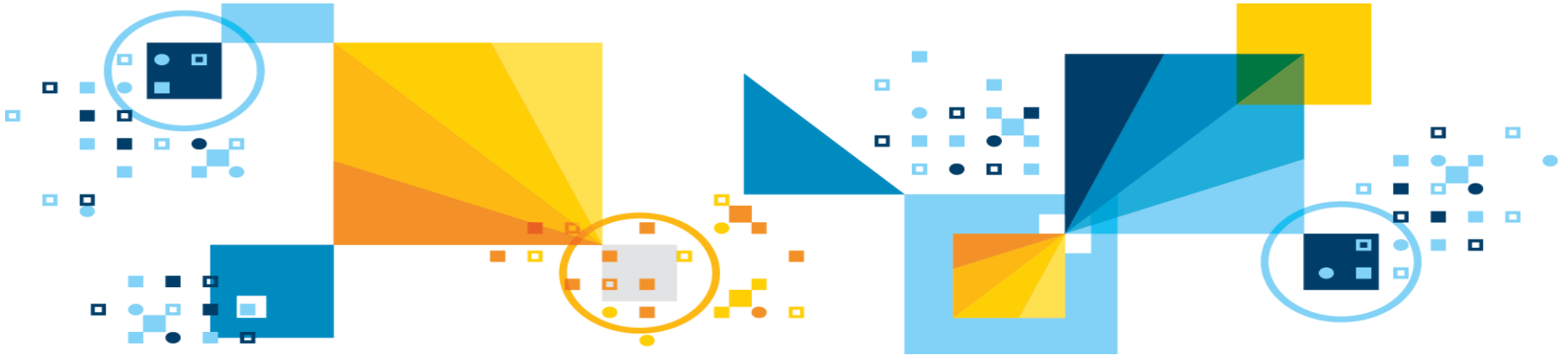


Les King
Director, Hybrid Data Management Solutions
lking@ca.ibm.com
ca.linkedin.com/pub/les-king/10/a68/426
March 28, 2019

Hybrid Data Management Strategy and Db2 Roadmap



Legal Notice

Copyright © 2019 by International Business Machines Corporation. All rights reserved.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation.

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at ibm.com/legal/copytrade.shtml.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER OR IMPLIED. IBM LY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. IBM makes no representations or warranties, express or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 1 0504- 785
U.S.A.

Hybrid Data Management Strategy: The Ladder to AI



Reimagine your workflows

TRUST: Achieve trust & transparency

Protect your data and insights

AUTOMATE: Apply ML Everywhere

ANALYZE: Scale insights on demand

Learn more from less data

ORGANIZE: Create a trusted analytics foundation

COLLECT: Make data simple & accessible



Modernize for simplicity & agility

Data of every type, regardless of where it lives

Hybrid Data Management Strategy - The AI Ladder

AI adoption is Accelerating but there is a set of unique challenges

94%

of companies believe that AI is key to competitive advantage

AI associated with CRM activities will boost global business revenue by **\$1.1T** from 2017 to 2021

Only **1 in 20** companies have extensively incorporated AI in offerings or processes

Top reasons for lack of AI Adoption

- **Skills**

Lack of requisite talent to drive AI adoption

- **Data**

Only 19% respondents strongly agreed that their organizations understand the data required to train AI algorithms. Data used is not of high quality or trusted.

- **Trust**

Only 35% of IT and Business decision makers had a high level of trust in their own organization's analytics. AI insights not well integrated into current processes

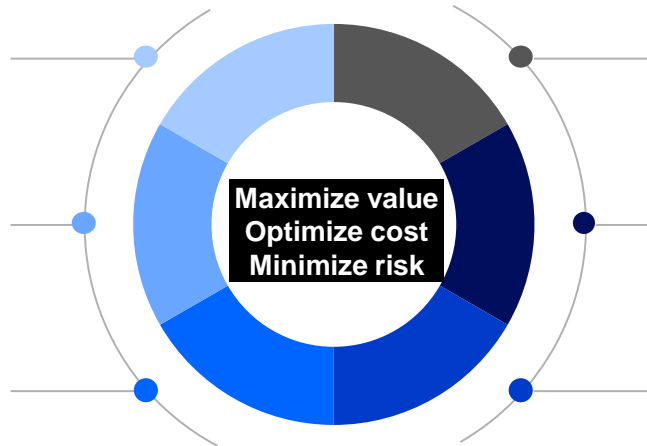
Hybrid Data Management Strategy - Comprehensive

Our hybrid data management design principles

Multi-Cloud Availability
Private & Public Cloud

All Data
Structured and
unstructured
Transactional & Analytical

At Rest & In-Motion
From Event-driven to
traditional systems

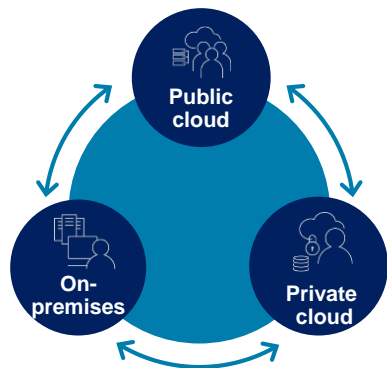


Self-service
Enable data-driven decision
for everyone

Cloud Native
Elastic and resource
optimization

Built-in AI
Smarter decisions and
optimized operations

Hybrid Data Management Strategy - Build Without Limits



- All data types
- Flexible consumption
- Fast and traditional data
- Deploy everywhere
- Unified experience from discovery to support
- Integrated Machine learning and smart capabilities



Db2 Family

Comprehensive suite for structured data



Open Source Market leaders



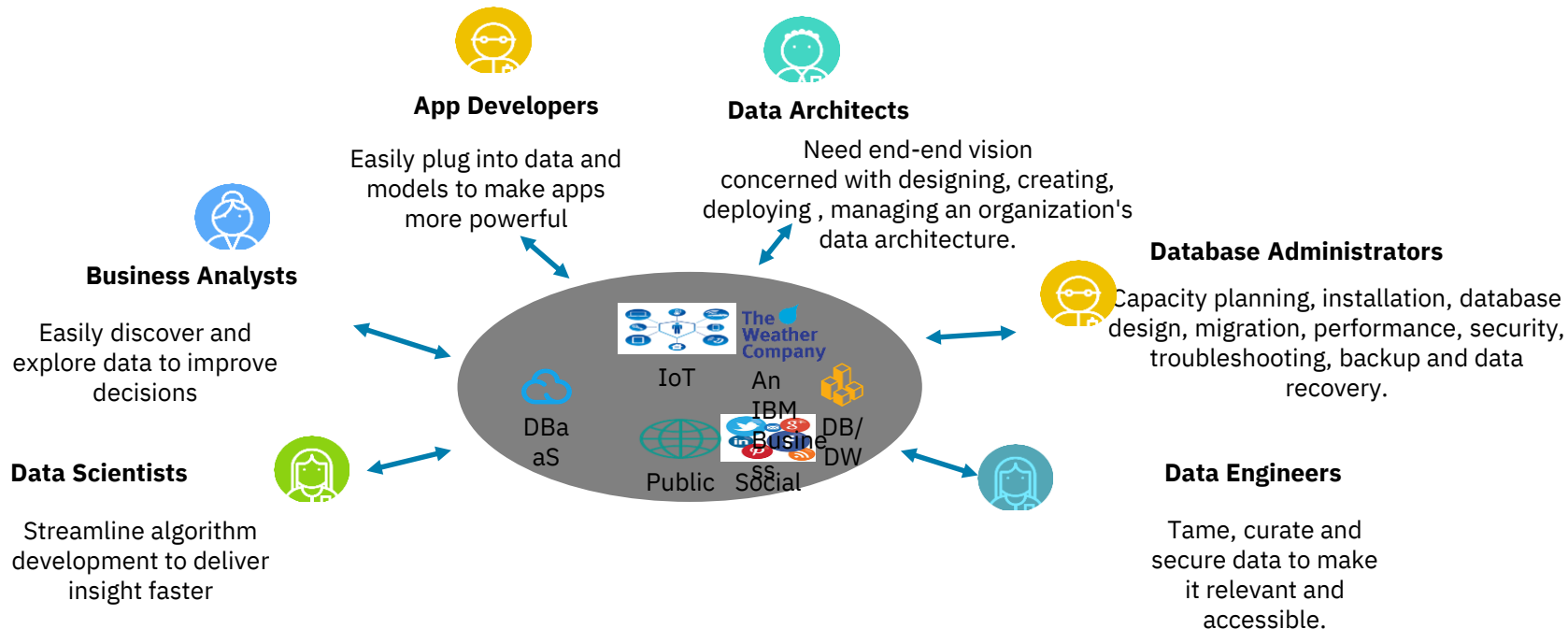
Informix Embed/IoT



HDP Hadoop for the enterprise

Hybrid Data Management Strategy – Needs of all Data Professionals

As data maturity increases, so does the number of data professionals who are hungry to put data to work



Hybrid Data Management Strategy – Address all Dimensions of Hybrid

NOT about Cloud **OR** On-premises

NOT about traditional relational **OR** open source

NOT about SQL **OR** NoSQL

NOT about structured **OR** unstructured data

NOT about data at rest **OR** data in motion

It's about Cloud **AND** On-premises

It's about traditional relational **AND** open source

It's about SQL **AND** NoSQL

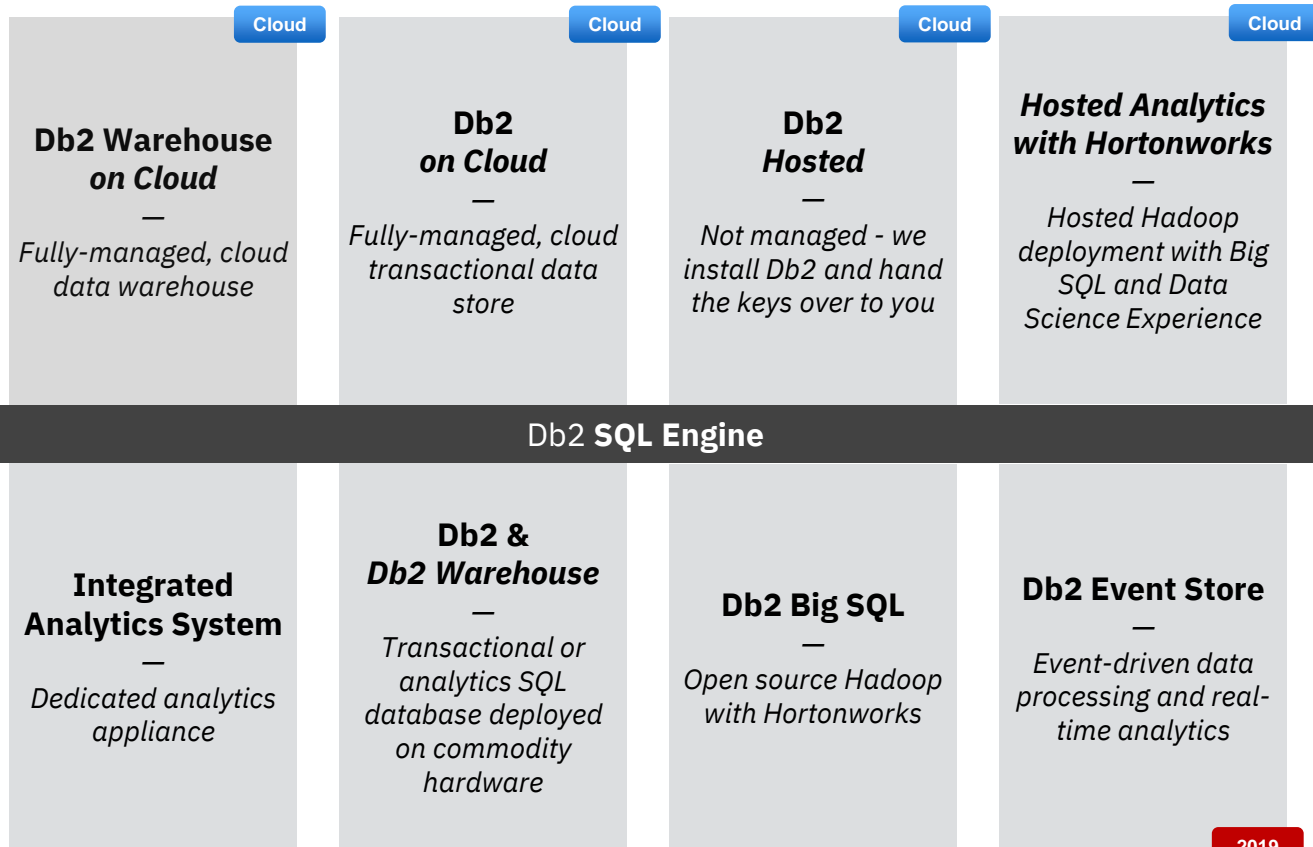
It's about structured **AND** unstructured data

