microcode update to DS8700



# DB2 pureXML Stores And Retrieves XML Data Faster Than The Competition

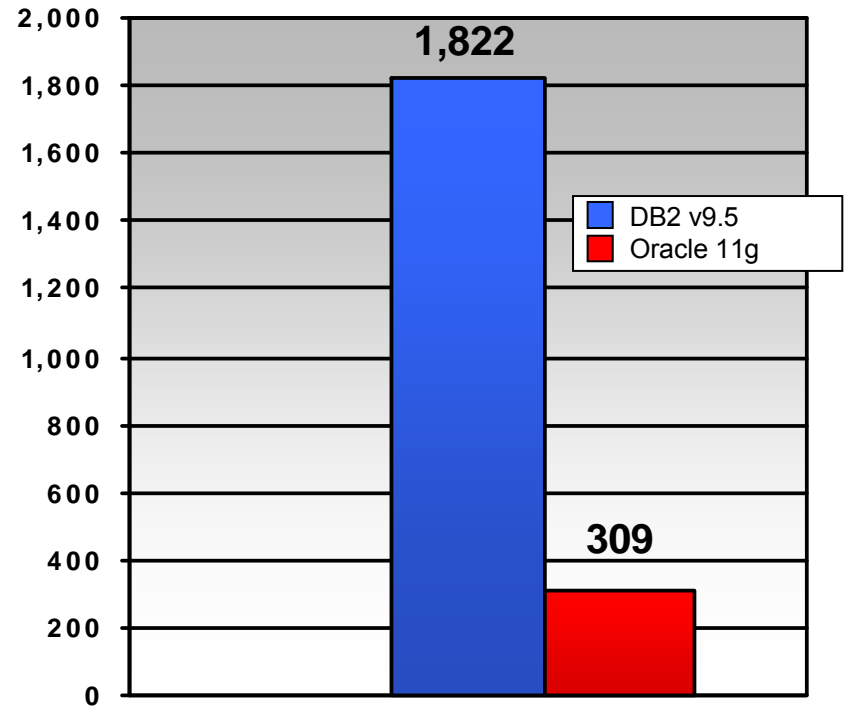
## XML Query Performance With Partial Index



Higher is Better

78x faster than Oracle 11g

## Insert XML Documents/Sec with Partial Index On 12 Elements



Higher is Better

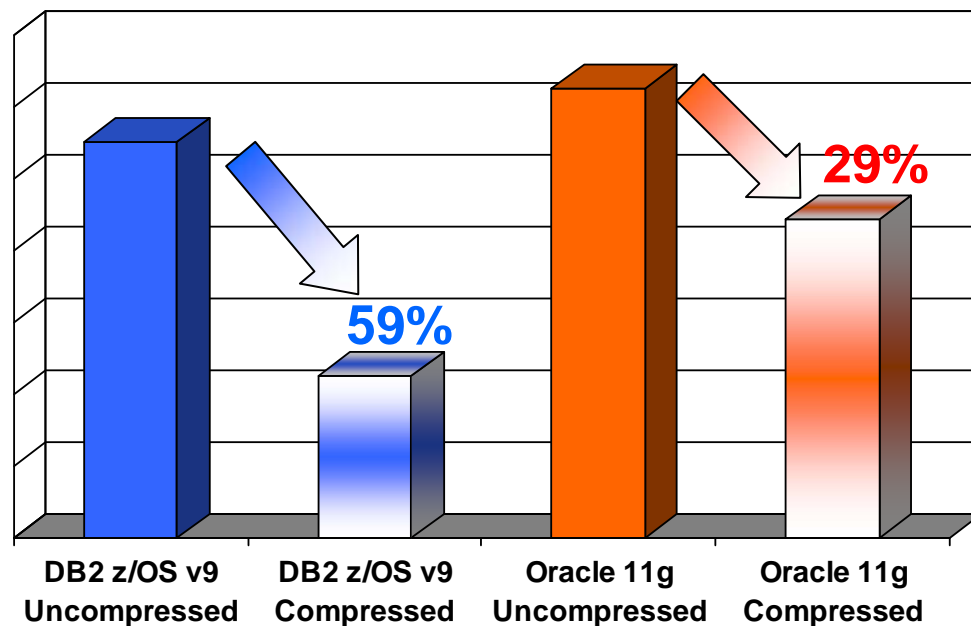
6x faster than Oracle 11g

SWG CPO study based on inserting TPOX benchmark data into DB2 v9.5 and Oracle 11g

# DB2 Compression Reduces Storage Requirements (And I/O Bandwidth Requirements)

- DB2 z/OS uses half the storage of Oracle
- Compression is included with DB2 z/OS
  - ▶ Compression is an extra charge feature with Oracle
- DB2 z/OS uses built-in hardware compression for data
  - ▶ Oracle compression is implemented in software adding to database processing overhead

1.2 GB Data Warehouse Compression Test



# Oracle Database Downtime Can Be Significant With 36 Security Patches Issued Last Year

## Oracle's Security Exposures

- Oracle.com – October 2010  
90 security patches, including **7** for the database
- Oracle.com – July 2010  
59 security patches, including **13** for the database
- Oracle.com – April 2010  
47 security patches, including **7** for the database
- Oracle.com – January 2010  
24 security patches, including **9** for the database
- In the last year, Oracle has issued 220 security patches, **36** for the database
  - ▶ 36 patches x 4 nodes = 144 possible System Freezes for 4 - node Oracle RAC

