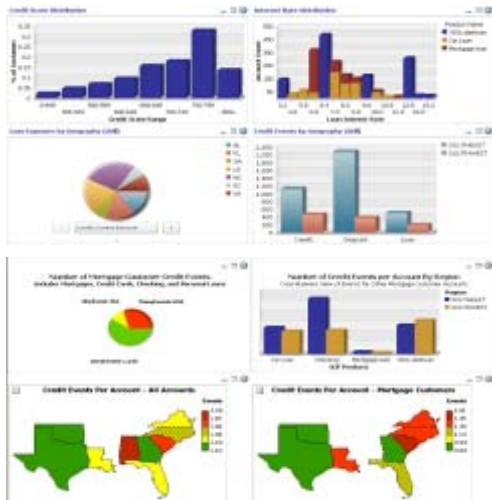# zEnterprise –
An Ideal Basis For
Smarter Computing

System z – Best Place For Business Analytics
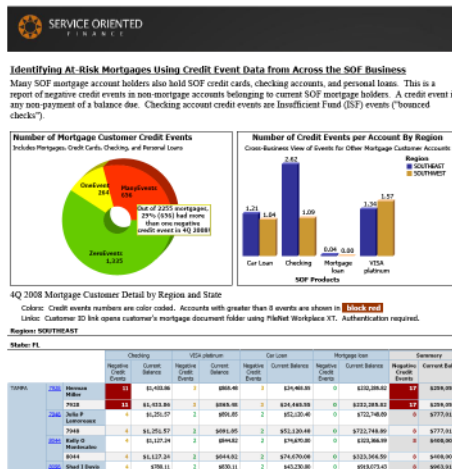
Businesses Analytics Answers Key Questions That Drive A Competitive Edge
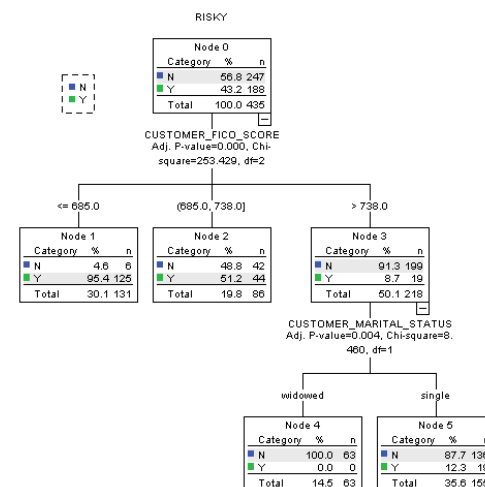
How are we doing?	Why are we on/off track?	What should we do next?

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

Businesses Benefit By Using An Analytic Approach Over Intuition

40% decline
in homicide rates

600% increase
in cross-sell
campaign

$13.8 Million
in cost savings

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

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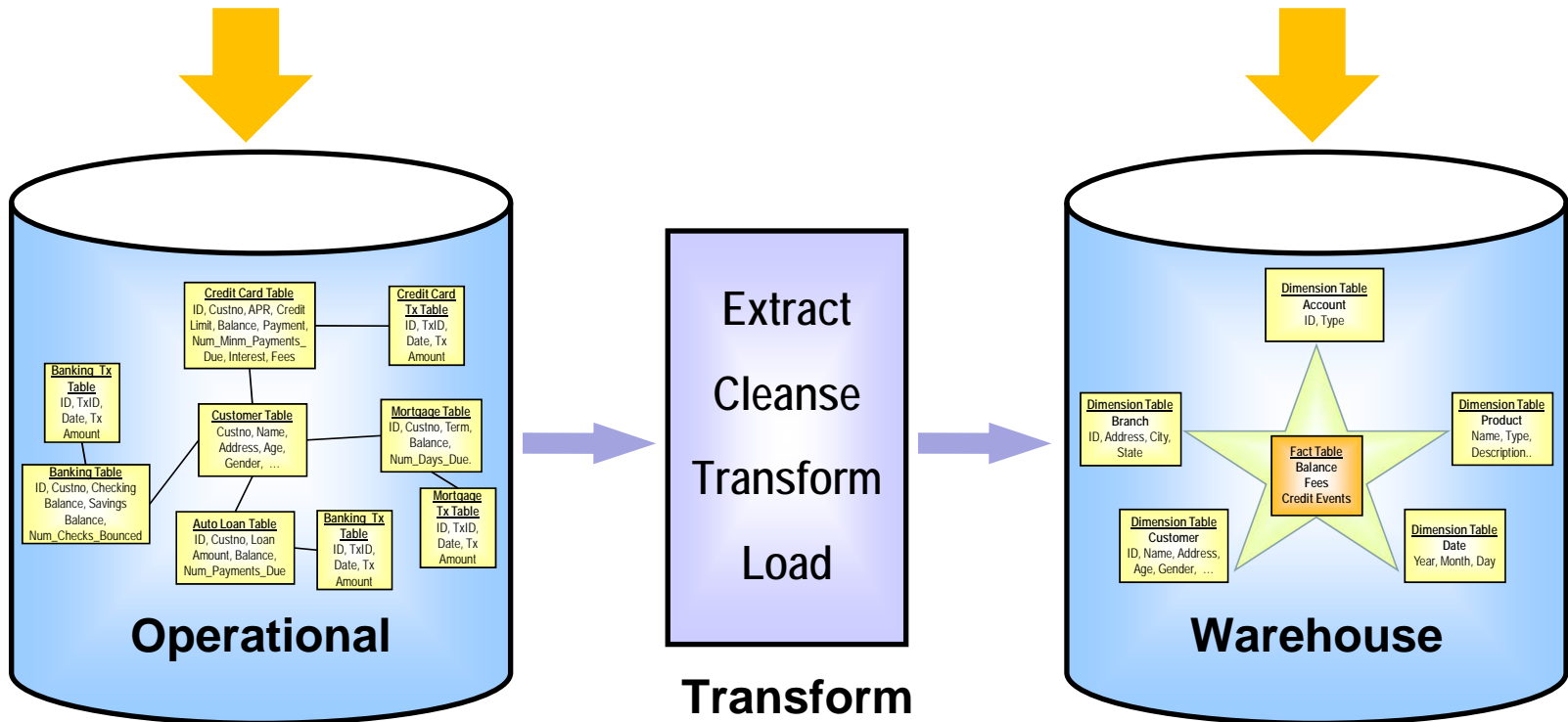Oracle e-business suite

$200 Million
increase in cash flow

The more analytics a business uses, the better it performs

There Are Two Approaches To Analytics

Operational Business Intelligence
Simple Queries
"How are we doing?"