It's about data at rest **AND** data in motion

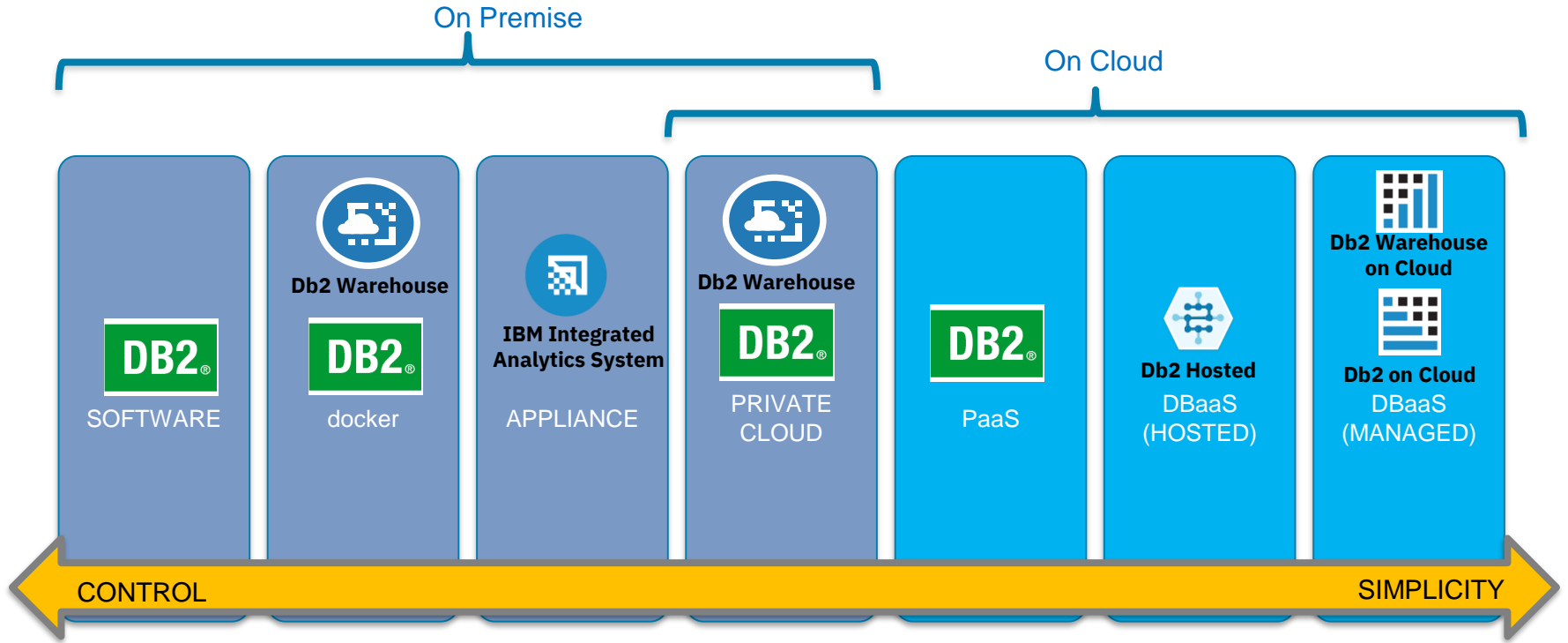
Hybrid Data Management Strategy - Common SQL Engine

Our family of **Hybrid Data Management** solutions built on the **Db2** common SQL engine

*Write your SQL once
deploy against any
form factor
run anywhere*



Hybrid Data Management Strategy - Flexibility of Deployment



Hybrid Data Management Strategy - HDMP

IBM Hybrid Data Management Platform

One Db2 family

One HYBRID entitlement



Available for purchase via two licensing models:

*(licensed **perpetually**)*

- ☐ Supports OLTP, warehousing, SQL on Hadoop, and Fast Data on private cloud and on-premises infrastructure
- ☐ Intended for customers who need to maintain ownership of their software licenses in perpetuity

*(licensed **monthly**)*

- ☐ Supports OLTP, warehousing, SQL on Hadoop, and Fast Data on public cloud, private cloud, and on-premises infrastructure
- ☐ Intended for customers who want to procure software and cloud services as an operational expense

Hybrid Data Management Strategy - Data Virtualization

Cornerstone of a Logical Data Lake Architecture

Query anything, anywhere.

Query **many diverse data sources** across cloud, on-premises and mobile with advanced analytics using the most popular languages and tool

1

SQL, Spark, R, Notebooks, Python, Data Science Experience (DSX), Cognos Analytics, Common Analytics tools, etc.

Query many sources as one with extreme simplicity.

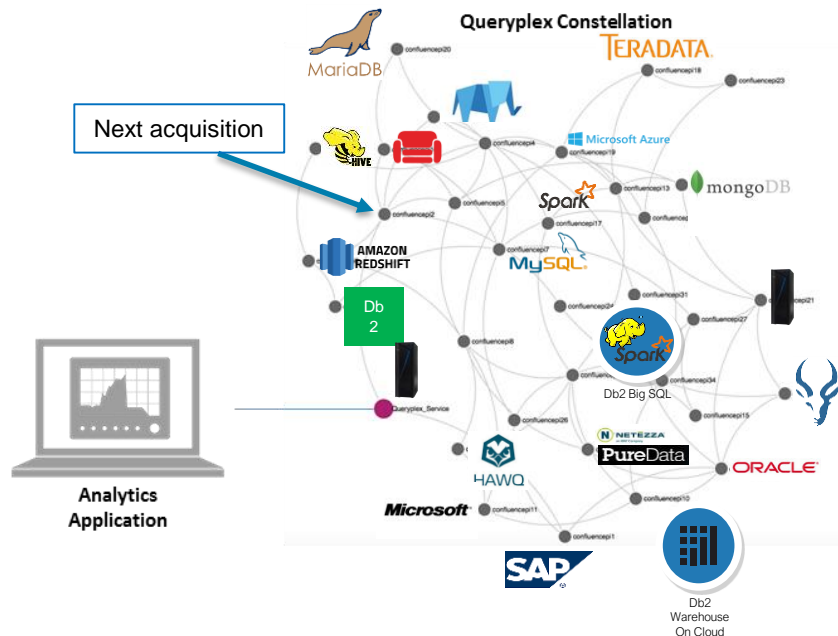
Connect **few to many devices and data stores** into a single self balancing constellation. Avoid the complexity of centralized copies. Data only persists at the source.

2

Massive speedup.

Many times acceleration using the power of every device.

3



Hybrid Data Management Strategy – Hybrid and Multi-Cloud



Competitive and deeply integrated IBM Cloud services



Fully-managed Db2 services on AWS

...support of all major cloud provider to deploy existing entitlements



Enable private cloud consumption

...ICpfD – all Db2 family and Open Source database

...ICp – limited to operational databases for application development


...use HDMP as the single entitlement supporting all consumption models



IBM Cloud Private for Data

Hybrid Data Management Strategy - IBM Cloud Private for Data

Making Your Data Ready for AI




IBM Cloud Private for Data

Make your data ready for AI – Access data anywhere, Lightning fast, No assembly required

✓ Instant, Pre-Assembled Provisioning


✓ Admin & Ops Dashboards

✓ Cloud-native Data Micro Services




IBM Cloud Private


The Cloud for Business – Run All Your Apps, AI-ready, Secure to the Core







IBM Middleware, Data, Analytics, Developer Services and Now Watson AI
Cloud enabled middleware, messaging, databases, analytics & cognitive services




IBM WebSphere



Core Operational Services
Simplify operations management, security, and hybrid integration
Provision apps across multi-cloud environments

 <p>Kubernetes-based Container Platform Industry leading container orchestration platform</p> 	<p>Cloud Foundry For prescribed app dev & deployment</p> 	<p>Terraform For multi-cloud management</p>  <p style="text-align: center; font-size: small;">(Cloud Automation Manager)</p>
---	---	--

Runs on existing IaaS:



Hybrid Data Management Strategy – Multi-Cloud Roadmap



IBM Cloud



Amazon Web Services



Microsoft Azure



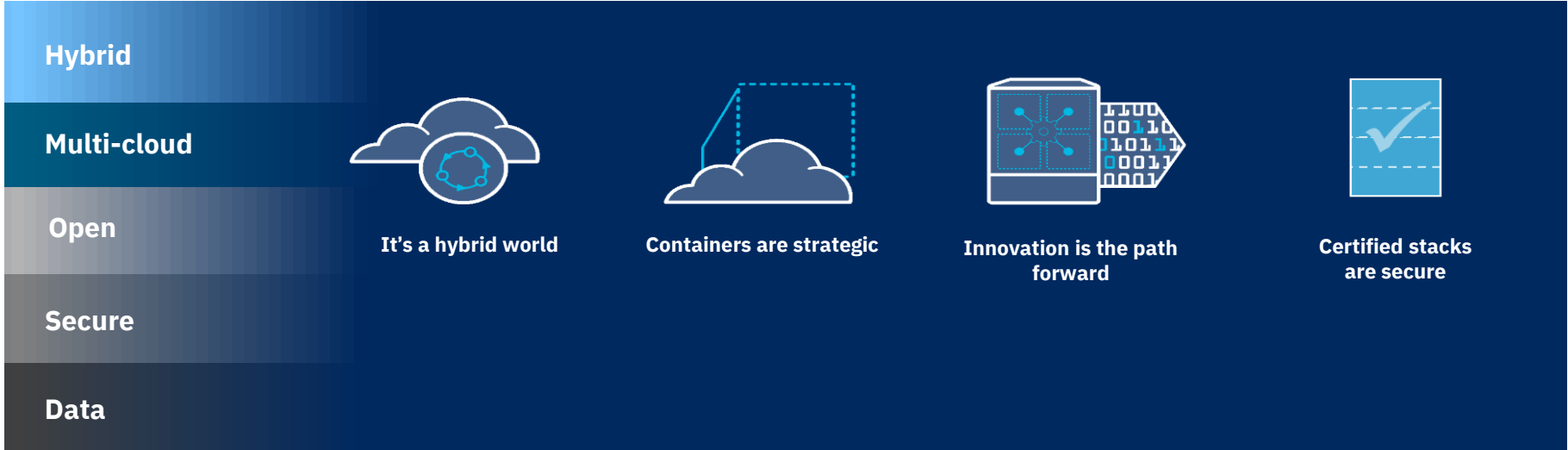
Google Cloud Platform



AliCloud

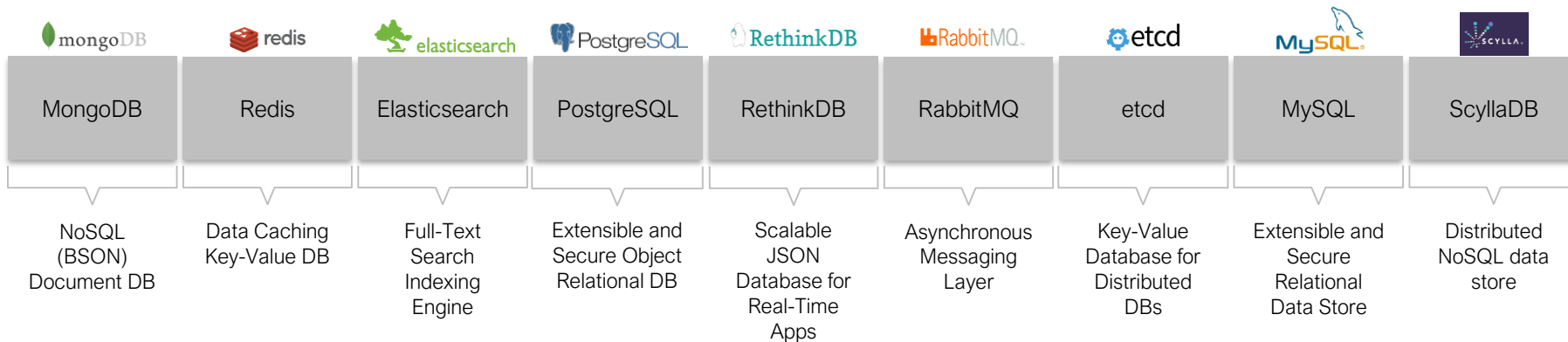
	IBM Cloud	Amazon Web Services	Microsoft Azure	Google Cloud Platform	AliCloud
BYOL	✓	✓	✓	✓	✓ *
Hosted	✓	Db2 - 2019			
Fully Managed	✓	Db2 Warehouse in Q1 Db2 in Q4			

Hybrid Data Management Strategy – Leverage IBM and RedHat

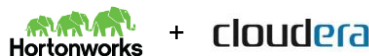
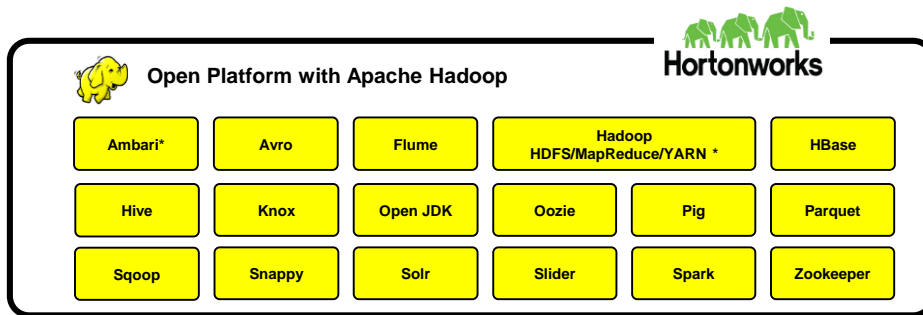


1	2	3	4	5
Hybrid	Multi-cloud	Open	Secure	Data
Enable enterprises across Traditional, Private, and Public environments	Manage other vendors' clouds, acknowledging the reality that client environments are heterogeneous	Build capabilities that are open by design, enabling client flexibility and reducing vendor "lock in"	Provide reliability and continuous security for the client's environment	Unlock the value of data in totally new ways and accelerate the journey to AI

Hybrid Data Management Strategy - Commitment to Open Source



PostgreSQL and MongoDB licenses and support available through IBM



Hybrid Data Management Strategy - Open-Source Database Strategy



One-stop-shop experience

...by OEMing market-leading capabilities



mongoDB

Integrate with HDM and ICPfD

...to unify the user experience and unlock the value of data



PostgreSQL

Offer consistent availability across private/public cloud

...enable hybrid data architectures and transition to public cloud

...unify user experience



Db2 Big SQL – Roadmap - Infuse the Power of Db2 in Hadoop



Accelerate time to market while modernizing your warehouse



Empower SQL users to operationalize ML models



Augment disparate data for deep analytics and AI



Enable BI analytics with high performance and enterprise security

Above all, bring stability to your applications even when the platform goes through updates.....

New News !!

Db2 Big SQL on HDP 3.0 will be available HDP 3.0 on Feb 28th, 2019

Has been announced and coming soon....

Introducing Db2 Big SQL support on Cloudera in 2Q 2019.

Make Db2 Big SQL the point of entry to Big Data irrespective of which platform has the data

Db2 (Common SQL Engine): The end state: *A Hyperconverged AI database engine* for a Hybrid Cloud World

Cloud native, Truly elastic

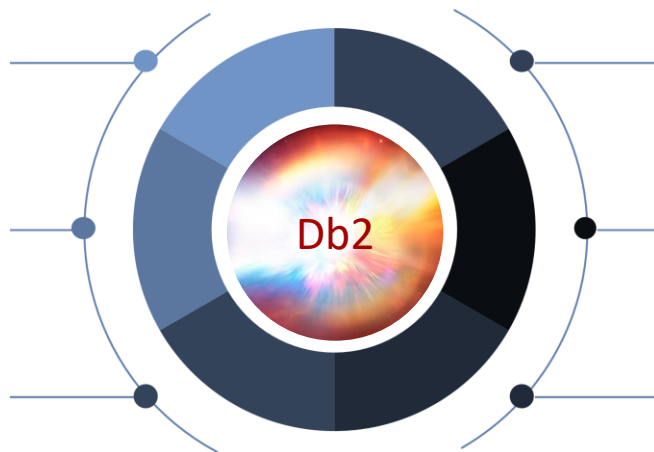
Pay for what you use
Highly resilient and infinite scalability

No more ETL

Minimize data movement
Boost distributed result processing

Hybrid & Multi-Cloud

Runs where you want to be
One view across your hybrid cloud environment



No lock-in

Open Parquet file format
All major languages supported

Powerful & Simple

OLTP, OLAP, No SQL, Graph, Blockchain
Automatic access optimization
One copy of data

AI-powered

ML-based SQL optimizer
ML-SQL
Embedded model training

Db2 (Common SQL Engine) - Roadmap



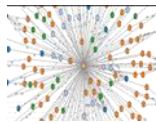
Confidence-based query results

leveraging ML-SQL



Faster data exploration in Natural Language

by using Natural Language Query and Augmented Data Exploration



Model Complex Relationships

by using Db2 Graph and SQL



Blockchain Ready

using Db2 Blockchain Connector



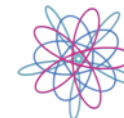
Up to 10x better query performance

powered by an ML-Optimizer



Build AI based applications

with Python, GO , JSON and Jupyter notebooks



No data movement & a single view on all data

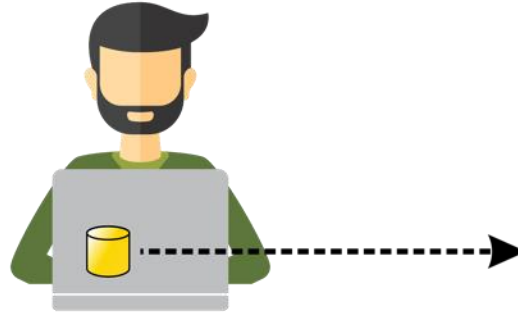
delivered by Data Virtualization

Built for AI and Powered by AI

Db2 Scales Easily and is Optimized for Maximum Performance

Db2 is the only database that offers exclusive features for maximum performance

Db2 is the only database that offers a shared disk and shared nothing architecture option for high-end optimization of all workloads



**Go from
gigabytes to petabytes ...
without compromising
performance**

- Industry leading cost optimizer
- ML Optimizer (New)
- Adaptive Workload Management (New)

Db2 - Only database with ML Optimizer and Cost Optimizer

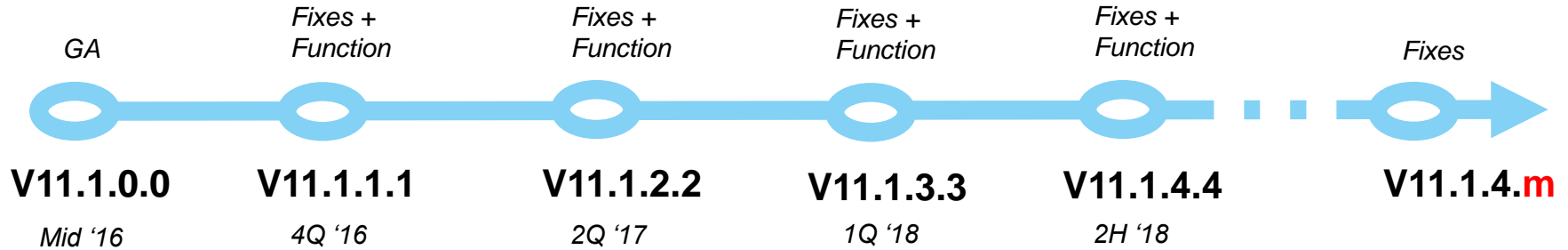


New

*Machine Learning
Optimizer that
improves from
experience to make
query optimization
simple, reliable and
stable*

- Number of rows flowing through the various operations impacts performance for most common issues and is calculated via cardinality estimates
- Basic tuning to improve cardinality estimation is high impact
- **Phase 1 – Cardinality Estimation**
 - Initial phase – Support equality & range local operators with no expressions
 - Future phases – Cardinality support for expressions, predicates & BLU, support for join enumerations,

The Db2 11.1 lifecycle



Regular fix/mod pack deliveries (every 6-9 months)

- With overriding focus on stability for immediate deployment to production environments
- Modifications will contain select functional enhancements (often off by default)

Interim fix (aka “iFix”) containing only critical fixes to last update released every 2-3 months until next update occurs

