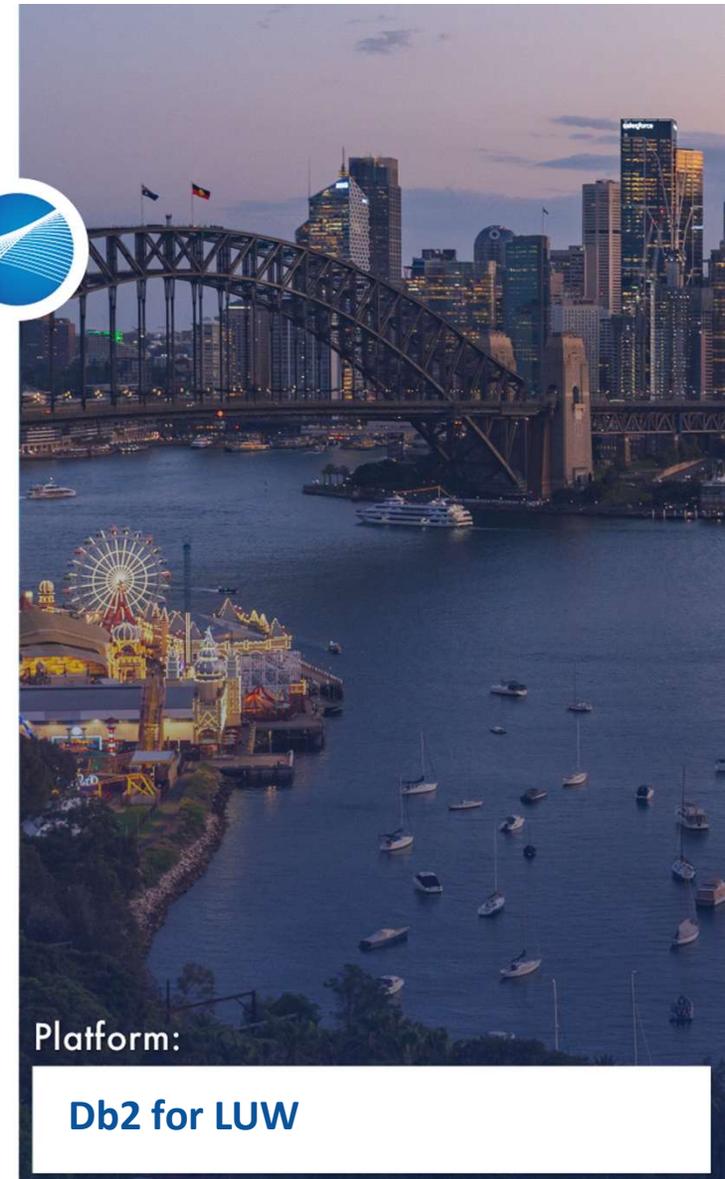IDUG

2026

Sydney | March 16 - 18

AU Db2
TECH
CONFERENCE

Db2 12 Latest News

David Kalmuk, *IBM*

Session Code: B03

Platform:

Db2 for LUW

AGENDA

- Db2 12 Highlights

- App Developer Updates

- Warehouse / Lakehouse Updates

- Tooling Updates

Db2 12 brings significant enhancements to Db2

Integration HA improvements: pureScale, DPF, HADR

Complete replacement of TSA with Pacemaker technology for cluster management in Linux, leading to significantly faster failure recovery times and better user experience

AI–powered query optimizer

Allows Db2 to continuously learn from customer's queries and achieve up to 3x query performance improvement over prior version

Name space separation with TENANT construct

Create a logical separation between one or more database schemas, easily isolating differing sets of tables from each other

Db2 infused with Generative AI

We're adding Gen AI capabilities to Db2. DB Assistant and more coming in 2025.

– **Improvements to backup performance** by initiating multiple threads to process a single table space

– **Mac M1/M2 driver support** for developers on macOS using Apple Silicon chip

– **HADR upgrade leverages RoS** for enhanced availability

– **Db2 pureScale HADR drop membe support (12.1.3)**

– **ADMIN_MOVE_TABLE** performance and availability enhancements

– **Vector Type (12.1.2/3)** Store AI Embedding and perform Vector Similarity Searches

– **Security enhancements** with AUDIT exceptions, Trusted Context and data masking (12.1) Disable World Writeable Files (12.1.2)

– **Data Mask at Read (12.1.2)**

– **Schema evolution with DROP and RENAME (12.1) ALTER DECIMAL PRECISION (12.1.3)** support for online schema updates to columnar tables

– **UPDATE and JOIN** performance enhancements for columnar tables

– **Federation enhancements** with support for Snowflake, Oracle 23c and performance improvements

12.1.3 Engine

- Vector Type Completion
- UUID Scalar Function
- DBMS_LOCK
  - update to leverage internal Db2 locking and support additional lock modes (transaction scoped)
- Access Plan Stability
- Columnar: Alter Decimal Precision
- Reduce the need to retrieve active logs files to include into an online backup

pureScale:

- remove SCP dependency
- move db2locssh out of /var
- Quiesce Tablespace (read-only tablespace)
- HADR Drop Member
- EFA performance improvement

Application Development

- Add support for Mongodb 8.x clients in IBM NoSQL/JSON Wire Listener
- Add support for keystash and keydb info into ldap registered nodes
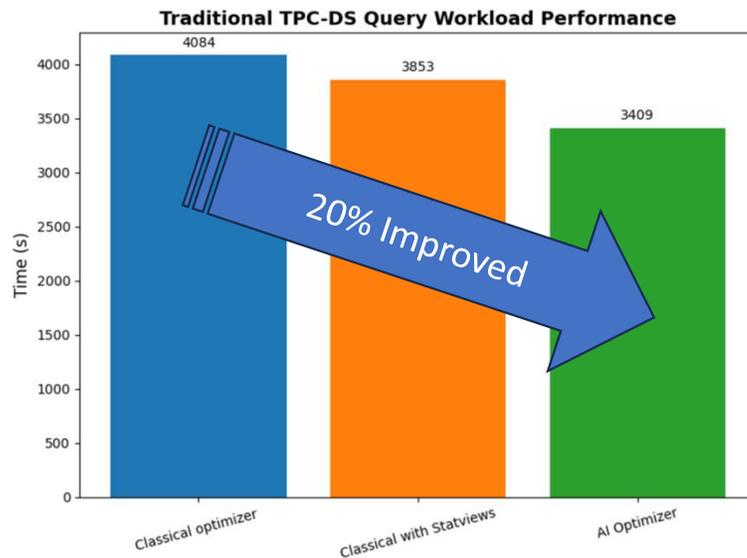
12.1.3 Container

- Support for Datalake table feature on Db2 Warehouse on Cloud Pak for Data
- Add support for external Iceberg REST Catalog for DATALAKE Tables
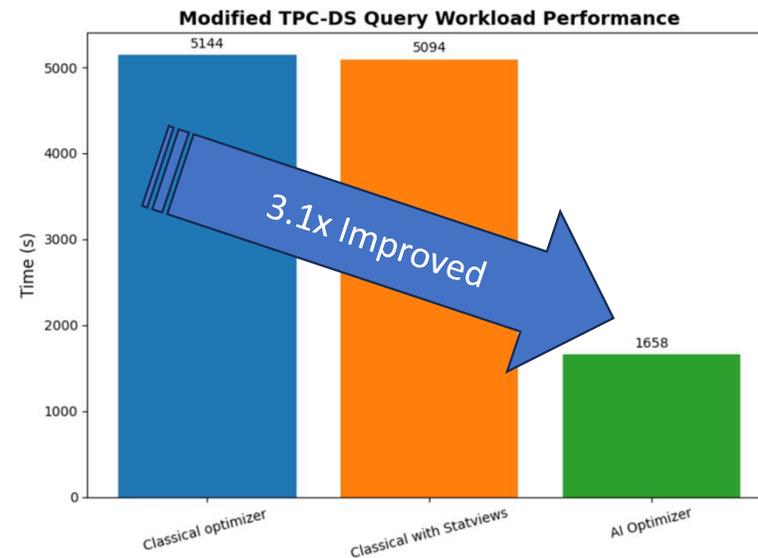- Support restricted-v2 SCC for Db2 standalone

Db2 12.1.3 – Performance Highlights

AI Optimizer JOIN Enhancements

Default 99 DS Queries

Traditional TPC-DS Query Workload Performance

20% Improved

"Join Intensive" Modified DS Queries

Modified TPC-DS Query Workload Performance

3.1x Improved

TPC-DS Workload Configuration

Scale Factor: 3TB. TPC-DS is a highly studied and tuned workload.

