Build a system health check for Db2 using IBM Machine Learning for z/OS

Jonathan Sloan
Senior Analytics Architect, IBM Analytics

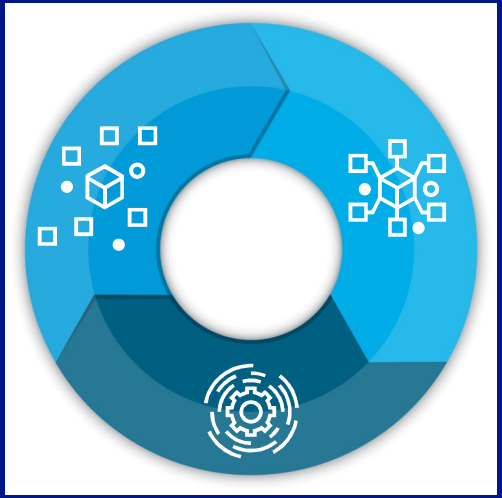
IBM | IBM z Analytics

Agenda

- A brief machine learning overview
- The Db2 ITOA model solutions template
- Customer experiences
- Db2 AI for z/OS
- Wrap up / Q&A

UNDERSTANDING
machine learning

Identify Patterns
not readily foreseen by humans



Score or Predict
Behavior with the deployment models

Build Models
of behavior from those patterns

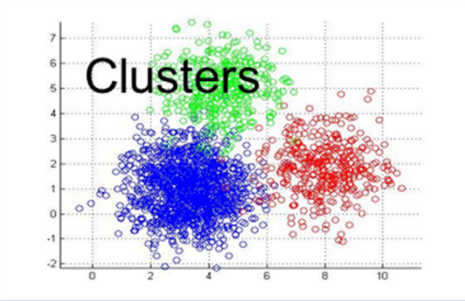
Traditional Programming

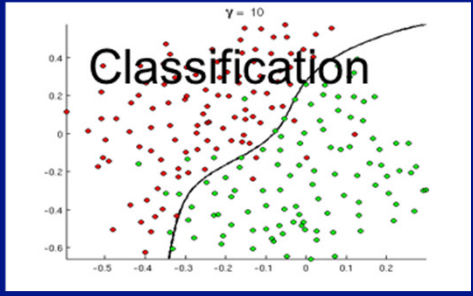
Machine Learning

The value machine learning to IT Operations

- Apply the same analytics technologies used in business for IT
  - Leverage latest reporting technologies
  - Achieve high speed reporting for SMF data
  - Leverage advanced analytics
- Improve operational productivity via IT Operational Analytics (ITOA)
  - Have overall view of the mainframe ecosystem over time
  - Monitor KPIs and System Resources
  - Achieve and maintain the highest levels of resiliency to meet business goals and prevent interruption of services

COMMON APPROACHES
to machine learning





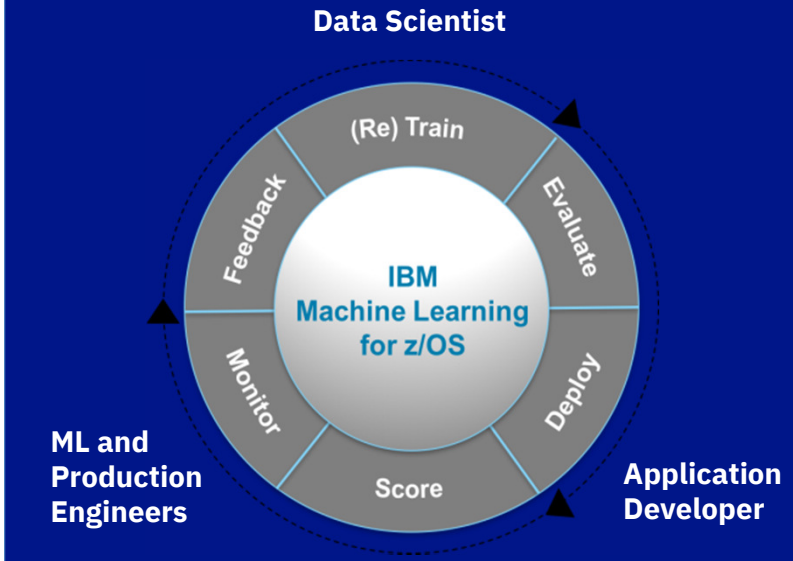
Unsupervised (Investigative) outcomes not labeled