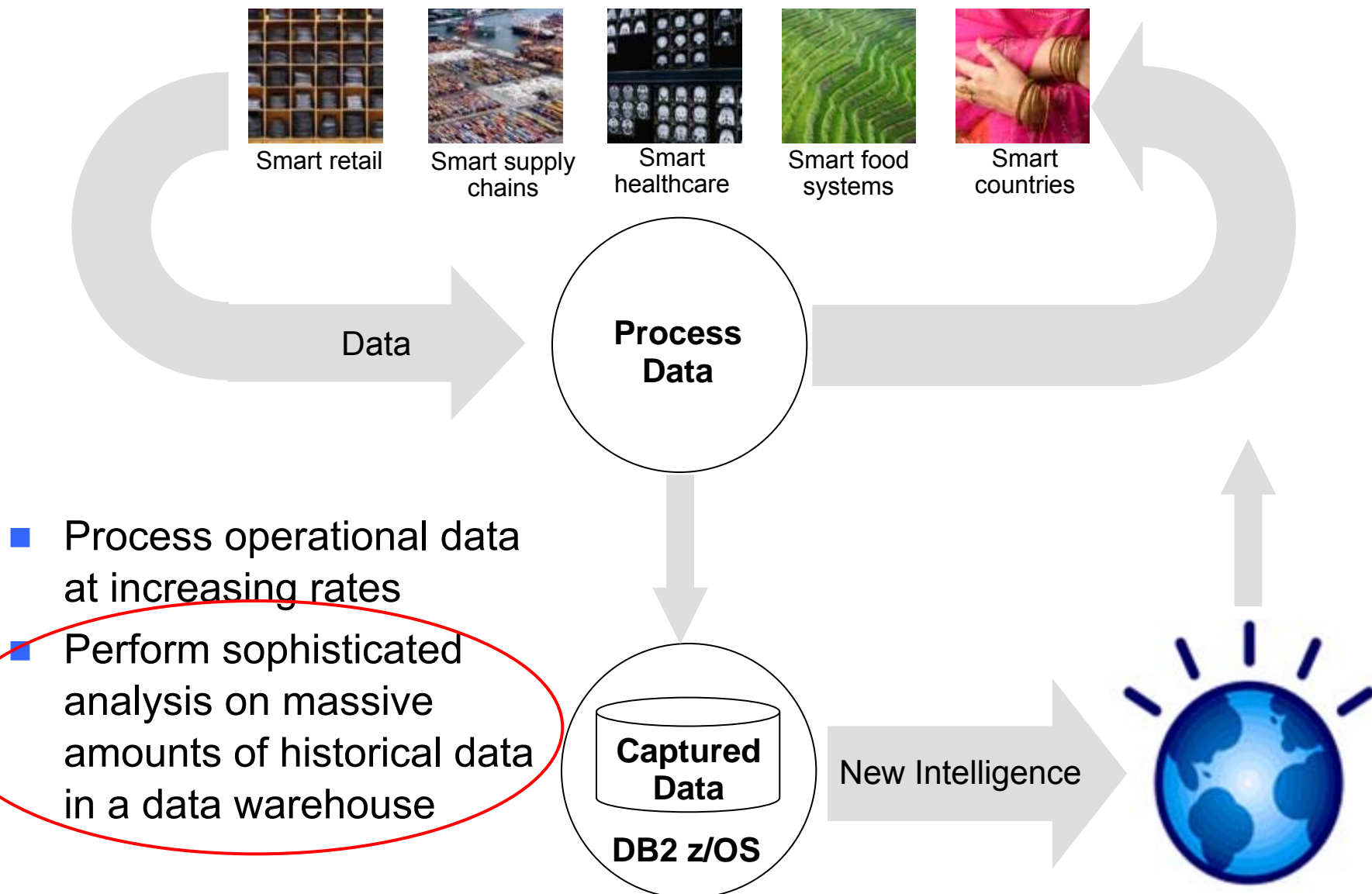
# DB2 For z/OS Provides Rock Solid Security

- Proven granular Multi Level Security leveraging RACF
  - ▶ Secures access of tables, views, rows, columns
- End-to-end encryption via hardware assist
  - ▶ CP Assist for Cryptographic Function (CPACF) and Crypto Express3
- Optim Test Data Management
  - ▶ Ensures anonymous access to data necessary for testing
- DB2 Data Archive Expert
  - ▶ Allows customers to easily archive and access data
- DB2 Audit Management Expert
  - ▶ Supports compliance requirements
  - ▶ Tivoli zSecure Audit for enterprise wide audit

