The New zEnterprise – A Cost-Busting Platform

World's Fastest Analytics

Getting analytics out to frontline workers is more critical than ever …

40% decline
in homicide rates

600% increase
in cross-sell campaign

$13.8 Million
in cost savings

The more analytics
a business uses,
the better it performs

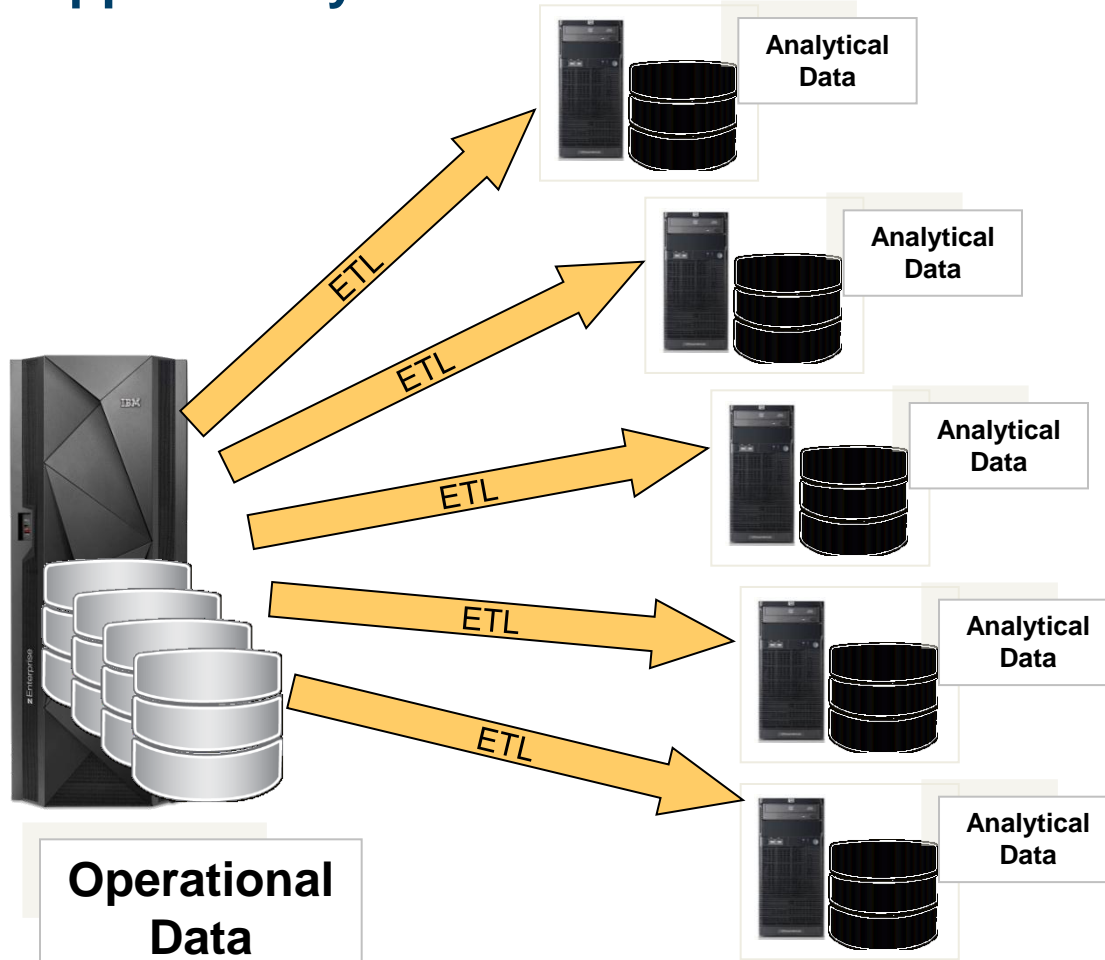**1000's of Reps** run their
daily business using IBM
Business Analytics

80% decrease
in reporting time on top
of Oracle e-business suite

$200 Million
increase in cash flow

Some customers choose a "mainframe quarantine" strategy to support analytics

Analytical Data

Analytical Data

Analytical Data

Analytical Data

Analytical Data

ETL

ETL

ETL

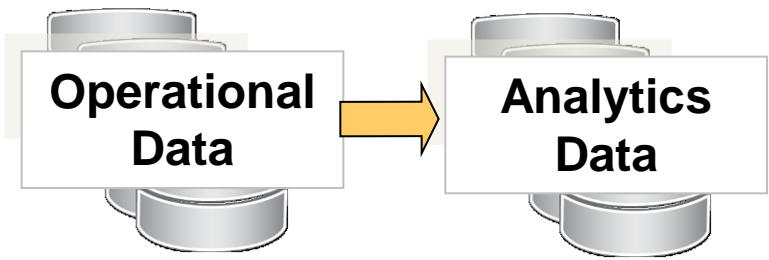ETL

ETL

Operational Data

A large European bank:

- 120 database images created from bulk data transfers
- 1,000 applications on 750 cores with 14,000 software titles
- ETL consuming 28% of total distributed cores and **16% of total MIPS**

A large Asian bank:

- One mainframe devoted exclusively to bulk data transfers
- ETL consuming 8% of total distributed core and **18% of total MIPS**

With this strategy, IT costs were growing faster than the business growth

The solution – add analytics workloads closer to the data

Business Analytics
Operational Analytics   Deep Analytics

Operational Data → **Analytics Data**

zBC12

IBM DB2 Analytics Accelerator

IBM zEnterprise Analytics System 9710

- Run analytics workloads in a separate 9710 LPAR

- Elect to include IBM DB2 Analytics Accelerator with 9710 offering

- Reduce data transfer costs
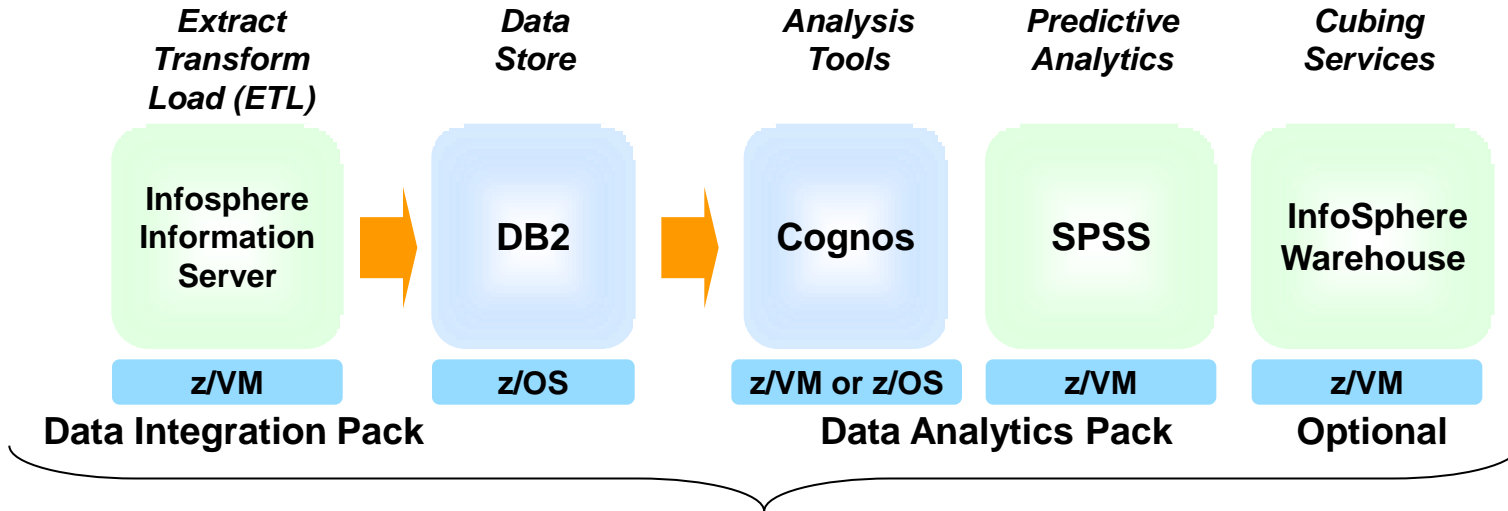
- Achieve lowest cost for analytic workload