Traditional Queries: normal 99 TPC-DS query set

Modified Queries: based on original queries with more joins, different driving fact tables, and more
                 number of tuples being processed by certain key joins

PureScale Batched XI (Improvement for EFA)

📌 Background

In version **12.1.2**, **EFA performance plateaued** at ~150 concurrent clients. Although an optimization was identified, it was deemed **too risky** to include late in **mod2**.
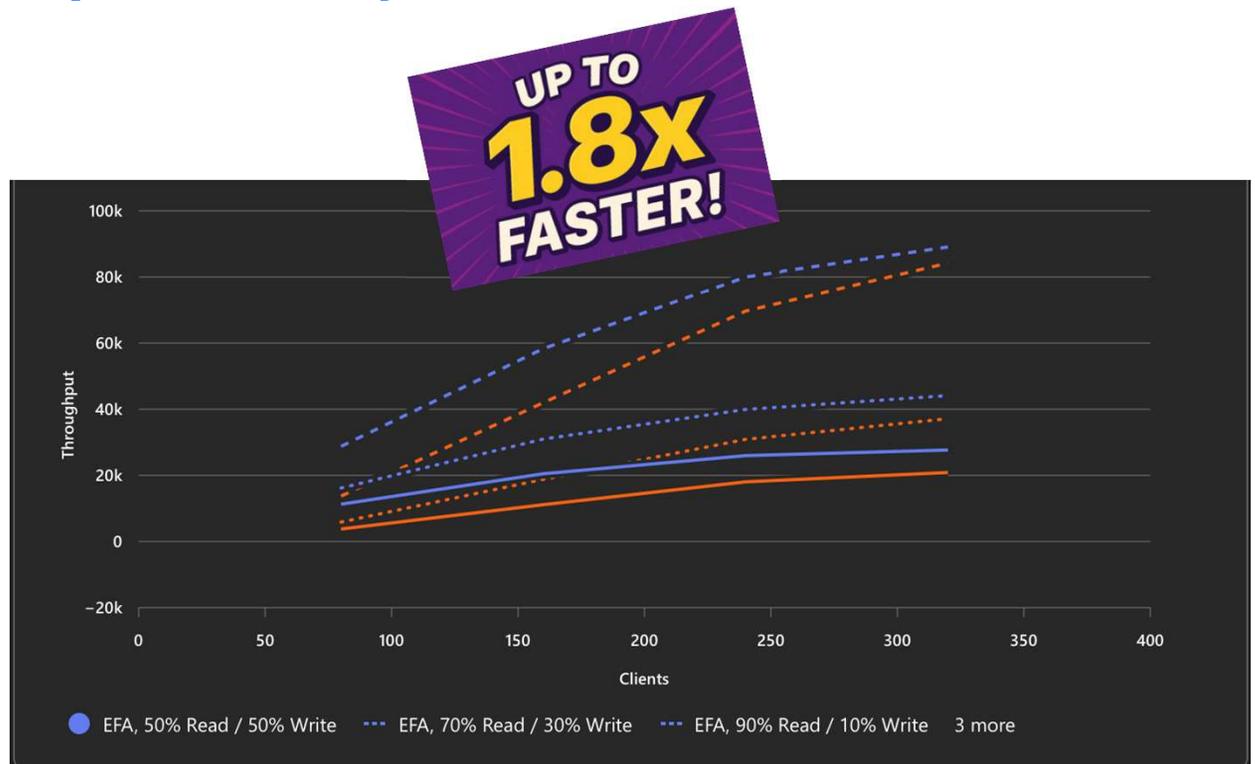

✅ Mod3 Batched XI Enhancement

🔧 **Optimization Implemented**
The optimization was successfully integrated into **mod3**.

📦 **Batched XI Improvements**
Throughput gains ranged from **1.1× to 1.8×**, depending on **Read/Write Ratio** & **Client Concurrency**
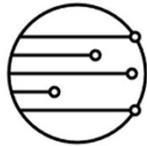
Db2 SaaS
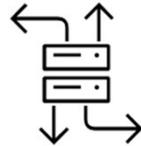## on IBM Cloud, AWS, Azure[new]

Fully Managed

Managed by Db2 experts so you can focus on what matters

Blazing Fast

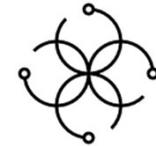Columnar-organized, memory-optimized data warehouse

Scalable & Elastic

Scalability through on-demand compute and cloud object storage

Durable & Reliable

Industry-leading continuous availability and disaster recovery

Open

Support for open formats and integration with watsonx.data

Db2 Priorities for 2026

Db2 Software

In-engine AI capabilities to help customers accelerate the creation of their AI apps

Db2 12.1.4 (Q1) **& 12.1.5** (Q2)

Additional capabilities to support AI builders with Db2, including:

- Vector indexing for performance
- LangChain, LlamaIndex, MCP Server for agentic AI
- Integration with watsonx.ai for in-Db2 LLM access
- Data Lake Integration within traditional software installations
- Columnar schema Evolution: Completion of alter data type
- More coming in 2H

Db2 SaaS

Expansion of fully-managed cloud services to Azure & GCP

Azure SaaS Db2 & Db2 Warehouse as fully managed services

Google Cloud BYOC Db2 & Db2 Warehouse delivered via bring-your-own-cloud model

Db2 Serverless (2026) Innovative deployment model for Db2 to support agentic development workflows

Db2 Tools

Products that help customers get the most of their Db2 investment

For DBAs: Agentic AI with Db2 Genius Hub (2026)

Stay tuned for future announcements

For Developers: VS Code Plugin (2026)

A completely rewritten plugin for Db2 providing key capabilities for developers
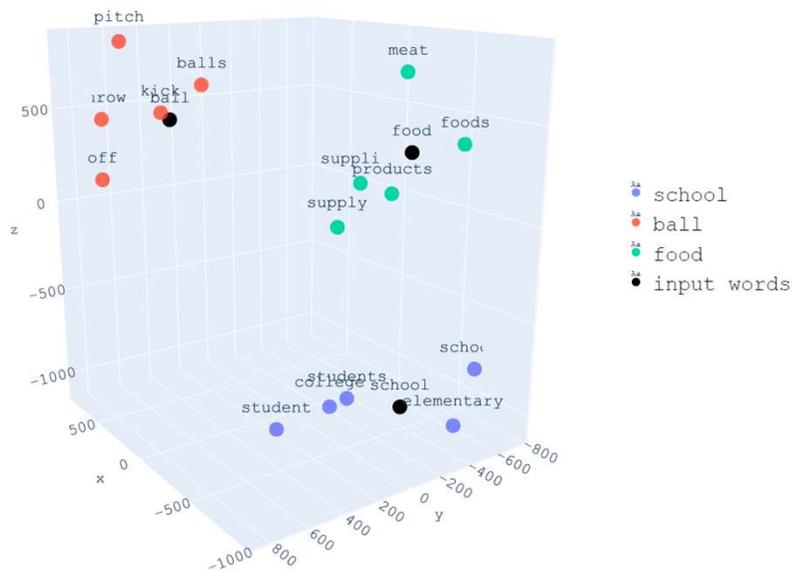
Analytics Appliances

Modernization of analytics appliance customers to next-gen SaaS or hardware

Db2 Warehouse SaaS + Db2 Warehouse on Power11

Default solution for IIAS/Sailfish customers modernizing to next-gen analytics platform, with choice of SaaS or on-premises Power11 infrastructure

You can vectorize almost anything

Eg. people, place, products, text, images, transactions

GenAI:
- Semantic document search for RAG

Financial:
- Fraud detection
- Credit risk assessment

Retail:
- Product recommendation
- Customer segmentation

Manufacturing:
- Anomaly detection
- Visual inspection

Government:
- Crime analysis
- Document similarity analysis

Healthcare:
- Personalized medicine
- Medical diagnosis

VECTOR type:
- FLOAT32 vectors
- INT8 vectors
- Max dimension:
  - 8168 (FLOAT 32)
  - 32672 (INT8)

VECTOR Functions:
- VECTOR_DISTANCE – between vectors
- VECTOR constructor: string -> vector
- VECTOR_SERIALIZE – vector -> string
- *VECTOR_NORM* – vector magnitude
- VECTOR_DIMENSION_COUNT – number of dimensions