- Clustering: Goal is to group data into clusters for better organization
  - Example: Categorize banking customers by behavior in order to understand how to market and what products to sell

Supervised (Predictive) outcomes are labeled

- Classification: Goal is to predict a category
  - Binary-classification (yes/no)
    - Examples: Fraud, Churn, Purchase, Spam email detection
  - Multi-classification (which of several items to recommend)
    - Examples: Netflix, Amazon recommendations, Ad recommendations for products
- Regression: Goal is to predict a value
    - Examples: Customer lifetime value, Stock prices prediction

IBM MACHINE LEARNING FOR z/OS



A Hybrid Cloud Approach to Model Lifecycle Management and Collaboration

- Platform agnostic model development
- Enterprise-grade, collaborative, extensible open source software
- Real-time insights embedded with transactions
- Insight from multiple platforms
- Reliability, availability and encryption in a security-rich environment

IBM Db2 IT Operational Analytics (Db2 ITOA) solution template

- Provides an example system health tree application
  - Applies machine learning technology to operational data, such as SMF records
  - Gains insight into the health of IBM Db2 for z/OS subsystems
  - Build dynamic base lines for Db2 for z/OS key performance indicators (KPI)
  - Visualizes and monitors the health status of Db2 for z/OS subsystems
  - Can provide real-time IT operational insight

IBM **z Analytics**

Db2 ITOA solution template

What it does

- Leverages machine learning and data science
- Ingests SMF data for model training and scoring
- Analyzes, monitors, and visualizes large amount of operational data
  - Builds a hierarchy health tree to represent the health status of the Db2 sub-systems, transactions and individual KPIs
  - Monitors the changes in health status over time
- Highlights abnormal KPIs in a timeline to assist root cause diagnosis
- Uses ML for z/OS functionalities to provide module life cycle management
- Provides real-time scoring capability by adopting SMF real-time interface

Db2 ITOA solution template
KPI selection

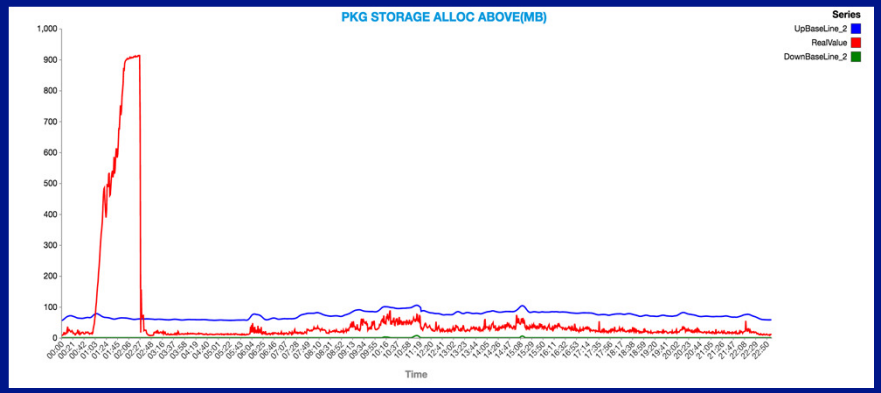
I believe elapse time and CPU time are important. I always start from them to check system issue.