Deep Analytics
Longer Queries
"Why?"

Operational

Extract

Cleanse

Transform

Load

Transform Schema

Warehouse

Schema organized to efficiently process simple queries
No Duplicated Data

Schema organized to efficiently answer longer queries
Data is Duplicated

Operational Business Intelligence Answers Simple Questions Quickly

- **Advantages**

  - ▶ Technologies designed for fast OLTP also work well for simple operational BI queries

  - ▶ Access most recent data with real-time queries

  - ▶ Does not require Extract-Cleanse-Transform-Load processes

  - ▶ No requirement to duplicate the data

- **Considerations**

  - ▶ Can impact OLTP performance

Demo: Use Dashboards To Answer The Question "How Are We Doing?"

- **Dashboards give quick access to information**
  - ▶ Present at-a-glance, high impact views of relevant scorecards and information
  - ▶ Use visual indicators such as gauges and dials
    - − Drill-down into details
  - ▶ No additional licensing requirements to create dashboards with Cognos BI
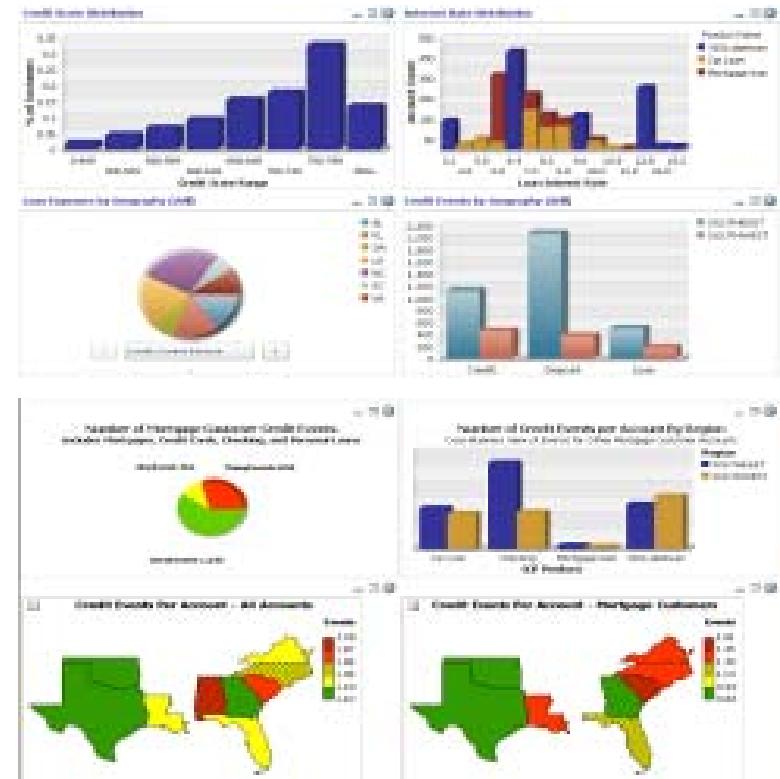
Deep Analytics Answers Deeper Questions

- Access historical data
- Use Extract-Cleanse-Transform-Load to unify customer information across OLTP silos into a Warehouse
- Queries are complex and long running
- Scan millions of records to compute aggregations
- Unique technologies to speed up complex queries

Operational Data Typically Stored In A Schema Designed For Fast Access To Small Amounts Of Data

SELECT Cust_Name, Monthly_Pymt FROM Customer, Mortgage WHERE Cust_No = '7928'

Customer Table				
Cust No	Cust Name	Cust Gender	Cust Age	….
7928	Herman Miller	M	29	…
8010	Hyong X Nolan	M	22	…
…	…	…	…	…

Mortgage Table				
ID	Cust No	Term	Monthly Pymt	….
LF236	7928	15	2263.09	…
LF318	8010	15	9432.96	…
…	…	…	…	…

Complete View Requires Joining Many Tables

Customer Table

Cust No	Cust Name	Cust Gender	Cust Age	….
7928	Herman Miller	M	29	…
8010	Hyong X Nolan	M	22	…
…	…	…	…	…

Mortgage Table

ID	Cust No	Term	Monthly Pymt	….
LF236	7928	15	2263.09	…
LF318	8010	15	9432.96	…
…	…	…	…	…

Car Loan Table

ID	Cust No	Term	Monthly Pymt	….
CL12	7928	35	463.05	…
CL38	8010	60	432.87	…
…	…	…	…	…

Credit Card Table

ID	Cust No	APR	Minm. Pymt	….
CC25	7928	12.99%	299	…
CC97	8010	5.99%	599	…
…	…	…	…	…

Data Warehouses Organize Data Differently For Efficient Comprehensive View

See All Activity For A Given Customer

| FACT Table						
CustNo	Acct_ID	Date_ID	Payment	Term	Credit_Events	
7928	LF236	365	2263.09	15	0	…
7928	CF12	365	463.05	35	3	….
7928	CC25	365	299	NULL	3	
…	…	…	…			…

DB2 For z/OS Is Optimized For Operational BI Queries

- Workload Management (WLM) optimizes resource sharing between OLTP and Analytical Queries

  - ▶ Minimize impact on OLTP performance

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

- Achieves Near-Linear Scaling And High Availability with Parallel Sysplex Clustering