Run a complete portfolio of analytics software on IBM zEnterprise BC12

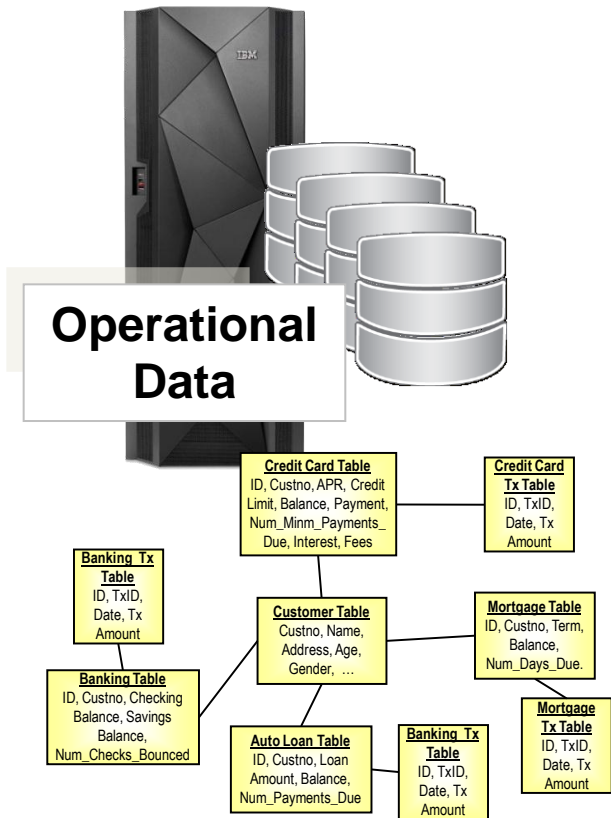| Extract Transform Load (ETL)	Data Store	Analysis Tools	Predictive Analytics	Cubing Services
Infosphere Information Server	DB2	Cognos	SPSS	InfoSphere Warehouse
z/VM	z/OS	z/VM or z/OS	z/VM	z/VM
Data Integration Pack		**Data Analytics Pack**		**Optional**

IBM zEnterprise Analytics System 9710 –
A comprehensive packaged solution including hardware, OS, and business analytics software

zBC12   Accelerator

FastStart Services Pack

02. World's Fastest Analytics

DB2 for z/OS is a first class platform for operational analytics

Operational Data
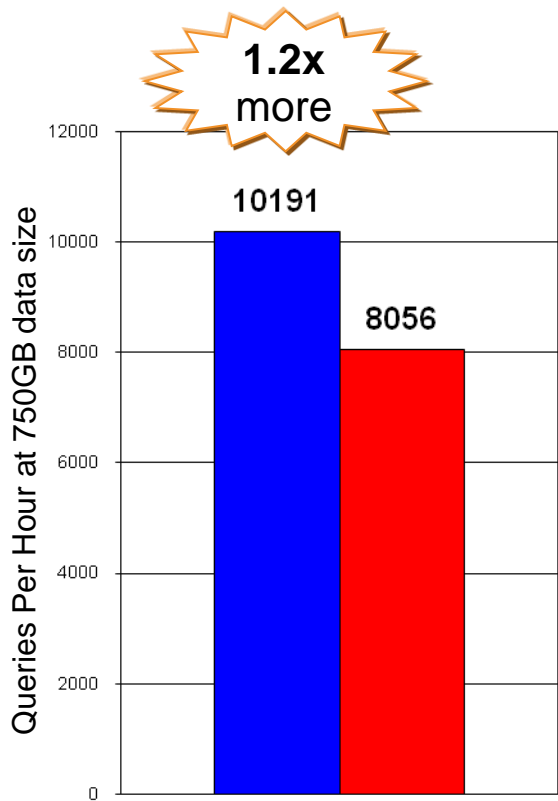
Credit Card Table
ID, Custno, APR, Credit Limit, Balance, Payment, Num_Minm_Payments_ Due, Interest, Fees

Credit Card Tx Table
ID, TxID, Date, Tx Amount

Banking Tx Table
ID, TxID, Date, Tx Amount

Customer Table
Custno, Name, Address, Age, Gender, …

Mortgage Table
ID, Custno, Term, Balance, Num_Days_Due.

Banking Table
ID, Custno, Checking Balance, Savings Balance, Num_Checks_Bounced

Auto Loan Table
ID, Custno, Loan Amount, Balance, Num_Payments_Due

Banking Tx Table
ID, TxID, Date, Tx Amount

Mortgage Tx Table
ID, TxID, Date, Tx Amount

Business Analytics
Operational Analytics   Deep Analytics

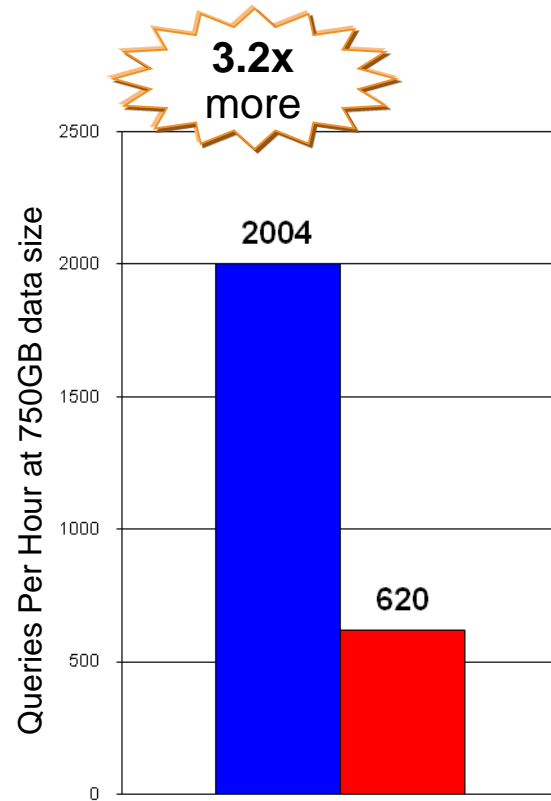*Many concurrent queries of varying complexity*

- DB2 Cost Based Optimizer provides best access path and query execution plan

- Parallel sysplex yields near-linear scaling and high availability

- z/OS Workload Management optimizes resource sharing to minimize impact on high priority transactional workloads

Customer performance study proves DB2 for z/OS operational analytics beats the competition
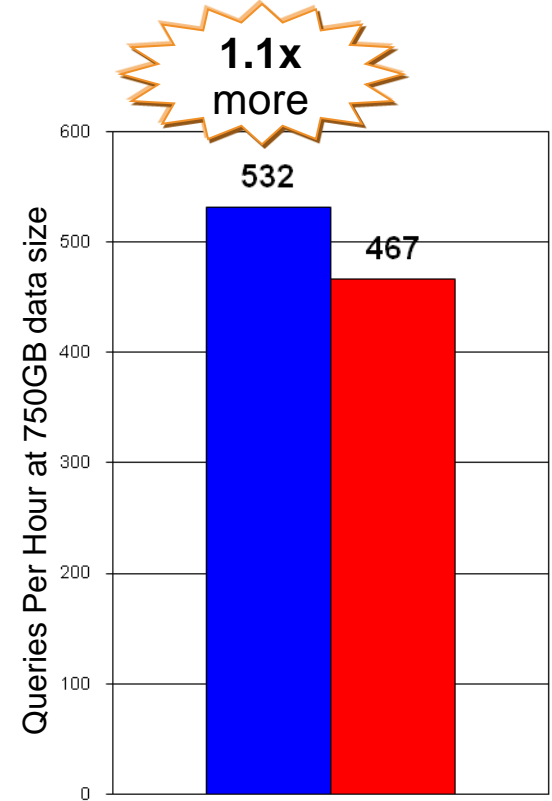