Db2 11.1.4.4 – New News

Continuous Delivery – 4Q18 New Capabilities

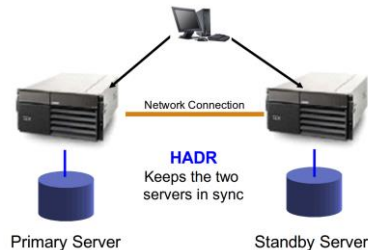
- Extent reclaim support for Db2 pureScale
- ALTER TABLESPACE ... REDUCE



- ISO JSON SQL functions – Part 1
- Following public standards SQL:2016



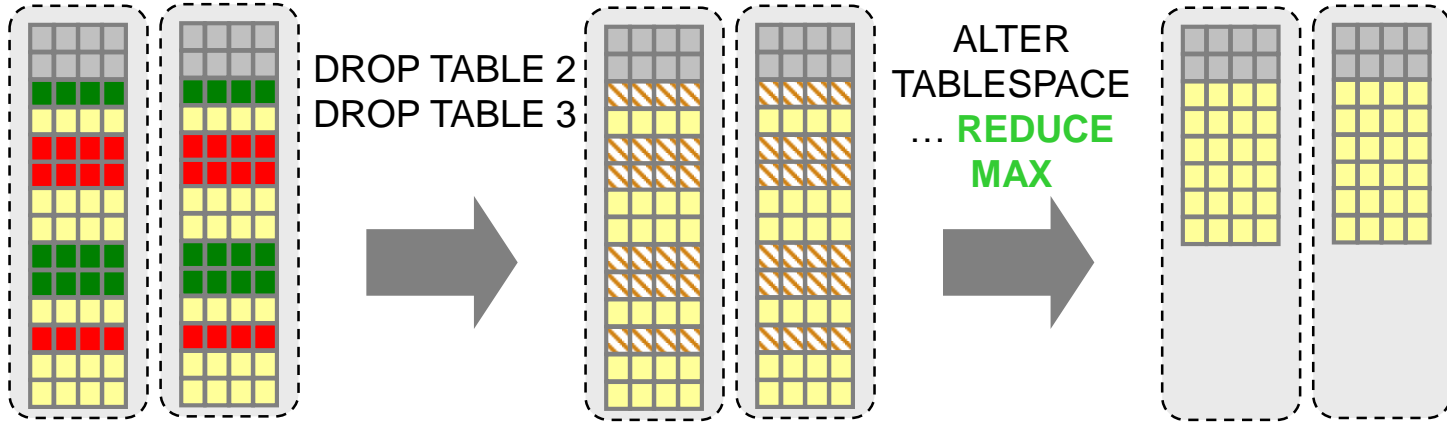
- HADR Read on Standby (RoS) enhancements








- Support for storage devices using 4K sector size
- Preview



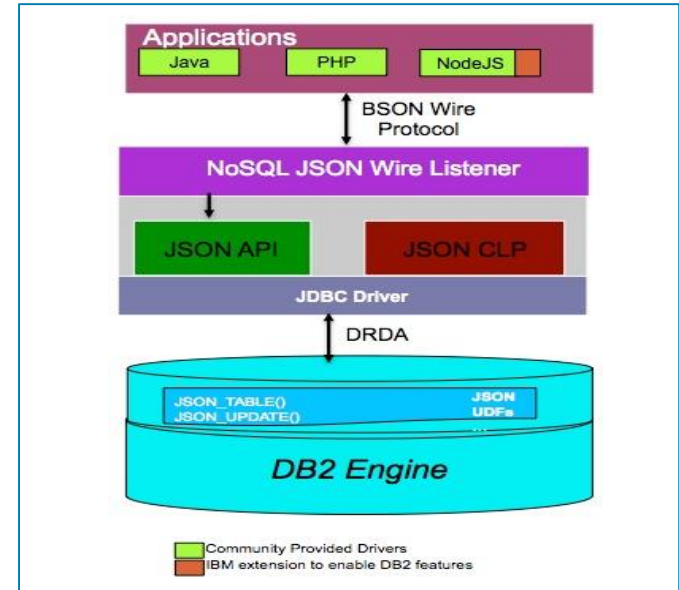
Db2 11.1.4.4 - Extent reclaim support for pureScale



-  Internal table space metadata extents
-  Table 1
-  Table 2
-  Table 3
-  Extent that is allocated to a table space, but not to a table

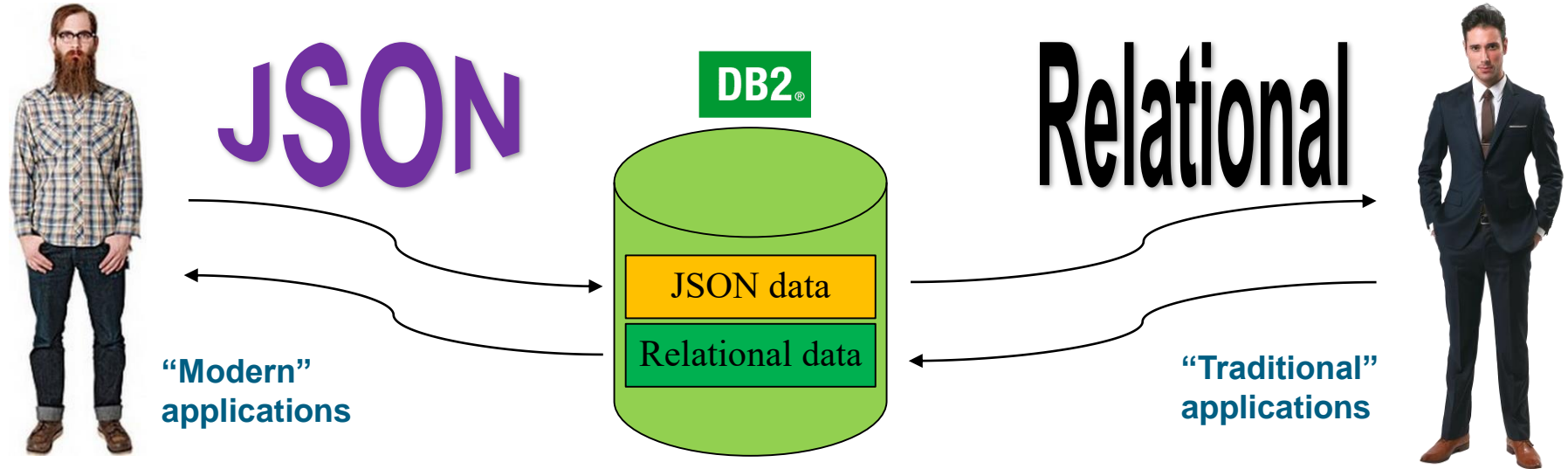
Db2 11.1.4.4 - A brief history of JSON Support (Part 1)

- **Db2 10.5 introduced JSON NoSQL support**
 - Focused on allowing Db2 to participate in the NoSQL paradigm
- **Customers asking for native SQL support of JSON**
 - Legacy applications accessing new JSON data
 - New JSON applications wanting to access legacy data
- **INDEX ON EXPRESSION allows for fast performing queries against JSON data**



Db2 11.1.4.4 - A brief history of Db2 & JSON (Part 2)

Our customers also asked for native SQL support for JSON in Db2



As a tactical response, we revealed our proprietary "SYSTOOLS" JSON SQL functions in DB2 11.1

Db2 11.1.4.4 - Proprietary JSON SQL functions

- To help simplify our customer's experience with the proprietary JSON SQL functions, we have now officially support them [Db2 11.1.2.2](#)
- The functions are automatically created in SYSTOOLS schema for a new database
- The functions are now documented in the knowledge center under a section called “SQL access to JSON documents”

Insert, Retrieve, and Document Validation

- BSON2JSON – Convert BSON into JSON
- JSON2BSON – Convert JSON into BSON
- BSON_VALIDATE – Ensure LOB is valid BSON

Field Retrieval

- JSON_VAL – Extract data from JSON into SQL data types
- JSON_TYPE – Returns data type of a JSON field

Array Retrieval

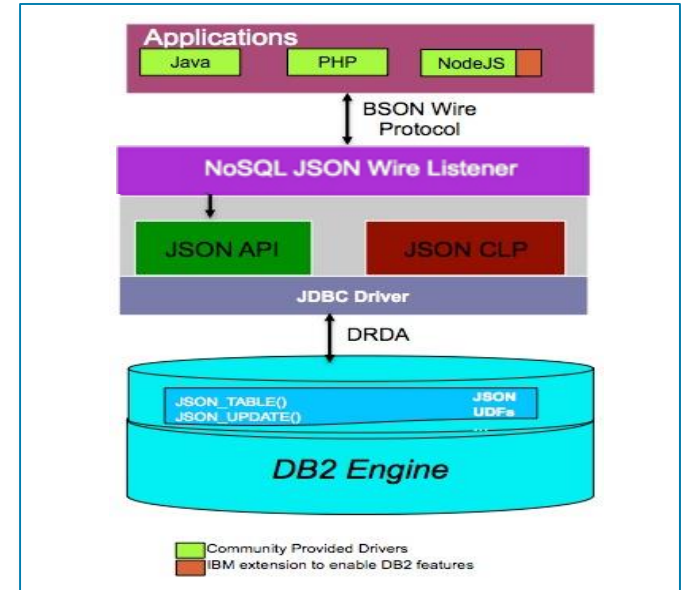
- JSON_TABLE – Returns a table of values from a JSON array
- JSON_LEN – Returns #elements in a JSON array
- JSON_GET_POS_ARR_INDEX – Find a value within an array

Document Update

- JSON_UPDATE – Update a field or document using set syntax

Db2 11.1.4.4 - Enhancements to **JSON NoSQL** support in **Db2 11.1.4.4**

- Db2 NoSQL JSON wire listener is enhanced to support latest mongo db client (3.6.3)
- Added support for Kerberos Authentication to wire listener
- Enhanced wire listener script to enable jcc and nosql trace using command line options.



Db2 11.1.4.4 - Enhancements to **JSON SQL** support in **Db2 11.1.4.4**

- **First wave implementation of new built-in JSON SQL functions based on ISO technical report for SQL support for JavaScript Object Notation (JSON)**
 - These JSON SQL functions follow a public “standard”
 - The (proprietary) SYSTOOLS JSON SQL functions will be de-emphasized but will continue to be supported

Schema	Name	Comments
SYSIBM	BSON_TO_JSON	Convert BSON formatted document into JSON strings
SYSIBM	JSON_TO_BSON	Convert JSON strings into a BSON document format
SYSIBM	JSON_ARRAY	Creates JSON array from input key value pairs
SYSIBM	JSON_OBJECT	Creates JSON object from input key value pairs
SYSIBM	JSON_VALUE	Extract an SQL scalar value from a JSON object
SYSIBM	JSON_QUERY	Extract a JSON object from a JSON object
SYSIBM	JSON_TABLE	Creates relational output from a JSON object
SYSIBM	JSON_EXISTS	Determine whether a JSON object contains the desired JSON value

Db2 11.1.4.4 - What makes these JSON SQL functions different/better?