- Provides Hash access path for even faster access to a single row of a table

- DB2 10 supports up to 10x more concurrent users and up to 20,000 concurrent connections per subsystem
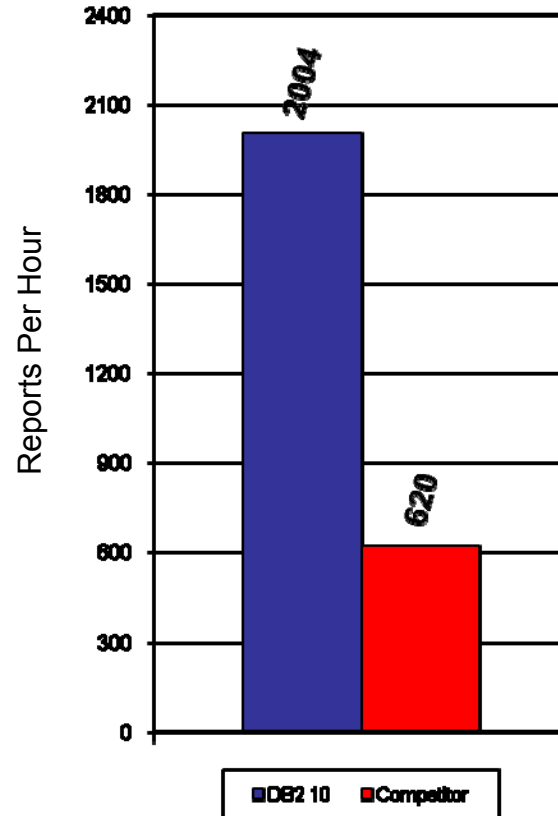
A Large US Financial Institution - Performance Comparison Of Operational BI
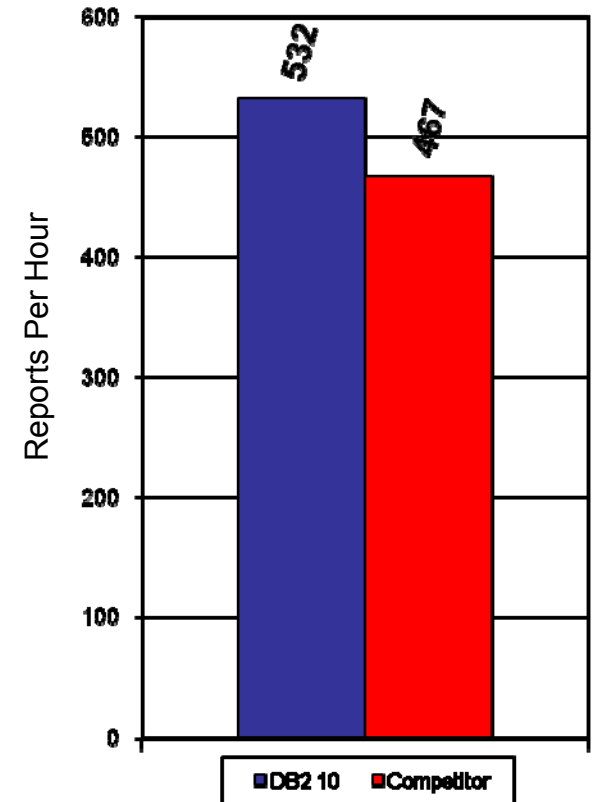
1.2X More Small Reports
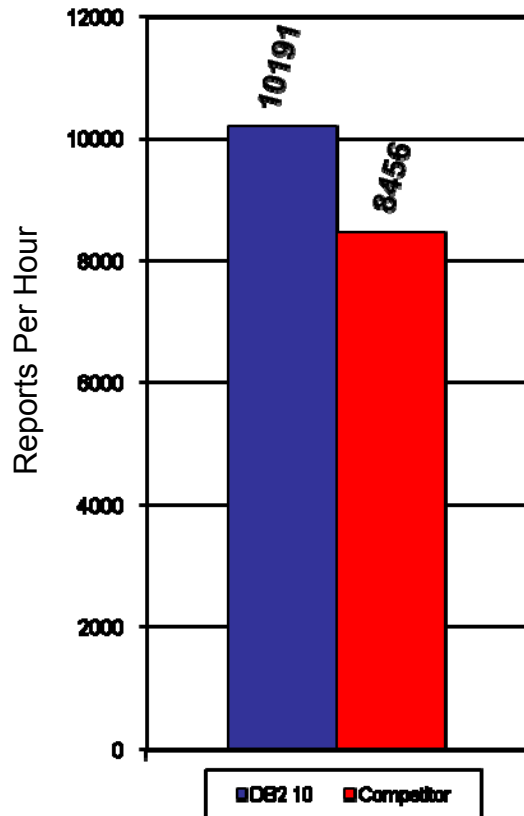
Reports Per Hour
at 750GB data size
(Higher is Better)

3.2X More Medium Reports

Reports Per Hour
at 750GB data size
(Higher is Better)

1.1X More Large Reports

Reports Per Hour
at 750GB data size
(Higher is Better)

DB2 For z/OS Is Also Optimized For Data Warehouse Queries

- Data Warehouse Workloads typically include a mix of simple, intermediate and complex queries

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

- DB2 for z/OS Cost Based Optimizer decides best execution plan for each query
  - ▶ Simple queries typically assigned to a single processing thread
  - ▶ Complex queries may be decomposed into operations that execute in parallel
  - ▶ Queries may be automatically rewritten to take advantage of pre-computed partial results in materialized query tables (MQT)

- Result: Optimum Throughput

Platform Capabilities Benefit Both Operation BI And Deep Analytics

- More processors, memory and cache than other enterprise servers provides high concurrency

- I/O offloaded to the Dedicated I/O Sub-system

- Parallel Sysplex clustering designed for near linear scaling

- Hardware compression

- Optimized resource sharing for mixed queries with WLM

- DS8000 delivers high storage bandwidth with caching

- Unmatched scalability

- Systematic Disaster Recovery

- Attractive pricing with IBM Smart Analytics Solution 9700

Add IBM DB2 Analytics Accelerator For Even More Optimization

- A workload-optimized, blade-based appliance with storage integrated into the hardware rack and based on Netezza Technology

- Pre-load data from DB2 for z/OS into IDAA at over 400GB/hr and update by partitions

- Deeply integrated with DB2 for z/OS and transparent to applications

- Significantly speeds up the response time for a wide variety of complex queries using massively parallel processing architecture and patented data filtering technology at streaming speed using Field Programmable Gate Arrays (FPGAs)

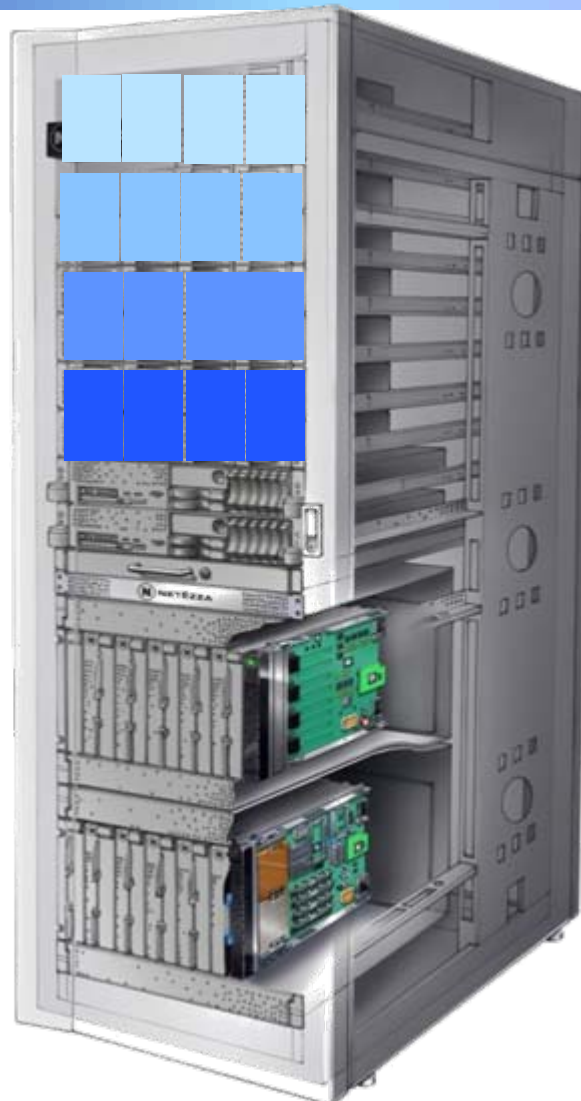- Drives down the costs of data warehousing and business analytics