Simple Queries

Intermediate Queries

Complex Queries

02. World's Fastest Analytics

DB2 for z/OS is also optimized for deep analytics

Analytical Data

Dimension Table
Account
ID, Type

Dimension Table
Branch
ID, Address,
City, State

Fact Table
Balance Fees
Credit Events

Dimension Table
Product
Name, Type,
Description…

Dimension Table
Customer
ID, Name, Address,
Age, Gender…

Dimension Table
Date
Year, Month, Day
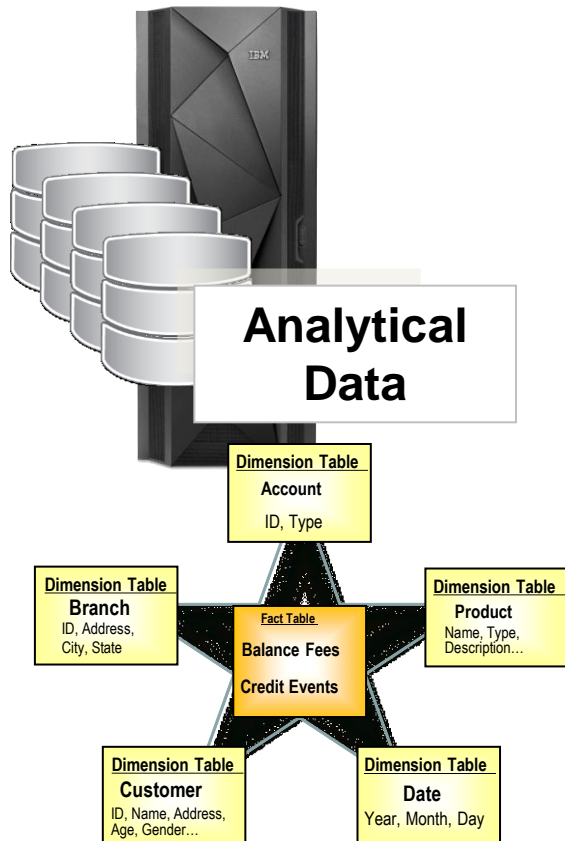
Business Analytics
Operational Analytics   Deep Analytics

Intermediate and complex analytical queries

- Data is partitioned to increase parallelism and compressed to increase I/O performance

- DB2 Cost Based Optimizer decides best execution plan for each query

  ▶ Complex queries may be decomposed into parallel operations

  ▶ Queries may be automatically rewritten to take advantage of pre-computed partial results in materialized query tables (MQT)

- Optional IBM Data Analytics Accelerator

Add IBM DB2 Analytics Accelerator to speed up deep analytics queries

- A workload-optimized, blade-based appliance based on Netezza Technology that runs queries in seconds versus hours

  - Storage integrated into the hardware rack

  - Eliminate table indexing and query tuning

- Integrated with DB2 for z/OS, and transparent to applications

  - Pre-load data from DB2 for z/OS into Accelerator at over 400GB/hr

  - Maintain a single copy of table or partition in Accelerator and update incrementally (High Performance Storage Saver)

  - System z workload management implemented across Accelerator

- Significantly speeds up the response time

- Drives down the costs of data warehousing and business analytics

Breakthrough technology enabling new opportunities

New V3.1!

DB2 Analytics Accelerator leverages massively parallel processing to speed up deep analytics queries

Data partitioned across CPUs and storage

Query distributed to CPU/FPGAs which decompress and filter data in real-time.
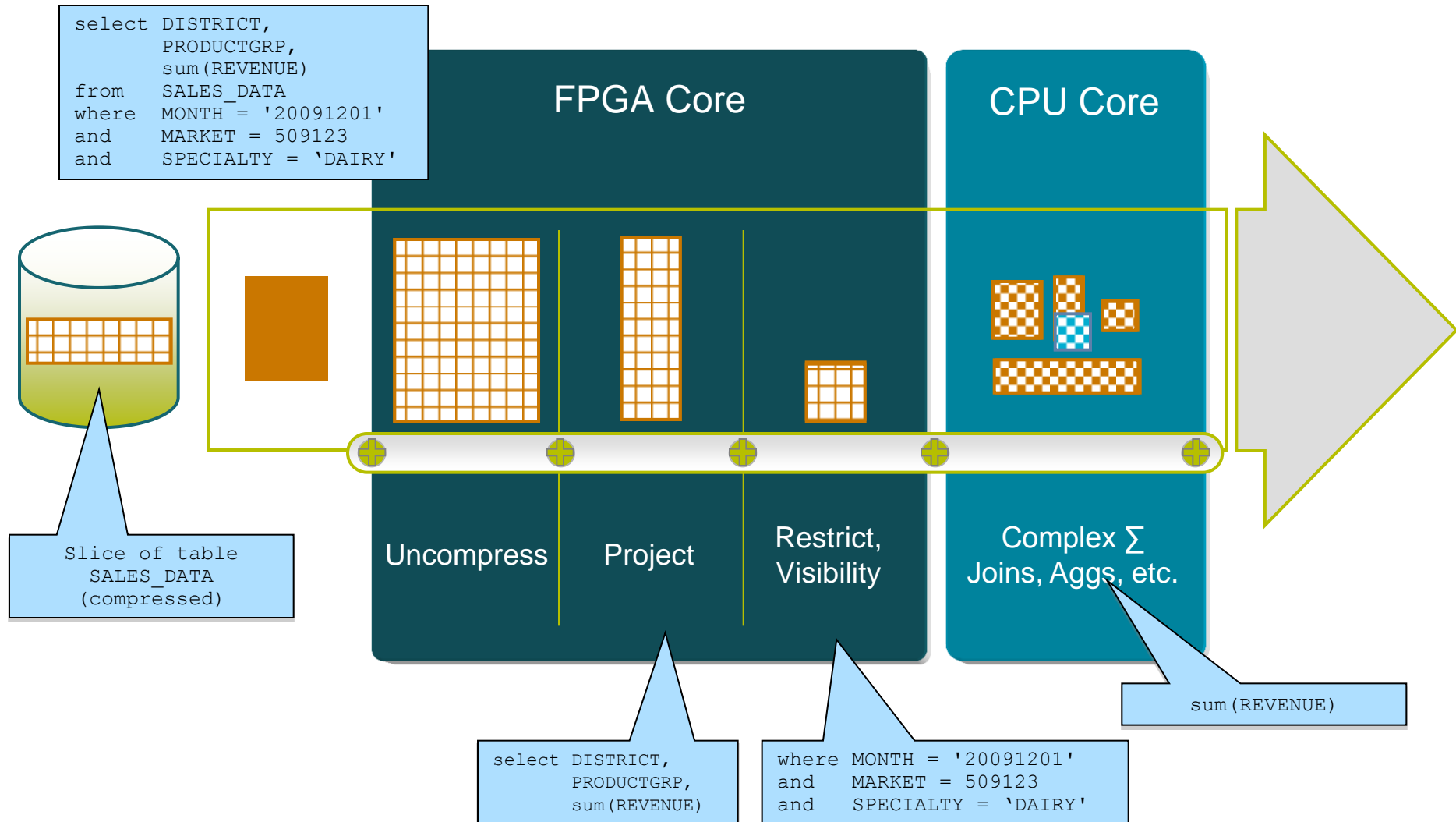
SMP Host assembles results and returns

Query

Storage

SMP Hosts

CPU
Field Programmable
Gate Array

"…when something took 24 hours I could only do so much with it, but when something takes 10 seconds, I may be able to completely rethink the business…" SVP, Nielsen

Key reason for world's fastest analytics – Field Programmable Gate Arrays

```
select  DISTRICT,
        PRODUCTGRP,
        sum(REVENUE)
from    SALES_DATA
where   MONTH = '20091201'
and     MARKET = 509123
and     SPECIALTY = 'DAIRY'
```

FPGA Core

CPU Core

Uncompress | Project | Restrict, Visibility | Complex ∑ Joins, Aggs, etc.

