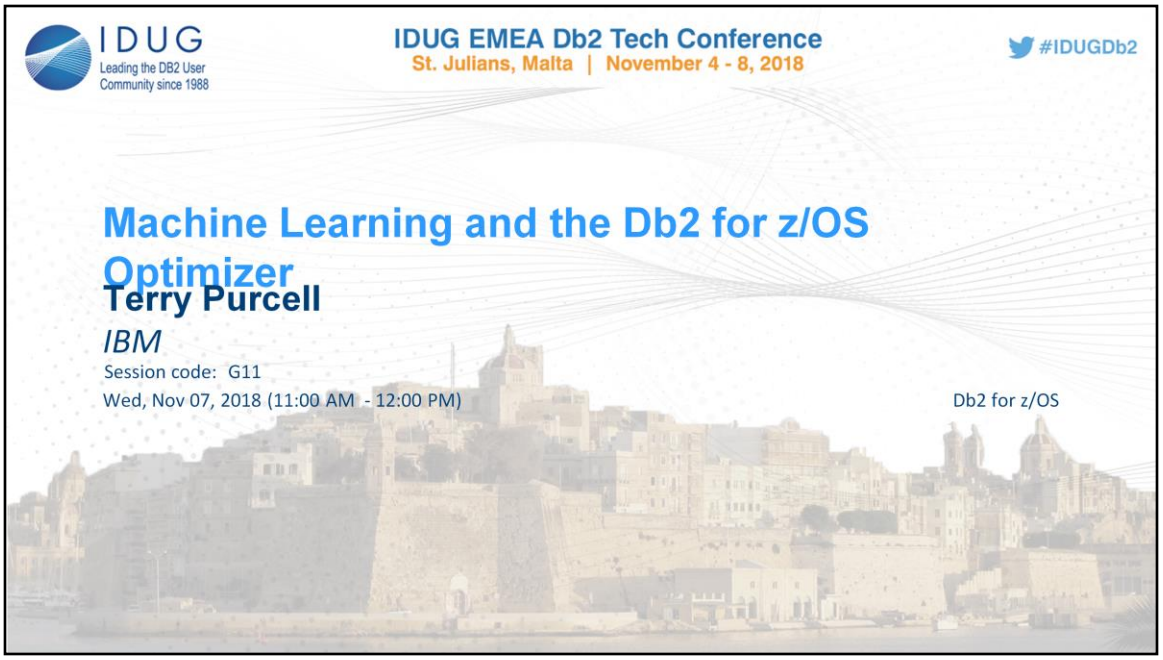


How many people are on V11 now? V12?

Agenda

- The query optimization challenge
- A brief overview of machine learning
- Db2 AI for z/OS: exploiting machine learning for query optimization

The query optimization challenge

One aspect of the optimization challenge

- Determining the degree of filtering for each predicate (i.e. filter factor) is a key input to the optimizer's access path formulas
- Question: how many of you were born before *parameter marker*?
  - It depends…
    - What if the substitution value is '1900-01-01'?
    - What if the substitution value is '2018-09-01'?

BIRTHDATE < ?

- With a range predicate with parameter marker or host variable…
  - …we use a default filter factor (these are documented)

The trouble with default filter factors

- They may not ALWAYS predict how an organization accesses its Db2 data
- Analogy: transportation recommendation for a local trip
  - Personal vehicle, bus, rail?
  - Answer depends not only on factors such as distance to be traveled and time of day, *but also on the particular area in which the trip will be taken*
    - Best answer for Boston may not be best for Atlanta, or Toronto, or Dallas
    - A location-specific answer would be nice, wouldn't it?

Tuning just ONE query: the user challenge

- Access path analysis can be a complex, time-consuming task
  - Depends on several factors (user skill level, available tooling, etc.)

- Solutions that we like to see:
  - Determine correct RUNSTATS options for targeted database objects
    - Sources: SYSIBM.SYSSTATFEEDBACK (system-level) or DSN_STAT_FEEDBACK (query-level)
  - Indexing – add or modify
  - Add OPTIMIZE FOR n ROWS to query ("paging-style" SQL, only part of result fetched)

Skill often missing

- Solutions that are less preferred:
  - BIND/REBIND with APREUSE or OPTHINTS
  - Query rewrite
  - "Fudging" catalog statistics
  - Query tricks (0=1 for example)

Organizations often revert to these

Large-scale REBIND can present another challenge for users

- Suppose you REBIND 10000 static packages using APCOMPARE(WARN)
  - After migration to new Db2 version, or applying maintenance
- If 10% (1000 packages) have access path differences, how do you validate that they are OK?
  - Review changes in REMARKS column of PLAN_TABLE if REBIND executed with EXPLAIN(YES)
    - Do access paths look better (e.g., increase in MATCHCOLS)? Any red flags?
  - If access path change looks OK, go with new instance of package
  - What if access path change does NOT look OK?
    - REBIND with SWITCH to revert to previous instance of package?
    - Use OPTHINT to force use of an alternate access path?
    - Try to tune affected queries?