Breakthrough Technology Enabling New Opportunities

IDAA Leverages Massively Parallel Processing To Speed Up Complex 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/FPGA

"…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

Outstanding Customer Results With IDAA

Query	DB2 (Secs)	DB2 + IDAA (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

Speed Ups Ranging From 13x to 1,908x

Actual customer results, October 2011

Customers Are Excited About Business Analytics On zEnterprise

DB2 Analytics Accelerator: "we had this up and running in days with queries that ran over 1000 times faster"

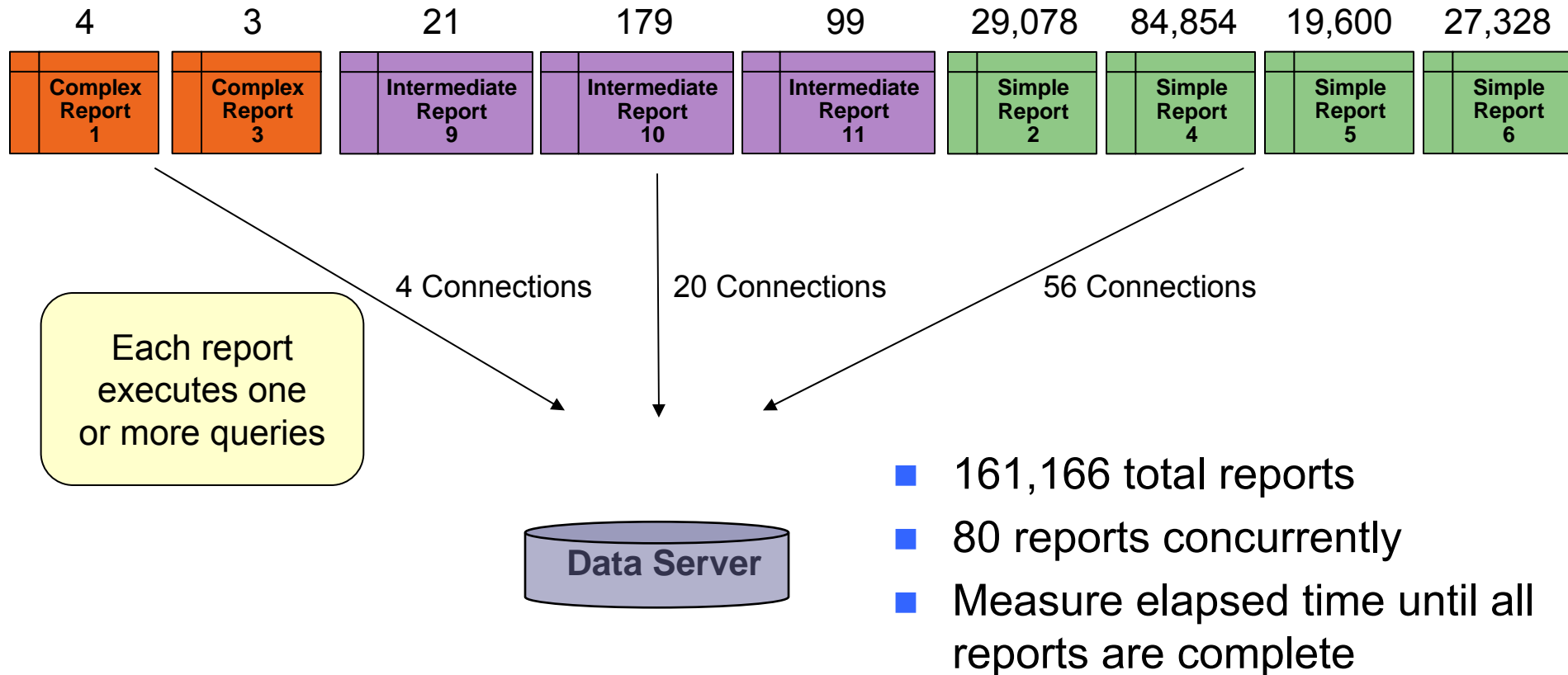**DB2 Analytics Accelerator: "we expect ROI in less than 4 months"**

Cognos BI for System z "We didn't have to justify a higher cost for putting this on the mainframe, *it was cheaper!*"

Analyst Claudia Imhoff "the industry pendulum in swinging towards centralization and there is no better platform than the mainframe"

Example: BI Day Fixed-Set Of Analytic Reports Run Concurrently On A Periodic Schedule

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

Each report executes one or more queries

Data Server

- 161,166 total reports
- 80 reports concurrently
- Measure elapsed time until all reports are complete

Note: Distribution of complex, intermediate, and simple workloads based on Forrester Research, Profiling the Analytic End User for Business Intelligence, 2004

BI Day Fixed-Set Analytic Reports Run Concurrently On A Periodic Schedule

- CASE 1:
  Run concurrently 160,860 simple reports and 306 intermediate and complex reports on Competitor ¼ Rack

- CASE 2:
  Run concurrently 160,860 simple reports and 306 intermediate and complex reports on IBM Smart Analytics System 9700 with dedicated 24-cores

- CASE 3:
  Run concurrently 160,860 simples reports on IBM Smart Analytics System 9700 10-cores and in parallel offload 306 intermediate and complex reports serially to IDAA (Twinfin 12)

