

IDUG

2026

Sydney | March 16 - 18

# AU Db2 TECH CONFERENCE

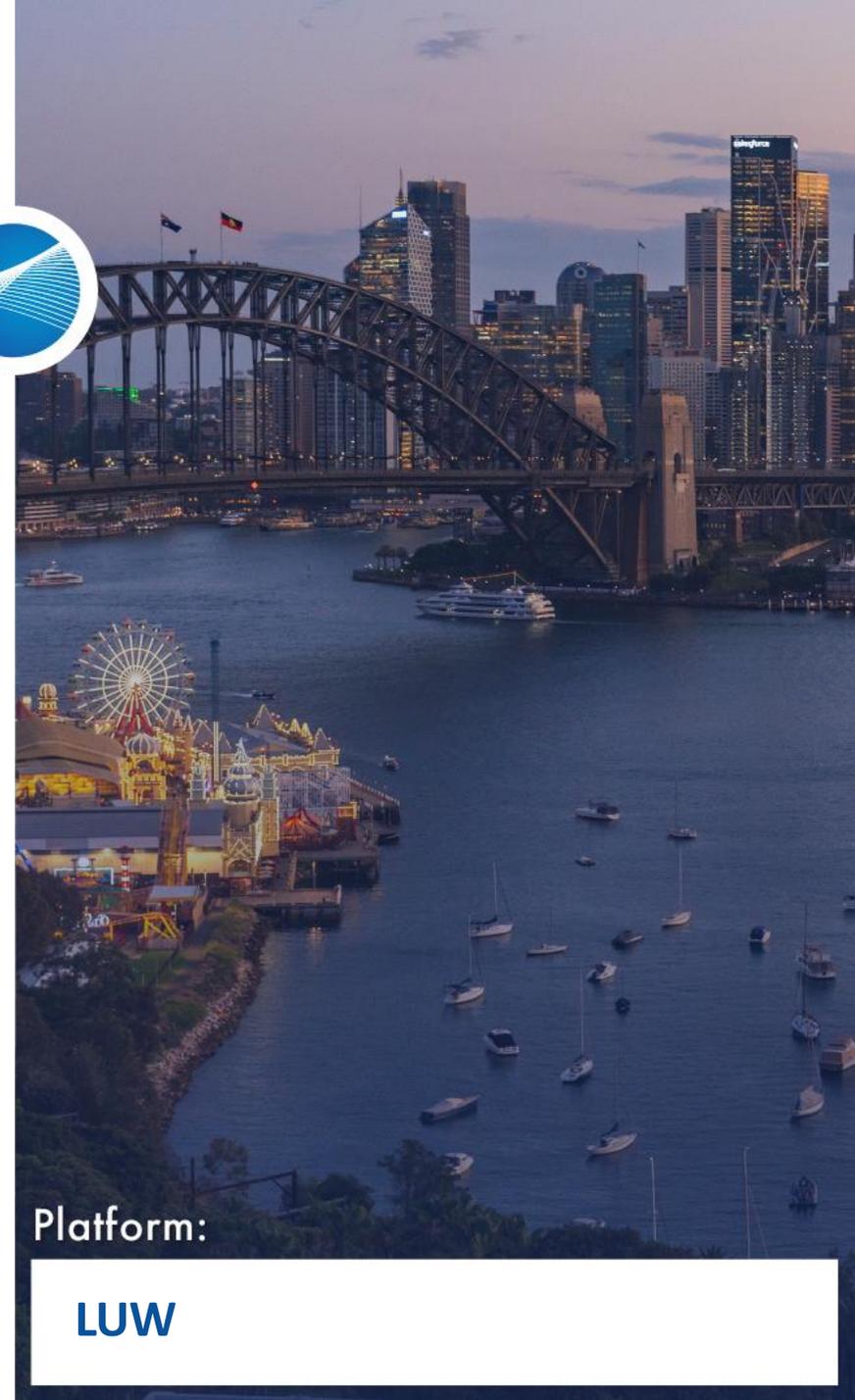
An Introduction to Db2 Serverless

Andrew Hilden, *IBM*

Contact: [ahilden@ca.ibm.com](mailto:ahilden@ca.ibm.com)

---

Session Code: B06



Platform:

**Luw**

# Notices and disclaimers

- © 2026 International Business Machines Corporation.  
All rights reserved.
- This document is distributed “as is” without any warranty, either express or implied. In no event shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.
- Case studies and client examples are presented as illustrations of how customers or IBM has used IBM products in production or test environments and the results they may have observed. Actual performance, cost, savings or other results in other operating environments may vary.
- Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM.
- Not all offerings are available in every country in which IBM operates.
- Any statements regarding IBM’s future direction, intent or product plans are subject to change or withdrawal without notice.
- IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corporation, 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: [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).
- Certain comments made in this presentation may be characterized as forward looking under the Private Securities Litigation Reform Act of 1995.
- Forward-looking statements are based on the company’s current assumptions regarding future business and financial performance. Those statements by their nature address matters that are uncertain to different degrees and involve a number of factors that could cause actual results to differ materially. Additional information concerning these factors is contained in the Company’s filings with the SEC.
- Copies are available from the SEC, from the IBM website, or from IBM Investor Relations.
- Any forward-looking statement made during this presentation speaks only as of the date on which it is made. The company assumes no obligation to update or revise any forward-looking statements except as required by law; these charts and the associated remarks and comments are integrally related and are intended to be presented and understood together.

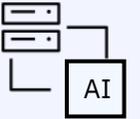
# Agenda

- Db2 Serverless Introduction and Overview
- Target Workloads
- Db2 Serverless and the Agentic Era
- SaaS Capabilities
- Brief Architecture Overview



# Introducing the Db2 Serverless, enabling AI workloads and AI development for a New Era

[Db2 Serverless](#), provides an infrastructure free SaaS experience and enables developers to create faster than ever before while enabling scaling and enterprise features for all database workloads.



## AI Ready

→ Built for AI development and applications enabling developers to create database applications and make modifications using powerful integrations and branching capabilities.

## Worry free

→ Complete autonomous SaaS workloads with enterprise capabilities freeing you up from the management of database resources

## Flexible

→ Dramatic cost savings for workloads while allowing you to meet the highest demands when needed



## Enterprise Grade SaaS

- Fully Managed

*Db2 Serverless Provides full SaaS Capabilities*

### Activities

- Managed of Patches and Updates  
Patches and updates are applied for you with zero or minimal downtime
- Monitoring and Availability  
Fully monitored to detect problems and make sure the database is always available.
- Health Check  
Proactive scans for common issues
- Backup & Recovery  
Automated backups with cross regional redundancy and point in time recovery to ensure your data is safe.
- Full support  
Automatic and full service support



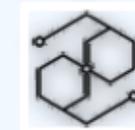
## Resource Free

No resources to worry about

*Db2 Serverless automatically scales up and down to meet your workload requirements..*

### Activities

- Scale up as workload grows  
Db2 Serverless identifies and rapidly scales up to ensure performance remains consistent
- Scales down to 0 to save costs  
Db2 Serverless scales resources down to 0 to ensure you are not paying for idle resources
- Pay for what you use  
Db2 Serverless is built to optimize costs by only paying for what you use.
- Configurable  
Configure minimum and maximum resource limits to ensure you are in control of your costs



## Ready For AI Workloads

Optimized for Context Engineering and AI Workloads

*Built with context engineering in mind to ensure allowing for rapid prototyping and feature delivery and capable for meeting the depends of AI machine driven workloads*

### Activities

- Build database enabled applications quickly  
MCP server allows AI tools like Bob and Claude to interact with Db2 Serverless directly accelerating productivity.
- Great for Ephemeral Development  
Quickly provision and delete instances in seconds allowing you to prototype and test with speed and without worry.
- Powerful Branching Capabilities  
Powerful branching capabilities increasing your velocity.
- Great for spikey and machine driven unpredictable workloads  
Machine driven workload require management that operates as the same pace

# IBM Db2 Serverless

Modern SaaS    Developer Friendly    Autonomous



- Always the latest generation of Db2 Software available via committed spend or pay as you go via [Cloud Marketplace](#)



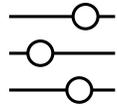
- [Scales to 0](#) when there is no workload (All connections are Idle for 5 mins)



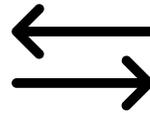
- [Incrementally scales](#) in granular increments mirroring workload requirements with min and max settings



- [API and MCP enabled](#) Allowing you to integrate into you AI Engineering and CICD processes with ease.



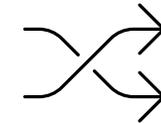
- [Near instant Provisioning](#) using Infrastructure as Code, UI, CLI or API



- By the second [consumption based](#) billing



- Provision a new database instances as [branches](#) of an existing database.