▪ Easier to use

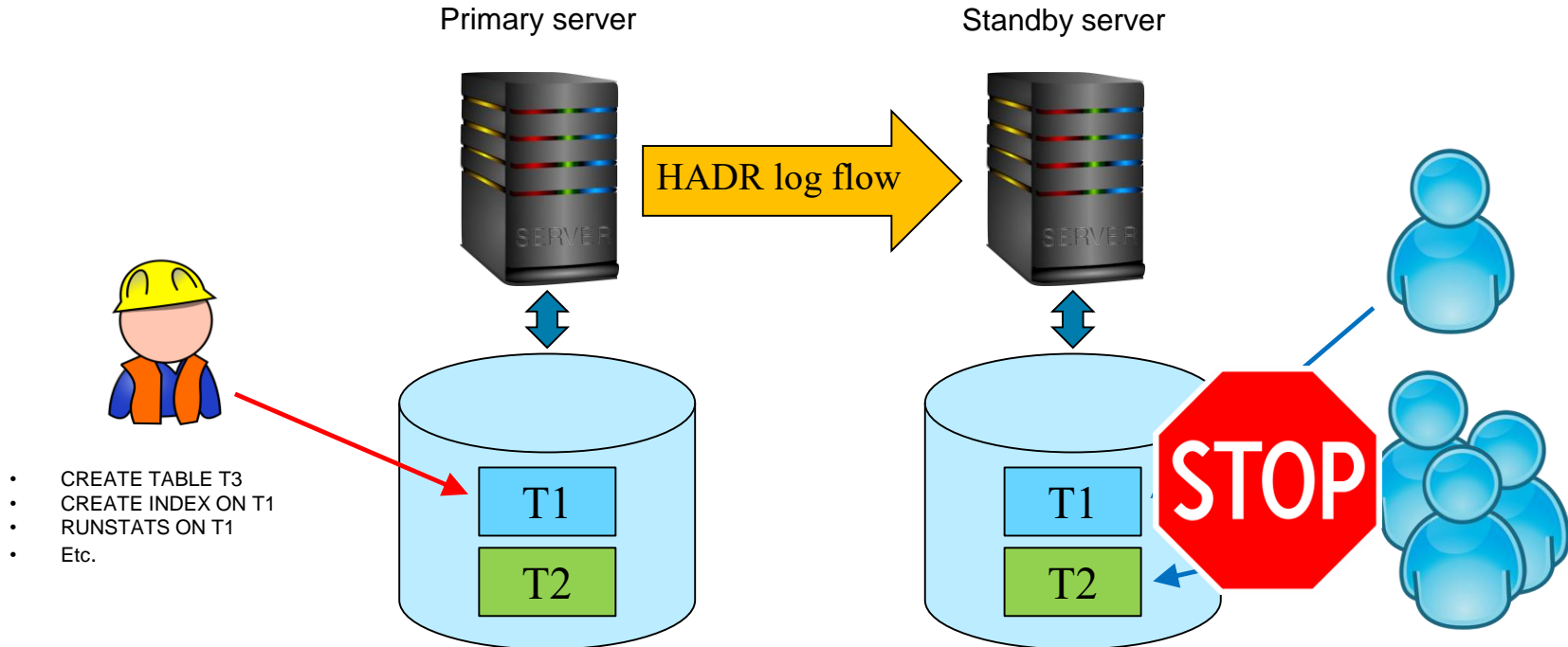
- No need to qualify call to function or add SYSTOOLS to function path
- No need to grant EXECUTE privilege

▪ Simpler and more flexible storage options

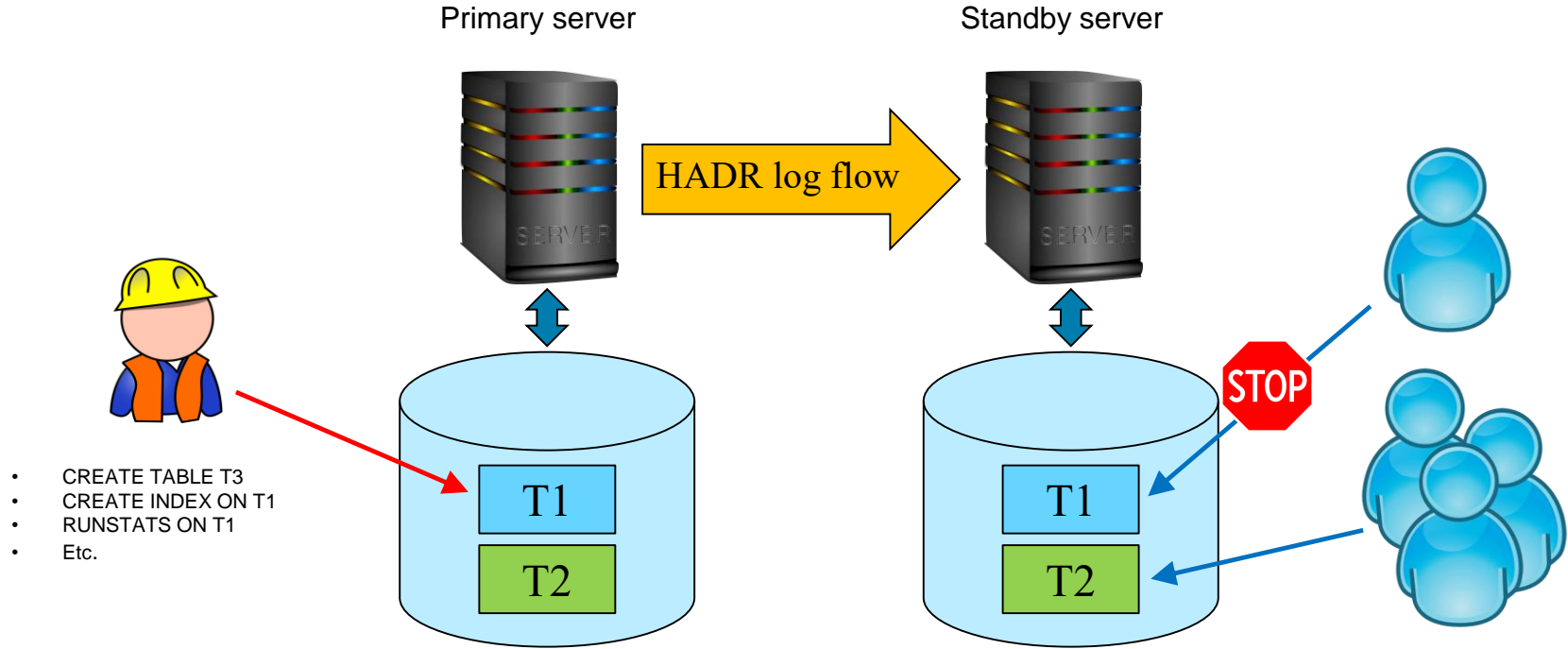
- You choose the stored format: JSON or BSON
- You choose the table organization: row or column
- You choose the column data type: BLOB, CHAR, CLOB, VARBINARY, VARCHAR

▪ BSON/JSON conversion functions are now optional not mandatory

Db2 11.1.4.4 - HADR Read on Standby (RoS) – Pre-Db2 11.1.4.4



Db2 11.1.4.4 - HADR Read on Standby (RoS) – New in Db2 11.1.4.4



*Not supported on pureScale at this time

Db2 11.1.4.4 - Support for storage devices using 4K sector size

- **Hard disk drive manufacturers have been moving the technology to a larger, more efficient sector size of 4096 bytes (4K sector)**
 - Legacy sector size (512 bytes) was supported through emulation and transition was transparent to disk users
 - Recent products now only offer 4K sector size support

- **This support can be enabled in Db2 by setting the DB2_4K_DEVICE_SUPPORT registry variable to ON**
 - Try it out and let us know!

- **Current limitations:**
 - This feature is not supported for production use at this time
 - This is a technology preview
 - Not available on pureScale
 - The use of DMS Raw containers is not supported
 - There may be a performance penalty accessing:
 - LOB data stored on 512-byte sector storage
 - Backup or load copy files created prior to the enablement of 4K device support

Db2 11.1.4.4 - Follow “Code Patterns” to Build Cognitive Applications

NEW!

CODE PATTERN | OCT

19, 2018

Train a cloud-based machine learning model using on-premise data

Get the Code



Analytics

Artificial Intelligence +

CODE PATTERN | OCT

22, 2018

Continuous learning with Watson Machine Learning and IBM Db2 Warehouse on Cloud

Get the Code



Apache Spark

Artificial Intelligence +

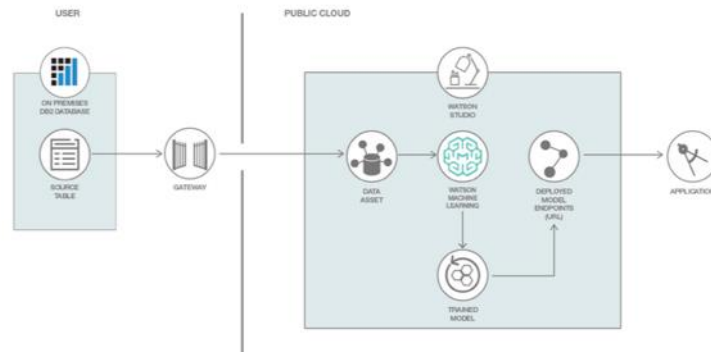
Train a cloud-based machine learning model from an on-premise database

build **fail0ng**

Many companies and individuals struggle to use their on-premises data — the kind of data that lives on a local machine, within your data center, behind your firewall — for machine learning in the cloud. It can be challenging to find a quick, easy, and secure solution for connecting resources in a protected environment to resources in the cloud.

With Watson Studio & Machine Learning, Db2, and Secure Gateway, it is possible to establish a secure, persistent connection between your on-premises data and the cloud to train machine learning models leveraging cloud computing resources like Spark, elastic environments, and GPUs.

In this guide, we will create an on-premises Db2 database on our local computer, populate it, and then connect it to Watson Studio via Secure Gateway. Then, we'll read Buildings Violations data from this database and build a model to predict the likelihood that a particular building will FAIL an inspection based on historical data from the City of Chicago. After we build the model, we will deploy it as an API endpoint with Watson Machine Learning that only authorized users can access.

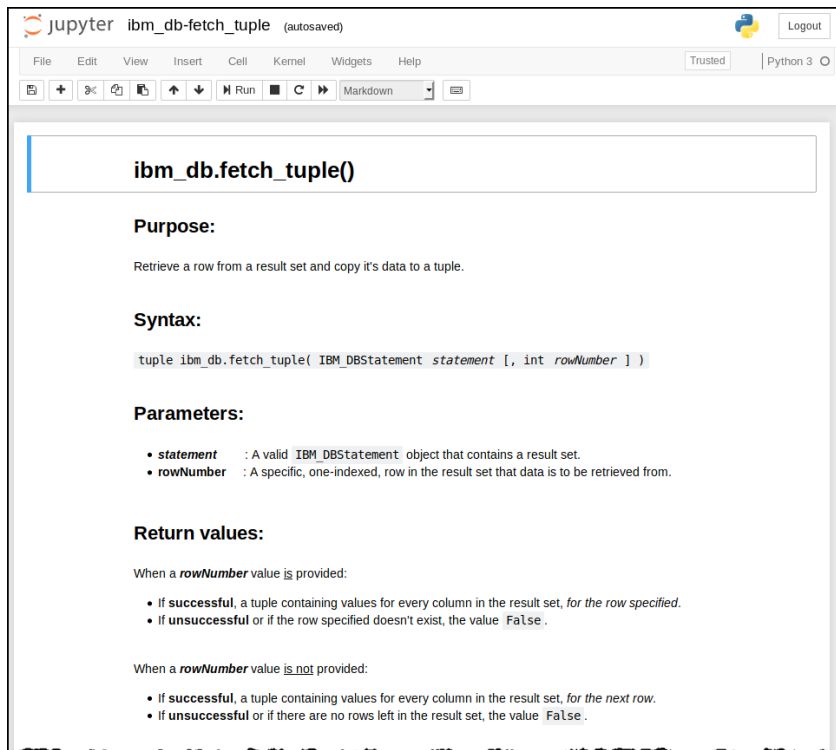


Db2 11.1.4.4 - An extensive set of Jupyter Notebooks

Kick start your Db2 development



NEW!



jupyter ibm_db-fetch_tuple (autosaved) Python 3

ibm_db.fetch_tuple()

Purpose:

Retrieve a row from a result set and copy its data to a tuple.

