Business Analytics on zEnterprise
High Performance Analytics & Integrated Attached Co-processors

Carl Parris, parris@us.ibm.com
STSM – System z Performance, Design, Strategy

Bill Reeder, breeder@us.ibm.com
WW IT Optimization and Cloud System z Sales Leader

IBM Systems & Technology Group
March 2011
zEnterprise Solutions Can Support and Integrate Data Like No Other Platform, Providing a Foundation for Other Analytic and Application Capability

- The only platform that can run nine commercial databases, supported at the same time
- Better align and synchronize data, for data integrity. Use the internal architecture to consolidate database communications
- Leverage internal networking between databases and applications
- Centralize management across entire enterprise

- Consolidation of databases
- Tighter integration of data to applications
- Business intelligence close to the data
These workloads have recognizable patterns

**Core Applications**

<table>
<thead>
<tr>
<th>Database (z)</th>
<th>Application (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2® for z/OS, IMS™</td>
<td>CICS®, COBOL, WebSphere®</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WebSphere</td>
</tr>
</tbody>
</table>

**Multi-Tier Web Serving**

<table>
<thead>
<tr>
<th>Database (z)</th>
<th>Application (z)</th>
<th>Database (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2 for z/OS</td>
<td>DB2 for z/OS</td>
<td>DB2 for z/OS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application (x86)</th>
<th>Application (Power / UNIX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WebSphere, Apache / Tomcat</td>
<td>WebSphere, JBoss</td>
</tr>
</tbody>
</table>

**SAP**

<table>
<thead>
<tr>
<th>Database (z)</th>
<th>Application (Power)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2 for z/OS</td>
<td>AIX®</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application (x86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linux for x86</td>
</tr>
</tbody>
</table>

**Data Warehouse & Analytics**

**Master Data Management**

<table>
<thead>
<tr>
<th>Database (z)</th>
<th>Application (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2 for z/OS</td>
<td>WebSphere MDM (AIX, Linux on z)</td>
</tr>
</tbody>
</table>

**Analytics**

<table>
<thead>
<tr>
<th>Database (z)</th>
<th>Application (Power)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2 for z/OS</td>
<td>WebSphere, JBoss</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application (x86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Presentation (x86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WebSphere, Apache / Tomcat</td>
</tr>
</tbody>
</table>

© 2011 IBM Corporation
zEnterprise with a System z Blade Extension (zBX)

- **System zEnterprise CPC**
  - z/OS
  - z/TPF
  - z/VSE™
  - Linux
  - z/VM
  - System z PR/SM™
  - z HW Resources
  - Support Element

- **Application Server Blades**
  - Linux on System x™
  - POWER7
  - Blade Virtualization

- **Optimizers**
  - Future Offering
  - Future Offering
  - Future Offering
  - Future Offering

- **Blade HW Resources**
  - zBX

- **Private Data Network**
  - Ensemble Management Firmware

- **Customer Network**
  - Private Management Network
  - Private High Speed Data Network

*All statements regarding IBM’s plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of General Direction is at the relying party's sole risk and will not create liability or obligation for IBM.
Cloud Service Lifecycle Management

Subscribe to Service
- Request a service
- "Sign" Contract

Offer Service
- Register Services and Resources
- Add to Service Catalog

Service Creation
- Scope of Service
- SLAs
- Topologies, Best Practices Management Templates

Deploy Service
- Request Driven Provisioning
- Management Agents and Best Practices
- Application / Service On Boarding
- Self-service interface

Manage Operation of Service
- Visualize all aggregated information about situations and affected services
- Control operations and changes
- Event handling
- Automate activities to execute changes
- Include charge-back

Terminate Service
- Controlled Clean-up
Hybrid Schema Mainframe and HPA Accelerator