LLM Apps Dev Support:
LangChain Connector

Vector Similarity Search Indexing Insights #1

- Db2 Vector Indexing <u>search performance improves as data set grows</u>
  - ✗ Exact search grows linearly with data set size, as it uses brute force approach that needs to scan all data and compute all distances
  - ✓ Vector Search grows logarithmically with data set size, as graph navigation absorbs data set growth, reducing the number of random IO operations and distance computations.
- Approximate search accuracy remains stable, with minor degradation as data set grows
- Resource usage limited to 5% of data set size
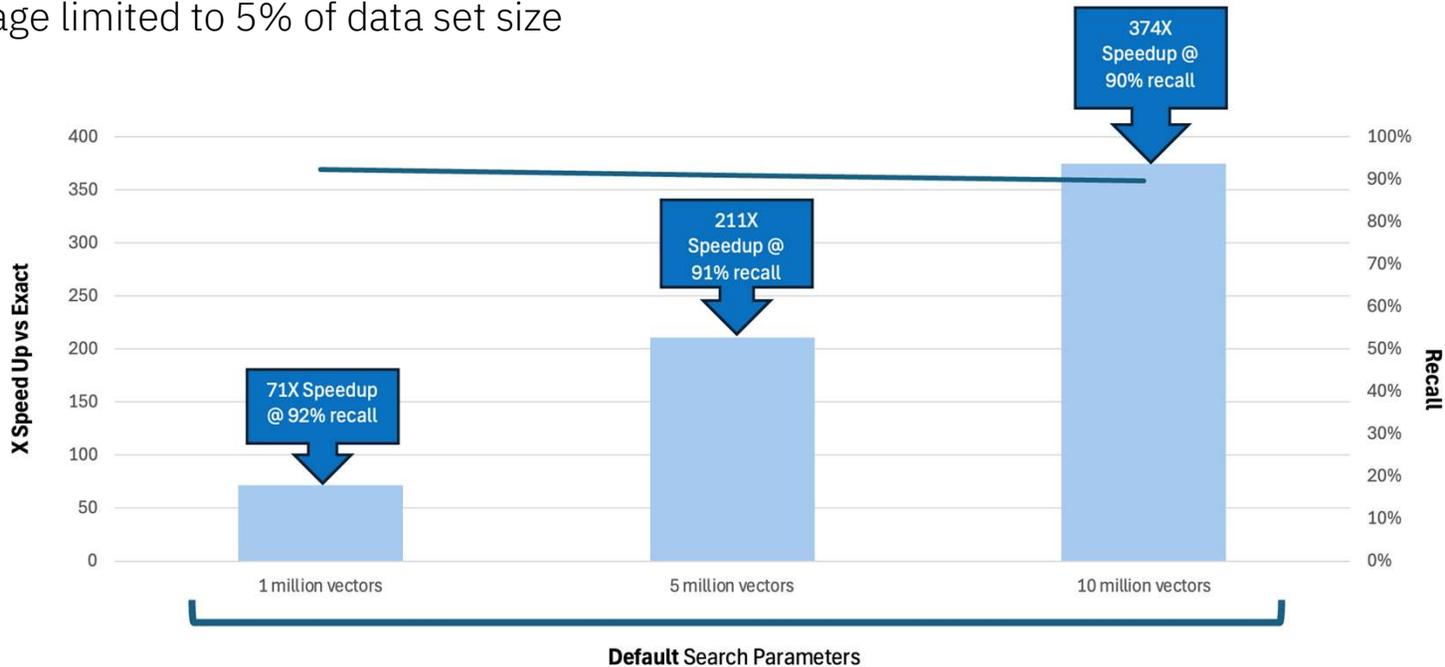
Entity Customer data set: 1, 5 and 10 million vectors, 768 float32 dimensions, embeddings generated using SLATE-125 model from JSON documents sourced from TPC-DS relational tables

SQL Based LLM Ecosystem Integration

SQL query with LLM functions

result

Db2

REST call

Response

LLM
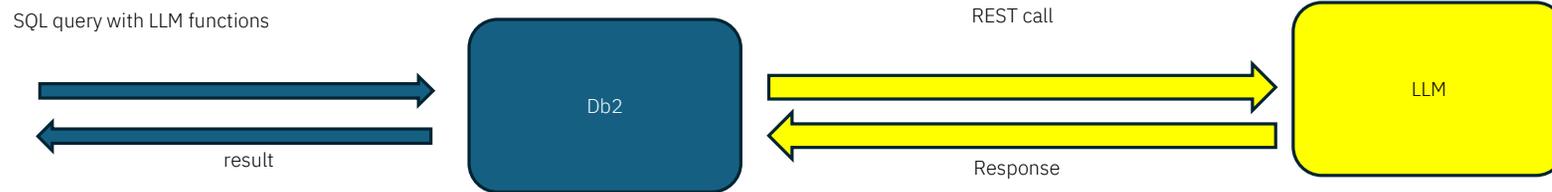
Embedding Generation

```
UPDATE DOCUMENT SET V1
= TEXT_EMBEDDING(doc,
TEXT_EMBEDDING_3_SMALL
)
```

Prompt

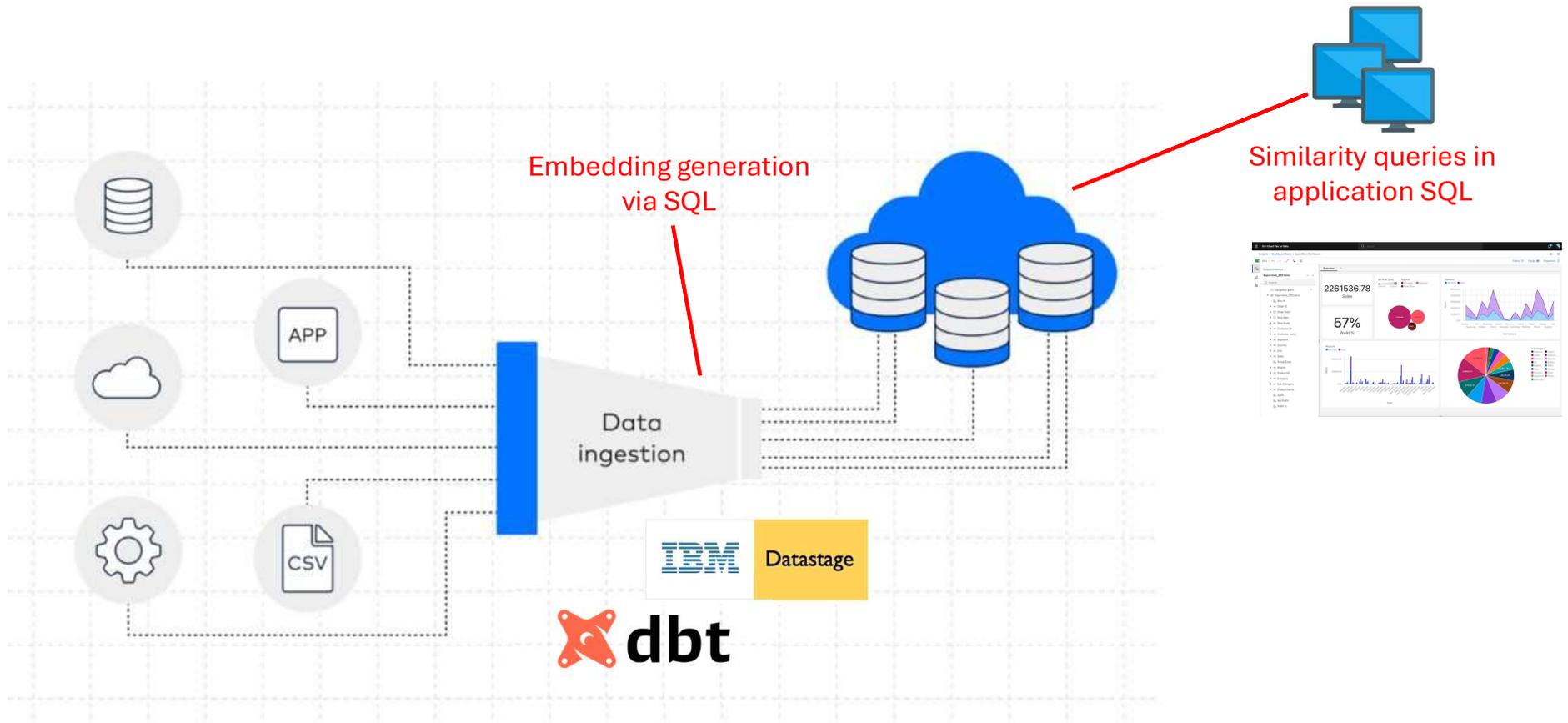Text summarization, text analysis, synthetic data generation

```
SELECT
TEXT_GENERATION('summa
rize the following
document
in 50 words: ' || doc,
MIXTRAL_8X7B_INSTRUCT_
V01) FROM DOCUMENT
```

Embedding Models

→ **watsonx.ai** models

→ **Built-in models** small language and text models

→ **3rd party providers** such as OpenAI, Cohere, more.