Slice of table
SALES_DATA
(compressed)

```
select  DISTRICT,
        PRODUCTGRP,
        sum(REVENUE)
```

```
where  MONTH = '20091201'
and    MARKET = 509123
and    SPECIALTY = 'DAIRY'
```

sum(REVENUE)

DB2 Analytics Accelerator speeds up complex queries

Query	DB2 (Secs)	DB2 + Analytics Accelerator (Secs)	Speed Up	Rows Reviewed	Rows Returned
Query 1	9,540	5	1,908x	2,813,571	853,320
Query 2	8,220	5	1,644x	2,813,571	585,780
Query 3	4,560	6	760x	8,260,214	274
Query 4	4,080	5	816x	2,813,571	601,197
Query 5	4,080	70	58x	3,422,765	508
Query 6	3,180	6	530x	4,290,648	165
Query 7	3,120	4	780x	361,521	58,236
Query 8	2,640	2	1,320x	342,529	724
Query 9	2,520	193	13x	4,130,107	137

N1001-10

"We had this up and running in days with queries that ran over 1000 times faster", AETNA

"We expect ROI in less than 4 months", Swiss Re

Now with "trickle" feed!

Actual customer results, October 2011

02. World's Fastest Analytics

DB2 Analytics Accelerator also speeds up SAP Business Warehouse

- Accelerates SAP Business Warehouse on DB/2 for z/OS

- Dramatic decrease in elapsed time for SAP Business Warehouse ad-hoc reporting

- Model N1001-10

No	Description	Records read	Records returned	DB2 [sec]	IDAA [sec]	Speed Up
1	Simple mass aggregation	17116647	21	117	0.78	150
2	Query #1 + 70% filter	11980812	21	94.2	0.86	110
3	Query #1 + 30% filter	5133708	21	54.8	0.82	67
4	Query #1 + 10% filter	1710293	21	17.6	0.87	20
5	Screwed data, low filtering	10790019	21	96.8	2.47	39
6	Screwed data, high filtering	24	14	7.28	0.83	9
7	Many restrictions	3805941	21	128	7.65	17
8	Navigational attributes	823646	21	17.1	1.27	13
9	Navigational attributes + selective condition	811	21	15.8	1.17	14
10	Open value ranges	2006	21	19.6	3.52	6
11	Hierarchy	1653981	21	17.6	0.97	18
12	Hierarchy + selective condition	55068	21	38.6	0.98	39
13	Restricted key figures on 2 dimensions	1314964	1948	207	7.22	29
14	Query #14 + hierarchy	132564	1499	> 1000	1.27	> 787
15	Calculated key figures (OLAP)	5321586	10	57.8	2.37	24
16	OR linked values	6212609	13	40.5	0.92	44
17	Non uniform data distribution	11016253	13	31.2	0.99	32
18	Selective line item	1724	1706	0.71	1.17	0.6
19	Non-selective line item	115481	68619	33.8	1.36	25
20	All together	3087692	468	87.7	4.42	20

Table 2: Dedicated Query Test on a 18-million-records InfoCube

Low Latency, High Capacity Update
Analytics Accelerator using N1001-10 hardware

02. World's Fastest Analytics

BI Day workload – A typical day's worth of analytics reports

BI Day Fixed Execution Test: Total Number of reports = 161,166

4	3	21	179	99	29,078	84,854	19,600	27,328
Complex Report 1	**Complex Report 3**	**Intermediate Report 9**	**Intermediate Report 10**	**Intermediate Report 11**	**Simple Report 2**	**Simple Report 4**	**Simple Report 5**	**Simple Report 6**

4 connections   20 connections   56 connections

- 80 concurrent connections to DB2 for z/OS

- All short running SQL queries processed by DB2 for z/OS

- Long running SQL queries directed to accelerator

DB2 for z/OS

Each report executes one or more SQL queries

Continuous improvements for analytics workloads – A typical work day

Elapsed Time To Complete BI Day Fixed Execution Test
(10 TB – 161,166 concurrent executing reports)
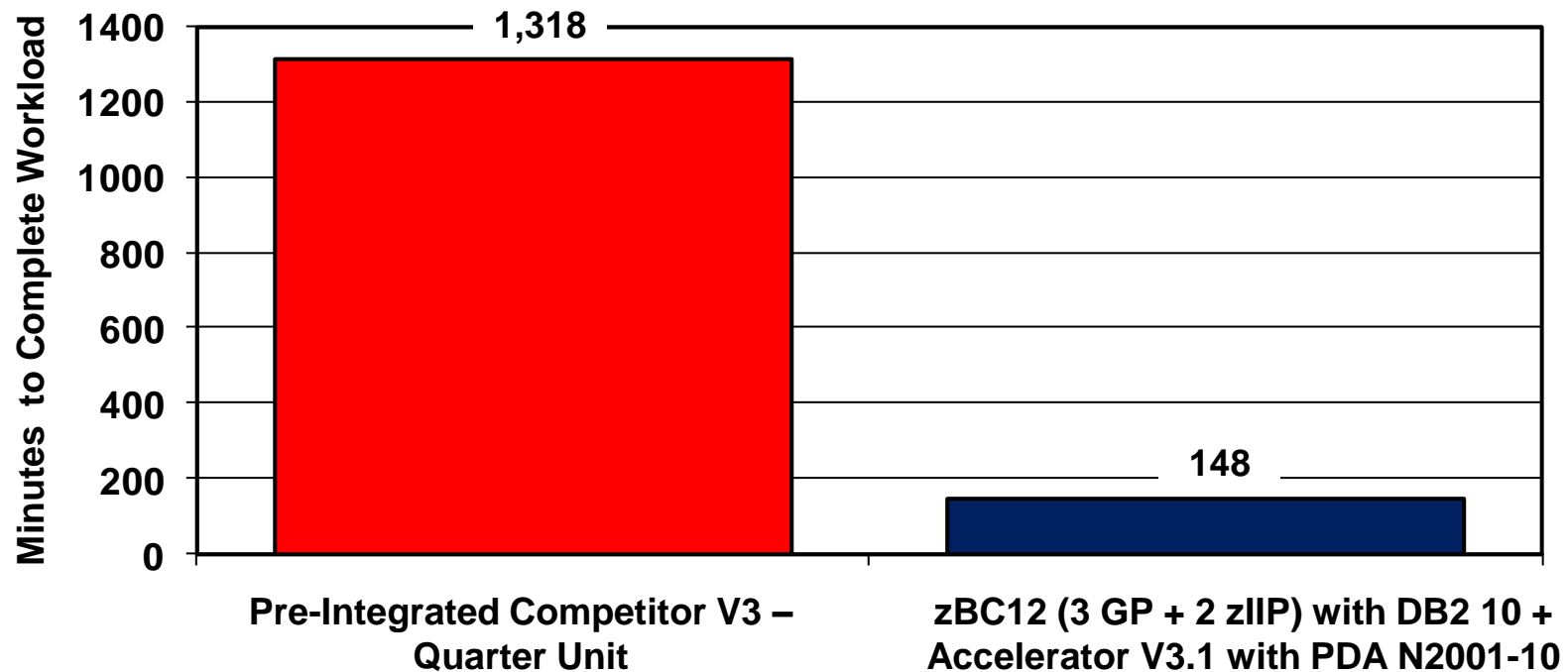
Based on BI Day Tests. Lower is better. Performance numbers may vary based on workload profiles.

Analytics Accelerator outperforms pre-integrated database competitor running BI Day intermediate analytics queries

http://www.youtube.com/watch?v=T3O6yJ_hdUU

IBM DB2 Analytics Accelerator (N1001-10)

Pre-Integrated Database Competitor V2
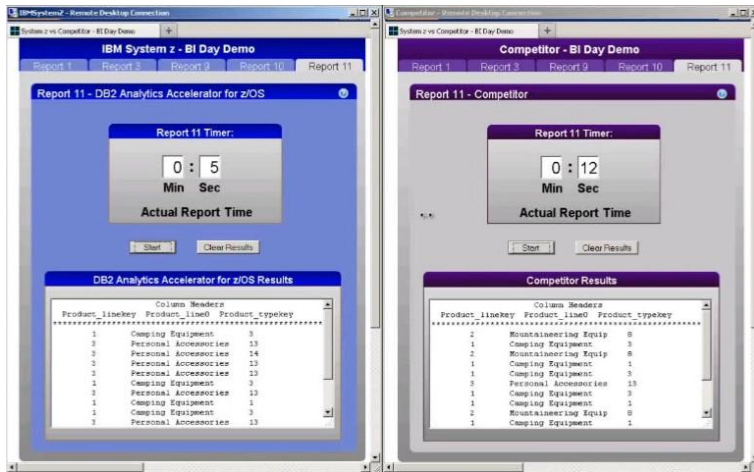
- Same queries, same data run on both platforms