# DB2 For System z Provides A Balanced Database For All Smarter Planet Solutions

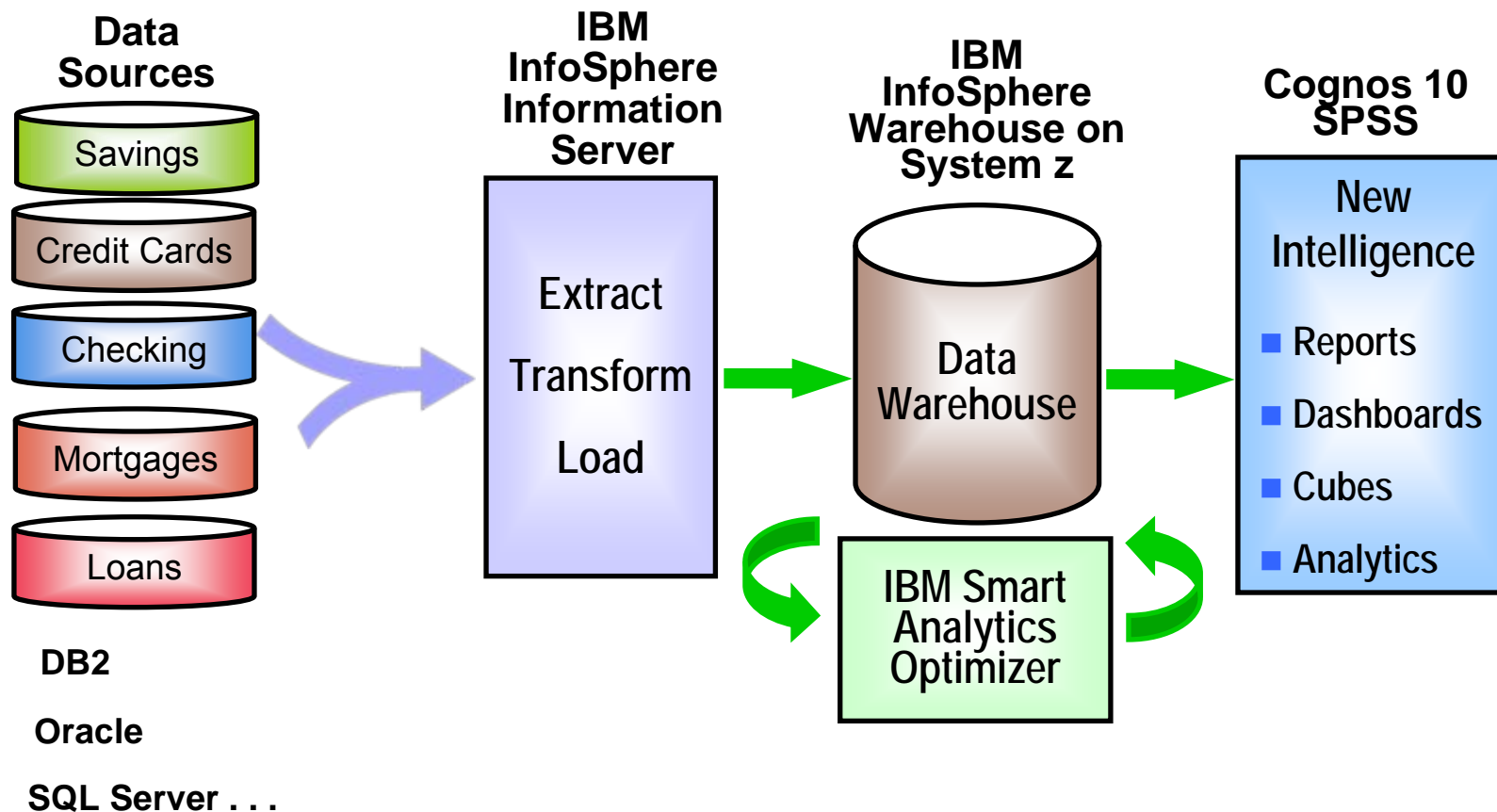
- Exploits System z Parallel Sysplex for availability and scale
  - ▶ RAC's distributed design difficult to scale, freezes during outages
- Dedicated I/O Hardware and Solid State Storage drives I/O throughput
  - ▶ Oracle running on distributed servers use precious general purpose processor cycles for I/O
- Superior Qualities of Service
- Competitive Price

# Data Plays A Key Role In Smarter Planet Solutions





# Load Operational Data From Silos Into A Unified Data Warehouse For New Intelligence



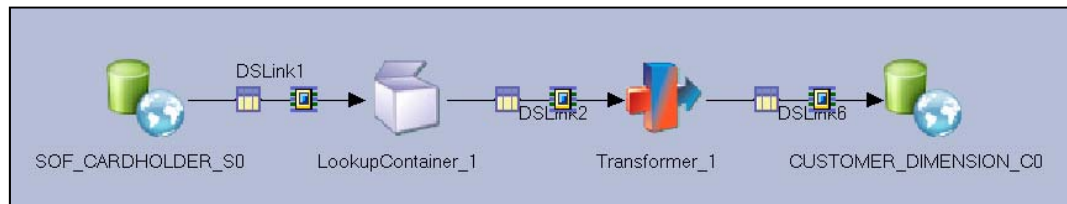
# InfoSphere FastTrack Creates Data Maps And Specifications For Your ETL Jobs

- Create simplified data maps and transformations using drag and drop
  - ▶ Automatically discover source and target columns
    - Uses database introspection and Web 2.0-style tagging
    - Use business terms to accurately match source to target
- Data analysts and developers share project specifications
  - ▶ Collaboration and reuse improve productivity
  - ▶ Use metadata common to all Information Server tools
  - ▶ Standard formats and centralized management for governance
    - Synchronize work across global teams
- Generate ETL code directly from job specifications
  - ▶ Reduces costs and errors in ETL job development

**Oracle doesn't offer any of these capabilities**

# InfoSphere DataStage Implements The ETL Jobs

- Graphical description of ETL jobs using hundreds of pre-built transformation and data quality functions
  - ▶ Allows easy reuse of integration work between projects
- Stores and retrieves metadata from Information Server
  - ▶ Dynamic partitioning and pipelining
  - ▶ Scale jobs across additional hardware without modification



Database Sources				
Classic Federation	DB2 UDB API	DB2/UDB Enterprise	DB2Z	Dynamic RDBMS
Informix CLI	iWay Enterprise	ODBC	Oracle Enterprise	Stored Procedure

Prebuilt Processing Functions			
Aggregator	Change Apply	Change Capture	Compare
Compress	Copy	Decode	Difference
Encode	Expand	External Filter	Filter
FTP Enterprise	Funnel	Generic	Join
Lookup	Merge	Modify	Pivot
Remove Duplicates	Slowly Changing Dimension	Sort	Surrogate Key Generator
Switch	Transformer		

Real Time Connectors			
Java Client	Java Transformer	Web Services Client	Web Services Transformer
WebSphere MQ Connector	WISD Input	WISD Output	XML Input
XML Output	XML Transformer		