- [No impact to connections](#) during scaling or most infrastructure maintenance

# Agenda

- Db2 Serverless Introduction and Overview
- Target Workloads
- Db2 Serverless and the Agentic Era
- SaaS Capabilities
- Brief Architecture Overview

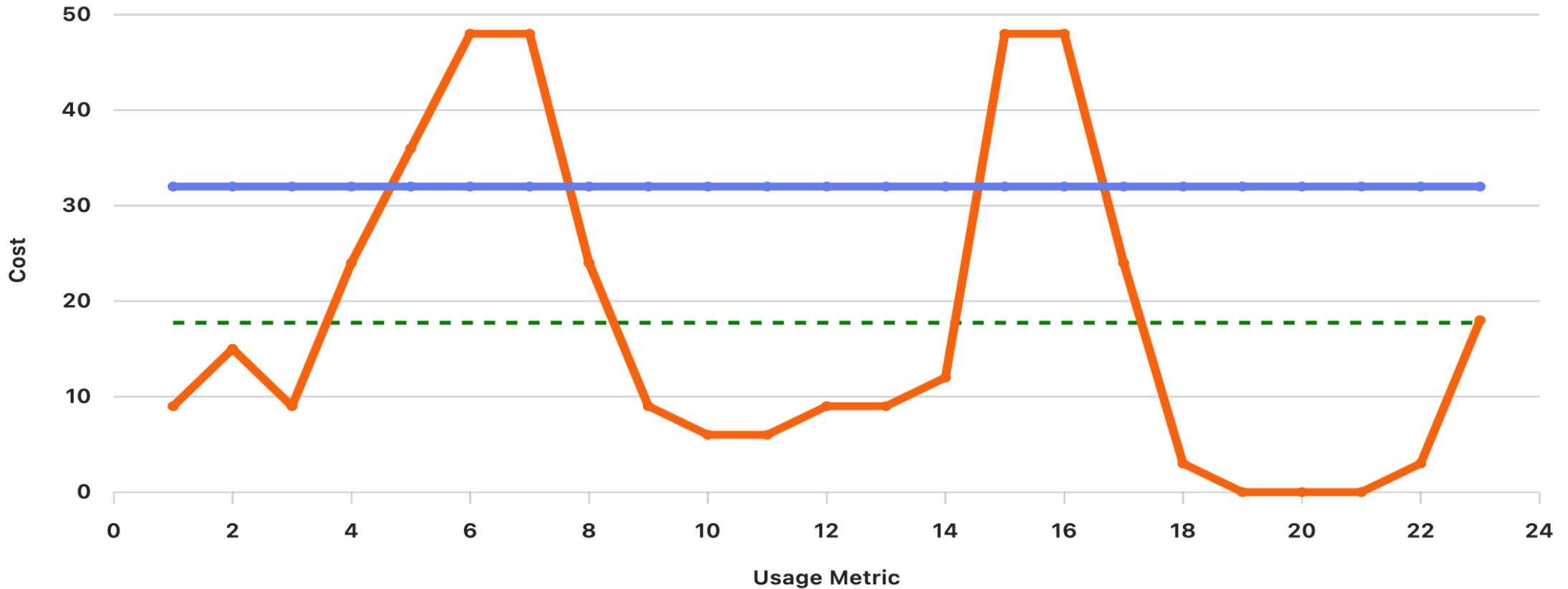
## IBM Db2 Serverless



**Built for  
Unpredictable and  
Stable Workloads**

- Initially focused on **transactional** and **single-node analytics** workloads
- Rapidly evolving to support high-volume, **large-scale analytics**
  - With full horizontal and vertical scaling
- **AI Workloads**
  - Vector
  - Ephemeral
  - Machine Driven
- Great for **unpredictable workloads**
  - Non-uniform
  - Spiky

# Example workload profile: Cost for Db2 Serverless versus Provisioned Capacity Costs



■ Provisioned Capacity Cost ■ Db2 Serverless Cost ■ Avg Provisioned Cost ■ Avg Db2 Serverless Cost

6 days a week – 9-5 PM – Peak workload required 8vCPU, 16GB of RAM, smaller workload from 5 PM until midnight and no workload after midnight until 7 AM

Note: **Fictual Units costs are for demonstration purposes only.** Actual ratios and Fictional Units are not reflected in dollars.

## Typical Database SaaS with provisioned capacity

Customer Acquires and Provisions 8 VCPU, 16 GB HA system.

