Mainframe & Analytics: Two Best Friend

Hélène Lyon

Distinguished Engineer & CTO, z Analytics & IMS for Europe

IBM Systems, Software Sales, Europe

IMS Technical Symposium 2015

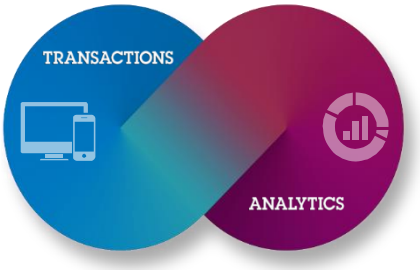
Agenda

- **Business Critical Analytics on z Systems**

- **A "proposed" split in four zAnalytics solutions areas**
  - Analytics for business applications
  - Analytics on "Transaction & Apps" Data and "IT" Data
  - Analytics on Reconciliation Zone
  - Reporting on Enterprise DW & DM

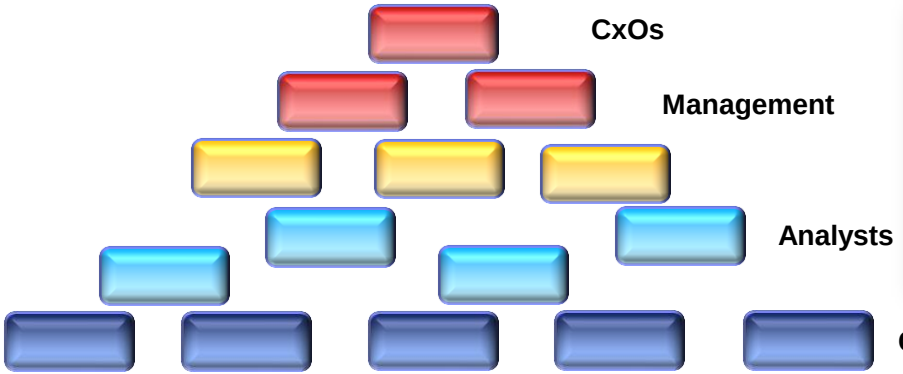
- **Focus on "Analytics for Business Applications"**
  - Business Critical Queries featuring IBM DB2 Analytics Accelerator
  - Real-time decision featuring SPSS Modeler, ODM & CPLEX
  - Hadoop & z integration featuring IBM BigInsights

Analytics is becoming the Keystone of every organization …

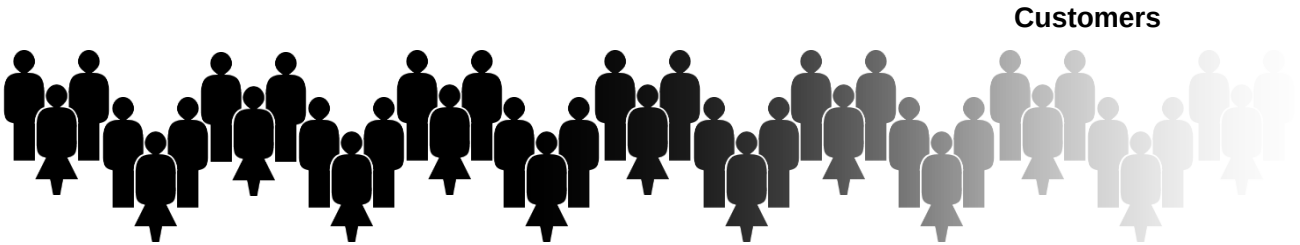


- **Analytics derive insight from data**
  - To help optimize business performance
  - To build new innovative services
  - To fight against fraud
  - To make all customer interaction personal!
  - …
- **Analytics become Business Critical!**
  - Analytics services are tightly integrated with business critical applications and data hosted on z/OS or copied from z/OS
  - Analytics is part of the flow of the business.
  - Decision processes have to be improved with new business insight derived from real time or near real time data.
  - Failure of these applications for any length of time can result in lost business or reputation.
- **Analytics are only as good as the underlying data foundation**
  - Data governance & Security & Performance

And more users across & beyond the organization want access to Analytics …



Increasing demands for analytics:
- More users
- More volume
- More complexity
- More critical to the business



… but IT remains aligned to the old way of doing business analytics.