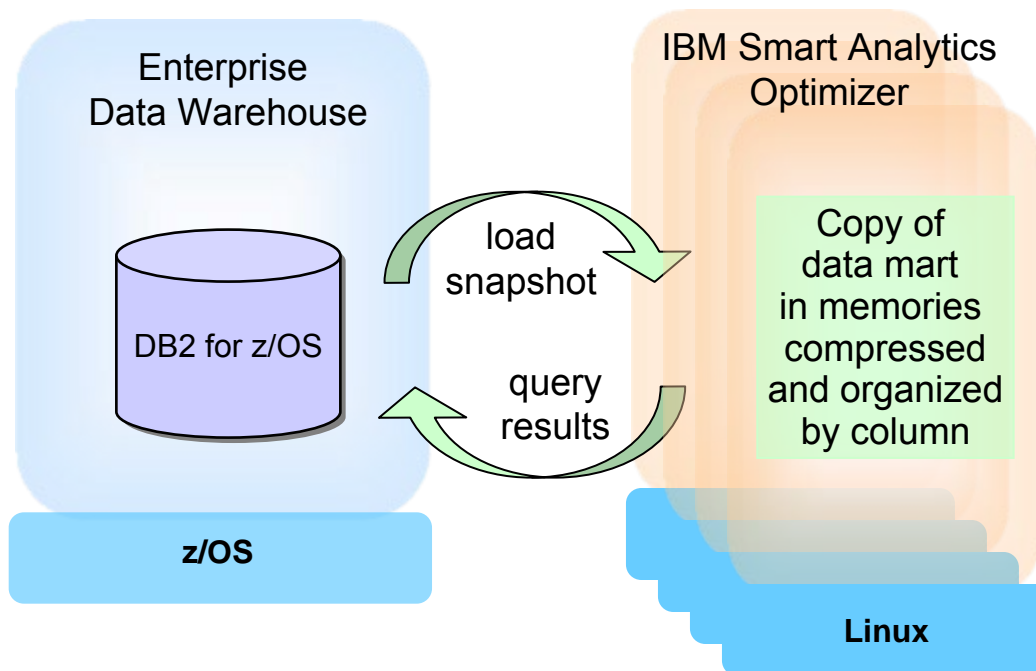
# IBM Leads In Data Integration

- Only InfoSphere Information Server delivers unified metadata across all tools for collaboration and reuse
  - ▶ Oracle Warehouse Builder and Oracle Data Integrator are two separate products that are not yet unified
- Model-driven design with FastTrack and DataStage speeds development
  - ▶ Oracle Warehouse Builder has no tools to help manage source to target mappings
- InfoSphere Information Server works in heterogeneous environments
  - ▶ InfoSphere gathers, processes, and cleanses more data from more sources than Oracle

"FastTrack enables our analysts to **capture more complete business requirements**. The ability to translate this information directly into DataStage jobs with up to 70 percent of the code completed will **significantly shorten our development lifecycle**."

*- Roderich Hofmann, project manager, WAVE, IT-Solutions provider of Bank Austria and member of UniCredit Group*

# IBM Smart Analytics Optimizer Enables Near Real Time Analytics On zEnterprise

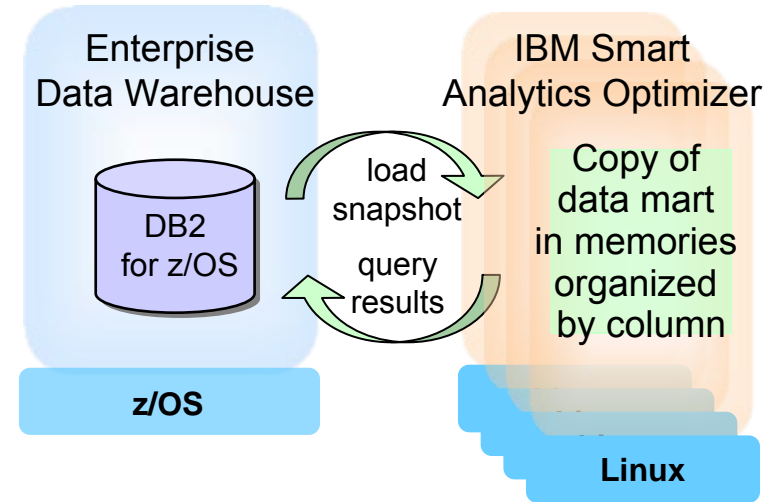


- Leverages blade memory and processors for OLAP-style queries
- Load snapshot then execute queries transparent to applications
- Queries automatically offloaded by DB2 for z/OS Optimizer

# Two Usage Scenarios For IBM Smart Analytics Optimizer

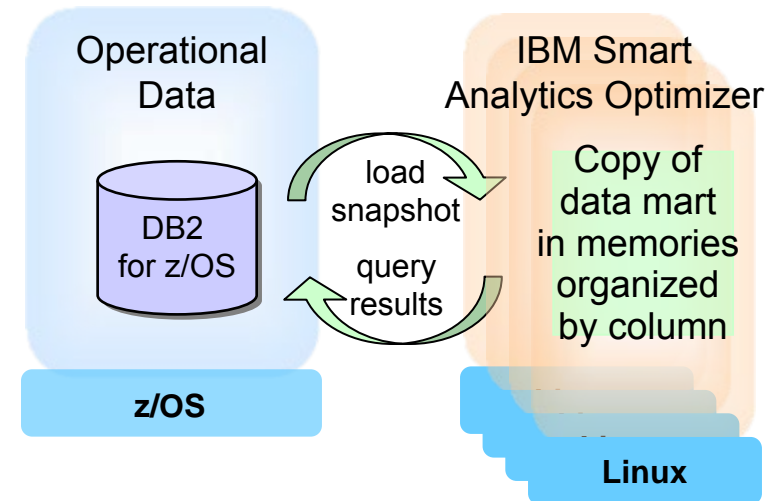
## ■ “Continuous use” scenario

- ▶ Load data mart once
- ▶ Run continuous jobs until load again
- ▶ Ad hoc queries, variations on reports, different drill downs
- ▶ Metrics: speed of each job, cost per job



## ■ “Real time analytics” scenario

- ▶ Load data mart, run one analytical job
- ▶ Repeat load as often as possible (i.e. sample rate)
- ▶ Continuous monitoring of some business process
- ▶ Metrics: Maximum possible sample rate