**Why Business Analytics on System z**
- Highest Frequency compute threads in industry z196
- Very good floating point performance z196
- Large Shared Resource Pool
  - Single point of resource management
  - Single point of operational control
  - Efficient use of underlying compute resources
  - Manage unpredictable loads between instances
  - Easy/fast provisioning
- Integration w/Commercial Business Processing
  - Security
  - Reliability
  - Availability
  - Auditing
  - Monetary Transactions

**Why Analytics on HPA Blade**
- Compute thread rich environment
- State of the art Vector/SIMD architecture

**Why Analytics on zGryphon**
- HPC enhanced commercial computing
- Single operational domain
  - Avoid standalone distributed cluster
  - Extend strengths of System z
Increasing prevalence of compute and data intensive parallel algorithms in commercial workloads driven by real time decision making requirements and industry wide limitations to increasing thread speed.  

Based on: Competing on Analytics, Davenport and Harris, 2007
Market Leading Business Intelligence & Analytics Software

Concentrating on this bit

- Cognos 10 BI
- Cognos Planning
- Cognos TM1

This is plumbing

- InfoSphere Information Server
- InfoSphere MDM Server
- InfoSphere MDM Server for PIM
- InfoSphere Foundation Tools
- Telco Data Warehouse & Other Industry Models
- Traceability Server

- Smart Analytics Systems

- Filenet P8
- ediscovery
- Content Manager
- InfoSphere Content Collector
- Records Management
- Content Integrator

- SPSS
- iLog

- Filenet BPM
- iLog

- DB2
- Informix
- IMS
- solidDB
- Optim
- DataStage
- Discovery
- Database tools
- InfoSphere Warehouse
- InfoSphere Streams
- Mashup Hub
- DB2 for z/OS

© 2011 IBM Corporation
Customers want to integrate analytics with Operational processes

New BI trends map well to core strengths of DB2 for z/OS and System z

Mixed workload performance - becoming single most important performance issue for DW/BI

Moving to a strongly centralized, shared infrastructure to improve economies of scale

New DB2 features, Cognos/SPSS/ILOG software offerings, new optimizations and improved solution packaging with ISAS/ISAO

Single view of enterprise, Continuous availability/DR, Security, Governance, Query prioritization

Virtualization and WLM enables consolidation of diverse DW and BI environments onto System z - zISAS

z196 performance w/ integrated zBX + technology providing new ways to integrate analytic solutions while managing costs – iSAO
Exploit Industry Trends that play to the strengths of System z

- Data Consolidation and creation of “Enterprise Database of Record”
- Operational Business Intelligence with z QOS requirements
- Operational trxs integrated with predictive analytics to provide additional insight

Leverage z Hybrid architecture, accelerators, multi-workload integration (zOS/zLinux)
System z Platform Direction: From Data hub to Analytics hub

- Exploit Industry Trends that play to the strengths of System z
  - Data Consolidation and creation of “Enterprise Database of Record”
  - BI/Analytics application consolidation and creation of enterprise single version of truth
  - Operational Business Intelligence with z QOS requirements
  - Operational trxs integrated with predictive analytics to provide additional insight
  - Superior end/end analytics life cycle integration
  - Analytics as a service in an internal or external cloud

- Leverage z Enterprise architecture, accelerators, multi-workload integration (zOS/zLinux)

---

Leverage System z Operational Data Store

- Operational Txnl data
- Operational Analytical Data

---

Data Integration

- Off Platform App
- Data
- Infosphere
- Infoserver
- InfoSphere Warehouse
- Websphere
- CICS/IMS
- SAP
- iFlex
- SPSS
- Rules
- ILOG....
- Scoring

---

Operational Data Store

- OLTP
- Cleanse Transform Warehouse
- Report
- Cognos
- Analyze
- SPSS
Business Analytics Life Cycle – Async and Distributed

Data Mining
Segmentation
Prediction
Statistical Analysis

Data Mover
Staging Area
Transformation Server

Enterprise Data Warehouse (RDBMS)

OLTP

Optimized Business Processes
- Customer Support
- Sales Effectiveness
- Claims Processing
- Fraud Management
- Underwriting
- Marketing