- **Some reluctances from the past**
  - Core business is primary, analytics is secondary!
    - On core business side: High volume transactions and batch processing running concurrently, shared everything DB
    - On analytics side: Low volume complex queries and batch reporting, shared nothing DB
  - Cost of running analytics on z … without looking at all hidden costs concerning data movement – latency, data governance, IT complexity
  - Impact on operational performance & security
- **Key drivers to change IT perception**
  - Awareness of z position as primary Systems of Records
  - Technology availability to build a fully-integrated, end-to-end system that executes intelligent business processes
  - Recognized business value of advanced real time analytics
  - Business leaders brainstorm to identify to rethink business process with HTAP influence
    - Gartner: Hybrid Transaction/Analytical Processing Will Foster Opportunities for Dramatic Business Innovation (2014)

Business Critical Analytics Systems with IBM z Systems

An Hybrid Vision

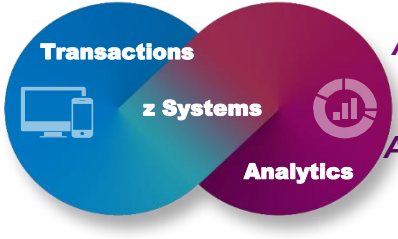
Minimize latency. Improve performance. Drive innovation.

Bring analytics to the data
- Reduced latency
- Reduced complexity
- Reduced cost

Deliver business critical analytics
- Timely, accurate, secure data
- Availability, scalability, performance
- Rapid deployment & expansion

Evolve with the business
- **Start with your top analytic requirement(s)**
- Grow without changing customer existing IT environment

Insights on ***every*** *transaction*

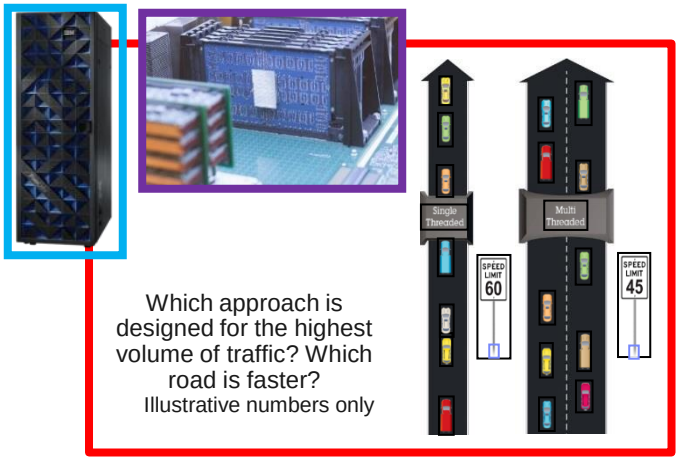


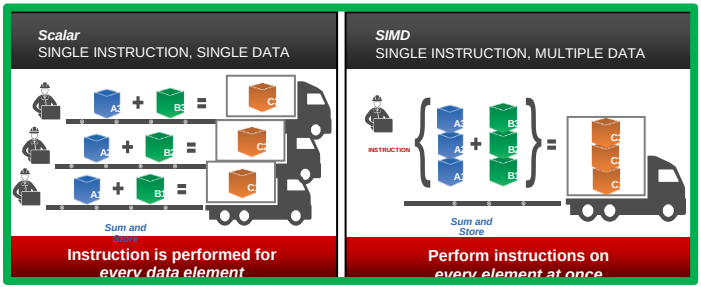
Analytics as part of the flow of business Analytics close to the data

IBM z13 designed for Analytics

Accelerate insight and simplify implementation

- **IBM DB2 Analytics Accelerator accelerates queries for faster insight**
  - New innovative use cases, such as in-database transformation and advanced predictive analytics
- **Large memory allows** new opportunities for in-memory computing
  - per system & per LPAR
- **SMT2 for increased zIIP & IFL cores capacity**
- **SIMD delivers accelerated analytics processing for complex queries**
  - Enable vector processing capabilities to z Systems
- **More**
  - Optimized math libraries and compilers that will speed up and simplify application development
  - Faster thread speeds
  - z Enterprise Data Compression (zEDC) to improve the economics of keeping data on z Systems