Hybrid Search

→ **Dense + Sparse** vector models

→ **Reranking Functions** for blending dense + sparse search results

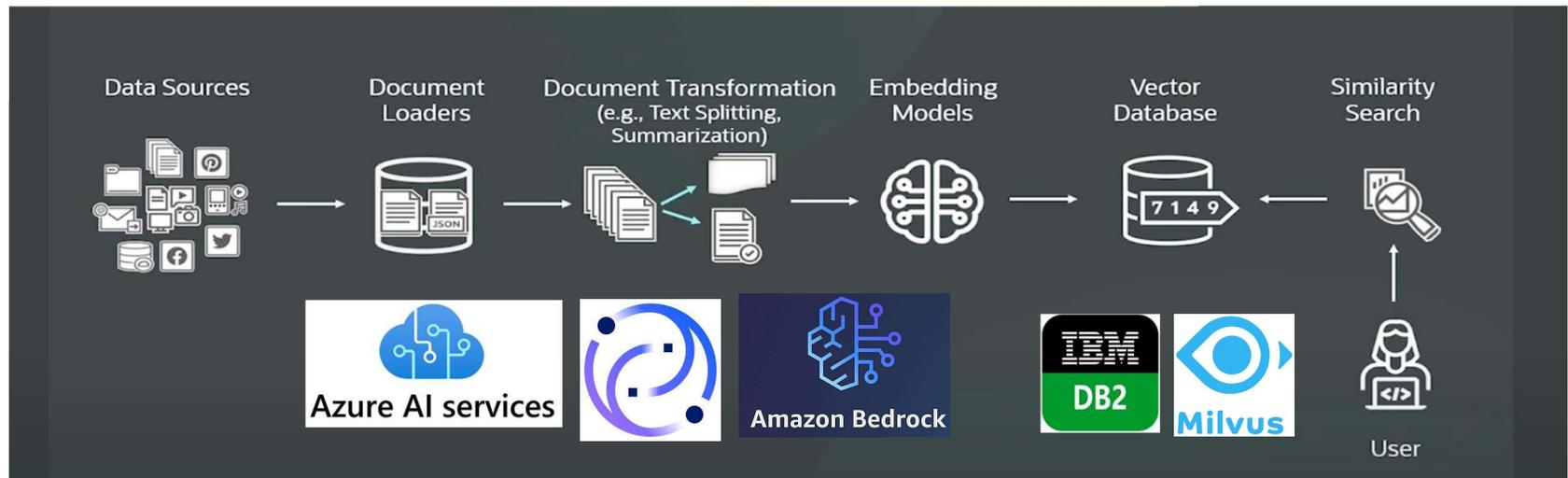→ **Models** such as RRF, Voyager

SQL App Dev Ecosystem

Embedding generation via SQL

Similarity queries in application SQL

Python App Dev Ecosystem

LLM Applications

RAG

Agentic RAG

Agents

LlamaIndex + Db2 →

IBM Db2 Vector Store for LlamaIndex

IBM Db2 for Developers
Elevating developer experience

Speed to productivity

Easy debugging /observability

Confidence in changes

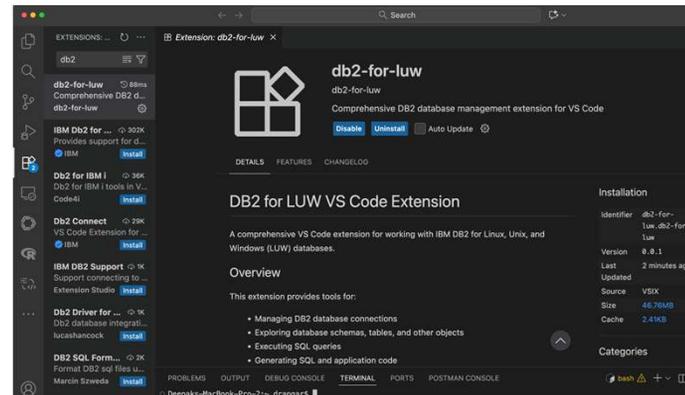**Toolchain integration**

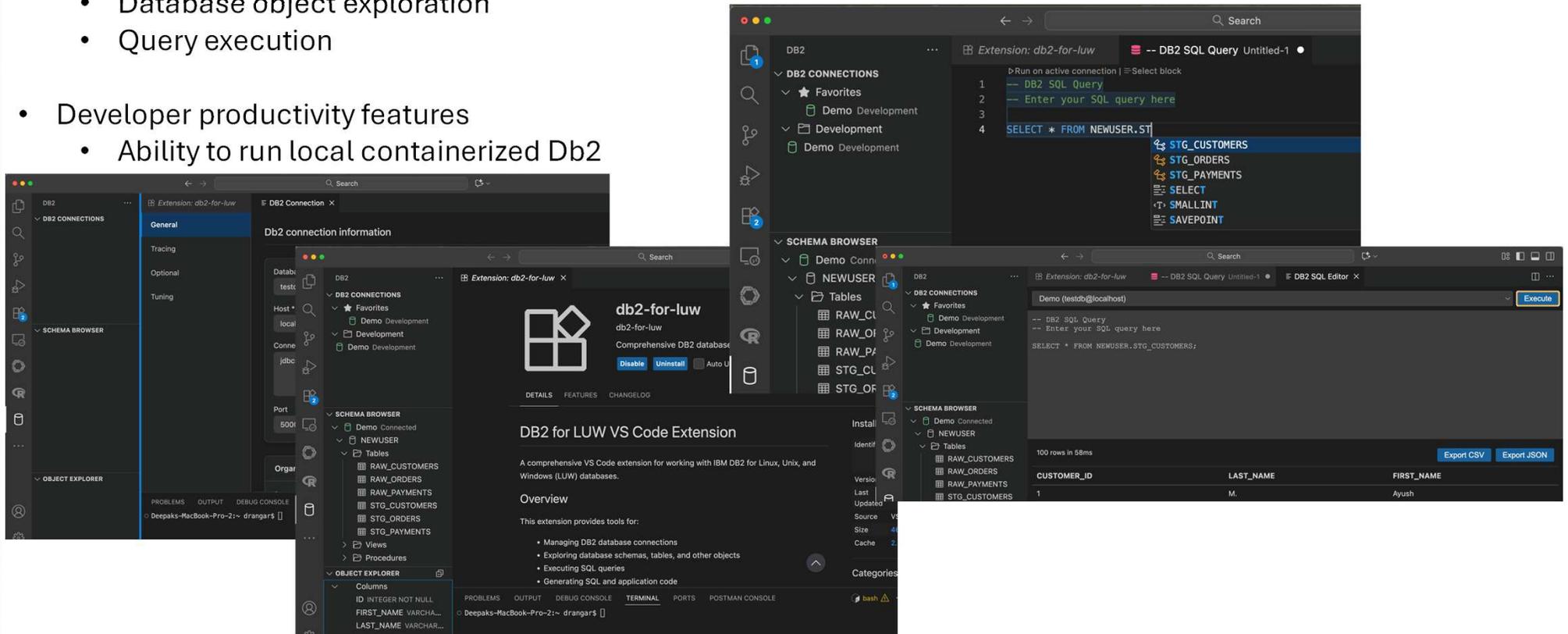**Fast, Local, Reliable:**
Db2 Community Edition for Developers

Db2 Meets VS Code:
Simplifying Database Development

Just Released!

- Features for existing Db2 databases
  - Connection profiles
  - Database object exploration
  - Query execution

- Developer productivity features
  - Ability to run local containerized Db2

Warehouse / Lakehouse Updates

Integrated Analytics
System (IIAS)

PureData System for Operational
Analytics (PDOA)

Modernization Path

like for like upgrade to....