IBM DB2 Analytics Accelerator Makes A Great Analytics Solution Even Better

Time To Completion For Fixed Size Concurrent Workload – 1 TB
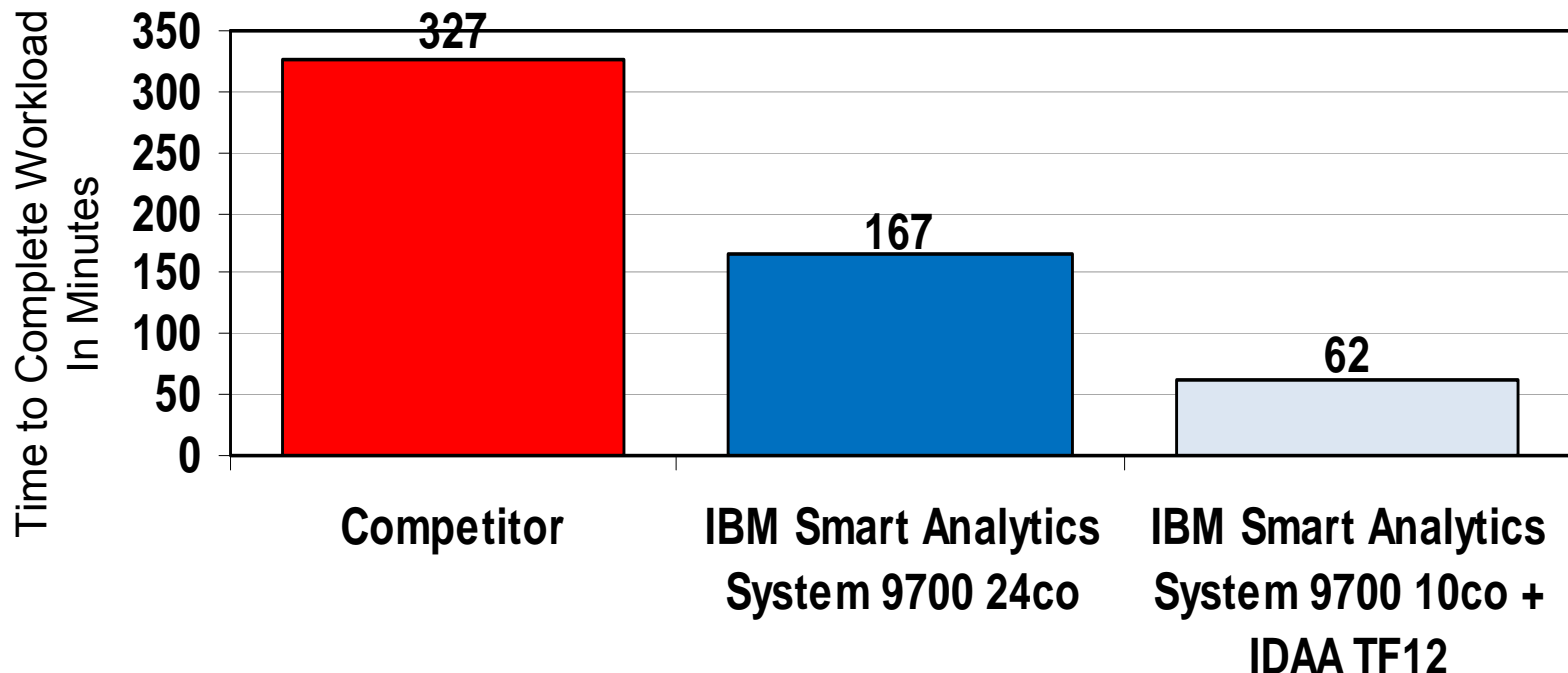

Lower is better

Test measures time to complete a fixed number of 161,166 concurrently executing reports

Based on BI Day Tests. Performance numbers may vary based on workload profiles. IBM Smart Analytics System + IDAA performance estimated from IBM Smart Analytics System running simple reports and NZ TF12 running intermediate + complex reports

Running Analytics On Optimized zEnterprise Platform Saves 75% Over Competition

Competitor

Quarter Rack
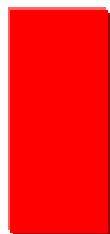
IBM Smart Analytics System 9700

| DB2 (IBM Smart Analytics System 9700) |
| z/OS 12 GP+12 zIIP |

IBM Smart Analytics System 9700 + IDAA

| DB2 (IBM Smart Analytics System 9700) |
| z/OS 5 GP+5 zIIP |

Netezza TwinFin 12

Unit Cost (3yr TCA) **$97/RpH**

RpH (Reports/Hour)	29,572
Competitor ¼ Rack (HW+SW+Storage)	$2,857,500

Unit Cost (3yr TCA) **$62/RpH**

RpH (Reports/Hour)	57,904
IBM Smart Analytics System 9700 24-co (HW+SW+Storage)	$3,600,000

Unit Cost (3yr TCA) **$24/RpH**

RpH (Reports/Hour)	154,893
IBM Smart Analytics System 9700 10-co (HW+SW+Storage)	$1,500,000
NZ TF12 (HW+SW+Storage)	$2,140,600

Based on IBM internal tests of Smart Analytics 9700 + IDAA solution compared to results of testing of a competitor's configuration (previous version; no longer available) executing an analytics workload in a controlled laboratory environment and a 3 year total cost of acquisition (based on US list prices). The cost calculation compares the average cost per report for 161,166 concurrently executing mixed complex, intermediate and simple report types. Intermediate/Complex reports offloaded to IDAA for serial execution. 9700+IDAA results are a projection based on actual data for simple reports on SA 9700 and complex/intermediate report times run on separate Netezza TwinFin 12. 3 year total cost of acquisition includes expected hardware, software, service & support. Results may not be typical and will vary 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.

5X performance
4X price performance

IBM Smart Analytics System 9700 – A Comprehensive Package For Business Analytics

- **Pre-configured appliance-like solution with aggressive solution edition pricing**
- **Hardware/OS**
  - ▶ IBM zEnterprise z196 technology
  - ▶ IBM System Storage DS8800 Intelligent Disk controller
    - – Large controller cache and 3 Tier disk offering
  - ▶ z/OS 1.12
- **Unique Software**
  - ▶ DB2 10 for z/OS
  - ▶ Cognos 10 BI (Linux on System z)
  - ▶ InfoSphere Warehouse (Linux on System z)
  - ▶ SPSS Modeler (Linux on System z)
- **Optional Components**
  - ▶ IBM DB2 Analytics Accelerator
  - ▶ Solid State drives, integrated within DS8800
    - – Easy Tier to identify and migrate "hot data" to SSD