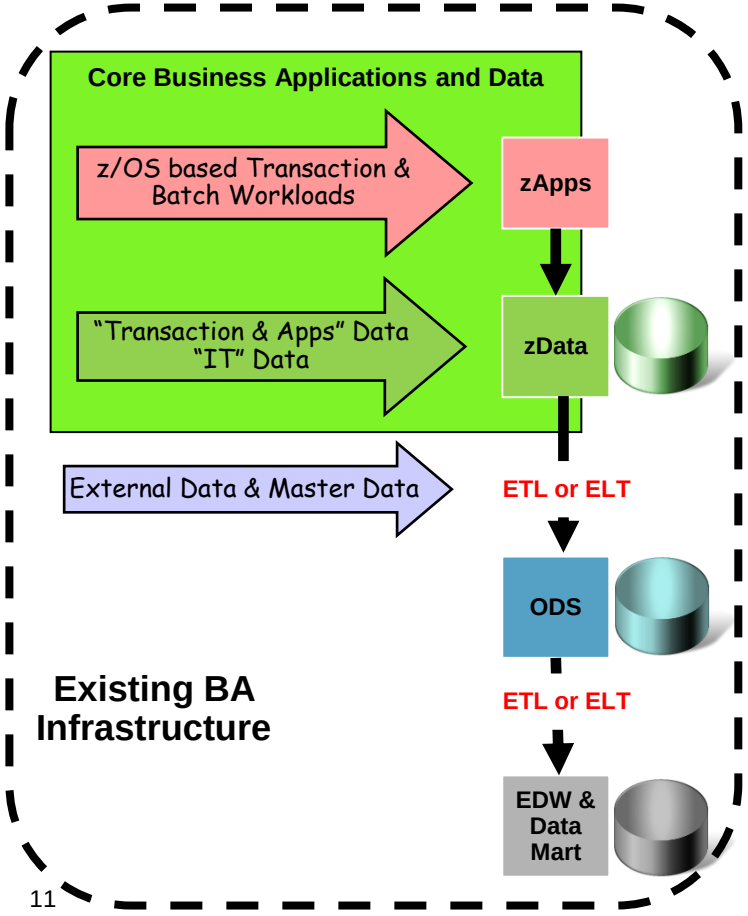
Agenda

- **Business Critical Analytics on z Systems**

- **A "proposed" split in four zAnalytics solutions areas**
  - Analytics for business applications
  - Analytics on "Transaction & Apps" Data and "IT" Data
  - Analytics on Reconciliation Zone
  - Reporting on Enterprise DW & DM

- **Focus on "Analytics for Business Applications"**
  - Business Critical Queries featuring IBM DB2 Analytics Accelerator
  - Real-time decision featuring SPSS Modeler, ODM & CPLEX
  - Hadoop & z integration featuring IBM BigInsights

4 z Systems Analytics Areas & associated "Buzz Words"



Business Critical Analytics

Add, Extend, Modernize Analytics Services

Analytics for Business Applications
business critical queries predictive analytics
scoring decision management
Real time analytics Hadoop integration
Optimization CPLEX ….
A

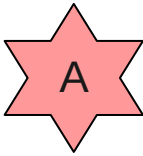
Analytics on "Transaction & Apps" Data
Ad-hoc Queries on zData Operational Reporting
Historical data non-DB2 data IMS VSAM
B

Analytics on "IT" Data
Hadoop on z machine data Log analytics
Capacity Management Software Cost Analyzis

Analytics on Reconciliation Zone
Single version of truth ETL simplification
data lakes Hadoop
C

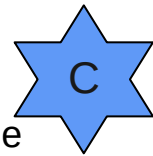
Reporting on Enterprise DW & DM
business intelligence data mining
ETL simplification predictive analytics
descriptive analytics Business reporti
D

z Systems Analytics Solutions Areas

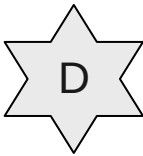


- **Analytics for Business Applications**
  - Just ReThink business applications with Hybrid Transaction/ Analytical Processing influence
  - Bring analytics in z/OS based applications and don't limit your thinking about what are analytics! You can really do things you couldn't do before!

- **Analytics on "Transaction & Apps" Data and "IT" Data**
  - Bring analytics to the data – Don't extend what has been done in the past.
  - Make it easily available for reporting tools, distributed applications or ERPs
  - Don't miss the Hadoop trend – Make it just relevant with zData without moving data off z



- **Analytics on Reconciliation Zone**
  - Harmonize & merge structured data and non-structured data
  - Includes Operational Data Zone or Exploration, Landing & Archive Zone from Watson Foundation



- **Reporting on Enterprise DW & DM**
  - Enhance and generalize business reporting
  - Provide a way to improve data governance

Agenda

- **Business Critical Analytics on z Systems**

- **A "proposed" split in four zAnalytics solutions areas**
  - Analytics for business applications
  - Analytics on "Transaction & Apps" Data and "IT" Data
  - Analytics on Reconciliation Zone
  - Reporting on Enterprise DW & DM

- **Focus on "Analytics for Business Applications"**
  - Business Critical Queries featuring IBM DB2 Analytics Accelerator
  - Real-time decision featuring SPSS Modeler, ODM & CPLEX
  - Hadoop & z integration featuring IBM BigInsights