MIS System
- Budgeting
- Campaign management
- Financial Analysis
- Selling Platforms
- Customer Profit Analysis
- CRM

Analyze

Rules

Scoring

Operational Systems

OLTP

Analytics Server

Multi-Dimensional Analysis

Report

Cleanse Transform Warehouse

Operational Systems

Batch Process

Batch Process

Departmental Data Marts

Online Queries & Reporting
Business Objects & Web Intelligence

Analytics Server

x/p server

Bulk

x/p server

x/p server

x/p server

x/p server

x/p server

x/p server

x/p server

x/p server

x/p server

x/p server
Business Analytics Life Cycle – zEnteprise (IBM Smart Analytics System)

Operational Systems
- Optimized Business Processes
  - Customer Support
  - Sales Effectiveness
  - Fraud Management
  - Marketing

OLTP

Analyze

Data Mining
- Segmentation
- Prediction
- Statistical Analysis

OLTP

Report

Cleanse

Transform

Warehouse

Analytical Foresight

Business Analytics Server

LPAR 2: z/VM
- zLinux
- IBM Risk Calc.
- Risk Calc.

LPAR 1: z/OS (OLTP)
- CICS
- DB2 z/OS
- IMS

LPAR 3: z/OS (DW)
- ODS/DW/EDS/DM (DB2 z/OS)
- High Perf Accelerator (ISAO/Netezza)

Analytical Foresight

Predictive Analytics Server

Batch Process

Staging Area

Transformation Server

Data Mover

Staging Area

x/p zBX
- Bulk
- Analytics Server

zLinux

zLinux

Pass
Steam
Merge

IBM Risk Calc.

Classic Federation Server

Multi-Dimensional Analytics
Online Queries & Reporting
Web Intelligence

© 2011 IBM Corporation
Evolution of OLTP

- Real time ‘transactional’ analytics
  - Credit Card Fraud Detection
    - Compute intensive ‘neural network’ calculations required off-load to alternative hardware
    - Batch runs overnight – business imperative for real time response. POC w/ ACI/PRM using z/OS and HPC.
      > Latency costs of offload negated compute advantages of HPC
  - Optimized on-board floating point architecture would re-host this application on z/OS
    - Eliminate network latency delays
    - Add value to OLTP transaction
    - Huge savings potential the sooner the act of fraud is detected

Bulk Data Analytics

- Batch and near real time
  - Risk Analysis (IBM Treasury POC)
    - Multiple repositories of operational data
    - Sophisticated numerical algorithms
      > Bayesian probability algorithms
      > Monte Carlo simulation
  - Batch and near real time good match for host/accelerator offload
    - High performance accelerator HW building block
    - High speed bulk data transport
    - Efficient data cleansing/transformation engines – ETL
    - Value added proprietary data mining algorithms
    - Open standard host/accelerator programming model
Customers demanding real time decision making

Enable real time transactional analytics w/embedded SPSS/iLOG scoring/rules in IMS, CICS, WAS - z196’s industry highest frequency compute threads, competitive floating point performance

Differentiated data flow from operational to DW to Analytic repositories, Event driven modeling/scoring refresh, And….

SPSS/iLOG algorithms on z196 with integrated attached zBX co-procs using thread rich P7 vector archoptimized algs, modeling embedded in DB2

Deeper integration of Cognos, SPSS, iLOG into z196 ecosystem. Operational Data Store w/platform mgmt, high speed connectivity, acceleration enable the zEnterprise Analytics Hub

Data Currency

Compute Intensive Modeling and Optimized algorithms

Integration with core online business applications and data with shared infrastructure to improve economies of scale
Predictive Analytics Use Case Scenarios – US Credit Union Example

A. Higher withdrawal limits to increase customer satisfaction
   - Many Neighborhood Financial Centers, ATMS, Kiosks do not have service personnel to override withdrawal limits.
   - Need real time method of scoring member to determine appropriate limit while limiting risk.
   - Built a scoring model and embedded it in credit union’s daily transaction processing system to automatically determine withdrawal limits.
   - Saved staffing costs, increased customer satisfaction, retention, enabled increased revenue generation with reduced risk.