- Both platforms returned same response set

- Competitor took from **2.5** to **9.5** times longer to complete these three runs

Query Response Times in seconds

	Report 11	Report 9	Report 3
DB2 Analytics Accelerator	5	6	162
Pre-integrated database competitor	12	58	1,427

DEMO: DB2 Analytics Accelerator

Compare DB2 BI Day query processing using the IBM DB2 Analytics Accelerator

02. World's Fastest Analytics

zEnterprise is optimized for operational analytics

Standalone
Pre-integrated
Competitor V3
Quarter Unit

IBM zEnterprise Analytics
System 9710

DB2 v10

z/OS
2 GP+ 1 zIIP

IBM DB2
Analytics
Accelerator
(with N2001-10)

zBC12

Unit Cost (3yr TCA) $481/RpH

Workload Time (mins)	1,318
Reports per Hour (RpH)	7,337
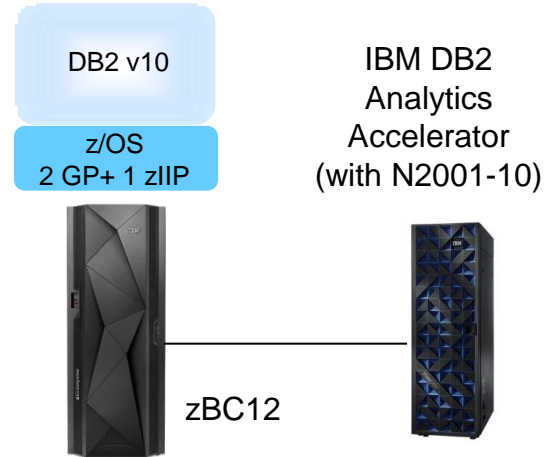
Unit Cost (3yr TCA) $46/RpH

Workload Time (mins)*	148
Reports per Hour (RpH)	65,338

9x performance
10x price performance!

Source: Customer Study on 10TB BIDAY data running 161,166 concurrent reports. Intermediate and complex reports automatically redirected to IBM DB2 Analytics Accelerator for z/OS. Results may vary based on customer workload profiles/characteristics. Note: Indicative 9710 pricing only internal to IBM, quotes to customer require a formal pricing request with configurations.

zEnterprise is best for analytics

Traditional Data Warehouse Competitor

6650H

(current generation)

IBM zEnterprise Analytics System 9700

DB2 v10

z/OS
2 GP+ 1 zIIP

IBM DB2 Analytics Accelerator (with N1001-10)

zEC12

(current generation)

Unit Cost (3yr TCA) **$330K/QpH**

Workload Time (secs)*	1,591
Queries per Hour (QpH)	9
Total Cost (3 yr. TCA) - Teradata 6650H 1-Node (HW+SW+Storage)	$2,946,046
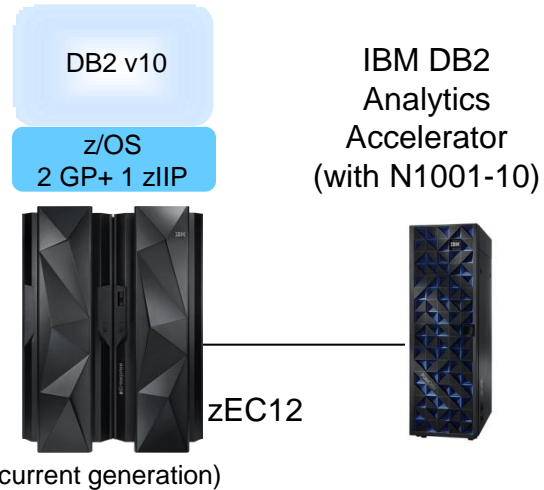
Unit Cost (3yr TCA) **$10K/QpH**

|	
Workload Time (secs)*	61
Queries per Hour (QpH)	236
Total Cost (3 yr. TCA) – 9700 : zEC12 (1 GP + 1 zIIP) + DB2 Analytics Accelerator (HW+SW+Storage)	$2,337,400

Source: Customer Study running 4 complex analytical queries across a 353GB Data Mart, Total User Data shared across 8 older generation Teradata Servers was 15TB. Customer workload was not from any benchmark applications, nor are they based on any benchmark standard. As such, customer applications, differences in the stack deployed, and other systems variations or testing conditions may produce different results and may vary based on actual configuration, applications, specific queries and other variables in a production environment. List Prices used in the comparison. Teradata 6650H price available from ftp://public.dhe.ibm.com/software/data/sw-library/infosphere/analyst-reports/ITG-ISAS-Exadata-Teradata.pdf . The 3 year total cost of acquisition includes expected hardware, software, service & support. Prices will vary by country. Results will v66ary based on actual configuration, applications, specific queries, and other variables in a production environment. Users of this document should verify the applicable data for their specific environment. Contact IBM and see what we can do for you. *Throughput projected for Teradata 6650H using TPerf data (http://www.monash.com/uploads/Teradata-Active-EDW-6660-6680.pdf : Rated TPerf per 6650H node 121.5) . Throughput for zEnterprise Analytics System 9700 was based on lab measurements for equivalent workload performance improvement on upgrading to DB2 10 for z/OS on zEC12

26x performance
33x price performance!

02. World's Fastest Analytics

IBM Blue Insight uses System z platform to deploy an internal private analytics cloud
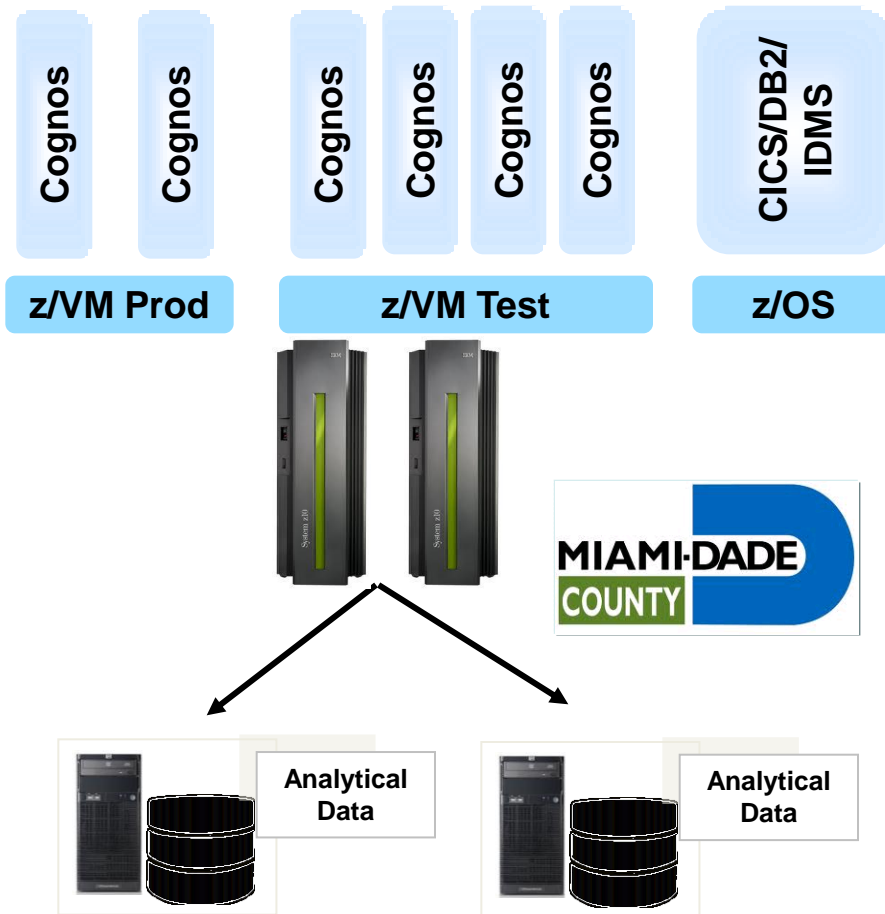
Project Scope