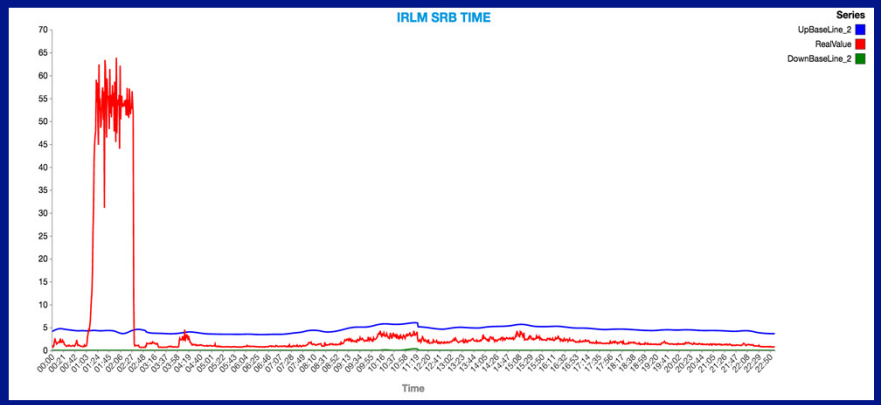
Correlation Matrix	CL2_ELAP_NONNESTED	CL2_NOT_ACCOUNT	CL3_TOT_LOCK_LATCH_DB2_IRLM	CL3_GLOBAL_CONTENTION
CL2_ELAP_NONNESTED	1.00	0.95	0.31	0.32
CL2_NOT_ACCOUNT	0.95	1.00	0.01	0.00
CL3_TOT_LOCK_LATCH_DB2_IRLM	0.31	0.01	1.00	0.84
CL3_GLOBAL_CONTENTION	0.32	0.00	0.84	1.00

- KPIs are selected from the SMF 100,101,102 records for Db2 subsystem and transactions
- Default collection is based on the input from domain experts
- Reduce the system cost and complexity of the health tree structure
  - Filter KPIs based on the study of their correlation
  - Cluster (group) KPIs together based on similar behavior
  - Select appropriate KPIs for monitoring