- System runs at 8vCPU and 16 GB of memory for 24 hours at a cost of 8 RU per hour

Total Cost Per Day – **192 Fictional Units**

## Db2 Serverless

Customer acquires Db2 Serverless instance (they optionally specify a minimum and maximum size)

- System runs at 4 vCPU for 4 hours at a cost of 6 RU per hour and at 8 vCPU for 4 hours at a cost of 12 RU per hour
- System runs at 1 vCPU and 2 GB of ram from 5 until midnight for a rate of 1.5 RU per
- System scales to 0 from midnight until 7 AM at a rate of 0 RU per hour

Total Cost Per day – **66.5 Fictional Units**

Used by millions of users to schedule and have other users purchase event tickets. The events are scheduled by any of its users and the workload is very unpredictable and may even be idle for long periods of time until major events are scheduled.

Note: Fictional Units costs are for demonstration purposes only. Actual ratios and amounts will differ.

## Database SaaS with provisioned capacity

ACME acquires and provisions a 64 vCPU, 128 GB HA system.

- System runs at 64vCPU and 128 GB of memory for 24 hours at a cost of 128 units per hour
- ACME throttles users to ensure the database is not overloaded and if it throttles for an extended time the system is scaled up by an administrator.

Total Cost Per Month –  $1536 \text{ RU} * 30 = 46,040$   
**Fictional Units**

## Db2 Serverless

Customer acquires Db2 Serverless instance (They optionally specify a minimum and maximum size)

- System runs at 4 vCPU normally
- System scales up to 64 vCPU during a major event for 4 hours twice a month
- System scales to 0 from midnight to 6 AM 50% of the nights

Total Cost Per Month = **4784 Fictional Units**

To ensure quality as well as efficiency they have an advanced pipeline that triggers up to 100s of times per to ensure changes going into the application are tested extensively..

Note: Fictional Units costs are for demonstration purposes only. Actual ratios and amounts will differ.

### **Database SaaS with provisioned capacity**

A database is setup per developer it is reused each time by the pipeline but requires cleanup each time and in between test cases.

Or:

A shared database is created with enough capacity and it is reused.

Or:

A new database is provisioned and that time is added to the execution of the pipeline.

### **Db2 Serverless**

The pipeline uses IaaS and provisions a new Db2 Serverless instance in minutes branching from a master branch.

- It provisions many databases to run a set of parallel tests to improve developer productivity while keep costs low due to low usage
- In the case of a failure the pipeline stops using the database and scales it to 0 until investigation can be done.

The workload is very stable and requires 64 vCPU and 64 GB of RAM and does not typically differ over the course of a day or week.

Note: Fictional Units costs are for demonstration purposes only. Actual ratios and amounts will differ.

## Database SaaS with provisioned capacity

ACME acquires and provisions a 64 vCPU, 128 GB HA system.

- System runs at 64vCPU and 128 GB of memory for 24 hours at a cost of 64 RU per hour

Total Cost Per Day – **1536 Fictional Units**

## Db2 Serverless

ACME provisions a Db2 Serverless system and specifies a min and max scaling value of 64 vCPU.

- System consistently scales and stays at 64 vCPU and 128 GB of RAM at a cost of 96 RU per hour

Total Cost Per Day – **2304 Fictional Units**

An application leverages LLMs to generate a non optimized machine-driven workload on the database that must run fast to give answers to users. These workloads cause the system to scale high and then are relatively idle for long periods of time.

Note: Fictional Unit costs are for demonstration purposes only. Actual ratios and amounts will differ.

### Database SaaS with provisioned capacity

ACME acquires and provisions a 64 vCPU, 128 GB HA system.

- System runs at 64vCPU and 128 GB of memory for 24 hours at a cost of 64 RU per hour

Total Cost Per Day – **1536 Fictional Units**

### Db2 Serverless

ACME provisions a Db2 Serverless system and specifies a max scaling value of 64 vCPU.

- System infrequently scales up to 64 vCPU
- Normal workloads are consuming very little Capacity