Syntax:

```
tuple ibm_db.fetch_tuple( IBM_DBStatement statement [, int rowNumber ] )
```

Parameters:

- statement** : A valid `IBM_DBStatement` object that contains a result set.
- rowNumber** : A specific, one-indexed, row in the result set that data is to be retrieved from.

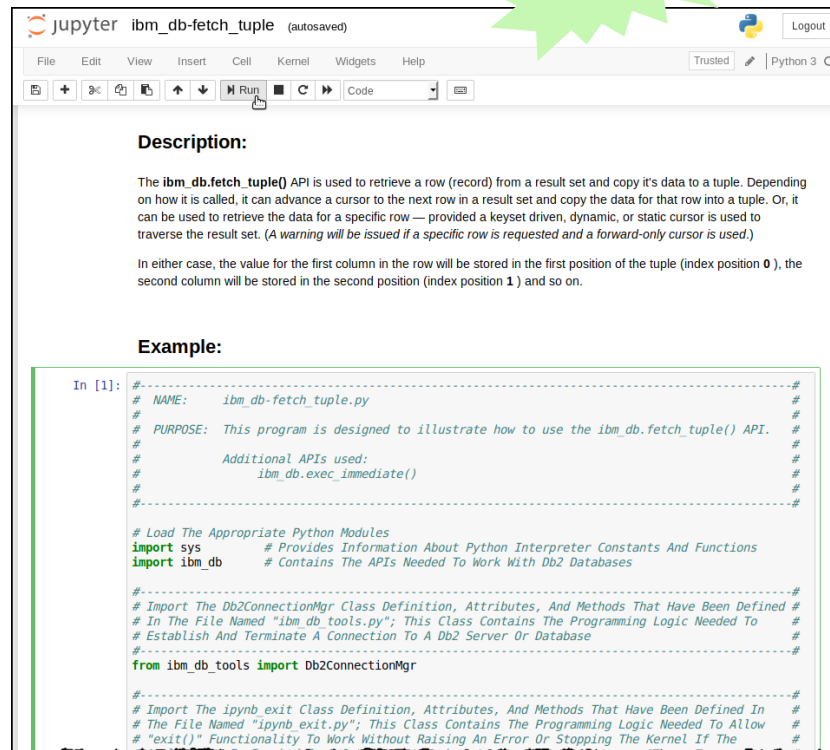
Return values:

When a **rowNumber** value is provided:

- If **successful**, a tuple containing values for every column in the result set, for the row specified.
- If **unsuccessful** or if the row specified doesn't exist, the value `False`.

When a **rowNumber** value is not provided:

- If **successful**, a tuple containing values for every column in the result set, for the next row.
- If **unsuccessful** or if there are no rows left in the result set, the value `False`.



jupyter ibm_db-fetch_tuple (autosaved) Python 3

Description:

The `ibm_db.fetch_tuple()` API is used to retrieve a row (record) from a result set and copy its data to a tuple. Depending on how it is called, it can advance a cursor to the next row in a result set and copy the data for that row into a tuple. Or, it can be used to retrieve the data for a specific row — provided a keyset driven, dynamic, or static cursor is used to traverse the result set. (A warning will be issued if a specific row is requested and a forward-only cursor is used.)

In either case, the value for the first column in the row will be stored in the first position of the tuple (index position 0), the second column will be stored in the second position (index position 1) and so on.

Example:

```
In [1]: #-----#
# NAME:      ibm_db-fetch_tuple.py
# PURPOSE:   This program is designed to illustrate how to use the ibm_db.fetch_tuple() API.
#           Additional APIs used:
#           ibm_db.exec_immediate()
#-----#

# Load The Appropriate Python Modules
import sys      # Provides Information About Python Interpreter Constants And Functions
import ibm_db   # Contains The APIs Needed To Work With Db2 Databases

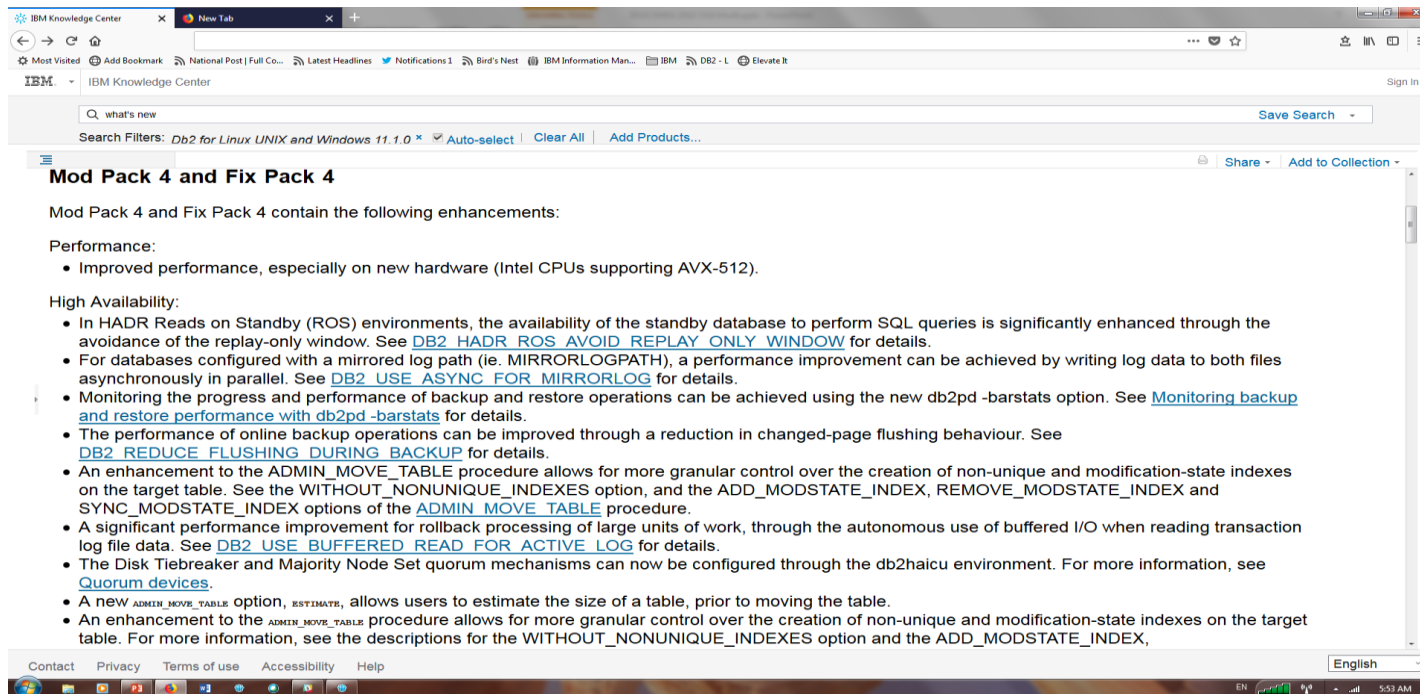
#-----#
# Import The Db2ConnectionMgr Class Definition, Attributes, And Methods That Have Been Defined #
# In The File Named "ibm_db_tools.py"; This Class Contains The Programming Logic Needed To #
# Establish And Terminate A Connection To A Db2 Server Or Database
#-----#
from ibm_db_tools import Db2ConnectionMgr

#-----#
# Import The ipynb_exit Class Definition, Attributes, And Methods That Have Been Defined In #
# The File Named "ipynb_exit.py"; This Class Contains The Programming Logic Needed To Allow #
# "exit()" Functionality To Work Without Raising An Error Or Stopping The Kernel If The #
```

Db2 11.1.4.4 - But wait, there's more!

Learn more about Db2 11.1.4.4 in the “Mod Pack and Fix Pack Updates” section

- https://www.ibm.com/support/knowledgecenter/en/SSEPGG_11.1.0/com.ibm.db2.luw.wn.doc/doc/c0061179.html



The screenshot shows a web browser window displaying the IBM Knowledge Center page for Db2 11.1.4.4. The page title is "Mod Pack 4 and Fix Pack 4". The content lists enhancements in Performance and High Availability.

Mod Pack 4 and Fix Pack 4

Mod Pack 4 and Fix Pack 4 contain the following enhancements:

Performance:

- Improved performance, especially on new hardware (Intel CPUs supporting AVX-512).

High Availability:

- In HADR Reads on Standby (ROS) environments, the availability of the standby database to perform SQL queries is significantly enhanced through the avoidance of the replay-only window. See [DB2 HADR ROS AVOID REPLAY ONLY WINDOW](#) for details.
- For databases configured with a mirrored log path (ie. MIRRORLOGPATH), a performance improvement can be achieved by writing log data to both files asynchronously in parallel. See [DB2 USE ASYNC FOR MIRRORLOG](#) for details.
- Monitoring the progress and performance of backup and restore operations can be achieved using the new db2pd -barstats option. See [Monitoring backup and restore performance with db2pd -barstats](#) for details.
- The performance of online backup operations can be improved through a reduction in changed-page flushing behaviour. See [DB2 REDUCE FLUSHING DURING BACKUP](#) for details.
- An enhancement to the ADMIN_MOVE_TABLE procedure allows for more granular control over the creation of non-unique and modification-state indexes on the target table. See the WITHOUT_NONUNIQUE_INDEXES option, and the ADD_MODSTATE_INDEX, REMOVE_MODSTATE_INDEX and SYNC_MODSTATE_INDEX options of the [ADMIN_MOVE_TABLE](#) procedure.
- A significant performance improvement for rollback processing of large units of work, through the autonomous use of buffered I/O when reading transaction log file data. See [DB2 USE BUFFERED READ FOR ACTIVE LOG](#) for details.
- The Disk Tiebreaker and Majority Node Set quorum mechanisms can now be configured through the db2haicu environment. For more information, see [Quorum devices](#).
- A new ADMIN_MOVE_TABLE option, ESTIMATE, allows users to estimate the size of a table, prior to moving the table.
- An enhancement to the ADMIN_MOVE_TABLE procedure allows for more granular control over the creation of non-unique and modification-state indexes on the target table. For more information, see the descriptions for the WITHOUT_NONUNIQUE_INDEXES option and the ADD_MODSTATE_INDEX,

Db2 “VNext”