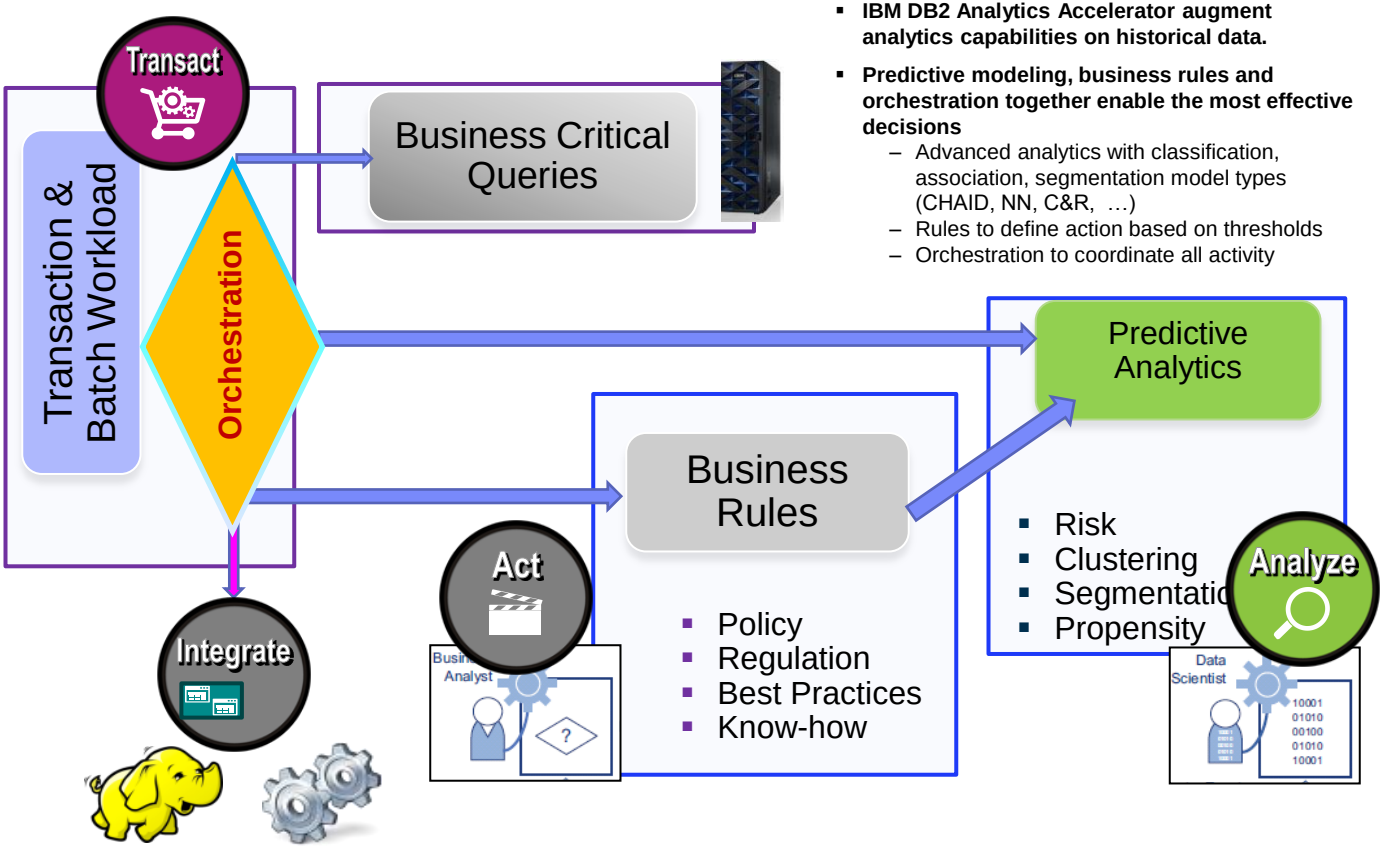
A – Analytics for Business Applications

- **Architects need new ways to ReThink business applications with HTAP influence**
- **IT architects need to accept mandatory changes in IT infrastructure to accommodate the new analytics needs without being afraid of any potential "bad impact" on production! It's not an option anymore!**

IT needs to support Business	Solution on z
Integrate business critical queries on demand in zApps • On DB2 and on non-DB2 data • On zApps data and historical data	• Use IBM DB2 Analytics Accelerator to accelerate business critical queries without extensive tuning needs. Performance improvements and cost reduction while retaining z/OS security and reliability • IDAA Loader provides optimized integration of non-DB2 data sources (from z/OS and distributed) into IDAA and DB2 concurrently
Move predictive analytics into transaction processing – "What could happen?" • **From Batch scoring to real-time scoring!**	• Direct integration in zApps of the SPSS Real-Time Scoring Adapter executed in DB2 for z/OS • Execution of the predictive model within a transaction, in-DB, and in real time
Automate real time decisions at the point of interaction with the customer!	• Use IBM Operational Decision Manager on z/OS • An alternative to rules imbedded in apps, in documentation, in people head ;) • Colocation brings performance! Externalization brings agility, traceability and auditability
Implement prescriptive analytics on demand by business - "What should we do?"	• Use IBM Decision Optimization solutions aka CPLEX, implemented directly on z/OS
Augment analysis thanks to non-z data like social media, machine generated data, e-mail stored in Hadoop clusters	• Use InfoSphere BigInsights on Linux on z Systems or on distributed • Combine SQL queries with Hadoop search inside zApps • Get Hadoop search results directly in DB2 for z/OS with DB2 V11

A – Analytics for Business Applications

- **IBM DB2 Analytics Accelerator augment analytics capabilities on historical data.**
- **Predictive modeling, business rules and orchestration together enable the most effective decisions**
  - Advanced analytics with classification, association, segmentation model types (CHAID, NN, C&R, …)
  - Rules to define action based on thresholds
  - Orchestration to coordinate all activity



Business Critical Queries on demand in zApps

- **Challenges**
  - IT refuses complex queries lasting too long because of impact on production workload.
  - DB administrator needs time and tools to tune the queries
  - So queries where executed on DW data and not in real-time on operational data at the time of transactions.