Approaches most commonly taken

- What if access path change validation work has to be done for several hundred packages?

Users want performance, but also stability

- We've been at this a long time
  - IBM invented cost-based SQL optimization
  - Db2 for z/OS optimizer has been enhanced over 35+ years
    - based on user feedback and internal analysis by the Db2 development team

- Stability and reliability have long been priorities for query optimization
  - Statistics
    - Current and accurate catalog statistics, provided by RUNSTATS (or inline statistics), provide quality inputs to the optimizer
  - Access path stability
    - a core characteristic of static SQL, also provided for dynamic SQL via prepared statement cache and (with Db2 12) dynamic plan stability
  - Furthermore, plan management (for static SQL) enables quick and easy switch to earlier instance of a package in the event of an access path regression

Optimizer development is guided by a twofold objective

Deliver new access path choices that enhance performance and stability…

…without compromising the quality and stability of existing access path choices

A brief overview of machine learning

What is machine learning?

A twist on traditional data processing

Traditional data processing

Machine learning

Put another way, it's about computers that learn without being explicitly programmed

Identify patterns
not readily discerned by humans

Build models
of behavior from those patterns

Score or predict
behavior with deployed models

Organizations can exploit machine learning technology

Machine learning…

- *Constantly learns and adapts*
- *Avoids making the same mistakes*
- *Provides faster, deeper, improved insights*

Resulting in…

- ✓ **More-favorable business outcomes**
- ✓ **Reduced risks and costs**
- ✓ **New opportunities**

Churn analysis helps identify cause of churn and facilitates implementation of effective retention strategies

Detect and understand life-threatening medical conditions and design more effective treatment programs

Learn, predict weather patterns and energy production from renewable sources and integrate into grid more effectively

Product recommendation, next purchase prediction, targeted offers – individually tailored shopping experience

Identify suspicious behavior, predict and prevent threats / fraud – continually reduce business risks and costs

How about Db2 optimizer exploitation of machine learning?

Machine learning…
- *Constantly learns and adapts*
- *Avoids making the same mistakes*
- *Provides faster, deeper, improved insights*

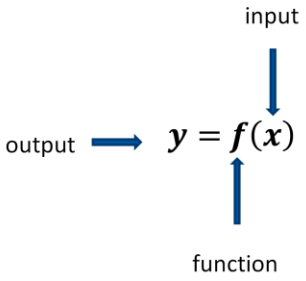
Resulting in…
- ✓ **More-favorable business outcomes**
- ✓ **Reduced risks and costs**
- ✓ **New opportunities**

How many optimizer "mistakes" can you tolerate while we learn?
NOTE: This is a rhetorical question – we know optimizer mistakes aren't tolerated
- Stakes are higher versus using machine learning to cut down on irrelevant online ads
- You wouldn't want a self-driving car to make a mistake…

Potential benefits:
- Enhanced reliability (less risk of performance regression)
- Improved performance

"Supervised" machine learning

- A model is trained from a set of known input data to a set of target output values
  - Input
    - Numbers (stock prices)
    - Text (car failure description)
    - Images
    - Audio files
  - Output might be a class…
    - Yes/no
    - Buy/not-buy
  - …or perhaps a numerical value…
    - House price
    - Annual energy consumption in kWh



Machine language terminology: features

- A **feature** is a piece of information that might be useful for prediction
  - Example: something that helps to predict the churn probability of a customer