B. Targeted campaigns to improve retention, revenue
   - Exported member data from CU’s BI system, applied analytic techniques such as regression to create member profiles to predict likelihood members will need additional products/services.
     - e.g. Home equity line of credit.
   - Combined member usage characteristics w/ census information (i.e. local home ownership).
     - Filtered out 30-40% of unlikely candidates. Focused on 60-70% most likely to respond.
   - Increased ‘lifted’ revenue generated per marketing $$ by 60-100%.
   - Analysts wrote queries for rules to assist customer service. Recommendations pop-up on monitors during customer calls for relevant offers.

C. Grow customer base while risk shrinks
   - Attract new customers w/ prior financial problems.
   - Used scoring models to control deposit loss.
   - Boosted CU bottom line and benefited customers avoiding check cashing services and payday lenders.

D. Identify new branch locations
   - Created predictive model to help identify new branch locations, operate existing branches more profitably, close sites.
   - Factor and regression analysis to identify composite performance based on new customers, deposits, loan distributions.

Predictive Analytics enabled getting more mileage of data. Saved over $1M annually, increased revenue and improved member satisfaction.
## Analytic Functional Areas

<table>
<thead>
<tr>
<th>Area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Sell</td>
<td>Analysis and exploitation of hidden relationships in data about existing customer behavior to predict efficient future activity (purchase of products)</td>
</tr>
<tr>
<td>Direct Marketing</td>
<td>Analysis of customer characteristics (demographics, responses) to predict the amount of variability and tailoring of a marketing campaign</td>
</tr>
<tr>
<td>Collection Analytics</td>
<td>Analysis of customer characteristics to predict ability to pay and optimization of resources to facilitate collection.</td>
</tr>
<tr>
<td>Portfolio Prediction</td>
<td>Analysis of a portfolio of items (patients, products, financials, stores, etc.) to predict (score) a future outcome (survivability, placement, profitability, etc.)</td>
</tr>
<tr>
<td>Customer Retention</td>
<td>Analysis of a customers past characteristics to predict the likelihood of a customer’s future action.</td>
</tr>
<tr>
<td>Risk Analysis</td>
<td>Quantitative analysis to numerically determine the probabilities of various adverse events and the likely extent of losses if the event occurs</td>
</tr>
<tr>
<td>Fraud Detection</td>
<td>Analysis of transactions to predict the likelihood of fraud usually based on a score or probability.</td>
</tr>
</tbody>
</table>
## Mapping industry requirements to analytic functions

**Example: FSS (Banking and Insurance)**

<table>
<thead>
<tr>
<th>FSS Analytics Trends</th>
<th>Industry Requirements</th>
<th>Relevant Functional Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Banking</td>
<td>Customer Insight</td>
<td>Customer Retention, Cross-Sell, Direct Marketing</td>
</tr>
<tr>
<td></td>
<td>Product Recommendations</td>
<td>Customer Retention, Cross-Sell, Direct Marketing</td>
</tr>
<tr>
<td></td>
<td>Fraud Detection and Prevention</td>
<td>Fraud Detection, Risk Analysis, Collection Analytics</td>
</tr>
<tr>
<td></td>
<td>Underwriting</td>
<td>Risk Analysis</td>
</tr>
<tr>
<td>Payments</td>
<td>Fraud Detection and Prevention</td>
<td>Fraud Detection, Risk Analysis, Collection Analytics</td>
</tr>
<tr>
<td></td>
<td>Anti Money Laundering</td>
<td>Fraud Detection</td>
</tr>
<tr>
<td></td>
<td>Underwriting</td>
<td>Risk Analysis</td>
</tr>
<tr>
<td>Financial Markets</td>
<td>Fraud Detection and Prevention</td>
<td>Fraud Detection, Risk Analysis, Collection Analytics</td>
</tr>
<tr>
<td></td>
<td>Portfolio Analysis</td>
<td>Portfolio Prediction, Risk Analysis</td>
</tr>
<tr>
<td></td>
<td>Product Recommendations</td>
<td>Customer Retention, Cross-Sell, Direct Marketing</td>
</tr>
<tr>
<td>Insurance</td>
<td>Cause and Effect Analysis</td>
<td>Portfolio Prediction, Risk Analysis</td>
</tr>
<tr>
<td></td>
<td>Underwriting</td>
<td>Risk Analysis</td>
</tr>
<tr>
<td></td>
<td>Fraud Detection and Prevention</td>
<td>Fraud Detection, Risk Analysis, Collection Analytics</td>
</tr>
</tbody>
</table>
## Mapping Trends and Requirements to Analytical Function
### Retail Sector