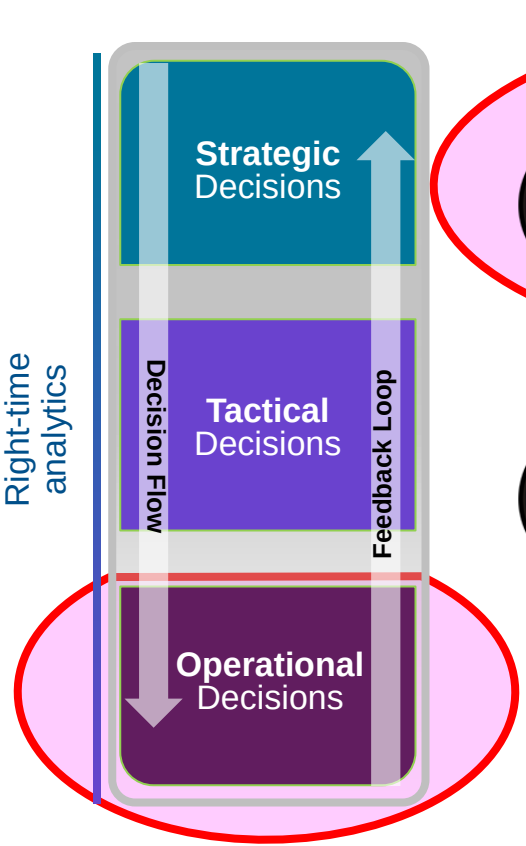
- **Solutions**
  - IBM DB2 Analytics Accelerator
    - Complex queries are now acceptable!
    - IDAA can be used as High Performance Storage Saver. No data in DB2 for z/OS!
    - IDAA Loader has several options for data refresh.
    - IDAA can be used for non-DB2 data (IMS DB, VSAM, …)
  - Workload management in z/OS
    - The optimized Mixed Workload environment
    - Performance goals assigned by workload based on SLAs

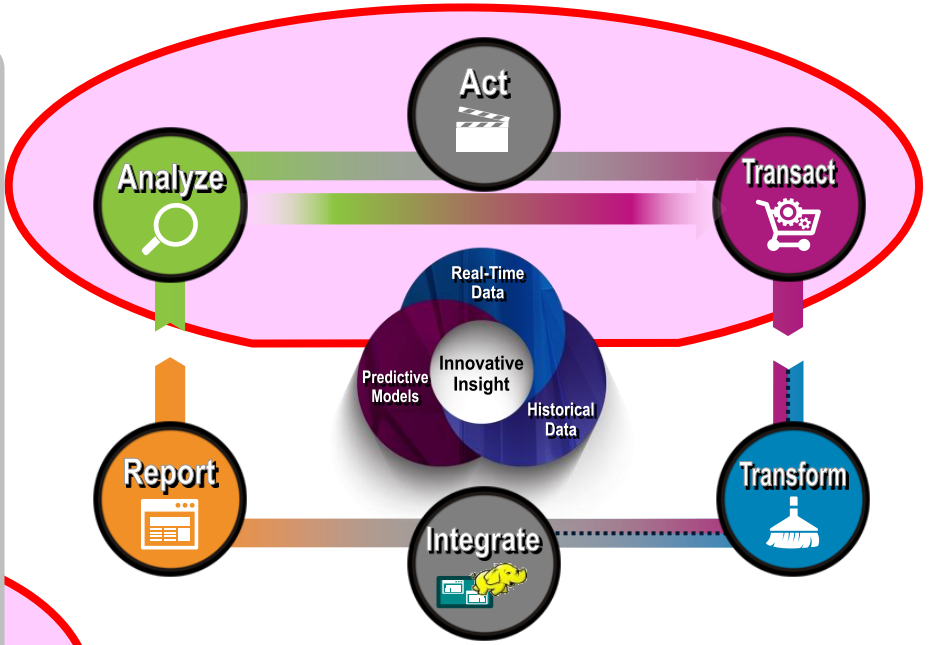
IBM DB2 Analytics Accelerator

Do things you could never do before!