– OR –

  - Inputs to a model used by the Db2 optimizer to provide a score for a given SQL

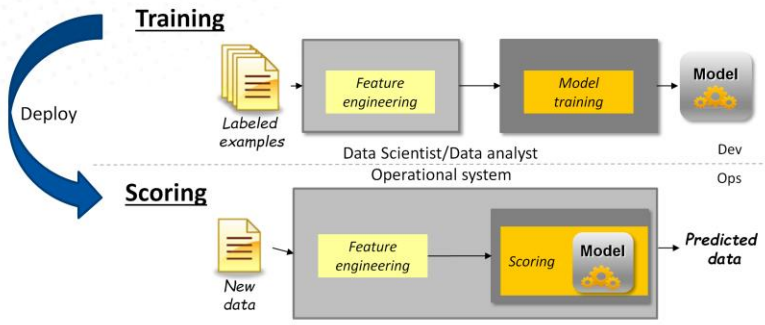
CUST_ID	AVG_DAILY_TX	EDUCATION	EDUCATION_GROUP	INVESTMENT	AVG_CR_AMT	CHURN_LABEL	AGE
1009530867	0.9178079962730408	5	Bachelors degree	114368	2090.3200683593	false	84
1009544000	0.9506850242614746	2	Bachelors degree	90298	2095.0400390625	false	44
1009534260	0.9205480217933655	2	Bachelors degree	94881	1723.4599609375	true	23
1009574010	0.9945210218429565	2	Bachelors degree	112099	1297.419921875	true	24
1009578620	0.9178079962730408	5	Doctorate	84638	1333.1799[illegible]0625	false	67

Feature

Feature

Model training, deployment, scoring

A TrainOps (DevOps) story



16

© 2015 IBM Corporation

Exploiting machine learning for query optimization

Machine learning in Db2 for z/OS – many possibilities

- Long-term goal: make Db2 a more self-tuning and self-managing data server
- Machine learning could be applied to…
  - Query optimization
  - Package rebinds
  - Memory and storage allocations (buffer pools, work files, RID pool, sort pool, etc.)
  - Utility execution (REORG, RUNSTATS,…)
  - Table space and index PCTFREE/FREEPAGE specifications, based on insert patterns
  - Fast traverse block (FTB) usage
  - Analytics Accelerator query offload decisioning
  - Application of maintenance (how, when, with what expected benefits?)
  - …

Internal Study – Access path selection

- Construct machine learning models to assist Db2 optimizer in choosing best access path for a query
- Models are based on actual costs observed for queries that are executed on a system
  - What that means: the optimizer is customized for the needs of a particular environment and workload
- Machine learning models stay current and adapt to changing conditions

Machine learning and access path selection – phase 1

- Initial focus on optimizer cost model – could it be improved via machine learning?
  - Analysis using supervised learning techniques (neural networks, random forest, linear regression)
- Analysis was used to validate the existing optimizer cost model
  - Several costing deficiencies found in optimizer base code – addressed via APARs
- After that work, focus shifted to using machine learning to improve the quality and effectiveness of inputs to the optimizer cost model
  - Machine learning does not replace optimizer (any more than IBM Watson replaces an oncologist)
  - Watson provides an oncologist with better inputs to patient care decisions – we use machine learning to provide the Db2 optimizer with better inputs to query access path decisions

Phase 2: focusing on what the optimizer doesn't know

- Consider access path selection factors that are currently unknown to the optimizer
- Two of the most important of those factors are:
  - The execution-time values used in place of host variables or parameter markers coded in query predicates
  - The number of result set rows that will actually be fetched by the query-issuing application

This phase two work has brought us to what's new…

New offering: IBM Db2 AI for z/OS (Db2ZAI)

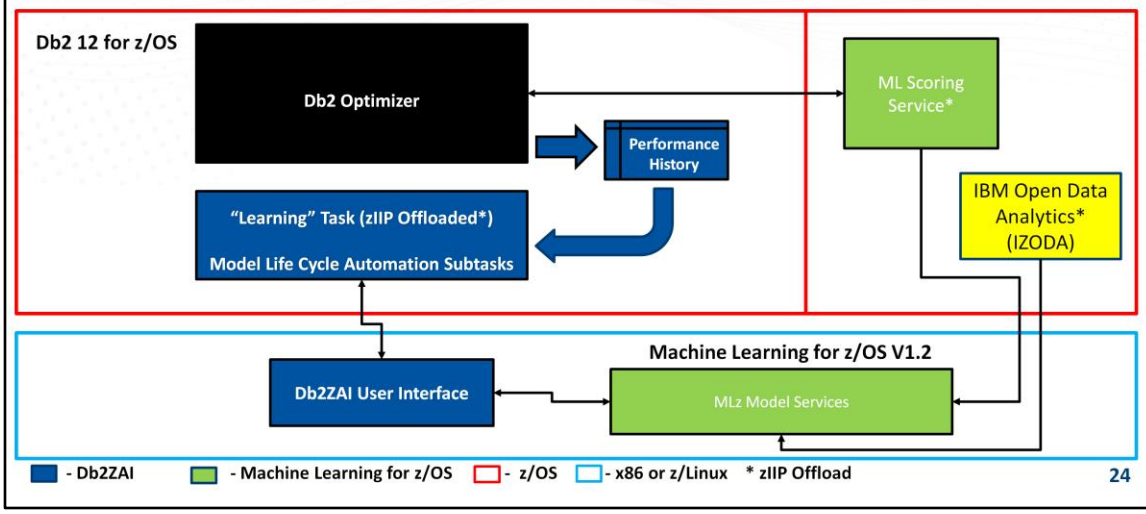
Db2ZAI enables the Db2 optimizer to determine the best-performing query access paths based on **your** workload characteristics, using machine learning