1. Db2 Warehouse SaaS

Fully managed, high performance, cloud-elastic
data warehouse
↳ Runs on IBM Cloud, AWS
↳ Available in US, EU, UK, JP, AP commercial regions

2. Db2 Warehouse Self-managed

Self-managed, data warehouse on cloud provider
managed Kubernetes or OpenShift infrastructure.
↳ Runs on Amazon EKS, Amazon ROSA, Azure AKS,
  Azure ARO
↳ Available worldwide
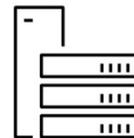
3. Db2 Warehouse on Power

On-premises, OpenShift high performance,
data warehouse deployable on Power 10 infrastructure
↳ Runs on Power 10

Db2 Warehouse on Azure
Introducing the third generation of our cloud data warehouse

Two deployment models

– **Bring-your-own-cloud (BYOC)** with Db2 Warehouse deployed in customer's account & VPC

– **Planned : Fully-managed SaaS** with Db2 Warehouse deployed in an IBM account with option for private deployment inside a dedicated VPC

Key features

– **Azure blob support** for table storage, significantly decreasing the cost of storing data without sacrificing performance

– **Query multiple open table & data formats** (Iceberg, Parquet, ORC, CSV and more) leveraging existing compute resources dedicated to the warehouse

– **Integration with watsonx.data** lakehouse service with sharing of data catalogs

Also includes

– **Multi-petabyte-scale**, multi-performant storage

– **Support for Azure Blob Storage and Azure Disk Storage**

– **Storage auto-increase for Block Storage** on set threshold ensuring you never run out of storage for your workloads

– **Pause compute** to reduce the cost of running dev/test or production environments

– **Public/private network** support providing options for maximum security/privacy of deployed warehouses

– **Self-service maintenance windows** for product and database engine updates

– **Cross-region snapshot backup** to Azure Blob for disaster recovery (coming post GA)

IBM Db2 / © 2025 IBM Corporation
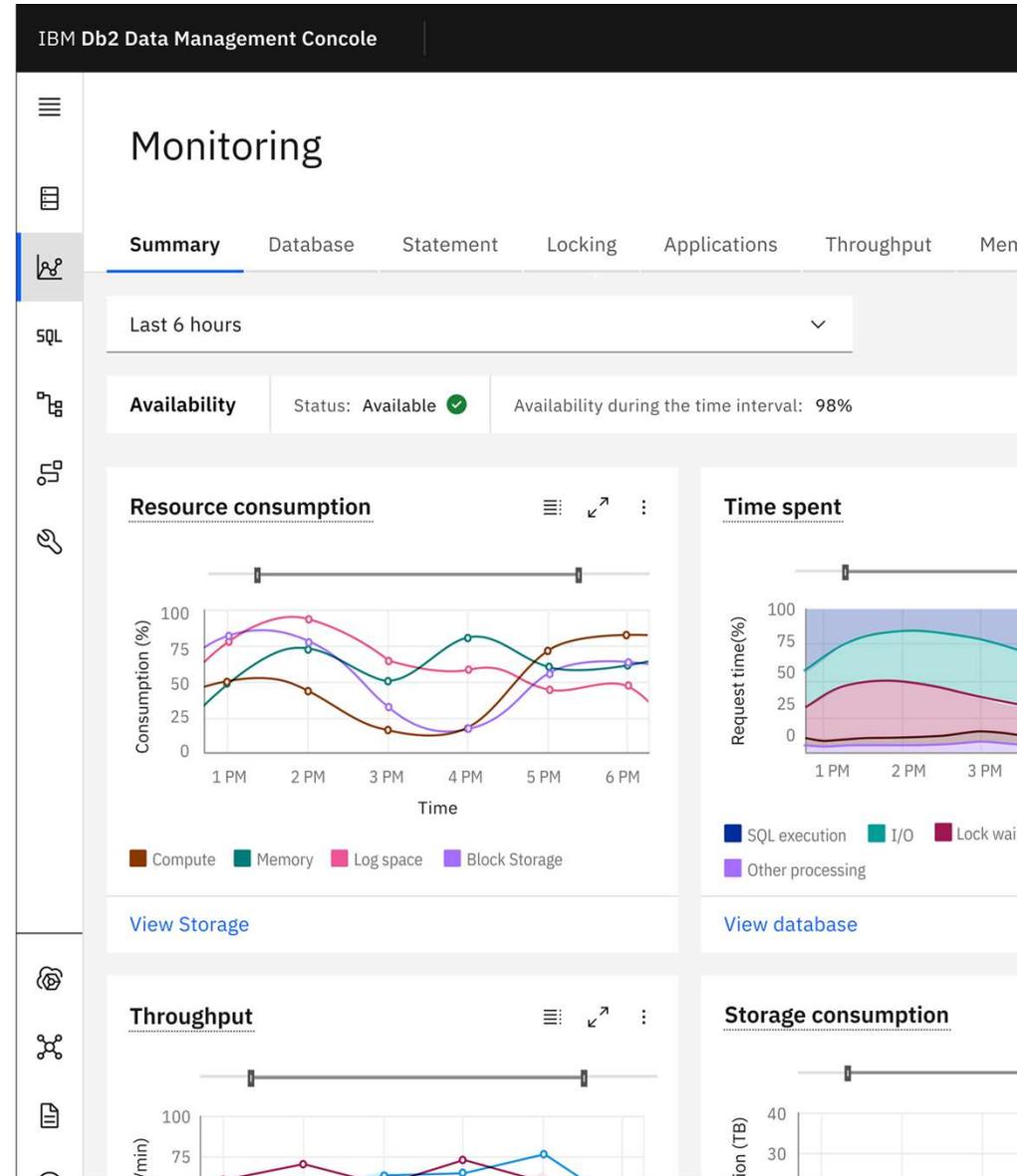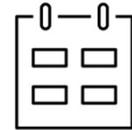
Reference Architecture

Prescriptive configuration of hardware, software and services

- **Powerful hardware and database engine**
- Optimized for performance ,scalability and reliability

- **Certified architecture design**
- Fully built and tested in IBM Lab
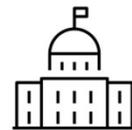
- **Optimized experience**
- Includes white-glove installation and Solution Support services with a single point of contact

PDOA/IIAS vs. P10 Cloud Rack for Db2 Warehouse

IBM Integrated Analytics System

- Power 8 2U Servers
- FS900 Flash Modules
- 4x10 Gb bonded network
- 16 Gb HBAs
- Db2 Warehouse

P10 Cloud Rack for Db2 Warehouse

- Power 10 2U Servers
- FS7300 Flash Modules
- 100 Gb bonded network
- 32 Gb HBAs
- Modernized Db2 Warehouse on OpenShift

IBM Integrated Analytic System

- 8 Data MLNs per Server
- 3 Power8 Cores per MLN
- 1.6TB Usable Tier 1 space per MLN
- 51GB Memory Per MLN
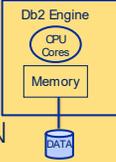
P10 Cloud Rack for Db2 Warehouse

- 8 Data MLNs per Server
- 4 Power10 Cores Per MLN
- 3TB Usable Tier 1 space per MLN
- 100GB Memory Per MLN

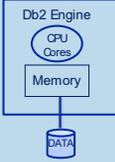# Db2 Warehouse On Power Deployment Options

	OCP Based	RHEL Based [NEW]
Software Stack	• OpenShift Container Platform • Storage Scale • Db2 Universal Container	• Red Hat Enterprise Linux • Db2 Software
Hardware Stack	• Power Cloud Rack • FS7300	• Power Cloud Rack • FS7300
File System	Storage Scale	XFS (for data) NFS (for home directory)
MLN per nodes	8	8
Resource allocation per MLN	• 3 Power10 cores • 4 TB Tier1 user space • 100 GB Memory	• 4 Power10 cores • 4 TB Tier1 user space • 124 GB Memory
High Availability	One standby node per rack	One standby node per rack
Cluster Management Solution	Kubernetes	Pacemaker
Monitoring	IBM Db2 Intelligence Center	IBM Db2 Intelligence Center
Deployment/Upgrade	Operator and Db2U auto script	Standard process with Db2 commands
Scalability & Elasticity	Operator and Db2U auto script	Standard process with Db2 commands
License	Db2 Warehouse Cartridge	Db2 Advanced Edition RHEL
Support	Single Point of Triage	Single Point of Triage