- Over 200K named users, 390 distinct Cognos BI reporting projects, over 2M reports/quarter
- 250 data sources - DB2, PowerCube, XML, Power, Linux on System z, z/OS
- Savings of over 74K sqft floor space, 30K MWh energy and cut per user cost from $250K to $13K
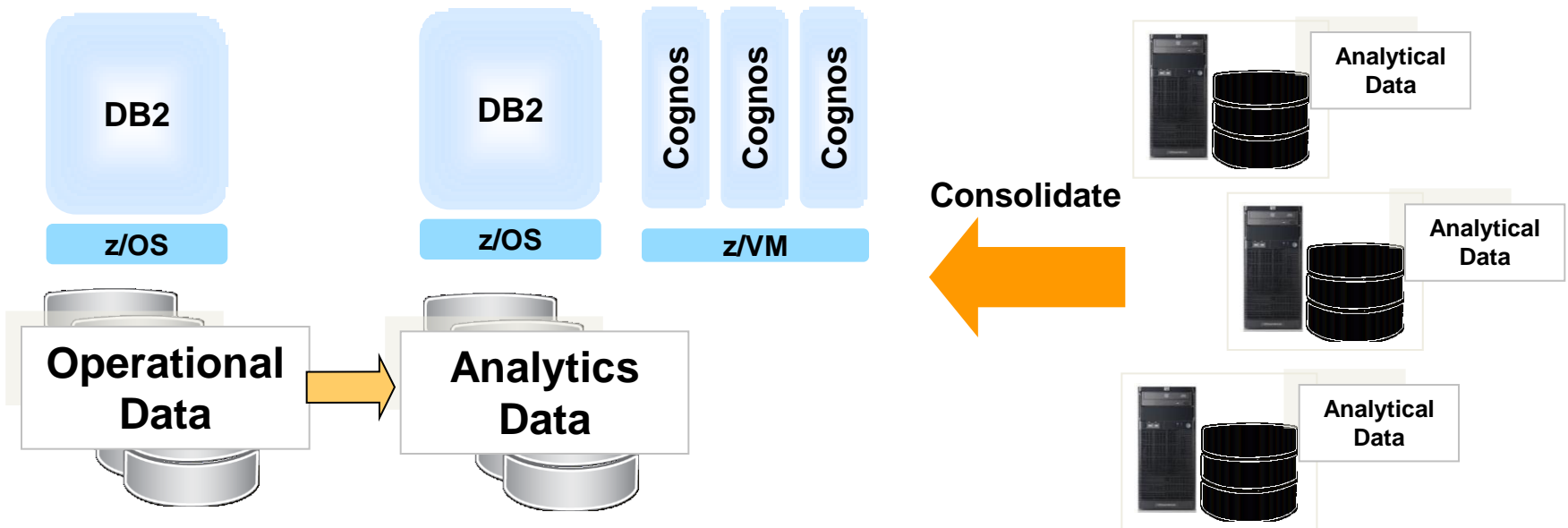
> **"** *Our commitment to informed decision-making led us to consider private **cloud delivery of Cognos via System z**, which is the enabling foundation that makes possible **+$25M savings over 5 years**.* **"**
>
> *-- IBM CIO Office*

02. World's Fastest Analytics

Miami-Dade County runs IBM Cognos on business class mainframes

Cognos **Cognos** | z/VM Prod

Cognos **Cognos** **Cognos** **Cognos** | z/VM Test

CICS/DB2/ IDMS | z/OS

MIAMI-DADE COUNTY

Analytical Data

Analytical Data

Business Benefit:

- Moved Cognos BI deployment from Intel servers to System z10 BC in 11 days
  - Consolidated multiple deployments to a single platform
  - Consolidated multiple disparate data sources
  - Single point for BI administration
  - Offer a complete disaster recovery plan
  - Additional green savings

- Easily met requirements for growth, 24x7 availability and TCO savings

- Upgrading to Cognos 10

"We have users from 25 County Service departments with almost 2000 users consuming and creating reports with stable environments on System z"

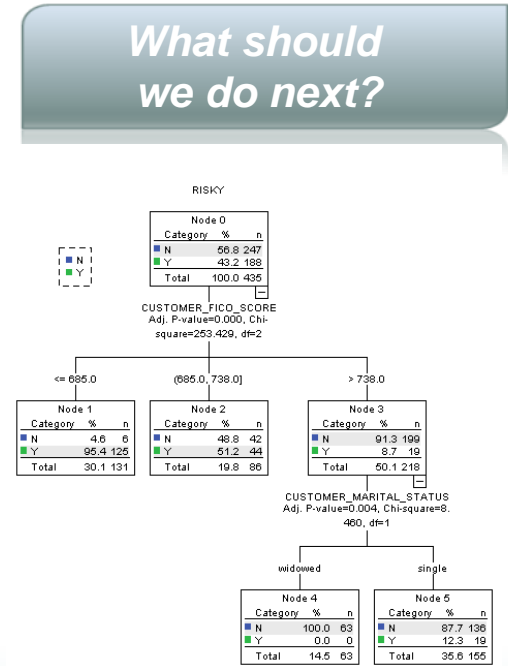*- Jaci Newmark, Miami-Dade County*

Add business analytics tools to answer key questions that drive a competitive edge

How are we doing?

Why are we on/off track?

What should we do next?

Dashboard
Operational or Strategic
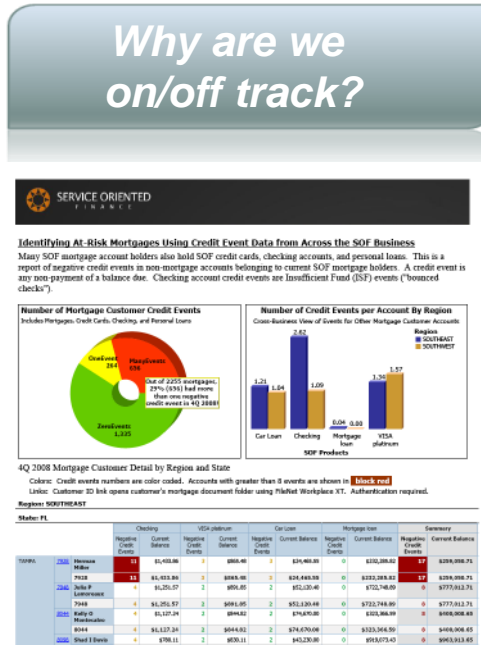
Query and Reporting
Exploratory Analysis

Analytics Statistics
and Predictive Analytics

Executive | **Business Manager** | **Casual Business User** | **Line Manager** | **Business Analyst** | **Financial Analyst**

Business analytics tools can be grouped into two areas

Descriptive Analytics:

- Provides reports/dashboards

- Insight into what has happened

- Drill-down on data using different dimensional attributes such as by date (month, year), geography (state, county, city), demographics (age, gender, race)

- Aggregate data

- Visualize data using interactive charts, graphs, maps and other objects

- IBM Cognos Enterprise

Predictive Analytics:

- Provides scores that helps in optimized decision support

- Predicts what might happen

- Build Models using historical data

- Some models provide rules that can be integrated into business processes

- Uses mathematical algorithms that can be unsupervised (such as clustering) or supervised (such as classification)

- IBM SPSS Statistics and Modeler

02. World's Fastest Analytics

Cognos can generate reports and dashboards for operational BI and deep analytics queries

People-centric

- Server based business analytics accessed via browser
- Consistent user interface for different analytic activities
- Reuse new intelligence assets
- Built-in collaboration and social networking
- Threaded discussions, activities, and notifications

Easy to deploy and manage

- Implemented in Java, runs on WebSphere
- Scales up and out across heterogeneous hardware and operating systems
- Runs on Linux on System z or z/OS