- Learns patterns from workload data collected in an organization's **unique** operating environment and uses derived insight in determining optimal access paths for SQL statements
- Built on top of the IBM Machine Learning for z/OS (MLz) stack
  - Leverages MLz services without requiring data scientist support – Db2 generates model training data, deploys and monitors and retrains models via MLz services
- Db2ZAI product ID: 5698-CGN

Db2ZAI – prerequisites

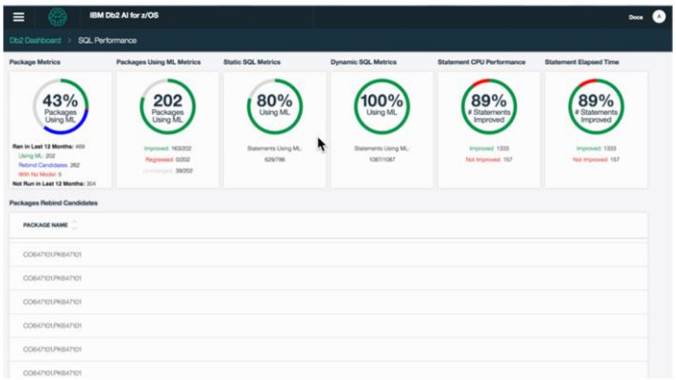
- Db2 12 for z/OS, with function level V12R1M503 (or later) activated (M503 functionality is provided via APAR PH00506)
- IBM Machine Learning for z/OS, V1.2 (5698-ML1)
  - With MLz APARs PI99400, PI99401, and PH02006 plus MLz 1.2.0.1 Linux packages
  - MLz uses IBM Open Data Analytics for z/OS (5655-OD1)

Db2ZAI – architecture



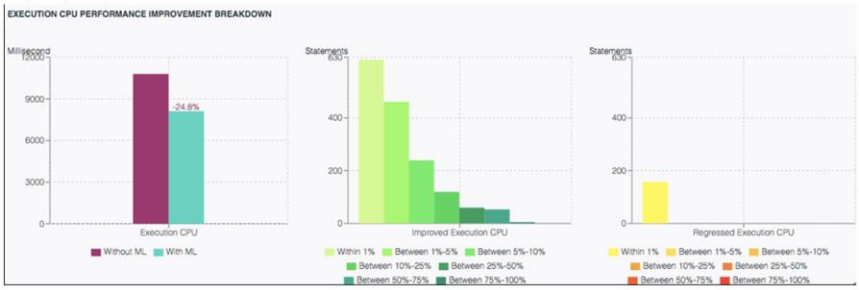
Db2ZAI – SQL performance dashboard

- Static (package) and dynamic SQL usage metrics and performance summary
  - Drill-down capability



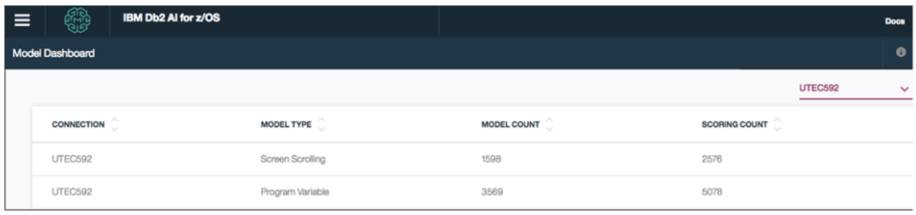
Db2ZAI – SQL performance dashboard: drill-down

- Provide before/after summary of SQL performance



Db2ZAI – model dashboard

- Since Db2ZAI manages (deploys, trains) models, model statistics are an FYI to show number of SQL statements using each model type



More on Db2ZAI

- Using filter factors customized for an organization's particular query workload, better-performing access paths can be generated by the optimizer
  - Internal IBM benchmarks have shown up to 25% reduction in CPU time when machine learning was employed via Db2ZAI for a set of Db2 queries compared to access paths selected by the Db2 optimizer without machine learning
- For static SQL, using machine-learning-enhanced optimization requires package rebind
  - Db2ZAI performance dashboard provides a list of packages for which model-augmented access path selection is available
  - **Note:** it is understood that improving the BIND/REBIND experience is important for enabling the exploitation of machine-learning-enhanced optimization for static queries – this is a parallel effort within Db2 for z/OS development

How do I know if my workload will benefit?