Experience: A balanced resilient architecture for maximum reliability and performance

Flexibility: The Data Warehouse starts small, but grows as you do

Efficiency: E-config enables the full stack pre-selected and configured as a "solution edition" set-up

Storage: IBM FlashSystem adds simplicity and consolidation, capable of delivering more work with fewer drives and lower cost
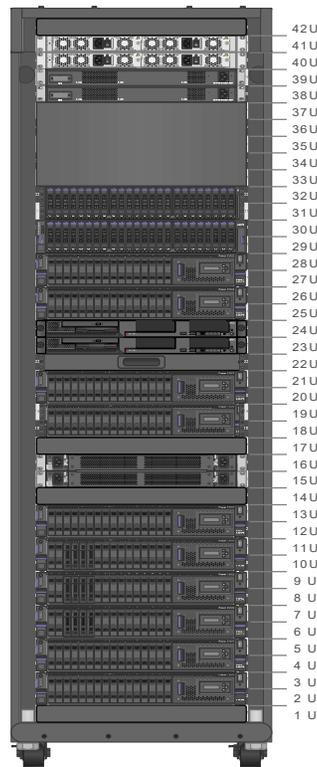
Hardware Stack

Servers:
- OpenShift Management: 3x Power 1022s
- Workers: From 3 to 80 - IBM Power 1022

Storage:
- FlashSystem 7300 for Tier-1
- FlashSystem 5300 for Tier-2

High Speed Networking
- 100 Gb Ethernet Switching
- Storage Networking SAN

Management:
- HMC

Software Stack

- RHEL OpenShift
- Db2 Warehouse on OpenShift
- PowerVC
- IBM Spectrum Scale Data Management Edition
- PowerVM Enterprise

IBM Services

Pre-sales:
- Size an appropriate configuration for customer's existing workload.
- Collaboration with business partners.

Install:
- Build the infrastructure hardware stack and pre-load the OCP/SS software stack in IBM MFG
- Gold-glove installation of the infrastructure hardware/software stack at customer's premises
- Remote installation/customization of Db2 Warehouse Software on the infrastructure at customer's premises
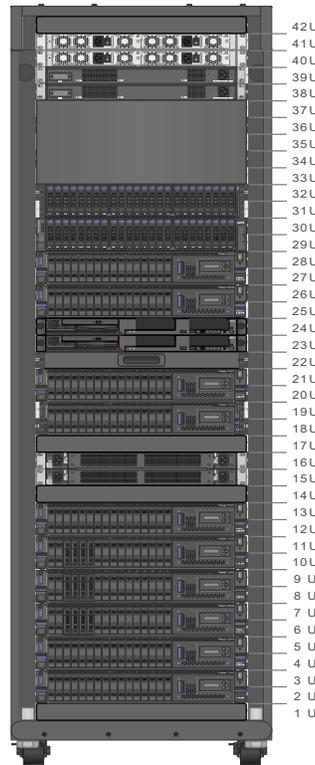- Migration Assessment session with the customer

Experience: A balanced resilient architecture for maximum reliability and performance

Flexibility: The Data Warehouse starts small, but grows as you do

Efficiency: E-config enables the full stack pre-selected and configured as a "solution edition" set-up

Storage: IBM FlashSystem adds simplicity and consolidation, capable of delivering more work with fewer drives and lower cost

Rack Units
42U
41U
40U
39U
38U
37U
36U
35U
34U
33U
32U
31U
30U
29U
28U
27U
26U
25U
24U
23U
22U
21U
20U
19U
18U
17U
16U
15U
14U
13U
12U
11U
10U
9U
8U
7U
6U
5U
4U
3U
2U
1U

Hardware Stack

Servers:
- Workers: From 3 to 80 - IBM Power 1022

Storage:
- FlashSystem 7300 for Tier-1
- FlashSystem 5300 for Tier-2

High Speed Networking
- 100 Gb Ethernet Switching
- Storage Networking SAN

Management:
- HMC

Software Stack

- RHEL OpenShift
- Db2
- XFS (data) and NFS(home directory)

IBM Services

Pre-sales:
- Size an appropriate configuration for customer's existing workload.
- Collaboration with business partners.

Install:
- Build the infrastructure hardware stack and pre-load the software stack in IBM MFG
- Gold-glove installation of the infrastructure hardware/software stack at customer's premises
- Remote installation/customization of Db2 Software on the infrastructure at customer's premises
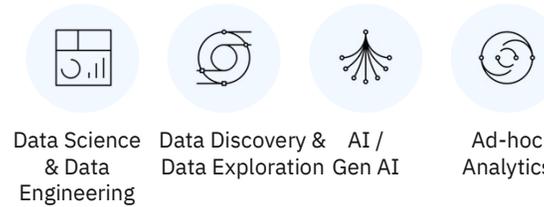- Migration Assessment session with the customer

Better Together – Db2/NZ Warehouse + WxD Lakehouse

Warehouse Ideal for:
Structured Datasets
Maximum Performance, Highly Concurrent Workloads
Streaming Data Ingestion and Real-time Analytics & Insights*
Low Latency Query*
Mixed Workloads and Operational Analytics*

Lakehouse Ideal for:
Structured, Semi-structured and Unstructured Datasets
Data Exploration and Data Engineering Workloads
Data Science and AI/ML
Generative AI
Consolidating very large datasets from multiple sources

Realtime Analytics

Operations Analytics

Production BI Reporting

Low Latency Query

Data Science & Data Engineering

Data Discovery & Data Exploration

AI / Gen AI

Ad-hoc Analytics

Warehouse
Optimized and/or Open Format analytics database on Block and/or Object Storage

← Dataset Interchange / Mirroring →

Lakehouse
Open Format & Flexible Multi-purpose Technology Stack on Object Storage

Object Storage

(*) These workloads are only supported by IBM Db2 Warehouse

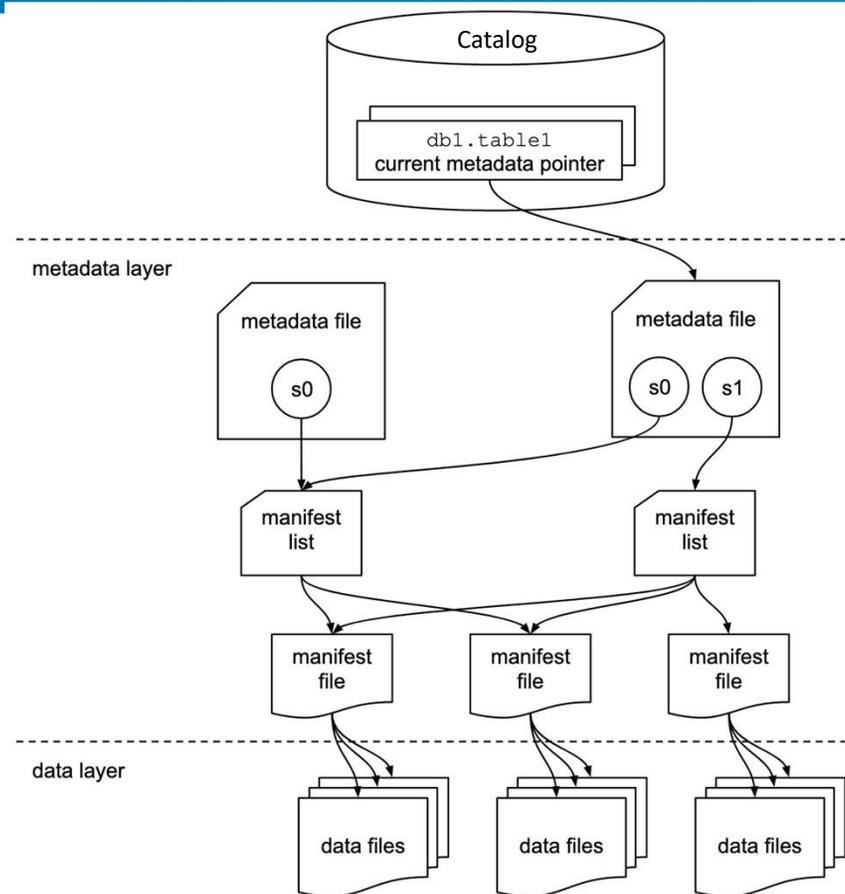## What does this data sharing enable you to do in your Db2 warehouse?