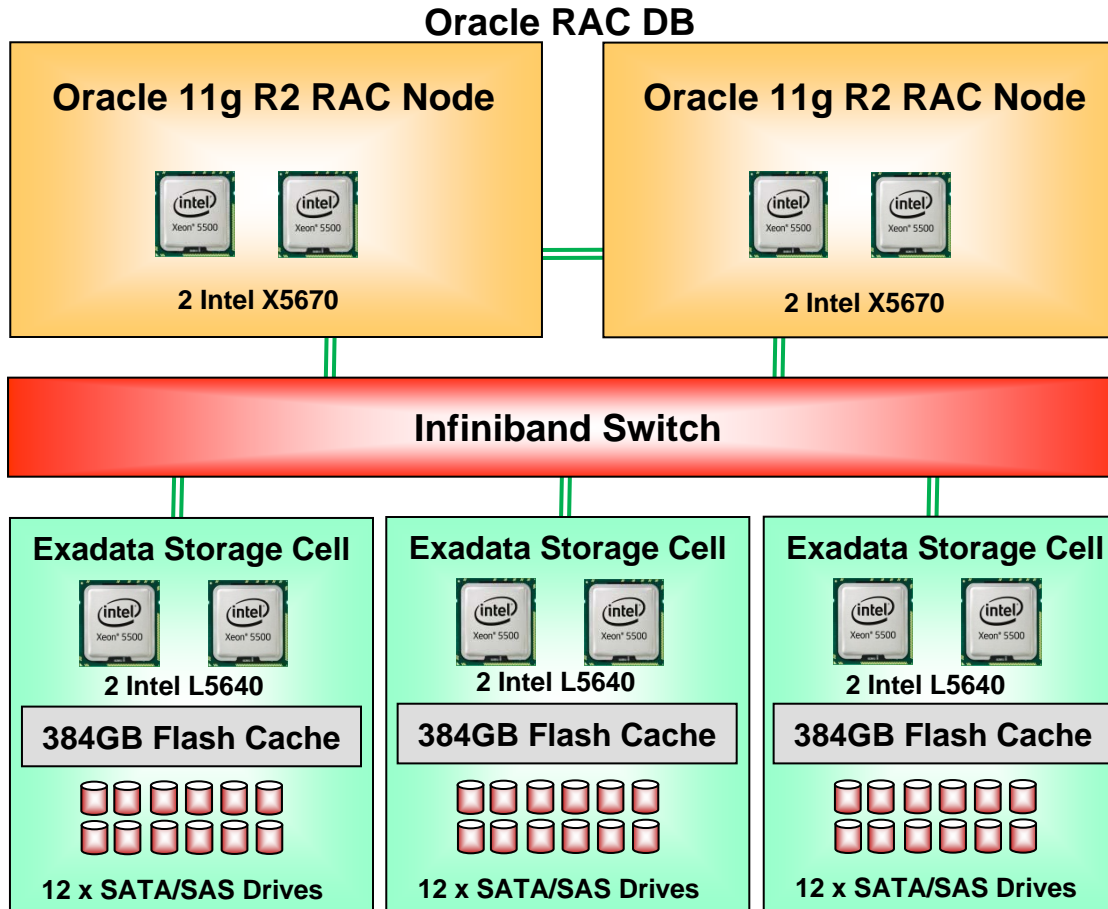
# Execute Query Jobs Faster With IBM Smart Analytics Optimizer

Query Job	Clock Time System z Only (sec)	Clock time with IBM Smart Analytics Optimizer (sec)	Job Execution Speedup Factor	IBM Smart Analytics Optimizer Load Time (sec)
Customer 1 (68 queries)	36,000	300	<b>120x</b>	360
Customer 2 (16 queries)	1,200	7	<b>171x</b>	216
Customer 3	15,854	5,246	<b>3x</b>	7,200

- Cost per job reduced by **1.4 – 1,838** times in continuous use
- Real time analytic sample rates increased to **7 – 387** samples per day

Note: Based on early customer results

# Oracle Exadata X2-2 System



## Depiction of smallest Oracle Database Machine Configuration (1/4 Rack)

2 DB Servers, 3 Exadata Storage Servers

- Very little processing is actually pushed to the Exadata Storage Cell
- RAC database nodes still do the majority of the “heavy lifting” processing
- RAC database nodes waste resources because of high RAC overhead
- Exadata Cells waste resources with typical low distributed CPU utilization



# Building A Data Warehouse On zEnterprise Costs Less Than Exadata

New Data Warehouse with 6 TB Storage + 1TB IBM Smart Analytics Optimizer

**Solution Edition  
Best Price applied  
to data warehouse  
on z/OS**



**Total 3 year  
cost of  
acquisition  
\$2.26M**

**MIPS total 2092  
zIIP 1  
Memory 32 GB  
Storage DS8000 6 TB**

**IBM Smart Analytics Optimizer "Small"  
1TB acceleration capacity  
14 HS22 blades (112 cores)**

Or add Oracle Database Machine (Exadata X2-2)  
1/4 rack with 6 TB SAS, and 1.1 TB SSD storage

**Oracle Database Machine 1/4 Rack  
6 TB SAS (36 HDD drives)  
1.1 TB SSD  
Two 12-core database servers  
Three 12-core storage servers**



**3 year  
cost of  
acquisition  
\$3.16M**

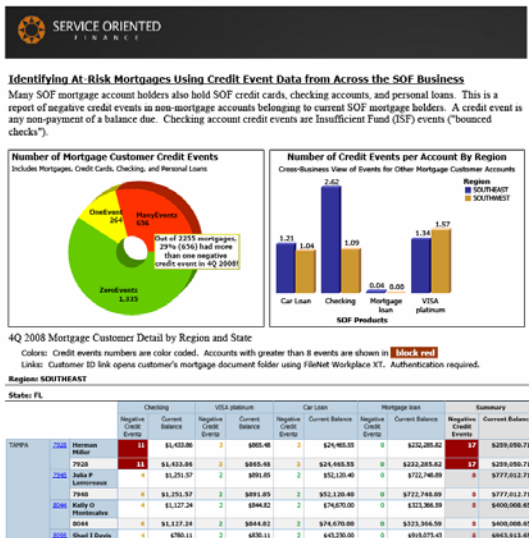
List prices for Oracle Database Machine and IBM Smart Analytics Optimizer. Solution Edition prices for System z and DB2 for z/OS

# Data Warehouse On System z Trumps Exadata Powered By Inefficient Oracle RAC

- z196 optimizes mixed query performance across OLTP and Warehouse data
  - ▶ Exadata storage nodes are designed for Warehousing and wastes CPU resources for OLTP
  - ▶ Works well for singleton queries but parallel query performance sags due to resource contention
- New analytics query accelerator makes complex queries much faster
- Application and database co-located on z196 sharing resources
  - ▶ Exadata over-provisions for peak, and supports growth through limited choice
- Exploits System z Parallel Sysplex for availability and scale
  - ▶ RAC's distributed design difficult to scale, freezes during outages
- System assist processors and I/O sub-system improve scale
  - ▶ Exadata I/O handled by x86 processors and architecture
  - ▶ Exadata performance falls dramatically as data exceeds smart cache size
- Hardware-based compression lowers costs of MIPs and storage
  - ▶ Oracle's advanced database compression software half as effective, hybrid columnar compression similar to what IBM Smart Analytics Optimizer provides
- System z central to disaster recovery and backup strategies
  - ▶ Locked-down Exadata does not fit easily into datacenter operations
- **System z Warehouse costs less** than Exadata