User Browser

IBM Cognos server

Query Report OLAP

Metadata Model

Operational   **Warehouse**
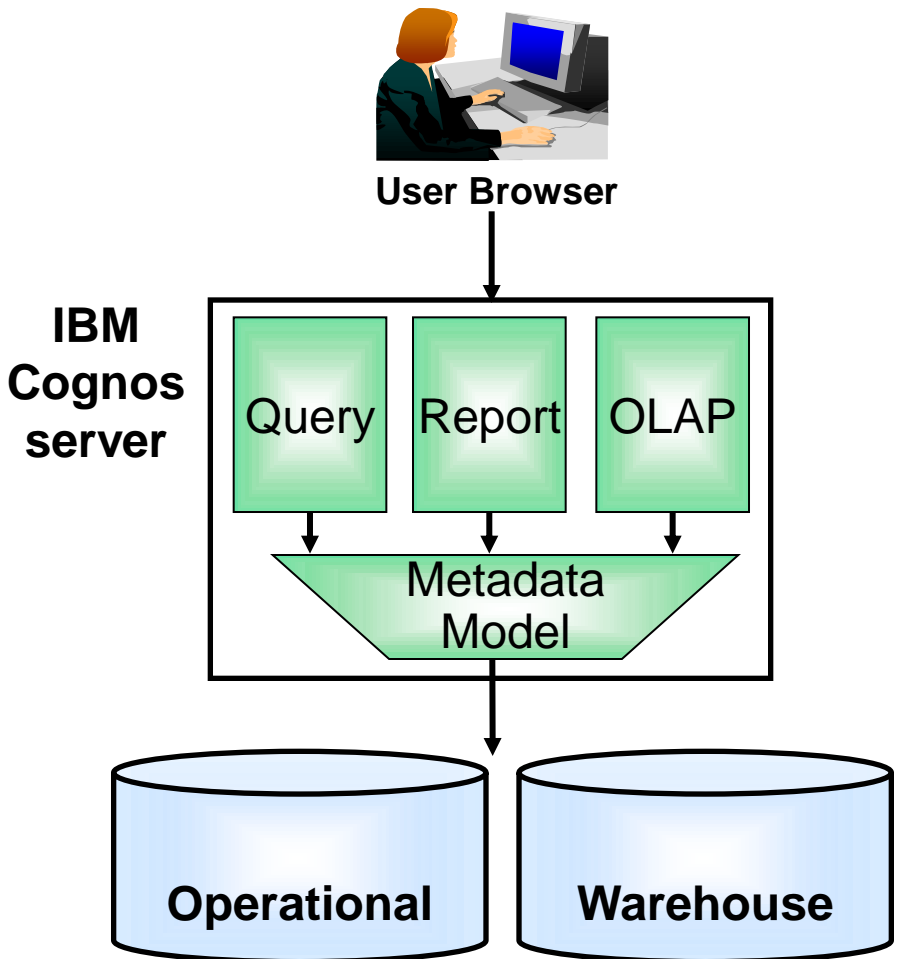
DEMO: Use Cognos to identify new business insights from the data warehouse

1. Reviewing Mortgage data provides false impression of credit risk

2. Report looking at negative credit events (bounced checks, missed payments) of customer across all accounts (Credit Card, Checking, etc.

3. Identify high risk mortgage customers

At risk customers are identified

02. World's Fastest Analytics

Predictive analytics helps a business run smarter

Turn a Call Center in a Profit Center.

A large Dutch financial services company generated **$30 Million in incremental sales**. Essentially, 1M calls generated 180,000 suggestions, reps made 60,000 offers generating 30,000 leads and 22,000 sales.

Prevent crime before it happens.

A large city in the US optimized deployment of police resources, **reducing homicides by 35%** year over year, and robberies by 20%.

Turn clients into advocates.

A large Swiss telco provider adopted a client retention approach based on satisfaction. And **reduced churn from 14% to 2%.**

Reduce the cost of claims.

A large US insurer maximized and accelerated the collections process achieving an **ROI of 403% with payback in 3 months**.
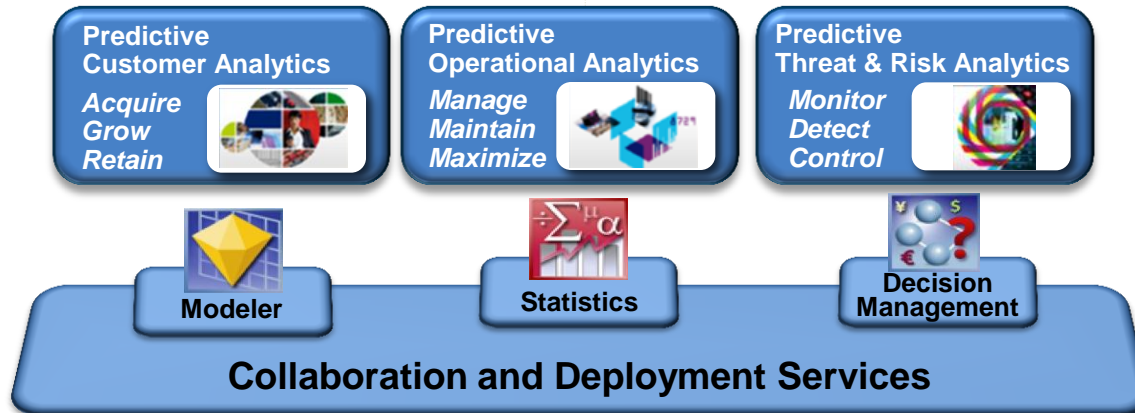
SPSS enables customers to predict future events and drive better business outcomes

SPSS Statistics for Linux on System z	SPSS Modeler for Linux on System z
▪ Version: Statistics v20 ▪ Apply math to decision making and research for commercial, government, and academic users	▪ Version: Modeler v15 ▪ Data mining tool used for generating hypotheses and scoring ▪ Text analysis for unstructured data to model consumer behavior ▪ In-Transaction Scoring with DB2 z/OS

NEW

Predictive Customer Analytics
Acquire Grow Retain

Predictive Operational Analytics
Manage Maintain Maximize

Predictive Threat & Risk Analytics
Monitor Detect Control

Modeler

Statistics

Decision Management

Collaboration and Deployment Services

What can we learn from historical data that would help evaluate new customers?

Historical Data

10% → Apply Rules to remaining 10% of data to evaluate accuracy of Decision Tree Model

90% → Build Decision Tree Model using 90% of randomly selected data to generate Rules

New Customer Data → Apply the Decision Tree Model Rules to score new customers

02. World's Fastest Analytics
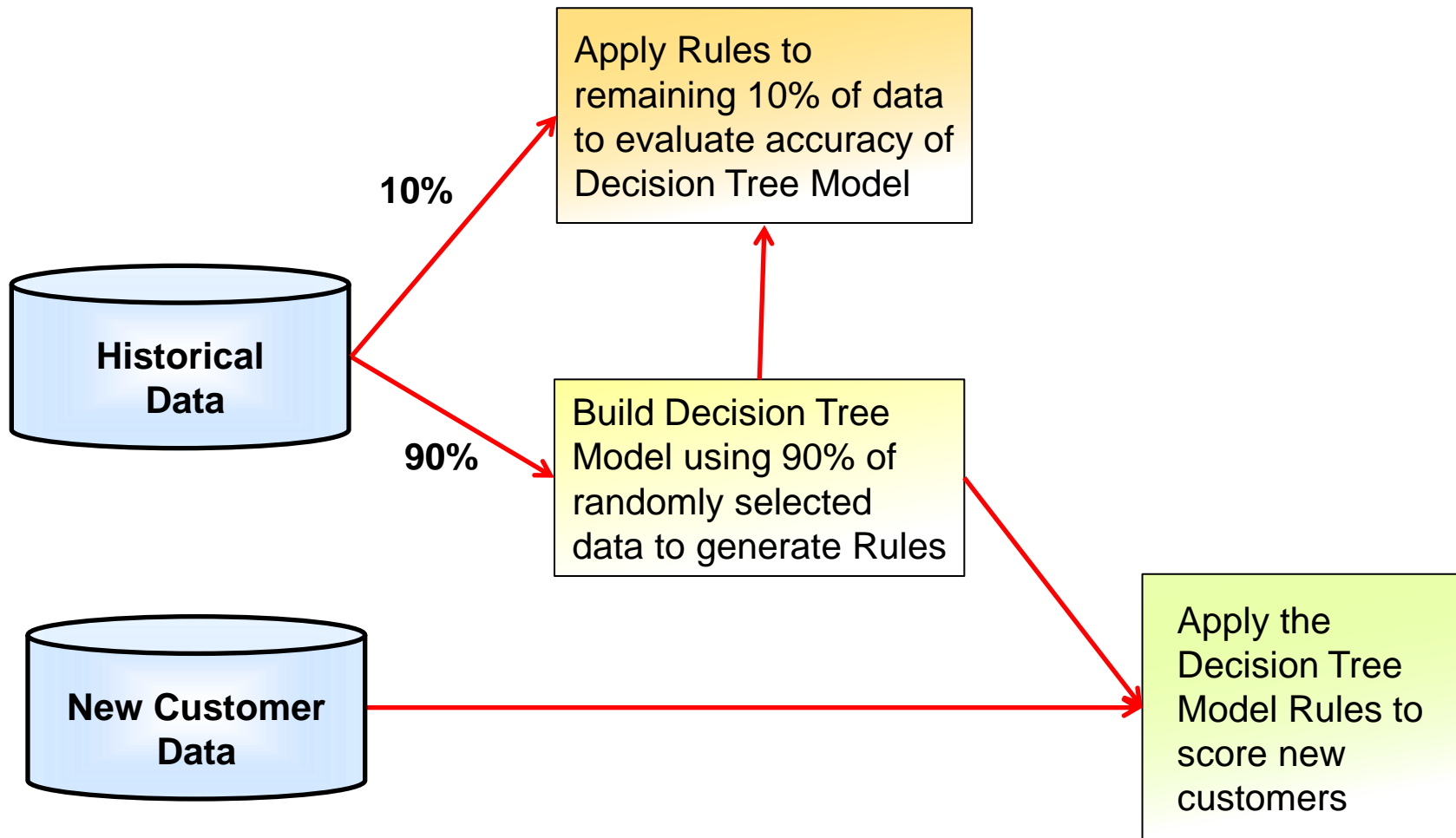
Example: Create a model for identifying risky customers for loan approval