<table>
<thead>
<tr>
<th><strong>Retail Trends in Analytics</strong></th>
<th><strong>Industry Requirements</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Optimization and Shelf Assortment</td>
<td>Merchandise Performance</td>
</tr>
<tr>
<td>Customer Driven Marketing</td>
<td>Customer Insight/Customer Churn</td>
</tr>
<tr>
<td>Fraud Detection and Prevention</td>
<td>Fraud Detection and Prevention</td>
</tr>
<tr>
<td>Integrated Forecasting</td>
<td>Merchandise Performance/Customer Insight</td>
</tr>
<tr>
<td>Localization and Clustering</td>
<td>Store and Channel Performance</td>
</tr>
<tr>
<td>Market Mix Modeling</td>
<td>Promotion Planning</td>
</tr>
<tr>
<td>Price Optimization</td>
<td>Merchandise Performance</td>
</tr>
<tr>
<td>Product Recommendation</td>
<td>Promotion Planning</td>
</tr>
<tr>
<td>Real Estate Optimization</td>
<td>Store and Channel Performance</td>
</tr>
<tr>
<td>Supply Chain Analytics</td>
<td>Supply Chain Optimizations</td>
</tr>
<tr>
<td>Workforce Efficiency Optimization</td>
<td>Store and Channel Performance</td>
</tr>
</tbody>
</table>
# Mapping Trends and Requirements to Analytical Function

## Telco Sector

<table>
<thead>
<tr>
<th>Telco Trends</th>
<th>Industry Requirements (from Sector Team)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Optimization</td>
<td>Customer Churn</td>
</tr>
<tr>
<td></td>
<td>Customer Retention</td>
</tr>
<tr>
<td></td>
<td>Product Cross Sell</td>
</tr>
<tr>
<td></td>
<td>Integrating Telco with retail sales</td>
</tr>
<tr>
<td></td>
<td>Social Networking Models</td>
</tr>
<tr>
<td></td>
<td>Behavioural Analytics</td>
</tr>
<tr>
<td>Network Analytics</td>
<td>Cell Tower Energy Management</td>
</tr>
<tr>
<td></td>
<td>Network Traffic Optimization</td>
</tr>
<tr>
<td></td>
<td>Capacity Planning</td>
</tr>
<tr>
<td>Revenue Assurance</td>
<td>Circuit Consolidation</td>
</tr>
<tr>
<td></td>
<td>Budget Forecasting</td>
</tr>
</tbody>
</table>
Mapping Trends and Requirements to Analytical Function
Healthcare Sector