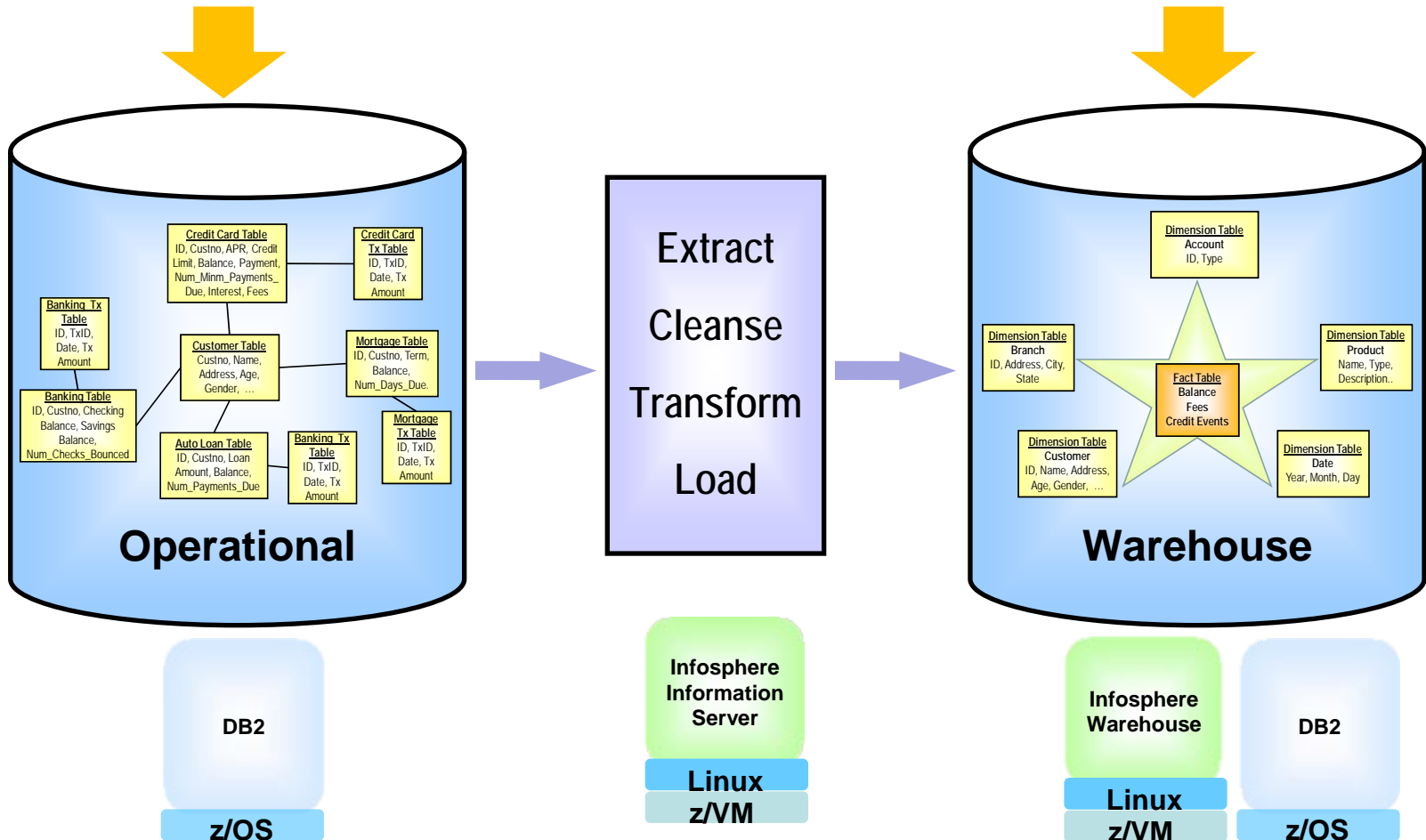| Infosphere Warehouse Cognos SPSS	DB2 for z/OS
z/VM	z/OS

Storage   Storage
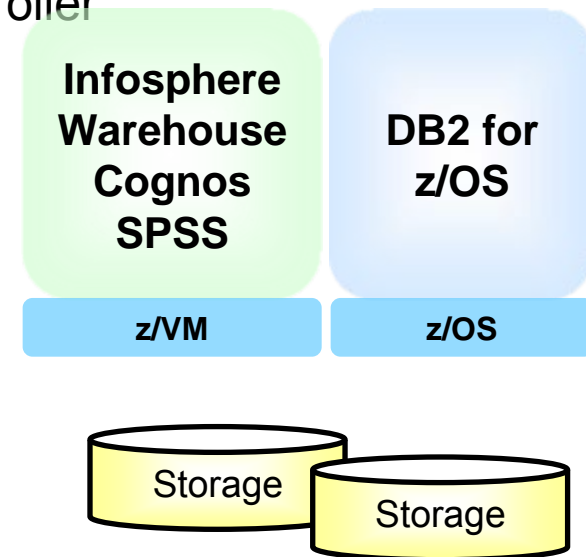
zEnterprise Now Provides All Of The Components You Need For A Complete Analytic Solution In One Platform

Operational Business Intelligence
Simple Queries
**How are we doing?**

Cognos
SPSS

Linux
z/VM

Deep Analytics
Longer Queries
Why?

Operational

Banking_Tx Table
ID, TxID, Date, Tx Amount

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

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

Extract

Cleanse

Transform

Load

Warehouse

Dimension Table
Account ID, Type

Dimension Table
Branch ID, Address, City, State

Dimension Table
Product Name, Type, Description..