# DEMO: Use Cognos To Identify New Business Insights From The Data Warehouse

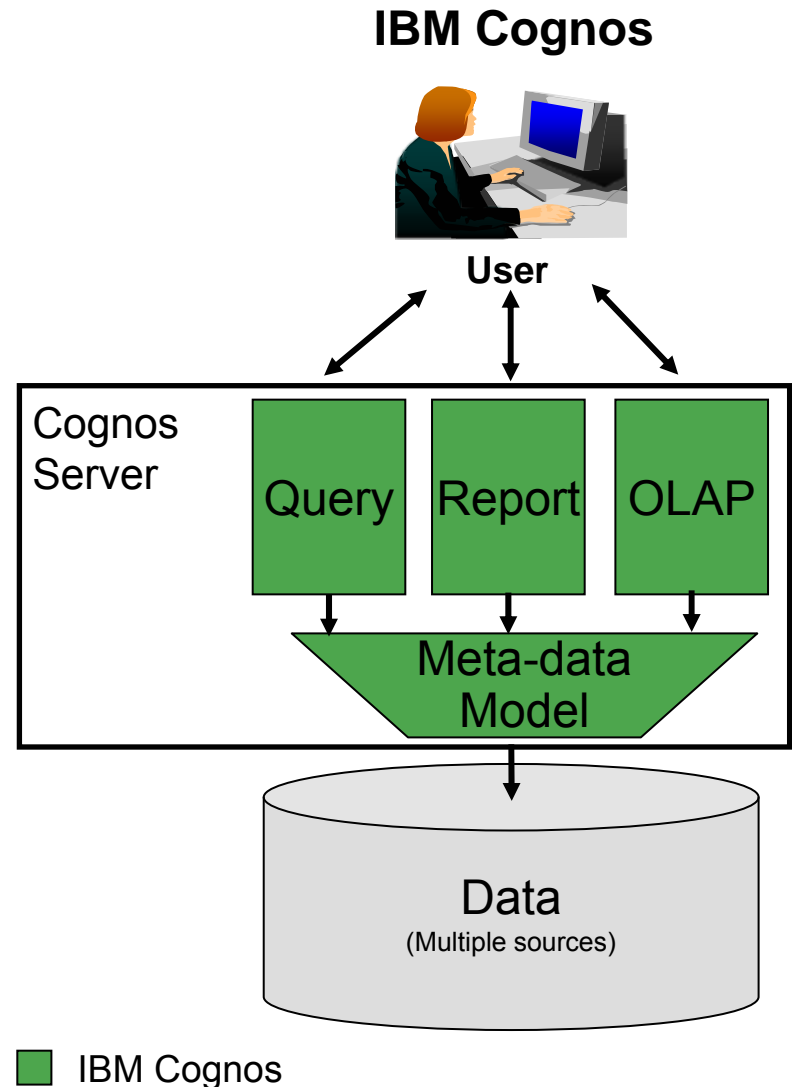
1. Show report generated in Cognos Report Studio in PDF format
2. Report identifies high-risk mortgages by looking at negative credit events in other customer accounts (CC, Checking, etc...)
3. Report uses both structured and unstructured data (link to mortgage data stored in FileNet)
4. Use Go! Dashboard to monitor the business operations