<table>
<thead>
<tr>
<th>Healthcare Trends</th>
<th>Industry Requirements (from Sector Team)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Sciences</td>
<td>Gene Pool Analysis</td>
</tr>
<tr>
<td></td>
<td>Drug Discovery</td>
</tr>
<tr>
<td></td>
<td>BioInformatics</td>
</tr>
<tr>
<td>Healthcare Payer</td>
<td>Insurance Fraud</td>
</tr>
<tr>
<td></td>
<td>Clinical Cause and Effect</td>
</tr>
<tr>
<td></td>
<td>Medical Record Management analytics</td>
</tr>
<tr>
<td></td>
<td>Network Management analytics</td>
</tr>
<tr>
<td></td>
<td>Employer Group Analytics</td>
</tr>
<tr>
<td>Healthcare Provider</td>
<td>Executive Analytics</td>
</tr>
<tr>
<td></td>
<td>Patient Access</td>
</tr>
<tr>
<td></td>
<td>Clinical Resource</td>
</tr>
<tr>
<td></td>
<td>Patient Throughput</td>
</tr>
<tr>
<td></td>
<td>Quality &amp; Compliance</td>
</tr>
</tbody>
</table>
## Mapping Functional Areas to Tasks

<table>
<thead>
<tr>
<th>Function</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Sell</td>
<td>Association</td>
</tr>
<tr>
<td>Direct Marketing</td>
<td>Classification, Clustering, Association</td>
</tr>
<tr>
<td>Collection Analytics</td>
<td>Clustering, Association</td>
</tr>
<tr>
<td>Portfolio Prediction</td>
<td>Prediction</td>
</tr>
<tr>
<td>Customer Retention</td>
<td>Classification, Estimation</td>
</tr>
<tr>
<td>Risk Analysis</td>
<td>Classification, Clustering, Prediction</td>
</tr>
<tr>
<td>Fraud Detection</td>
<td>Anomaly Detection</td>
</tr>
</tbody>
</table>
## Mapping Tasks to Techniques/Algorithms

<table>
<thead>
<tr>
<th>Task</th>
<th>Technique/Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>Association Rules(Apriori), Decision Trees, Minimum Description Length</td>
</tr>
<tr>
<td>Classification</td>
<td>Decision Trees, Neural Net, Naïve Bayes, Support Vector Machines</td>
</tr>
<tr>
<td>Clustering</td>
<td>Clustering, Attribute Analysis, K-Nearest Neighbor</td>
</tr>
<tr>
<td>Estimation</td>
<td>Logistic, Regression, Discrete Choice Models</td>
</tr>
<tr>
<td>Prediction</td>
<td>Linear Time Series, Non-linear Time Series, Exponential Smoothing</td>
</tr>
<tr>
<td>Anomaly Detection</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>Procedure Family</td>
<td>Procedure</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>LINEAR</td>
<td>ALM</td>
</tr>
<tr>
<td>LINEAR</td>
<td>ANOVA</td>
</tr>
<tr>
<td>LINEAR</td>
<td>DISCRIMINANT</td>
</tr>
<tr>
<td>LINEAR</td>
<td>MEANS</td>
</tr>
<tr>
<td>LINEAR</td>
<td>ONEWAY</td>
</tr>
<tr>
<td>LINEAR</td>
<td>REGRESSION</td>
</tr>
<tr>
<td>LINEAR</td>
<td>T-TEST</td>
</tr>
<tr>
<td>LINEAR</td>
<td>UNIANOVA</td>
</tr>
<tr>
<td>LINEAR</td>
<td>GLM</td>
</tr>
<tr>
<td>LINEAR</td>
<td>2SLS</td>
</tr>
<tr>
<td>LINEAR</td>
<td>WLS</td>
</tr>
<tr>
<td>LINEAR</td>
<td>CSGLM</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>GLMM</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>PLUM</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>PLS</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>COXREG</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>GENLIN</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>GENLOG</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>HILOGLINEAR</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>LOGLINEAR</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>MIXED</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>VARCOMP</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>CNLR</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>LOGISTIC REGRESSION</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>NLR</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>NOMREG</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>PROBIT</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>CSCOXREG</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>CSLLOGISTIC</td>
</tr>
<tr>
<td>NON-LINEAR</td>
<td>CSORDINAL</td>
</tr>
</tbody>
</table>