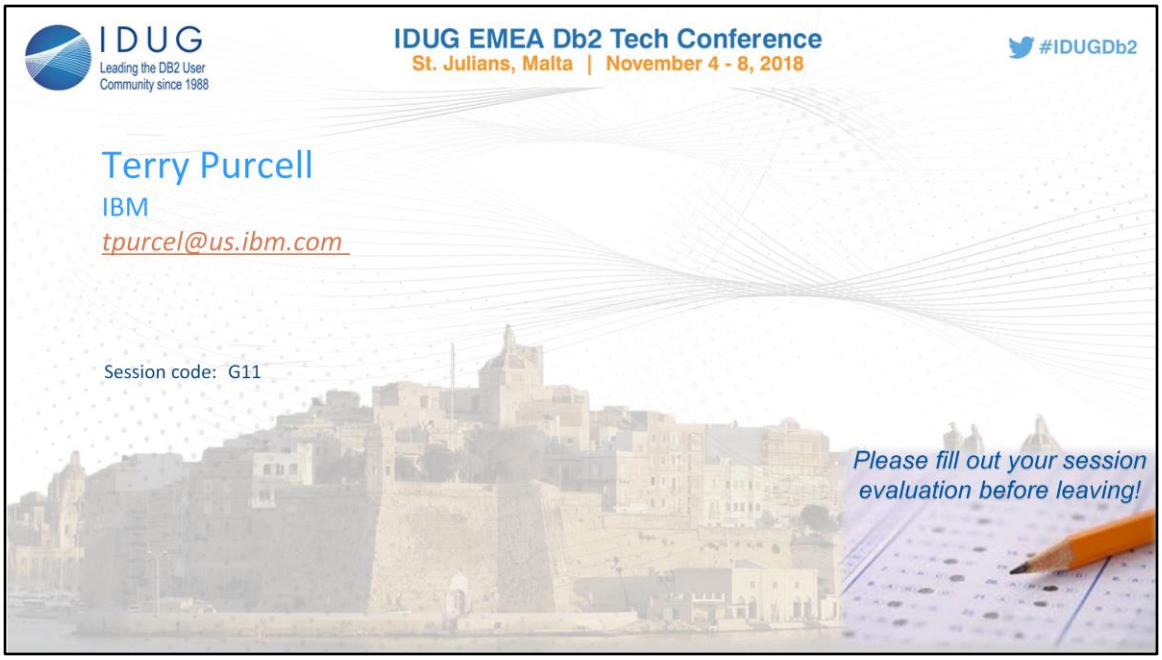
- If you're asking "Can you tell me what the Db2 Optimizer doesn't know before Machine Learning has learned what the Db2 Optimizer doesn't know?"
- Db2 development has SQL script that can identify the number of
  - SQL statements in your system with risk of a sub-optimal access path due to the combination of
    - Host variables/parameter markers
    - WHERE clause predicates with data skew or range predicates
  - AND
    - Index design which exposes the optimizer to index choices that have above risk
    - OR multi-table queries where risk of choosing the wrong join sequence or method exists
  - Contact your IBM rep for the Db2ZAI Analyzer

What's next?

- The next targeted use cases are being evaluated
  - *That includes soliciting customer input*
- This is a journey
  - Along the way, Db2 for z/OS development is committed to the safe and cost-effective exploitation of machine learning technology

Where to find more information on Db2ZAI

- The announcement letter:
  https://www-01.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_ca/1/897/ENUS218-341/index.html&request_locale=en
- The IBM Knowledge Center:
  https://www.ibm.com/support/knowledgecenter/SSGKMA_1.1.0/src/ai/ai_home.html
- Overview video:
  https://www.youtube.com/watch?v=t5fTNxfehQA


IDUG
Leading the DB2 User Community since 1988
IDUG EMEA Db2 Tech Conference
St. Julians, Malta | November 4 - 8, 2018
#IDUGDb2
Terry Purcell
IBM
tpurcel@us.ibm.com
Session code: G11
Please fill out your session evaluation before leaving!