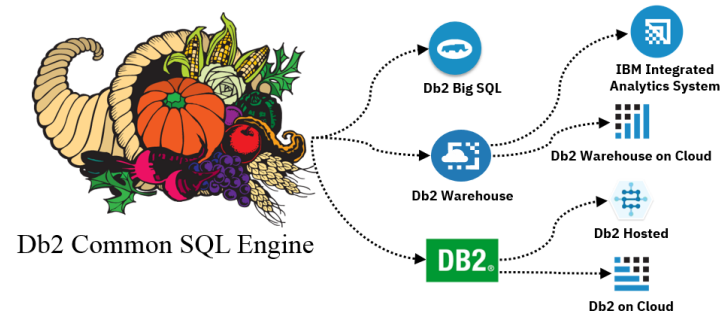
- **We plan to release a new version of Db2 in 2019**
 - Db2 11.1 will be 3 years old in 2019
 - Db2 10.5 is EOS as of April 30, 2020

- **We will continue to follow the “continuous delivery” paradigm started with Db2 11.1**
 - Steady release of function through GA and subsequent modification updates

- **It will be still be version 11 – with a new release number – Db2 11.?**

Db2 “VNext” GA – Key Objective

- With the introduction of Db2 VNext in 2019, all Db2 products will be based on the same level of the Db2 common SQL engine



- **Benefits for Db2 customers:**

- Existing enhancements in Db2 Warehouse branch will become available to them
- Reduced latency between feature appearance in different Db2 products as development will all be on same level
 - Any enhancement that is “ready to go” for Db2 environments will appear in next Db2 update

Db2 “VNext” GA - Summary

▪ **Planned content:**

- Functionality available in the Db2 Warehouse family today:
 - External tables
 - Grooming for columnar tables
 - Reduced logging for columnar tables
 - Additional enhancements of interest

- New functionality
 - Simplified log management
 - Enhanced support for Currently Committed (CC) isolation on pureScale
 - Support for LOBs in columnar tables

Db2 “VNext” GA - External Tables

- **A simple mechanism that lets you access external “files” within an SQL statement just like a table**
 - Can also be used to load from or unload to external files
 - Can be used to define a permanent external table or directly within a SQL statement

- **Currently supports CSV/delimited files and local/remote sources**

Example:

```
create external table ext_orders(order_num INT, order_dt TIMESTAMP)
    USING(dataobject('/tmp/order.tbl') DELIMITER '|');

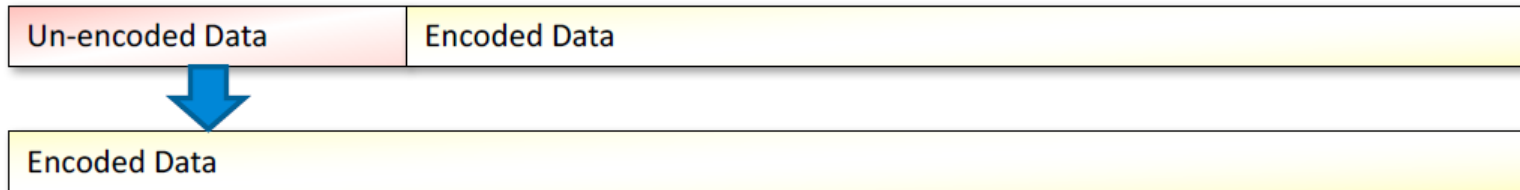
insert into orders (select * from ext_orders);
```

Example:

```
insert into orders (select * from external '/tmp/orders.txt' using(REMOUTESOURCE GZIP
delimiter \, '));
```

Db2 “VNext” GA - Grooming for Columnar Tables

- **When Automatic Dictionary Creation (ADC) is used, some portion of the data will be inserted before the dictionary is created**
 - Grooming is focused on these uncompressed first portions of the data
- **Grooming simply applies the compression dictionary to the first portion of the data that was uncompressed initially**



Db2 “VNext” GA - Reduced logging for columnar tables

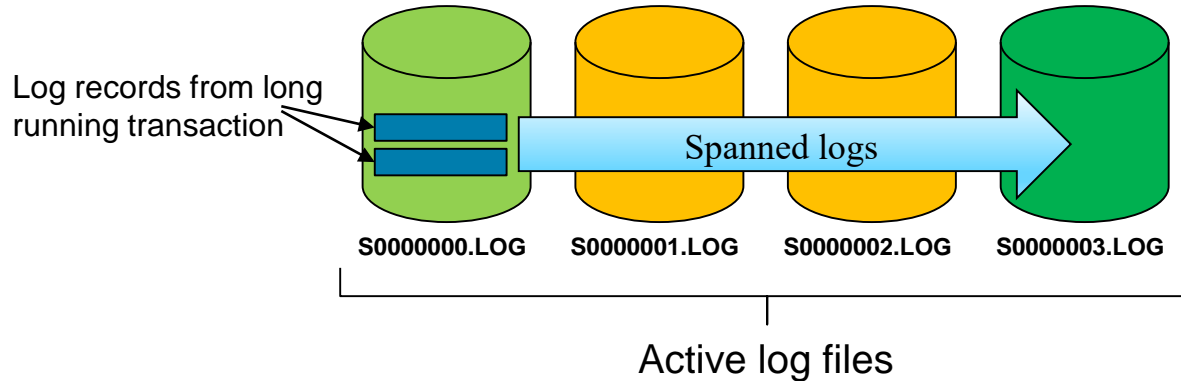
- **An enhanced solution to the same problem addressed by “Not Logged Initially” (NLI) tables**
 - Inserting large amounts of data into a table without swamping the transaction logs
 - Table contents are preserved during rollback, crash recovery, and roll-forward recovery to END of online backup
- **Unlike NLI, reduced logging is not implemented as a table property**
 - It is associated with any operation driving a large amount of insert work against a table and kicks in automatically after a certain amount of insert work has already occurred
 - Relevant operations include: INSERT, INGEST, MERGE, and UPDATE statements as well as CREATE TABLE AS DDL, etc.
- **With this behaviour active, very large INSERTs are possible without running out of log space**
 - Up to 95% less data logged!

Db2 “VNext” GA - Additional Db2 Warehouse Enhancements

- **Numerous performance and footprint enhancements for columnar table queries**
 - The result of many Db2 Warehouse and IAS customer engagements
- **Increased SQL compatibility for Netezza applications**
 - The result of building the IAS appliance
- **Built-in database support for spatial analytics**
 - Better performance from built-in spatial procedures, functions, and views

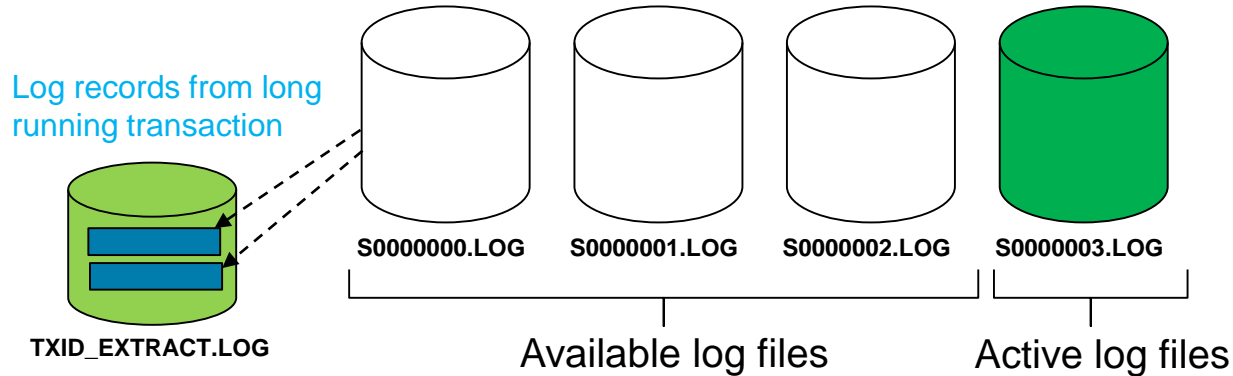
Db2 “VNext” GA - Simplified Log Management

- **As part of a focus on simplifying log management, the first delivery will result from a focus on avoiding “log full” scenarios caused by long running, low volume transactions hold up log space**
 - E.g. transactions that span multiple log files with little content in the intermediate files



Db2 “VNext” GA - Simplified Log Management

- **Solution is to extract log records for long running active transactions to a separate file and allow intermediate log files to be closed, archived, and reused**



Db2 “VNext” GA - Additional New Function of Interest

- **Support for LOB data type in columnar tables**

- Will probably be exposed in Db2 Warehouse first prior to Db2 VNext GA

- **Enhanced support for Currently Committed (CC) isolation on pureScale**

- No lock wait when log record on different member

Db2 “VNext” Modification 1 - Summary

▪ **Planned content:**

- Functionality available in the Db2 Warehouse family today:
 - Schema level authorization
 - Adaptive WLM

▪ **Plus a number of proposed items...**

Db2 “VNext” Modification 1 - Schema Level Authorization

Database

Administration = DBADM
Security = SECADM, ACCESSCTRL
Access = DATAACCESS
Load = LOAD

Schema

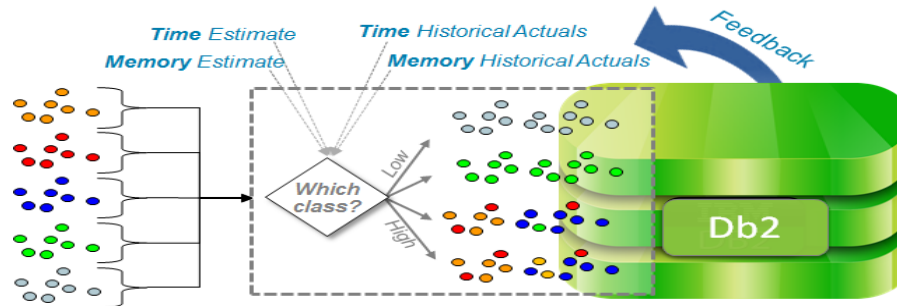
Administration = SCHEMAADM
Security = ACCESSCTRL
Access = DATAACCESS, SELECTIN, INSERTIN, UPDATEIN, DELETEIN
Load = LOAD

Object

Administration = Owner/CONTROL
Security = Owner/CONTROL
Access = SELECT, INSERT, UPDATE, DELETE
Load = LOAD

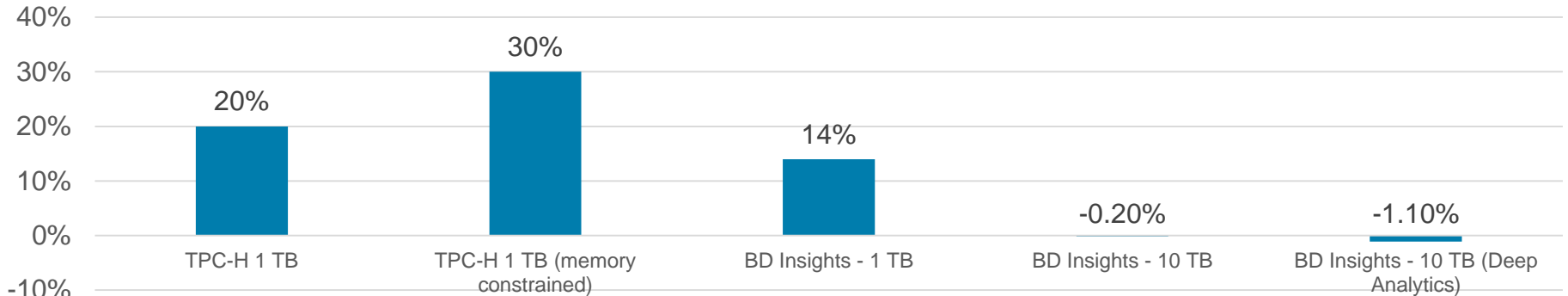
Db2 “VNext” Modification 1 - Adaptive workload management (WLM)