Db2 ITOA solution template

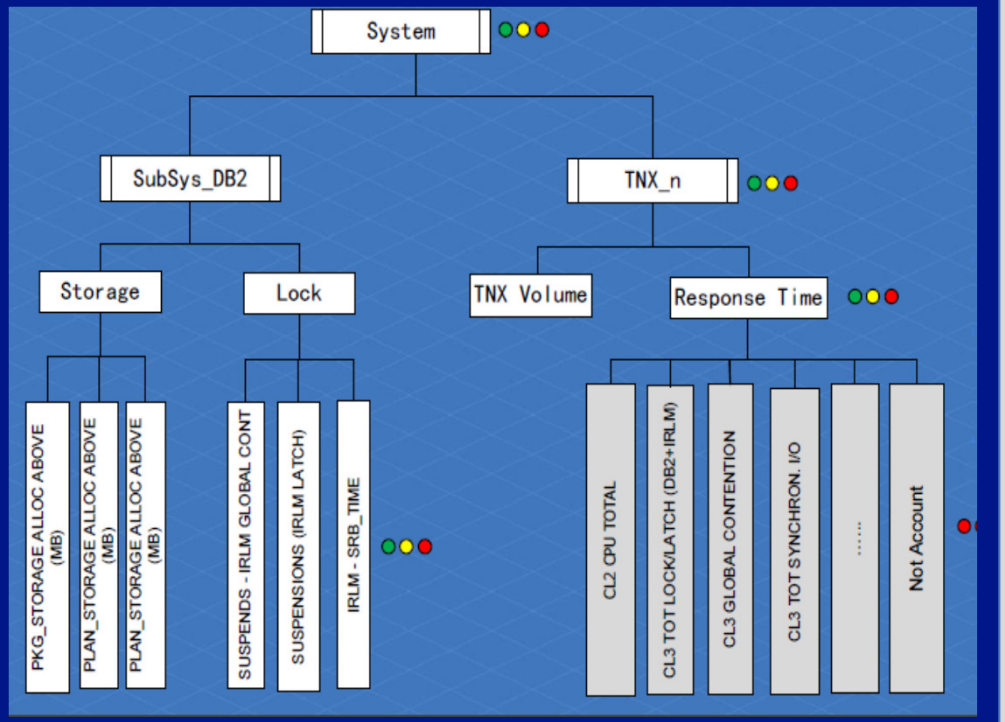
Building baseline model





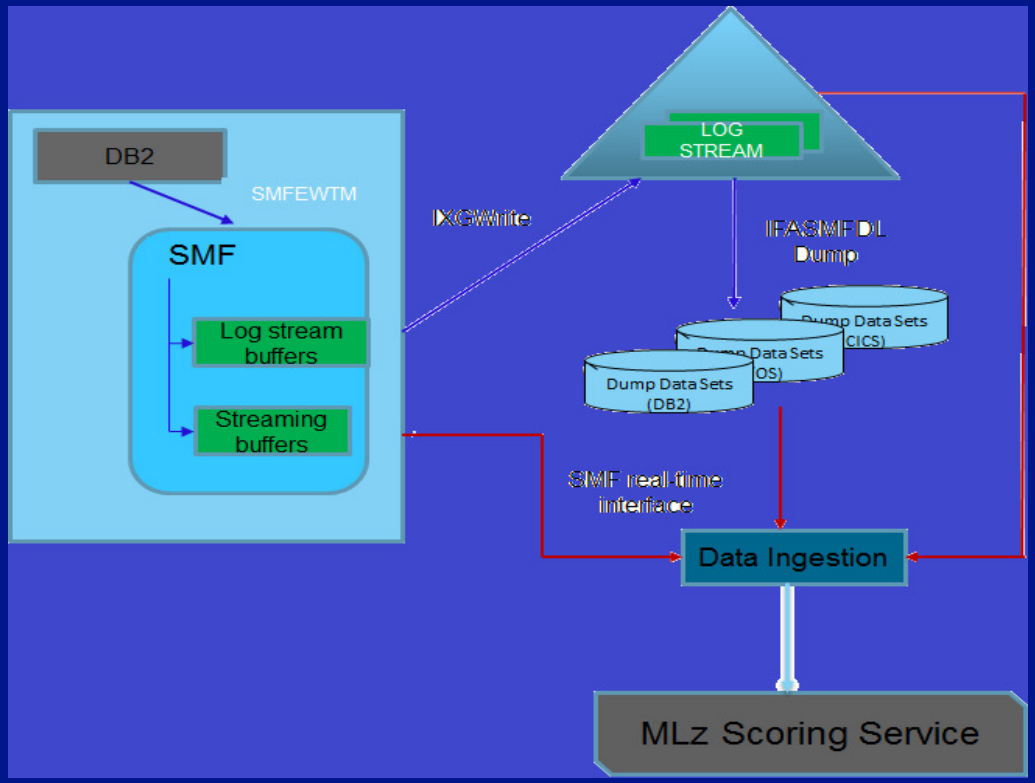
- Use customized statistical methods and time series models to build baseline model for each selected KPIs
- Factor-in the impact of the calendar (workday, weekend, special business day, etc.)
- Baseline models define the normal system behaviors, used to detect anomalies
- Baseline models identify the normal, warning and critical ranges on the timeline for each KPI

Db2 ITOA solution template
Health Tree structure



- Uses a traffic light approach to represent the health status of KPIs and groups of KPIs
  - Green for normal, yellow for warning and red for abnormal
  - Indicates how long a KPI has been in “red” state
- Determines system health by the status of the Db2 subsystem and Db2 workload
  - For Db2 subsystems, KPIs are grouped by category
  - For transactions, volume per minute and transaction response times are monitored
  - For response time issue diagnosis, you can drill down to key factors