Fact Table
Balance Fees Credit Events

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

Dimension Table
Date Year, Month, Day

DB2

z/OS

Infosphere Information Server

Linux
z/VM

Infosphere Warehouse

Linux
z/VM

DB2

z/OS

Cognos Can Generate Reports And Dashboards For Operational BI And Deep Analytics Queries
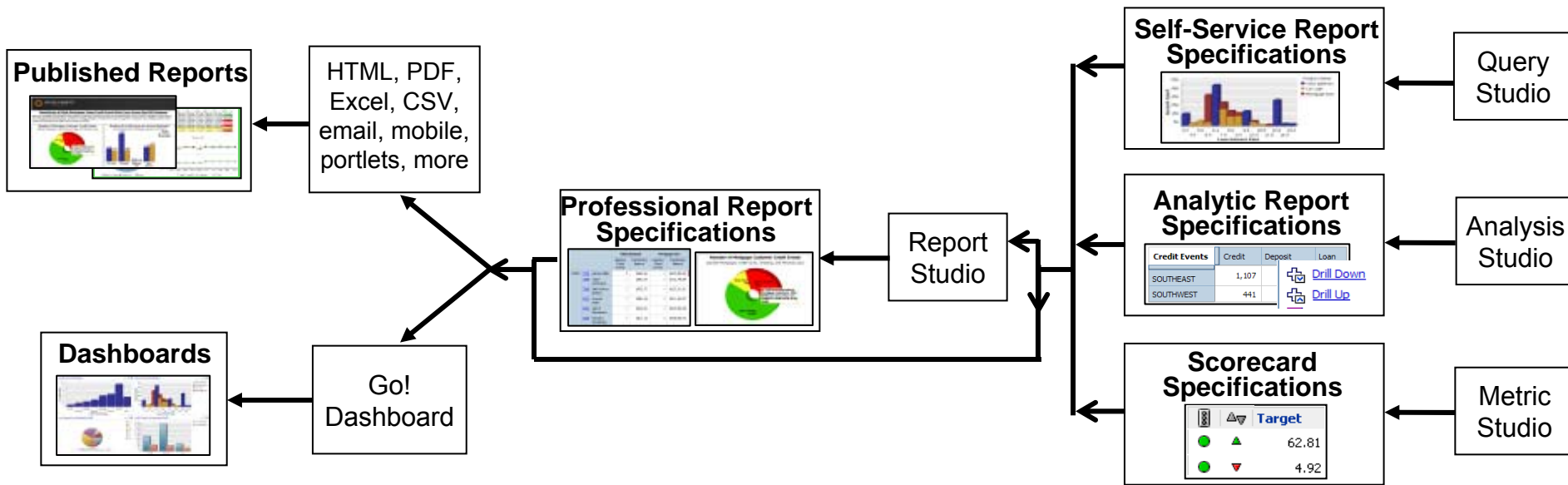
- **People-centric**
  - ▶ Server based business analytics accessed via browser
  - ▶ Consistent user interface for different analytic activities
  - ▶ Reuse new intelligence assets
  - ▶ Built-in collaboration and social networking
  - ▶ Threaded discussions, activities, and notifications

- **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

User Browser

IBM Cognos server

Query | Report | OLAP

Metadata Model

Operational   **Warehouse**
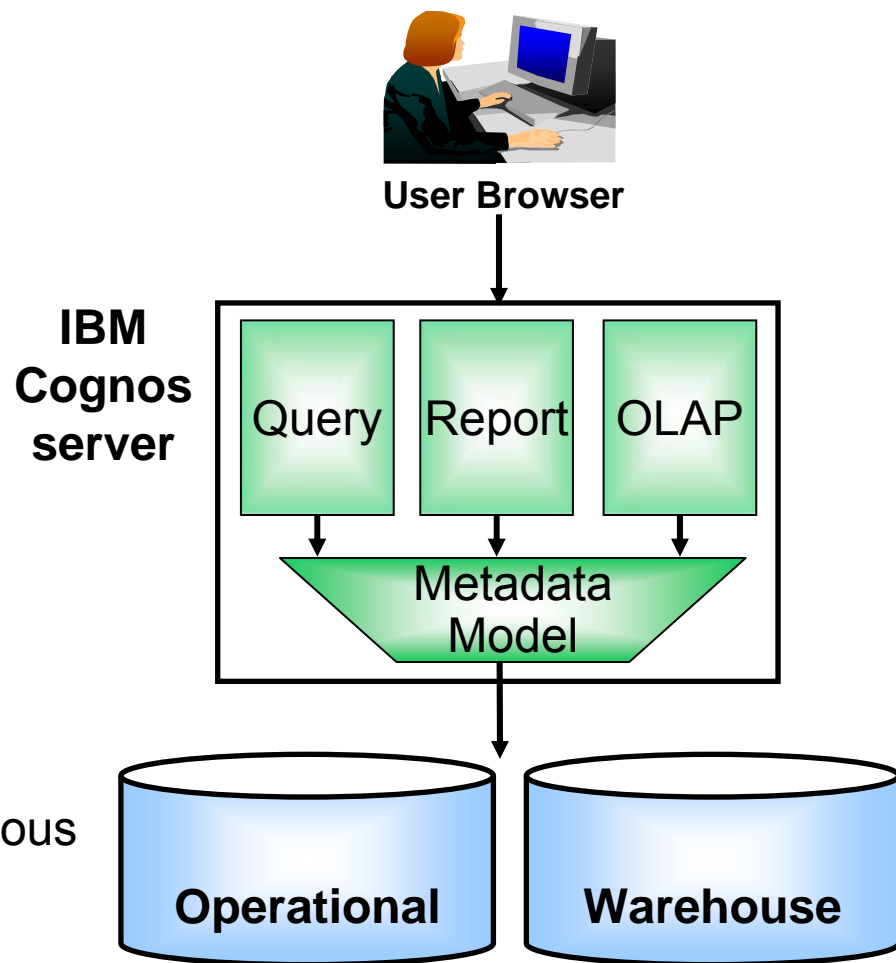
Reuse Prior Cognos Reports In New Deliverables

- Author once, consume anywhere

- All analytic assets share a common metadata model and a common multilingual report specification

- Ensures consistent information and enables reuse across platform functions

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…), identifies high-risk mortgages

3. Assess risk of flagged customer using FileNet case management data (Appraisal) linked to the Report

At risk customers are identified

Miami Dade County Runs Cognos On Linux On System z

Requirements:

- Demand for BI has really taken off
  - ▶ New Federal reporting requirements
  - ▶ Every new system, every new solution, every new application is having a business intelligence component
- Multiple Cognos 8 BI deployments
- Wanted an enterprise BI standardized solution, but
  - ▶ Needed higher capacity – grow from approx 400 to 1000 users
  - ▶ Do more with less - less researchers, less software, less hardware, same staff
  - ▶ Had available IFL's on System z

Results:

- 11 days to move from distributed to System z deployment model for Cognos 8 BI
  - ▶ Quickly and easily meet new requirements
- Consolidate multiple BI deployments on to a single platform
- Single point for BI administration
- Consolidate multiple disparate data sources
- Ensure 99.999% availability
- Offer a complete disaster recovery plan
- Additional green savings

SPSS Enables Customers To Predict Future Events And Drive Better Business Outcomes

Capture

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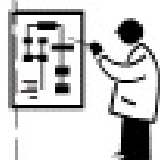
IBMSPSS Data Collection