- **What is it?**
  - The IBM DB2 Analytics Accelerator is a workload optimized, appliance add-on to DB2 for z/OS to drive exceptional business value
- **What does it do?**
  - Accelerates complex queries, up to 2000x faster
  - Lowers the cost of storing, managing and processing historical data
  - Minimizes latency
  - Reduces storage capacity requirements
  - Improves security and governance
  - Reduces operational costs and risk
  - Complements existing investments

Inject insights into Operational Decisions





Consistent, accurate view of data for uniform decision results

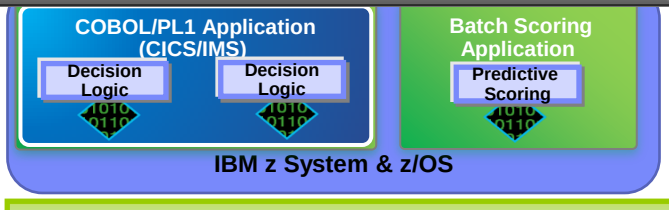
Richness of capability to support every decision type

Range of responsiveness for time-to-value from days to msecs

Real-Time Decision - Challenges

- **Enabling real time analytics for mission critical business application aims to move predictive analytics into transaction processing and automate real time decisions at the point of interaction with the customer.**
- **Mainframe customers might be struggling with the following challenges according to their context:**

Their decision logic is still imbedded in the CICS/IMS applications and the predictive scoring is managed as a batch process (not real-time scoring)



- Rules written in COBOL or PL1 and cannot be read by business people
- Hard coded rules are difficult to change
- Rules intertwined within applications cannot be reused by other systems
- Lack of decisions ownership for business stakeholders
- No real-time scoring

The business rules and the predictive scoring have been externalized to a distributed environment and both are called from z/OS through web service calls .

- Potential performance issues:
  - ✓ Latency due to the network communication between the z/OS and distributed environments
  - ✓ Call out web services from CICS/IMS
- Continuous Availability of the overall solution might be fairly complex to achieve
- Confidentiality of the data in transit might be challenging to guarantee

Real-Time Decision - Solutions

- **Solutions**
  - IBM DB2 Analytics Accelerator
    - Fast access to historical data
  - SPSS Real-Time Scoring Adapter
    - Execution of the predictive model within a transaction, in-DB, and with real time data as input.
    - Support of PMML language (within limitations)
  - IBM Operational Decision Management
    - Rules execution server on z/OS next to the transactional or batch systems.
    - No need to call distributed services
- **Industry Solutions Focus**
  - Counter Fraud Management – Detection phase
  - Predictive Customer Insight
  - …

- **Video on Decision Management on System z – IBM Client Center Montpellier**

- **Customer POC**
  - Bank in Turkey
    - Predictive Model created by SAS and deployed in SPSS Real-Time Scoring Adapter
- **Customer Reference**
  - FNB, South-Africa
    - Easy integration of decision management with mainframe Apps
    - Rules sharing between distributed and z

Business View – All Industries need a Competitive weapon!

Business Rules and Predictive Analytics Are Complementary

- **Real-Time Decision enable organizations to automate, optimize and govern repeatable business decisions to improve the value of customer, partner and internal interactions.**
  - Learn from the experts: Author a rule-based model capturing expert knowledge
  - Learn from the facts: Build automatically a predictive model by self learning from data
- **Real-Time Decision is now a competitive weapon with a rapidly expanding range of uses.**
  - Credit scoring: Usage of information from the loan applications to predict the risk of taking a loss and accept or reject the loan consequently.
  - Customer relationships: Customer characteristics and behavior are strongly predictive of attrition. Attrition or “churn” models help companies set strategies to reduce churn rates via communications and special offers proposed by business rules.
  - Fraud detection: Many types of fraud have predictable patterns and can be identified using statistical models for the purpose of prevention or for after-the-fact investigation and recovery.

Business View - Fraud Management

Clients and .. "non-clients"

Benefits from Integration with Transactions

Analysts

Investigators

Channel → Monetary Transactions → Fraud Detection → Fraud Analyst → Fraud Investigator

Trusteer

Prevent malware and potential fraud at the gate, before transactions can even be initiated

Bank's Existing Systems

Generates financial service transactions

Orchestrate

Rules

Stops known bad transactions, allows known good and takes appropriate actions on scored TX.

Predictive Models

Scores TX that passed the list filtering, so that rules can take appropriate action depending on the scoring.