Db2 ITOA solution template
Real-time capability



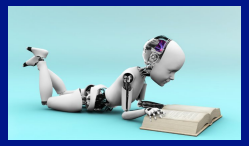
- Using the SMF real-time interface, real-time SMF data is fed into the ML for z/OS data ingestion service
- The ML scoring service provides real-time system health status
- Depicts problems with only a few minutes delay

Db2 ITOA solution template

Predict potential issues in near future

- Prediction model will only be triggered when applicable KPIs are in “yellow” status
  - Not suitable for all KPIs
  - Will continually improve the prediction function
- Predict KPI trends in the next five minutes
- Health Tree UI displays the predicted status for KPIs

Db2 ITOA solution template
Customer Experience 1



Business Background

- A large bank was impacted by hanging transactions resulting in a significant production outage lasting almost 3 hours
  - 9:00 am - 10:29 am: Transactions started hanging
  - 10:30 am: CICS region was restarted and system started to recover and transactions resumed
  - 11:55 am: System resumed normal status

Goal

- Demonstrate how ML for z/OS could detect the transaction issues earlier
- Help DBAs identify the root cause of the anomaly with granular accuracy

Db2 ITOA solution template
Customer Experience 1

Results

- Training: Used historical SMF 100, 101 records to build a dynamic baseline for 62 KPIs (Covered DB2 sub-system and transactions)
- Scoring: Detected abnormal behavior for 17 KPIs on the day of outage
- The Health Tree Application
  - Detected the transactions hung at 9:04 am, about 11 minutes earlier than the customer's rules-based monitoring system
  - Visualizations helped DBAs to more easily discover the abnormal situation and focus on the abnormal KPIs in Timeline feature
  - Identified the root cause at 9:16am (in comparison, Level 2 took an entire day to identify the same issue within a post-outage analysis)