Create a new dependent variable **RISKY**
• "Y" - 5 or more credit events
• "N" – less than 5 credit events

→

Pick 90% of the customers randomly to build a Decision Tree model. Choose:
• **RISKY** as dependent variable
• Add other independent variables such as **PROFESSION, STATE OF RESIDENCE, FICO_SCORE, TOTAL_BALANCE_OWED**

→

Apply the Decision Tree rules for remaining 10% of the customers to test the accuracy of predicting **RISKY**

02. World's Fastest Analytics

# 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 Decision Tree to discover rules for characterizing risky customers

4. Evaluate if Herman Miller is classified as "RISKY" by applying the Decision Tree rules
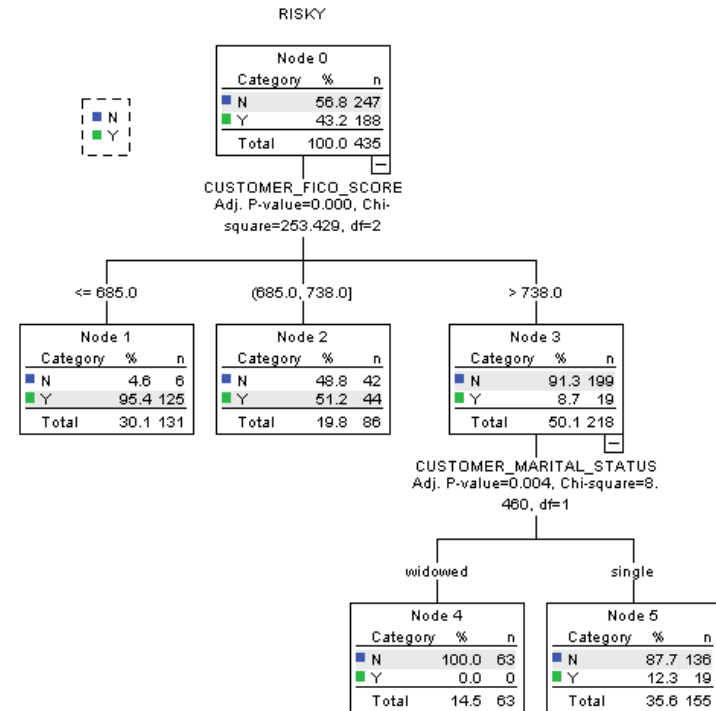
- Credit Limits identified for characterizing risky customers
- Use these credit limits for automated loan approval process

Improve business outcomes with SPSS Modeler in-transaction scoring

- Instantaneous and accurate decision based on real-time information or events

- Reduce risk by putting high risk customers on "watch"

- Increase satisfaction of valued customers by providing the "next-best offer"
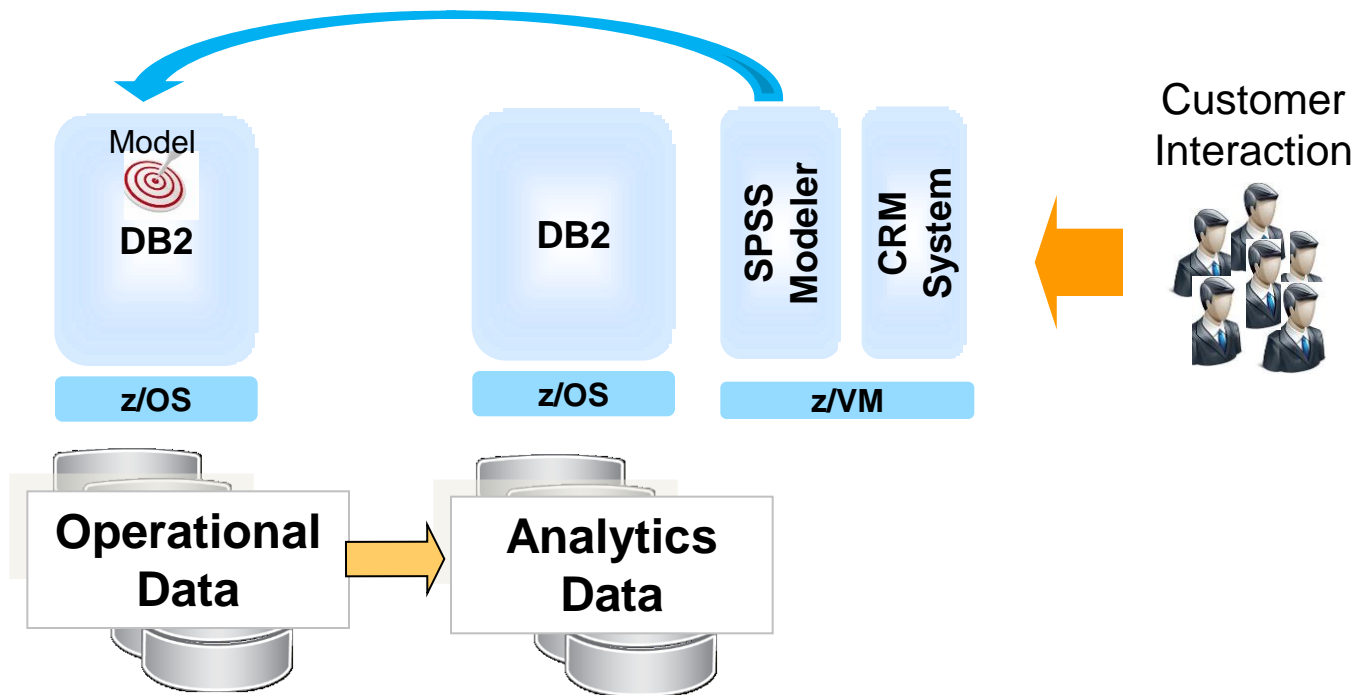
Run end-to-end analytics on zEnterprise to reduce costs and improve reliability

- 60-70% of operational data resides on System z*

- zEnterprise offers a fully integrated, optimized solution on one platform

  - From operational data to business analytics

- Consolidating data warehouses on zEnterprise with Analytics Accelerator can reduce costs over 90%

- Cognos and SPSS add unmatched predictive intelligence

* Source http://www.ibmsystemsmag.com/mainframe/trends/whatsnew/The-Mainframe-at-a-Crossroads/

02. World's Fastest Analytics