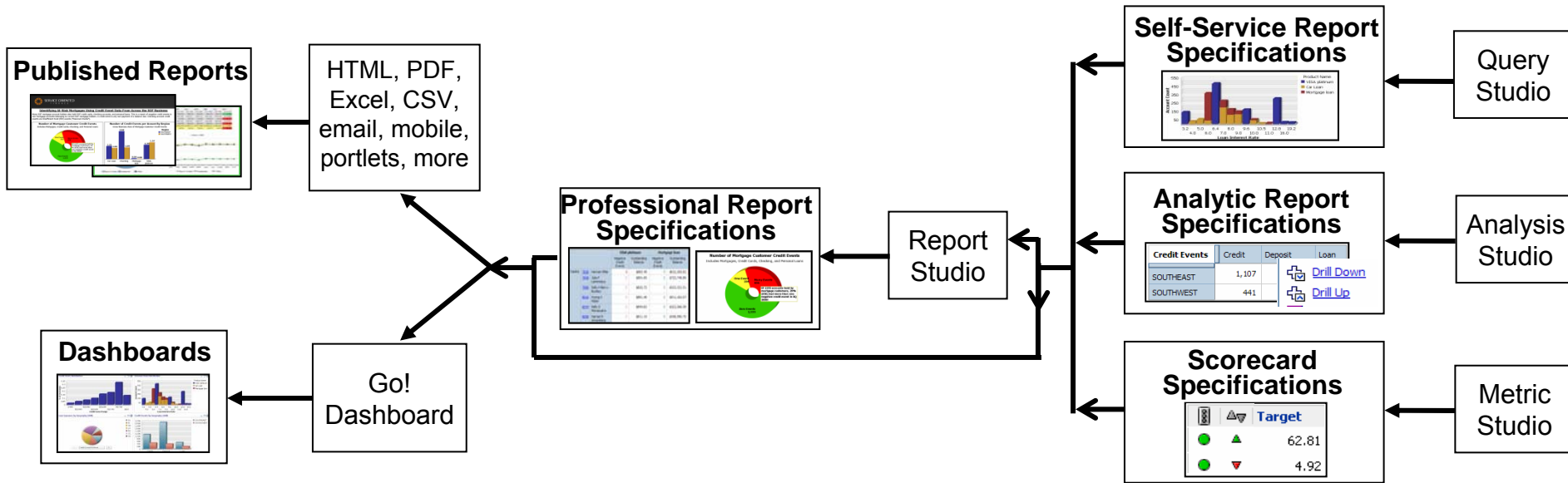
■ At risk customers are identified

# IBM Cognos Is An Integrated Platform For Smart Analytics

- Implemented in Java, runs on WebSphere
- 100% browser based access
  - ▶ Server side business intelligence
  - ▶ Users can access new intelligence from anywhere
- Easiest for IT to deploy and manage
  - ▶ Scales up and out across heterogeneous hardware and operating systems
  - ▶ Unified security
  - ▶ Unified administration
- Consistent user interface across tooling
  - ▶ Greater user satisfaction and increased business agility with lower IT costs
- Common meta data model
  - ▶ Author new intelligence assets once, consume anywhere
  - ▶ Common view enables open data strategy
  - ▶ Supports Unicode and multilingual features without recreating reports



# Reuse Trusted New Intelligence Assets Across The Cognos Platform



- All new intelligence assets share a common metadata model and common report specification
- Author Once – Consume Anywhere
- Ensures consistent information and enables reuse across platform functions

- Oracle has multiple metadata models depending on source type
- Oracle has multiple different report formats
- Oracle cannot reuse assets between tools

# IBM Smart Analytics System 9600: Delivers Enterprise-wide Analytics On z196

- **Extend the qualities of service, inherent in the z/OS environment to ensure the availability and security of data.**
  
- **Hardware/OS**
  - ▶ z/OS 1.12
  - ▶ IBM zEnterprise technology
  - ▶ IBM System Storage DS8700 Intelligent Disk controller
    - Large controller cache and 3 Tier disk offering
  
- **Unique Software**
  - ▶ DB2 for z/OS
  - ▶ Cognos included in base offering (Linux on System z)
  - ▶ InfoSphere Warehouse (Linux on System z)
  
- **Optional Components**
  - ▶ Solid State drives, integrated within DS8700
    - Easy Tier to identify and migrate “hot data” to SSD
  - ▶ IBM Smart Analytics Optimizer



# SPSS Enables Customers To Predict Future Events And Proactively Act Upon That Insight To Drive Better Business Outcomes

## Capture

*Data Collection delivers an accurate view of customer attitudes and opinions*

IBM SPSS  
Data Collection

## Predict

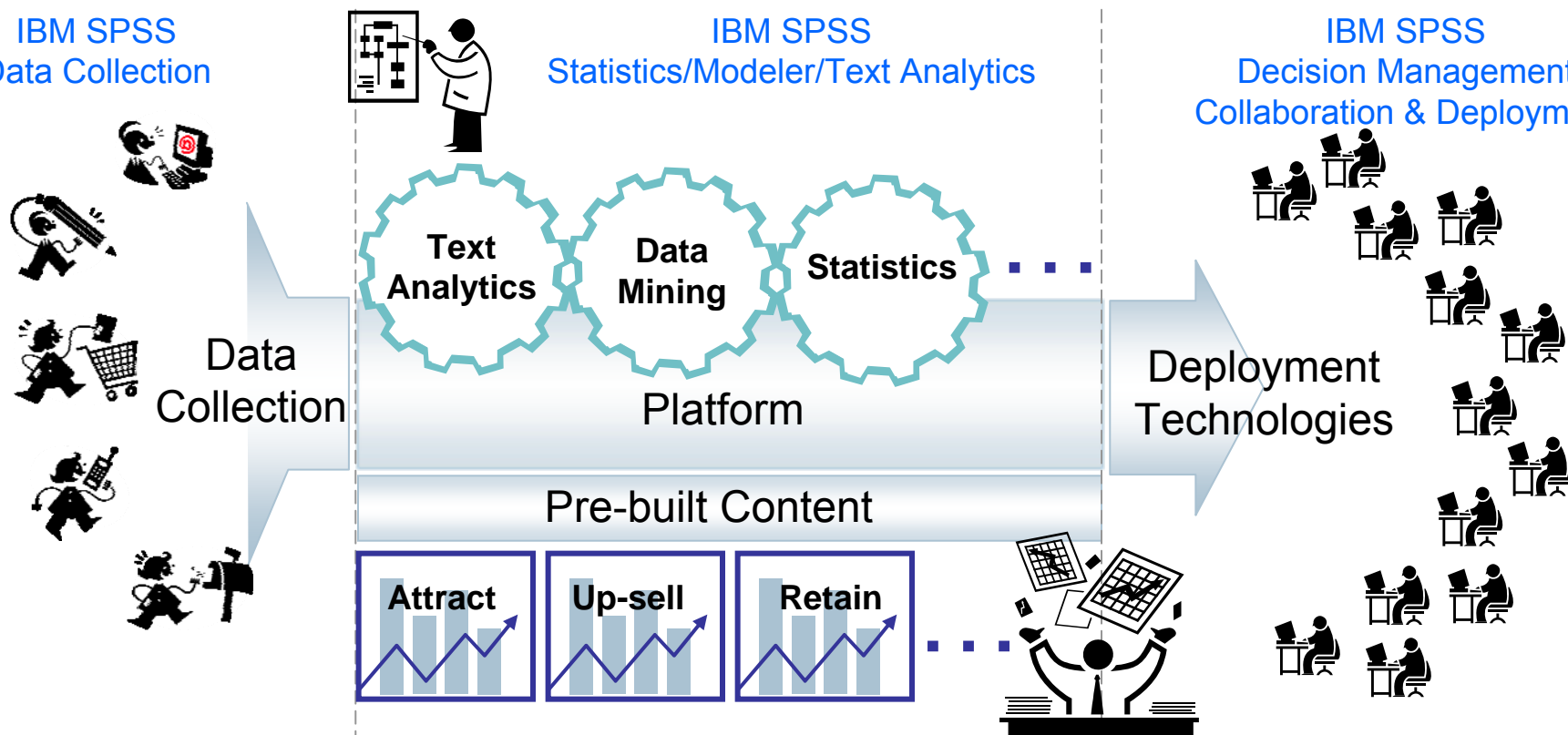
*Predictive capabilities bring repeatability to ongoing decision making, and drive confidence in your results and decisions*