Db2 ITOA solution template
Customer Experience 2

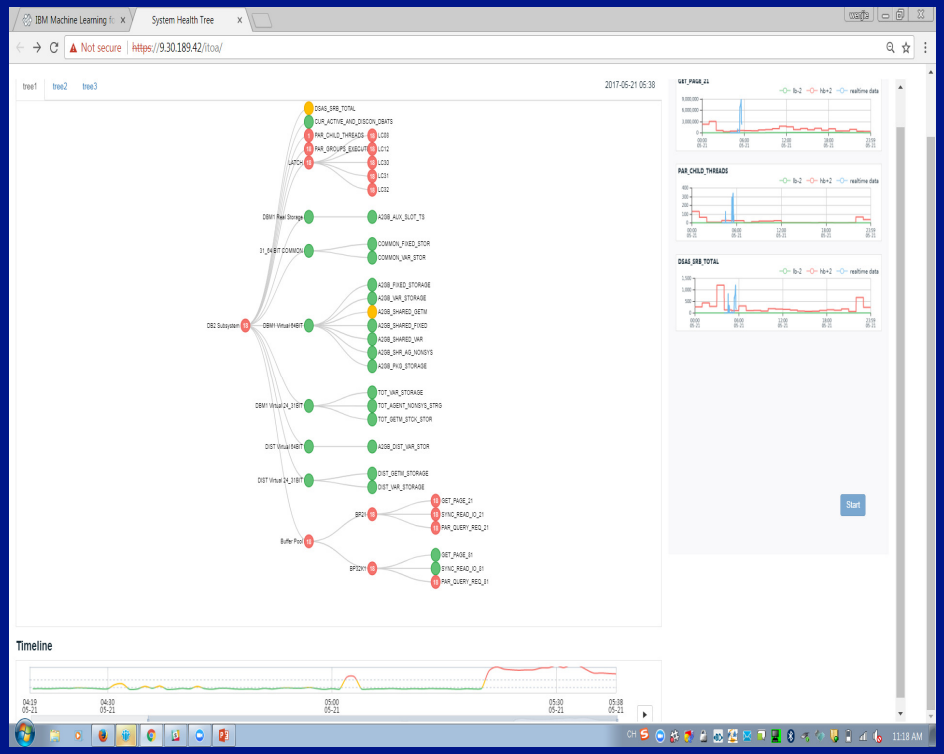
Business Background

- May 21, 4:00 AM: Implemented a DDF application change.
- A new application issued non-optimal SQL statements, which caused a Db2 performance issue, slowing down the Db2 workload
- May 22, 11:08 AM: Db2 issued the message DSNV508I indicating a below-the-bar memory constraint

Goal

- Demonstrate how early ML for z/OS can detect the change
- Enable DBAs to quickly and easily pinpoint the root cause by correlating the anomalous behavior to resource usage and constraint in Db2
- Showcase model training, deployment and scoring functionalities of Machine Learning for z/OS

Db2 ITOA solution template
Customer Experience 2



Results

- Training: used historical SMF 100 records to build a dynamic baseline for 148 KPIs
- Scoring: abnormal behavior for 53 KPIs were detected on the day of outage
- The health tree application: detected the system departure from normal on May 21 at 5:20 AM (more than 24 hours before the system performance was impacted)
- DB2 domain expertise: was able to readily pinpoint the root cause using the health tree

Traditional challenges of a rules-based ITOA system

All rules are based on the human experience

- *typically rules don't get reviewed as often*
- *as a result the % of false alerts could be high*

There are so many KPIs, it is impossible to set rules for each KPI

With outages, its difficult to identify which KPIs appear abnormal first

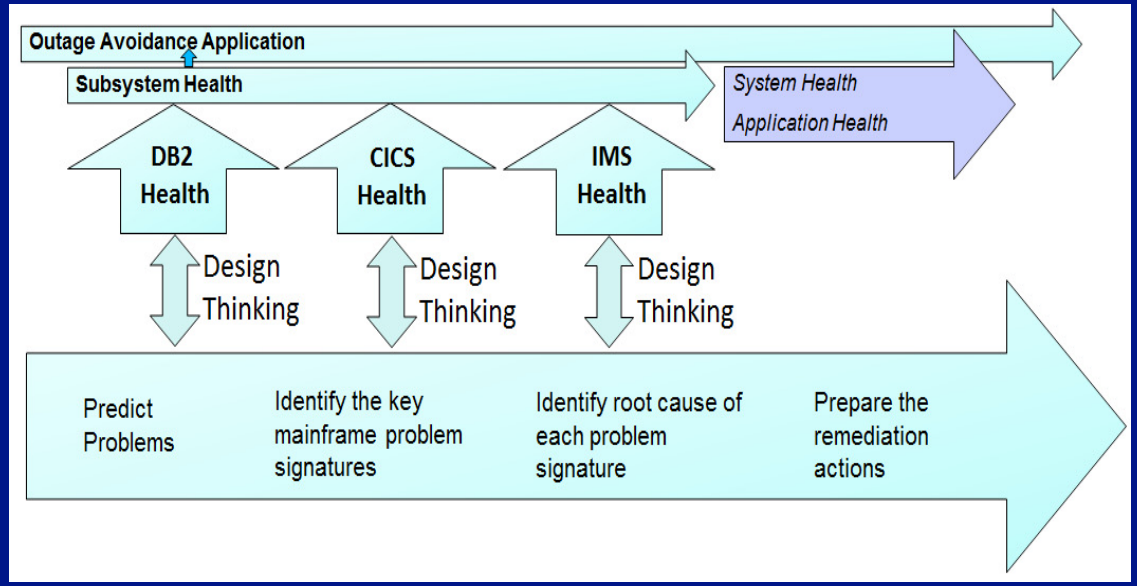
Db2 ITOA solution template

A powerful use case for ML for z/OS

- Monitor systems and transactions at a granular level
- Be sensitive to data characteristics
- Find hidden patterns from the data
- Avoid human bias and limitations in experience
- Use self-monitor, self-evaluate, and self-retrain functions, to minimize concern about when data patterns change and when model performance regress
- Indicate when to re-train, to gain more accurate outage predictions and uncover how system monitoring will be achieved
- Gain flexibility as changes and updates occur