Predict

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

IBM SPSS
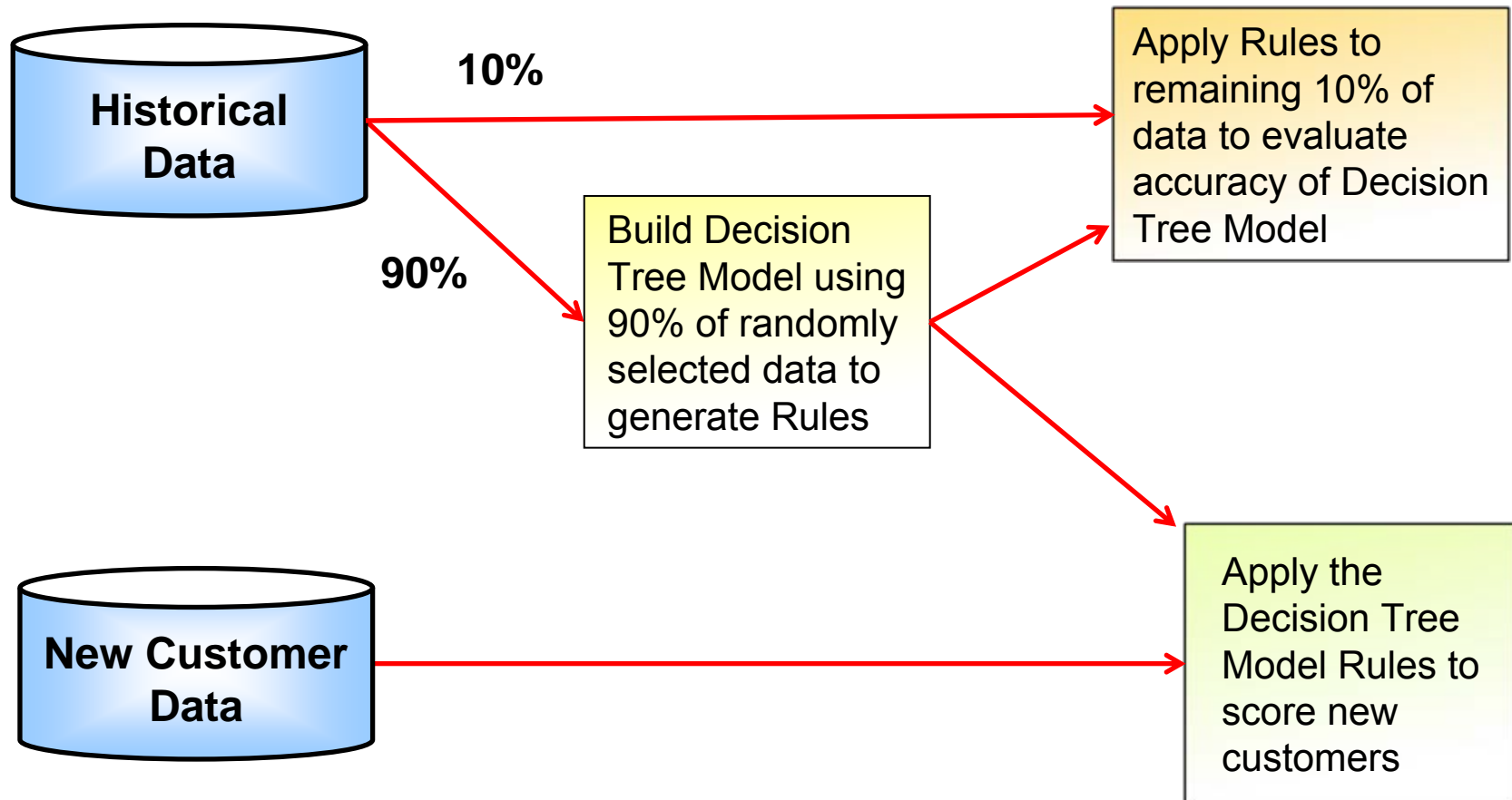Statistics*/Modeler*/Text Analytics

Act

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

IBM SPSS
Decision Management
Collaboration & Deployment *

Text Analytics — Data Mining — Statistics

Platform

Pre-built Content

Attract | Up-sell | Retain

* Runs on Linux on System z

What Can We Learn From Historical Data That Would Help Evaluate New Customers

Historical Data

10%

90%

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

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

New Customer Data

Apply the Decision Tree Model Rules to score new customers

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 using
• **RISKY** as dependent variable
• **GENDER, PROFESSION, MARITAL STATUS, STATE OF RESIDENCE, FICO_SCORE** and **TOTAL_BALANCE_OWED** as independent variables

→

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

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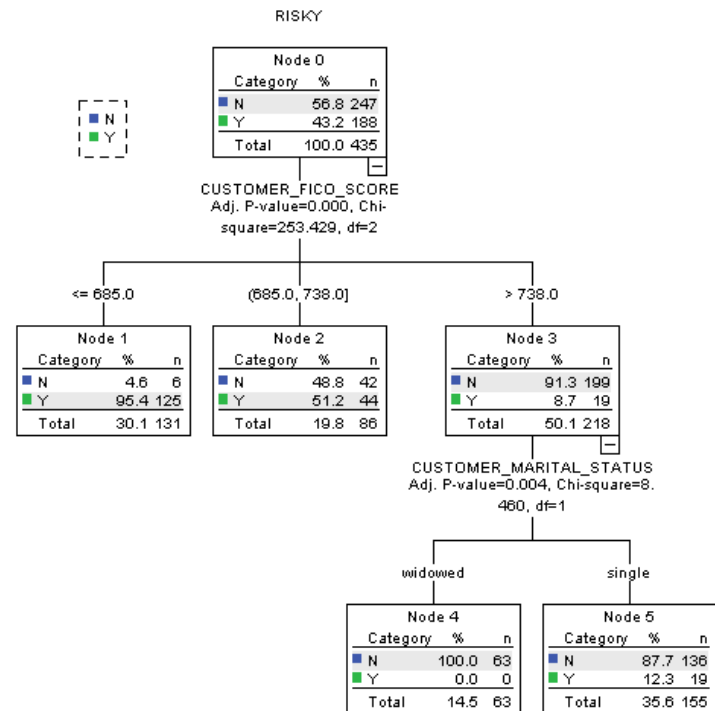
- Credit Limits identified for characterizing risky customers
- Use these credit limits for automated loan approval process

Run End-To-End BI On zEnterprise To Reduce Costs And Improve Reliability

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

- zEnterprise offers a fully integrated, optimized solution from operational data to business analytics in one platform

- Consolidating data warehouses and data marts on IBM Smart Analytics 9700 with IDAA can reduce costs by 75%