IBM SPSS  
Statistics/Modeler/Text Analytics

## Act

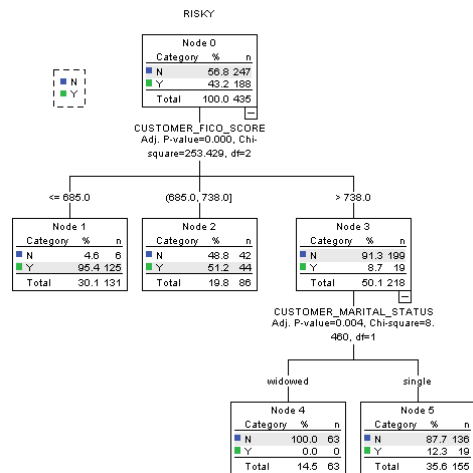
*Unique deployment technologies and methodologies maximize the impact of analytics in your operation*

IBM SPSS  
Decision Management  
Collaboration & Deployment



# DEMO: Discover Rules For Identifying Risky Customers Using SPSS Statistics

1. Load data from Data Warehouse on DB2 for z/OS into SPSS Statistics
2. Pre-process the data to create new attributes for quantifying negative credit events across different product lines and create a risk flag for mortgage
3. Run Comparison of Means and Decision Tree to discover rules for characterizing risky customers



We can now predict our risky customers!

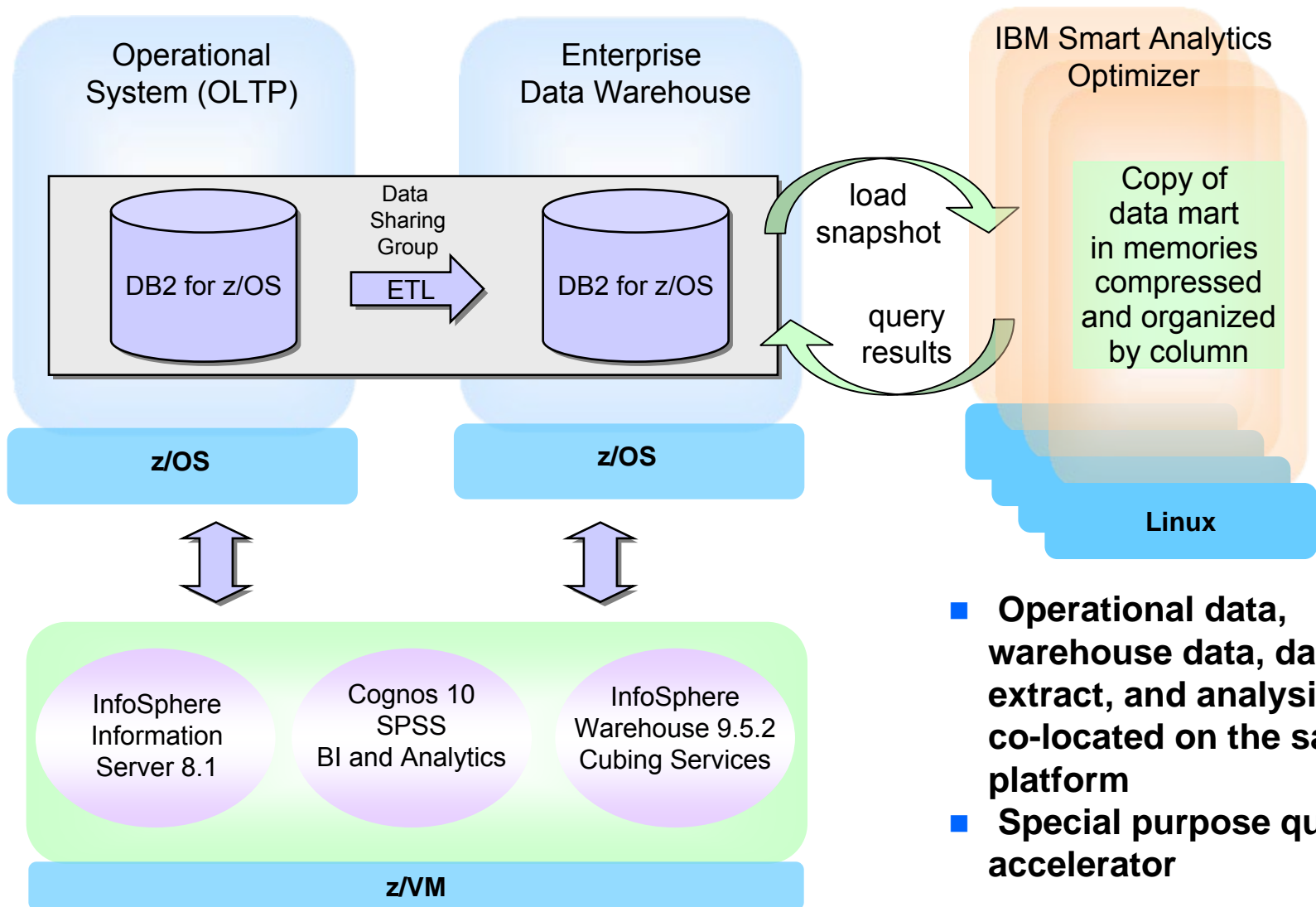


**Mortgage  
Line of Business VP**

- Credit Limits identified for characterizing risky customers
- Use these credit limits for automated loan approval process



# Consolidate Complete Business Intelligence Solution On zEnterprise



- **Operational data, warehouse data, data extract, and analysis all co-located on the same platform**
- **Special purpose query accelerator**

# zEnterprise Is An Excellent Base For Your Data Warehouse And Business Analytics

- Operational and warehouse data co-located on z196
- Exploits System z Parallel Sysplex for availability and scale
- Cognos supports a common metadata model and report specification and provides 100% browser based access
- SPSS predictive analytics provides actionable insights versus hindsight
- Systematic disaster recovery and backup strategies
- Qualities of Service
- Competitive Price