- Offload cold/historic data or data with lower performance requirements from the warehouse to the lakehouse
  - Reduces warehouse costs (less high-end infrastructure, less management effort, smaller/faster backups)

- Data warehouse workloads can leverage other data in the lakehouse, including business data that doesn't live in the warehouse, weather data, systems of engagement data, and much more

- Offloaded data can be made available to a larger community of users without impacting data warehouse SLAs
  - Business analysts, data scientists, data engineers, and more

- Offload transformation workloads from the warehouse to the lakehouse

- Enables a "medallion" architecture where bronze/silver data lives in the lakehouse and silver/gold data in the warehouse
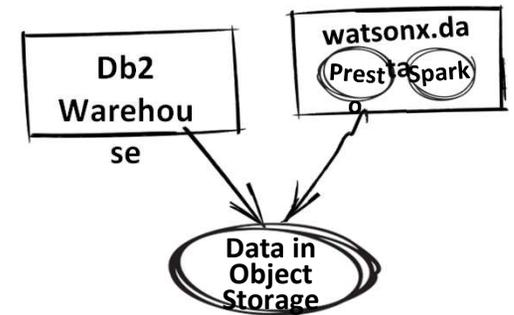
Open-source data table format that helps simplify data processing on large dataset stored in data lakes

People love it because it has:

- SQL access — Build the data lake and perform most operations without learning a new language

- Data Consistency — ACID compliance (not just append data operations to tables)

- Schema Evolution — Add/remove columns without distributing underlying table structure

- Data Versioning — Time travel support that lets you analyze data changes between update and deletes

- Cross Platform Support — Supports variety of storage systems and query engines (Spark, Presto, Hive, +++)

What is a metastore?

- Manages metadata for the tables in the lakehouse, including:
  - Schema information (column names, types)
  - Partitioning
  - Location and type of data files

- Similar in principle to the system catalogs of a relational database

- Shared metastore/catalog ensures query engines see schema and data consistently

- May be a built-in component of a larger integration/governance solution

IBM watsonx.data Metadata Service

- Open and flexible access, with industry standard interfaces
- Enables comprehensive governance and control over diverse data assets
- Tight integration with IKC

AWS Glue Data Catalog

- Component of AWS Glue integration service
- Inventories data assets of AWS data sources
- Includes location, schema, and runtime metrics

Microsoft Purview Data Catalog

- Component of Microsoft Purview data governance solution
- Helps manage on-premises, multicloud, and SaaS data
- Offers discovery, classification, and lineage

Databricks Unity Catalog

- Provides centralized access control, auditing, lineage, and data discovery across a Databricks lakehouse
- Contains data and AI assets including files, tables, machine learning models, and dashboards

Unity Catalog

Gravitino

Nessie

Lakekeeper

Polaris

Microsoft Purview

One Protocol Many Catalogues

| Open Metastore Protocol |
| Open Table Format · ICEBERG |
| Open Table Data · Parquet |
| Object Storage · Amazon S3 · ceph · 10 01 · IBM Cloud Object Storage |

Db2 Bridge

One Data Movement tool for all your mission critical workloads.

"What"

Move any workloads with Db2 being the target

➢ **IBM Db2 (source/target)**

➢ **IBM Db2 Warehouse (source/target)**

➢ **IBM Integrated Analytics System (source)**

➢ **Power 10 Cloud Rack (target)**

➢ **Db2 Data files (source/target)**

➢ **Oracle (source)**

➢ **Teradata (source)**

"Where"

Support all possible deployment options for Db2

➢ **On-prem to On-Prem**

➢ **On-prem to Cloud**

➢ **Cloud to On-Prem**

➢ **Cloud to Cloud**

"How"

Extensive set of features to accommodate complexities of data movement

➢ **Schema-level selection of objects**

➢ **Move massive number of databases or large databases with hundreds of Terabytes or Petabytes of data**

➢ **Enable High Performance Unload for faster migrations**

➢ **Support:**

- **Row/Columnar organization**

- **Little/Big Endian formats**

- **All object types, Security labels Privileges, Authorizations**

35

🚀 Key Innovations for this Release

Teradata to Db2 Warehouse Migration

A streamlined two step blueprint for migrating from Teradata,

- **Schema Conversion** -- Automatically convert Teradata objects into Db2 ready DDLs.

- **Data Movement** -- Monitor real time logs and track every row as it transitions to Db2 Warehouse.

This approach ensures compatibility, performance and operational continuity reducing migration effort and risk.

Db2 Bridge Health Check

A powerful new module that enables organizations to

- **Scan the Db2 environments** for upgrade and cloud migration readiness.

- **Identify risks and deprecated features** that could impact modernization success.

- **Receive actionable recommendations** via a unified dashboard.

The Daily Reality for DBAs

Too Many Tasks, Too Little Time

Troubleshooting is complex

62%* say it is the #1 time consuming activity.
Finding the right documentation, commands, and fixes takes too long.

Query tuning requires deep expertise

Performance optimization is an **ongoing activity.**
Indexing, query tuning, and workload balancing requires constant attention.

Tools are fragmented and manual

68%* use 3-4 management tools daily.
Disconnected tools make it harder to get a unified view of database health.

DBAs are reactive instead of proactive

$14k/min is the average cost of unplanned downtime.
Reactive intervention is not just inconvenient, it's expensive.

Repetitive tasks consume bandwidth

Only **36%*** have sophisticated task automation.
Most teams rely on local scripts, manual handoffs, or outdated runbooks.

* Based on survey of 30+ expert Db2 database administrators

What if...

|	
Troubleshooting is complex	✅ DBAs had **direct access to the right answers** - no more hours lost in documentation?
Query tuning requires deep expertise	✅ AI made query tuning **faster, smarter, and less effort-intensive**?
Tools are fragmented and manual	✅ DBAs had a **unified platform** instead of juggling disconnected scripts and dashboards?
DBAs are reactive instead of proactive	✅ DBAs had a tool that **monitors your database 24/7** and flags issues before they impact performance?
Repetitive tasks consume bandwidth	✅ Repetitive tasks like backups and patching **ran with minimal intervention**, freeing up time?

Db2 Genius Hub
AI-powered database management solution for Db2 admins
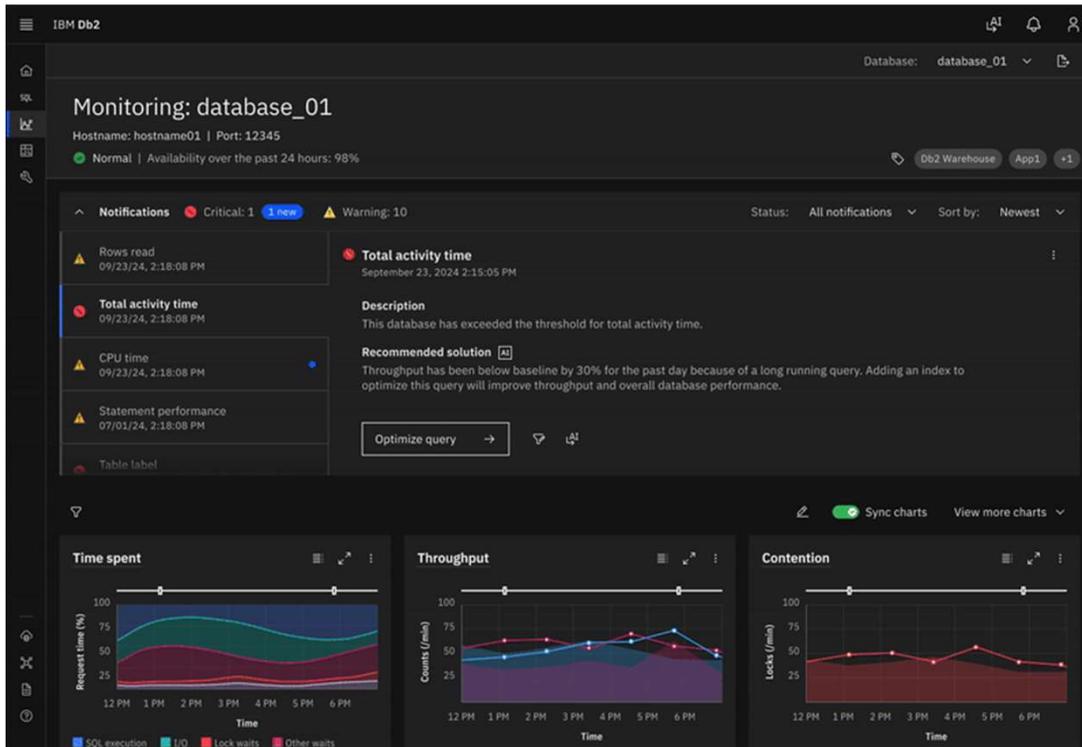
20%
reduction in management costs[1]