Case Handling

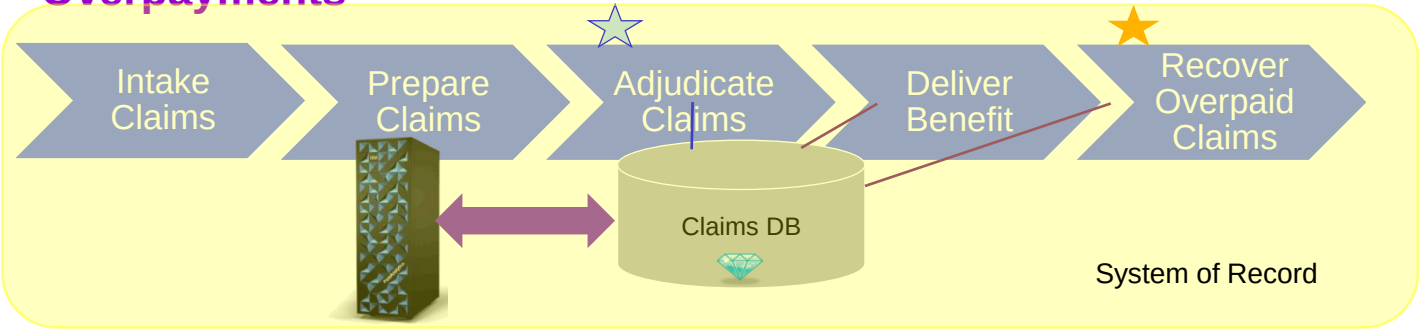
Provides Operator Uis for deciding on whether to let a flagged TX through, or to start an investigation.

Investigative & Other tools

Tools to help investigators working with a case, to understand the specific transaction context and help build evidence.

Counter Fraud Management scope

Business View – Insurance: Minimize Loss from Claims Overpayments



Challenge 1: Complex reports for overpaid claims not completing on time, result is monetary losses

Solution: Integrated optimized analytics of IDAA with overpayment reporting transactions

Benefits:

- Up to 2000x improvement in speed of overpayment reports
- LOB users enabled to respond with more agility to overpayment trends
- Informed decisions at the right time

Challenge 2: Stop improper payments *prior* to payment, avoid pay & chase, meet SLAs

Solution: Integrate predictive analytics into claims adjudication for analytics in place

Benefits:

- Very efficient scale for analytics
- Scale requirements only achievable with analytics as part of transaction flow
- Expected results of efficient in-transaction analytics is multi-million dollars per year

Prescriptive Analytics - Optimization and Simulation

Math Programming
CPLEX Optimizers
(Simplex, Barrier, Mixed Integer)

Constraint Programming
Constraint-based scheduling
CPLEX CP Optimizers

- **Answering the Question: "What should we do?"**
- **Challenges**
  - Mathematical algorithms were implemented by specific LOBs for their own needs
  - Performance on z could be seen as an issue
- **Solutions**
  - CPLEX Optimization Engine from IBM Decision Optimization on z Systems
    - z/OS as well as Linux for z Systems
    - Based on C routines
    - Possible integration in zApps thru a simple API

Taking advantage of the SIMD (Single Instruction Multiple Data) accelerator on z13, IBM CPLEX Optimizer on z/OS 12.6.1 8-way Barrier demonstrates up to 75% higher throughput than IBM CPLEX Optimizer 12.6.0.1 8-way Barrier running on zEC12.

- **Customer - A Large Central Bank in Europe**
  - CPLEX used as core technology for an optimized settlement system
  - Business Results
    - Settling more trades at lower cost will increase liquidity and capital flow.
    - Allowing the bank to respond more quickly to new constraints as legislation and customer behavior changes.
    - Freeing up hundreds of millions of euro worth of collateral used to back up trades.

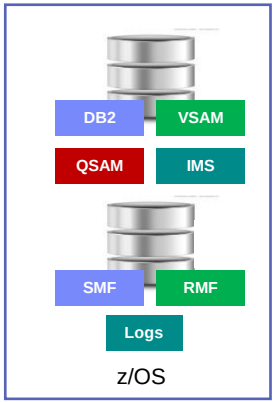
The volume of trades is expected to be high (over a million transactions), and finding the best set of trades to execute each night in a short time window is an extremely challenging technical problem.

Business View – Finance: Optimization Problems

- **Classic Applications**
  - Portfolio Optimization
  - Trade Matching and Timing
  - Asset-Liability Management
  - Cash Management

- **Novel Applications**
  - Loan Configuration and Lending
  - Derivatives Pricing
  - Workforce scheduling/dispatch
  - Ad scheduling
  - Targeted Marketing
  - Collateral management
  - Trade Settlement – Netting
  - Intra-day liquidity for wholesale bank operations

Hadoop & z Systems Integration – 2 Use cases

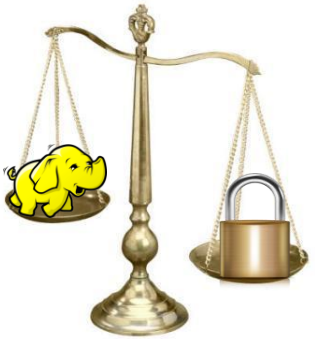


1 - Mainframe clients want to incorporate sensitive mainframe data into exploratory analytic models

What has been holding them back?