DATA MINING

- Bayes Network
- NaiveBayes
- SVM
- MLP
- RBF
Categories of Optimization Problems Covered by ILOG Technology

<table>
<thead>
<tr>
<th>Mathematical Programming (NP-complete)</th>
<th>Continuous Optimization</th>
<th>Discrete Optimization (NP-hard)</th>
<th>Constraint Programming (Combinatorial Optimization)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linear Programming (LP)</strong></td>
<td><strong>Quadratic Programming (QP)</strong></td>
<td><strong>Mixed Integer Programming (MIP)</strong></td>
<td><strong>Vehicle Routing</strong></td>
</tr>
<tr>
<td>- linear objective function</td>
<td>- quadratic objective function</td>
<td>- one or more non-continuous variables</td>
<td></td>
</tr>
<tr>
<td>- linear constraints</td>
<td>- quadratic constraints</td>
<td>- includes MILP, MIQP, and MIQCP</td>
<td><strong>Job Scheduling</strong></td>
</tr>
<tr>
<td><strong>Quadratically Constrained Programming (QCP)</strong></td>
<td></td>
<td></td>
<td><strong>Custom Search</strong></td>
</tr>
<tr>
<td>- quadratic constraints</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Major iLOG Algorithms of Mathematical Optimization

- Optimisers
  - Simplex
    - Dual and primal simplex
    - Dual simplex is often the best choice
    - Problems where both dual and primal simplex perform poorly are rare
    - Research literature of running simplex on GPUs exists
  - Barrier
    - Suitable for large, sparse problems
    - The only optimizer for QCP problems
    - Parallel version available
  - Network
    - Suitable for network-flow problems
  - Sifting
    - Suitable for problems with large column/row ratios
    - Extension of simplex

- Search strategies
  - Branch and cut
    - Search tree with nodes being subproblems
    - Parallel version available
  - Dynamic search
    - A variation of branch and cut
Data Warehousing And OLTP Co-Located On zEnterprise

- Operational data moved to warehouse via ELT
- DB2 for z/OS centrally manages warehouse and data marts

ODS – Operational Data Store

- ISAO accelerates query execution
- Transparent to applications
Summary

Business Analytics exploits operational data to try to operate your business better.

Fully integrated solution: HPC + algorithms + transactions + data => insight

- Cognos, SPSS, ILOG, Infosphere WH with DB2/zOS provide the base for powerful new integrated Business Analytics Solutions with real time OLTP applications

- Emerging host/accelerator programming models will facilitate the ease of exploiting co-processors without specific accelerator architecture knowledge with cross-vendor portability

- zEnterprise with integrated attached co-processors provides a unified combination of scalability, aggressive single thread performance and Power based throughput computing threads and vector processing
Questions
SPSS Predictive Analytics Models Available on System z

SPSS on Linux for System z supports over 30 models,

- The 8 popular models support database push back for scoring in DB2 z/OS.
- 5 popular models now available listed below:

1. Logistic regression, Trees (Algorithm names Include CHAID, Quest, C&R Tree)
   - Finance-Used in banking to predict which customers are credit worthy. Which customers should I make a loan to?
   - Finance, Retail, Insurance, Entertainment-Used in marketing departments to determine which customers are going to respond to an offer
   - Insurance-Used in insurance to determine which claims are legit vs. Fraudulent
   - Telecommunication -Predicting customer churn

2. Cluster Analysis (Algorithm names Include K Means, Kohonen, Two Step)
   - Finance, Banking, Insurance -Used in marketing departments across industries to better understand customer segments
   - Customer attrition analysis

3. Market Basket Analysis (Algorithm name" Apriori)
   - Retail -Product assortment planning

4. Time series analysis/forecasting
   - Retail -forecasting catalog sales, forecasting demand, sales planning

5. Cox Regression
   - Retail, Telecommunications -Predicting the time for customer churn
   - Healthcare -determining the efficacy of a drug