Productizing predictive analytics with sponsor customers

- IBM is working with sponsor users to leverage the ML for z/OS toolkit and productize IT ops use cases such as the Db2 ITOA solution template
- Focused on a relatively small number of subsystems to ensure critical mass of problem signatures



The benefits of implementing IBM Machine Learning for z/OS

Advantage of leveraging your enterprise resources

- Gain insight at the point of transaction
- Keep sensitive data in place, encrypted, secure
- Leverage best of open source and unique innovation

The value to data scientists

- Increase data scientist productivity
- Collaborate across data science teams
- Deliver better insight to more business teams

The value to the organization

- Optimize processes
- Improve employee productivity
- Increase customer satisfaction

IBM z Analytics Marketing
Db2 AI for z/OS

IBM | IBM z Analytics

What is the Optimizer? An analogy....

#Db2ZAI

- The optimizer is responsible for
  - Choosing the most efficient method of accessing the data for a given SQL statement

- Think of your transportation choices
  - Start/end location, time of day, construction, traffic, available options/routes
    - All can impact the “quickest” route

Announcing IBM Db2 AI for z/OS

#Db2ZAI

IBM Db2 AI for z/OS (Db2ZAI) empowers the optimizer in your Db2 for z/OS engine to determine the best-performing query access paths based on your workload characteristics using machine learning

- Learns the patterns from the collected data from workloads in customer's unique operating environment and determine the optimal paths for SQL statements entering Db2 for z/OS

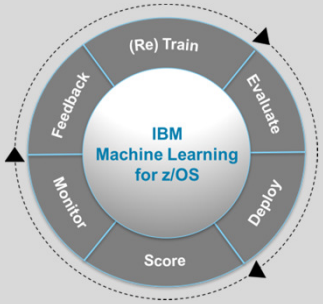
- Built on top of the Machine Learning for z/OS stack
  - leveraging all the services without requiring data scientist support
  - Db2 generates the training data, deploys and monitors/retrains models with MLz

What is the Db2 Optimizer?
It's job is to determine the best way to return data results from a query in Db2