4x
more productive DBAs[1]

Unified Management
Single point of management across your entire Db2 estate

Agentic AI Database Assistant
GenAI-powered expert at your fingertips

Monitoring Drill-downs
Find and fix database problems instantly

Query Optimization
Enhance query and database performance with advanced tuning
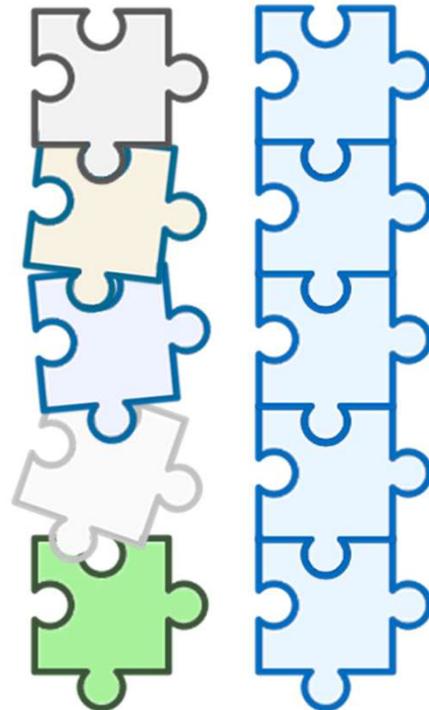
Tune Configurations
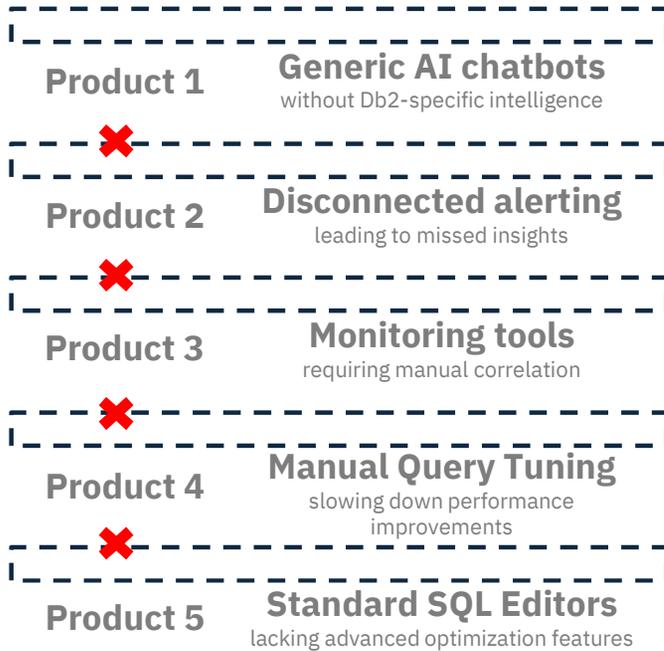Optimize Db2 deployments for peak efficiency

Advanced Automation
Streamline operations with powerful scripting, scheduling, and automation

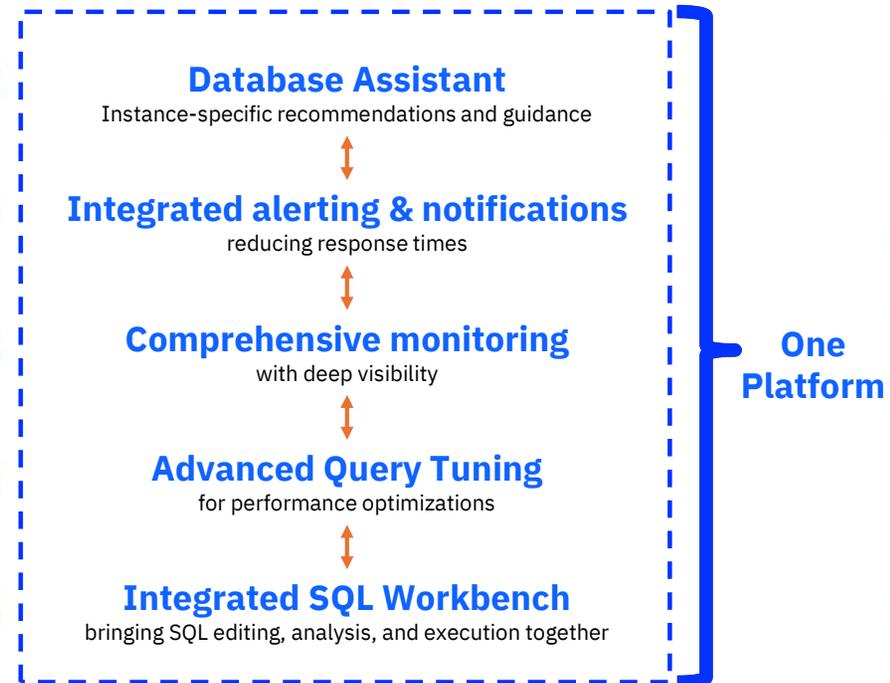[1] Innovation Insight: How Generative AI Is Transforming Data Management Solutions

IDUG

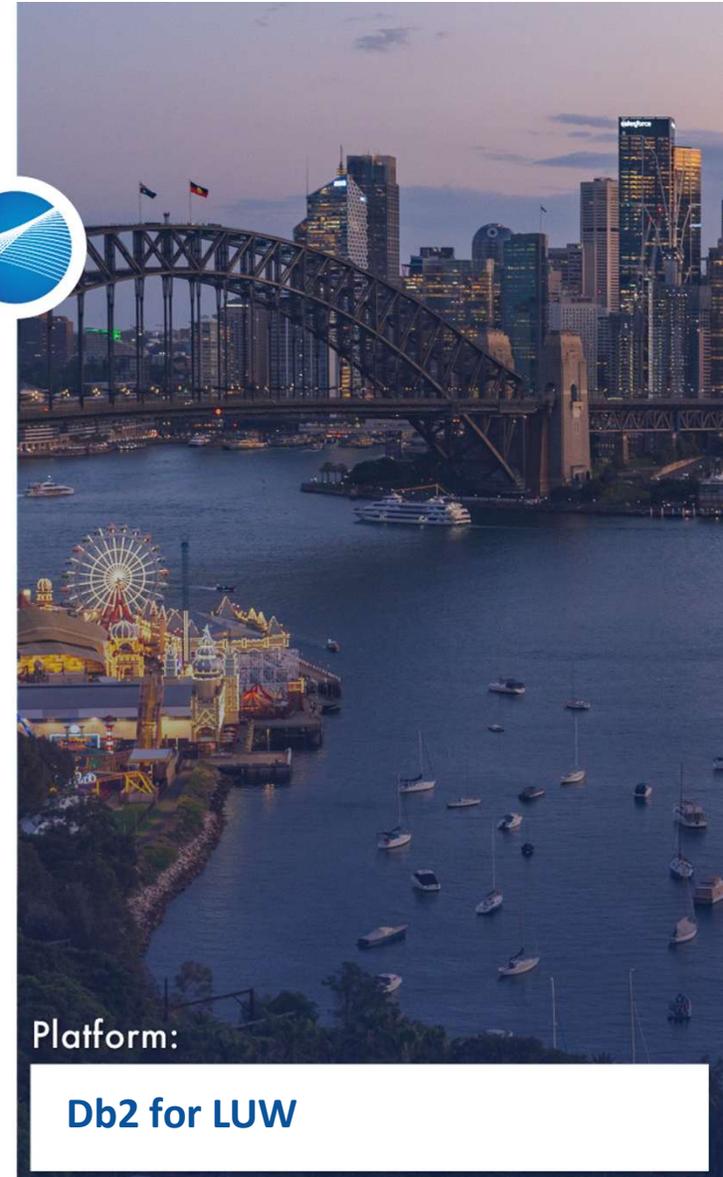2026

Sydney | March 16 - 18

AU Db2 TECH CONFERENCE

Db2 Latest News

David Kalmuk, *IBM*

Contact: dckalmuk@ca.ibm.com

Session Code: B03

Platform:

Db2 for LUW

IDUG

2026 Australia **Db2** Tech Conference