There is risk associated with having copies of sensitive data existing outside the mainframe

C



2 - Mainframe clients want to incorporate into zApps analytics based on non-z data like social media, machine generated data, e-mail

What has been holding them back?

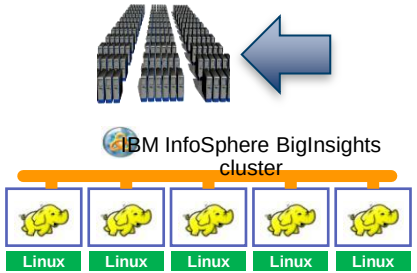
Performance & Integration are key inhibitors for real-time analytics.

A

Hadoop & z Systems Integration – Use case 2: "Augmented Analysis"

- **Very large amounts of non-relational data originate outside of z Systems**
  - e.g. e-mails sent by customers, tweets, posts to company Facebook page
- **Analyze sentiments and identify customers who are dissatisfied with company**
  - Words 'cancel', 'terminate', 'switch' or synonyms thereof
  - Names of competitors
- **Gather names and e-mail addresses of customers at risk**
- **Join these results with operational data**
  - Alert agents of at-risk customers
  - Agents work with customer and offer a promotion to stave off defection

All of the operational applications & data are here



There is also potentially relevant data here

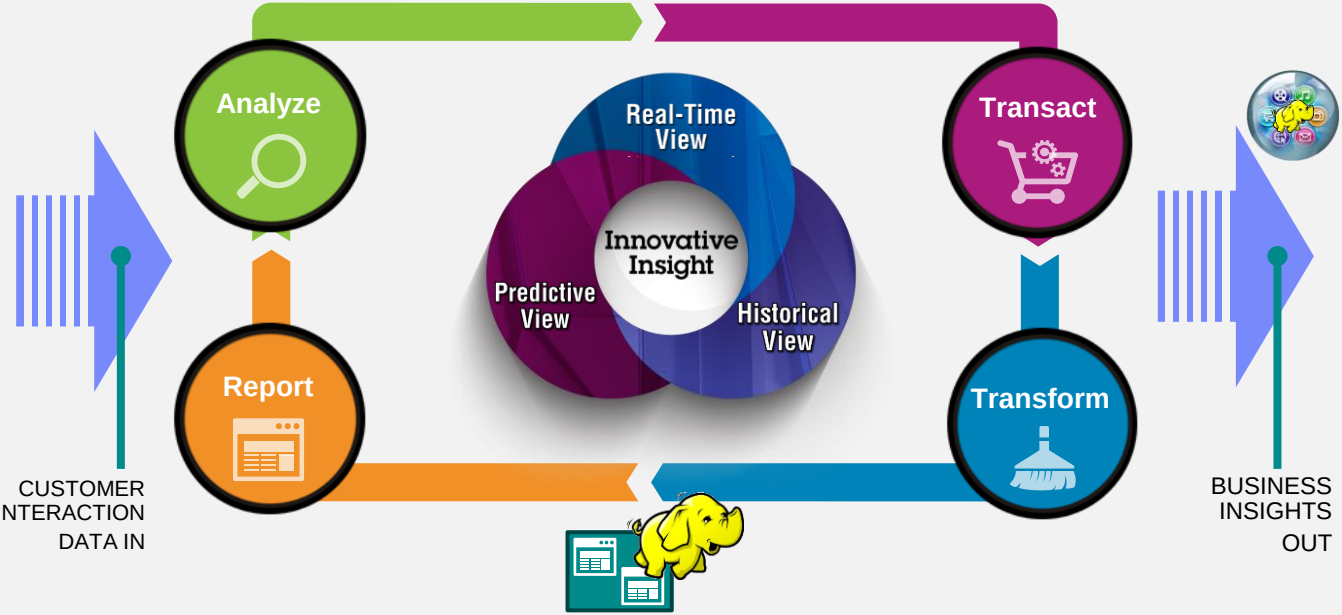
Agenda

- **Business Critical Analytics on z Systems**
- **A "proposed" split in four zAnalytics solutions areas**
  - Analytics for business applications
  - Analytics on "Transaction & Apps" Data and "IT" Data
  - Analytics on Reconciliation Zone
  - Reporting on Enterprise DW & DM
- **Focus on "Analytics for Business Applications"**
  - Business Critical Queries featuring IBM DB2 Analytics Accelerator
  - Real-time decision featuring SPSS Modeler, ODM & CPLEX
  - Hadoop & z integration featuring IBM BigInsights

Does that Help?
Are z & Analytics Best Friends?

A Complete Workload-Optimized System

Integration of operations and business-critical analytics into one streamlined, end-to-end data lifecycle



Better business response

Reduced data movement, reduced complexity, reduced configuration resources

More accurate, more secure, more available