IBM Db2 AI for z/OS – Business Value

#Db2ZAI

Db2z infused with ML



MLz

IBM Z

Up to 25% CPU Savings*

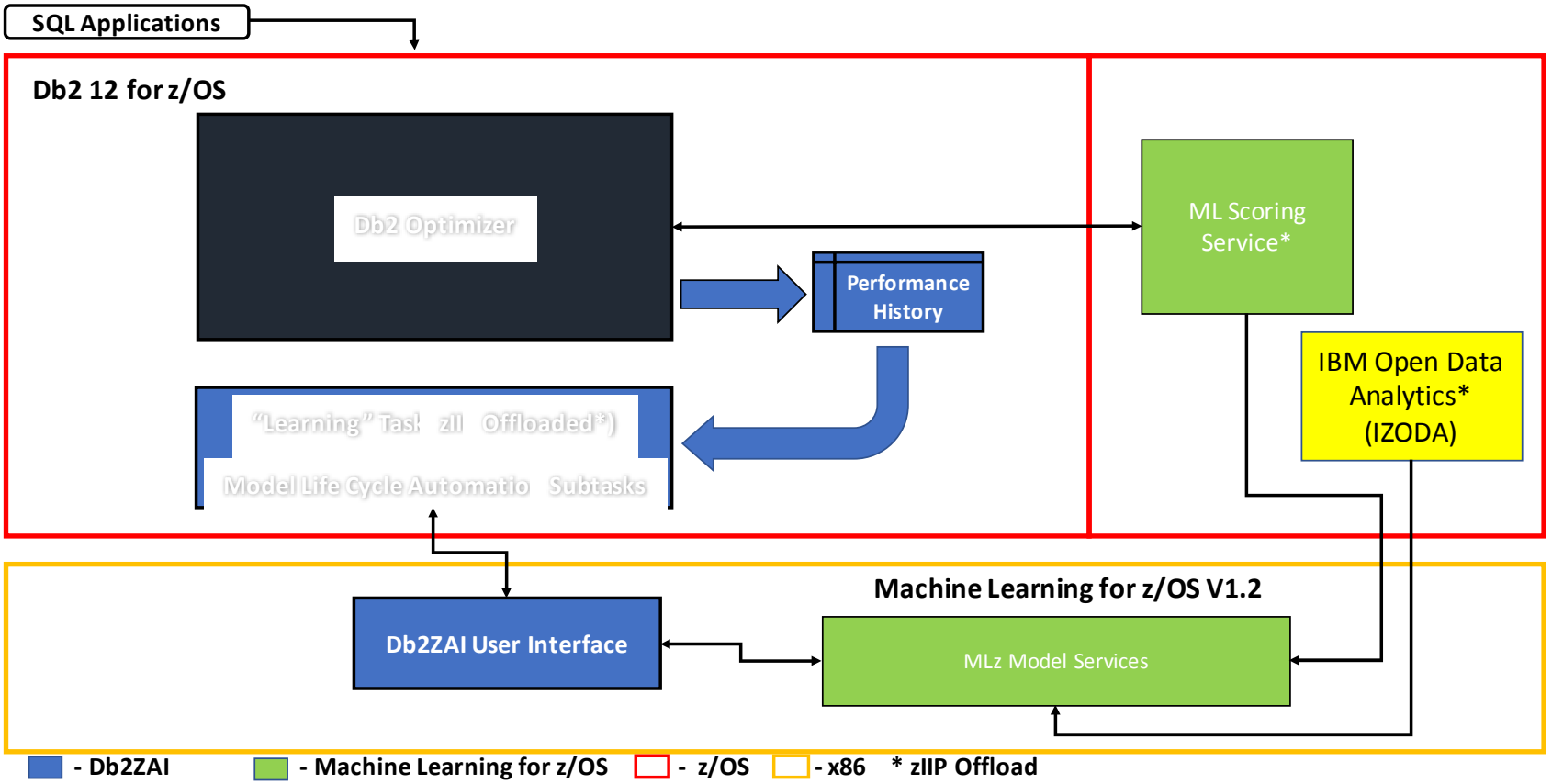
Value Proposition

Enable Db2 for z/OS optimizer to leverage Machine Learning for z/OS (“MLz”) services and IBM Z

- Reduce CPU consumption and IT cost through optimization for best query access paths
- Improve Db2 application performance
- Rapid model learning specific to the data/application behavior per subsystem without requiring data science skills

**CPU savings metric based on IBM internal benchmarks; actual savings will vary according to customer workloads and environment*

IBM Db2 AI for z/OS – Architecture




IBM Db2
AI Makes Db2
Better, Smarter,
Faster
IBM Db2 AI for z/OS
#Db2ZAI
IBM

Next Steps

- IBM Db2 for z/OS Announced 11th September 2018
- IBM Db2 AI for z/OS General Availability 21st September
- Join the Launch webcast - available on replay http://ibm.biz/Db2ZAIWebcast

Q&A

Please use the Q&A block on your console to submit your questions

LEARN MORE...

IBM z Analytics
www.ibm.com/analytics/z-analytics

IBM Machine Learning for z/OS
www.ibm.com/us-en/marketplace/machine-learning-for-zos

Test Drive IBM Machine Learning for z/OS
https://ibm.biz/BdZ38G