- The “next generation” of Db2 workload management (WLM) whose objective is to deliver automatic workload management within Db2 that ensures system stability and responsiveness with zero tuning
 - Make sure that the system is well-utilized but don’t let it become overcommitted
 - Schedule jobs appropriately to ensure fairness and appropriate responsiveness
- Adaptive WLM is currently being used as the default system behaviour within the entire Db2 Warehouse family including IAS
 - The ability to customize with user definitions to be introduced to Db2 Warehouse in 2019

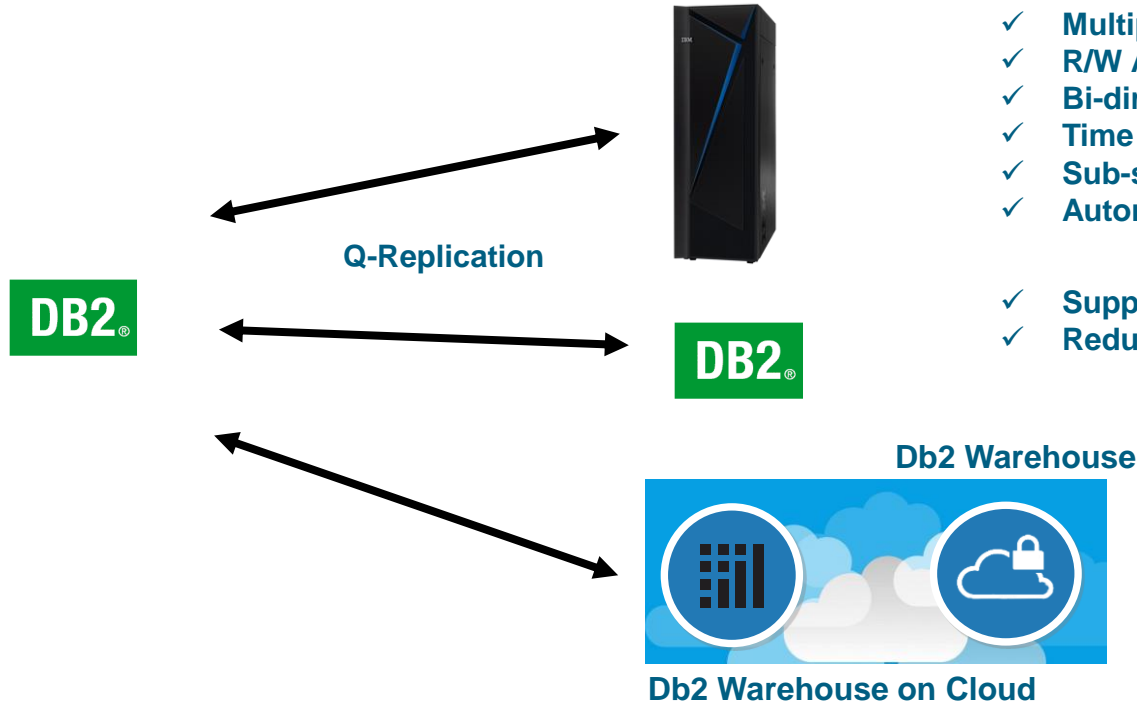


Db2 “VNext” Modification 1 - Adaptive WLM Performance Numbers

- **Analytical workloads (BLU) performed at par or better with Adaptive WLM compared to default WLM configuration**
 - Since work is admitted into the system based on matching of CPU and memory needs versus current usage, processing of work is generally more efficient due to smoother, steadier consumption patterns
 - System stability is maintained throughout



Db2 “VNext” Modification 1 - Disaster Recovery and Replication Solution



- ✓ Multiple stand-bys
- ✓ R/W Active-Active
- ✓ Bi-directional R/W Active-Active
- ✓ Time Delay
- ✓ Sub-setting
- ✓ Automatic Client Reroute

- ✓ Support for columnar
- ✓ Reduced logging is automatically handled

Note: Removal of bundled restricted licenses of IIDR occurs in “VNext”

Db2 “VNext” Modification 1 – Additional Content

- Completion of JSON ISO SQL function enhancements
- Support for triggers on columnar tables
- Support for system maintained temporal columnar tables
- And more being considered

Db2 “VNext” – Other Contributions

- **Adaptation and Integration of new technologies into the Db2 common SQL engine:**
 - Leveraging Machine Learning (ML) as part of Db2 optimization
 - Exploring Natural Language Query (NLQ) within Db2

- **Focus on making life better for application developers work with Db2**
 - Introduction of support for new languages
 - Introduction of server side REST API support
 - Better integration with common developer environments

Db2 “VNext” - Db2 Graph Capability

Real time visualizations and modelling



Coming

Support graph queries and SQL on the same data

- A graph layer between graph applications and Db2
- No change to Db2 runtime for relational data
- Support open-source graph with Gremlin language & Tinkerpop framework
- Store graph data in tables, so that SQL engine can query them
- ACID transactions update the graph real-time without disturbing (huge number of) existing relational applications
 - Ability to view the graph real-time (as and when transactions happen)
- Also provide a graph view of data with Gremlin API
- Run SQL analytics on graph data, as SQL is fully operational on graph data

Db2 “VNext” - Natural Language Query (NLQ) Support

New tool add on for data analysis



New!

*Eliminate the
need to know
ANSI SQL syntax*

- Conversational analytics is the future...
 - Alexa, Google Home, etc. prove trend
 - Natural extension of conversation to data
 - It is fast, accommodative, and can work with any schema / table
 - Covers most use cases for business users
- Free tool available to download from Developer Works
 - Currently in Beta
 - Seamlessly plugs and plays with Db2

Db2 “VNext” - Augmented Data Exploration Support

New tool for rich and deeper insights



New!

*Get to the
insights faster –
explore
unfamiliar data
sets*

- Data scientists often face the challenge of exploring data sets that are not familiar to them
 - Time consuming
 - Not sure what to look for
 - Need some knowledge transfer or understanding of what is in the data set
- Need an intuitive data exploration tool
 - Provide multiple views of the data with minimal touch
- Free tool available to download from Developer Works
 - Currently in Beta
 - Seamlessly plugs and plays with Db2

Beta - Try out the new Db2 Augmented Data Explorer

<https://www.ibm.com/us-en/marketplace/db2-augmented-data-explorer>

Augmented Data Discovery

Q

×

Steps to get started

*Set up
your database*

Tell us where to find the data that you want to discover. [Database Setup](#)

*Ask questions
to your data*

Using the search bar, start asking questions containing column names or field values

“ Average Sales by Region ”

“ Average Profit for Northwest region where Sales > 10,000 ”

“ Max Profit for all region and department ”

*Discover
and Refine*

Discover through Visualization and Natural language insights

Refine your search through the interactive pills in the search bar

Beta – Try out the new Db2 Augmented Data Explorer

<https://www.ibm.com/us-en/marketplace/db2-augmented-data-explorer>

Table: Monthly Sales 2018 - Last Cached: 2 days ago

Sales (Avg) by Product

Average Sales by Product by Region

Average Sales by Product by Age

Average Sales by Product by Gender

Quick View

245K
Avg Sales

1.3 % Increase
Month over Month

Sales Contribution by Product

9,204

Avg Sales by Product

Avg Sales of my Products by customer age and customer satisfaction

No match found

Answer known and unknown questions

No single data set can answer your question. Create one now ?

Create

Quick View

You might be interested in this ?

- [Average Sales by Product by Region](#)
- [Average Sales by Product by Gender](#)
- [Average Sales by Product by Year](#)

Sales Contribution by Product

9,204

Avg Sales by Product

More Information - Follow our plans using Aha

<http://ibm.biz/AnalyticsRoadmaps>

- **We revisit development priorities frequently (e.g. every quarter) in response to customer and market demand/feedback**
 - As a result: some items move up, some down, some in, and some out.
- **We have committed to keeping our core roadmaps visible to the public eye using Aha**

Db2 Roadmap

Distributed Platforms



More Information - Try out the new capabilities before they GA!

<http://ibm.biz/DB2-EAP>



Search



IBM Analytics > Technology > Data and content management > Db2 > Early Access Program >

Db2 Early Access Program

↓ Download Now

The Early Access Program offers a sneak peek at next generation IBM Db2 technology. Take this opportunity to shape the future of Db2, test integration with your existing infrastructure, and be part of the community that makes Db2 great.



Shape the future of Db2

Provide feedback and shape the world's most advanced database software. We are constantly improving and innovating new Db2 capabilities for the future of enterprise. You

Be Part of a Community

Join our community of business leaders, developers, and enthusiasts that is the Db2 Early Access Program community. Discuss ideas, share knowledge, and help each other out.

Gain Early Access

Be one of the first customers to use next generation Db2 software. Whether you are upgrading from Db2 10.x or 9.x, or have purchased Db2 before, this is something for you.

More Information - IBM Db2 Developer Community Edition

<http://ibm.biz/Db2devc>

- **Docker-based offering for small PoCs and “trial” production systems**
 - Also offered as a traditional Db2 installation: Db2 Developer-C Community Edition

- **Highlights:**
 - Free, fully-functional version of Db2
 - Includes all Db2 features such as compression, BLU Acceleration
 - Supports all Db2 configurations including pureScale and DPF
 - You can use it for development, test, or production !

- **Primary restrictions:**
 - No support/warranty
 - Environment limited to:
 - 4 cores with 16GB of memory
 - 100GB per database



More Information - Stay current with Db2 updates

▪ Passive

- Browse the list of available fix packs
 - <http://www-01.ibm.com/support/docview.wss?rs=71&uid=swg27007053>
- Security Vulnerabilities, HIPER and Special Attention APARs fixed in DB2 for Linux, UNIX, and Windows Version 11.1
 - <https://www-01.ibm.com/support/docview.wss?uid=swg21994955>

▪ Proactive

- Go to the IBM Support page and sign up for “My notifications”
 - <http://www-01.ibm.com/software/support/einfo.html>

Sign up to stay informed of critical IBM software support updates with **My notifications**.

- Take a proactive approach to problem prevention.
- Receive support content tailored to your needs, delivered directly to you!
- Receive immediate notifications of Security Bulletins and Flashes.
- Receive daily or weekly notifications of technical support information such as downloads, tips, technical notes, and publications.

Questions?

Les King
Director, Hybrid Data Management Solutions
lking@ca.ibm.com
ca.linkedin.com/pub/les-king/10/a68/426
March 28, 2019

Hybrid Data Management Strategy and Db2 Roadmap