Total Cost Per Day – **384 Fictional Units**

# Agenda

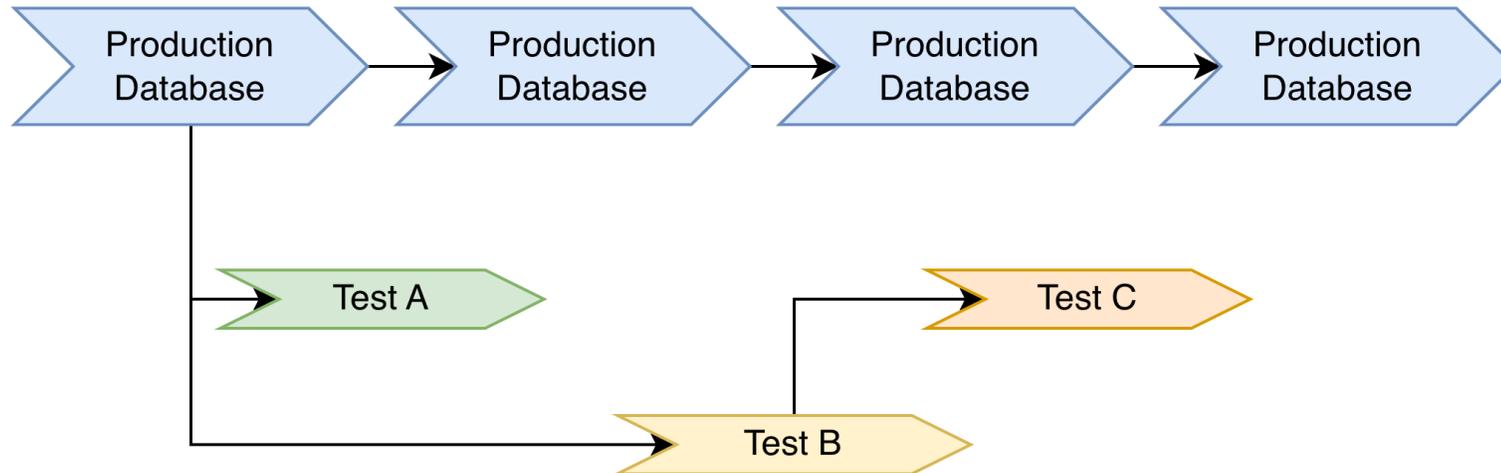
- Db2 Serverless Introduction and Overview
- Target Workloads
- Db2 Serverless and the Agentic Era
- SaaS Capabilities
- Brief Architecture Overview

## *From Prompt to Production – Db2 Serverless*



And More.....

Imagine if every time you wanted to test an application on your computer you needed to reinstall your operating system. **Why should you have to rebuild your database just to add or test a new feature?**



- Quick Branching
- Ephemeral
- No infrastructure
- Dev to production

- Db2 Serverless exposes a **Rich API** for integration with tools and processes
- Db2 Serverless provide **MCP Servers** that help supercharge your engineering abilities
- Db2 Serverless provides **branching capabilities** that allow you to integrate ephemeral instances into your processes for testing, reporting or other use cases
- Db2 Serverless **elastically grows and shrinks** with your workload. You need not worry about these details.

Control Db2 Serverless via Prompts

Create a database and call it IDUG. Create a schema that represents the attendees at the conference and the sessions they attended.

Automatically creates a schema for you with optimizations

Create a database and call it IDUG. Create a schema that represents the attend... 33k / 200k 0.08

IDUG database created successfully with the following schema:

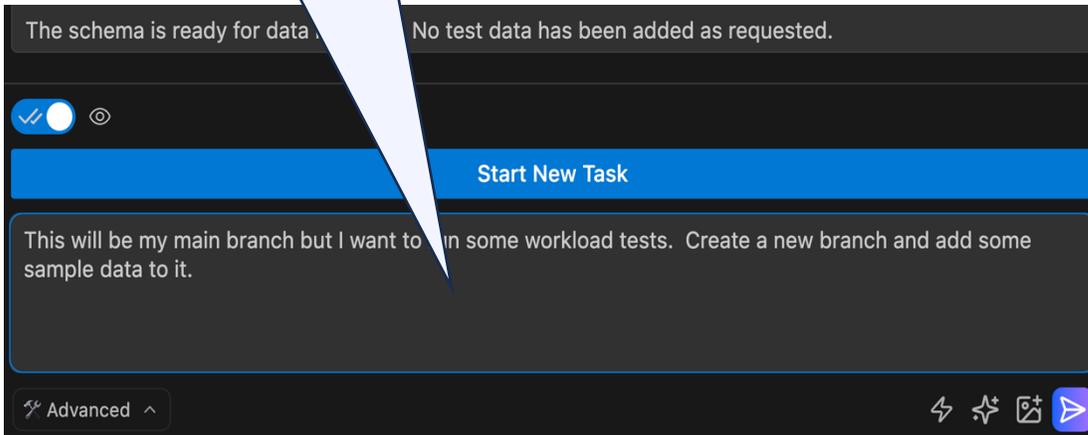
Tables Created:

- attendees** - Stores conference attendee information
  - attendee\_id (Primary Key)
  - first\_name, last\_name, email (unique)
  - company, job\_title
  - registration\_date, created\_at, updated\_at
- sessions** - Stores conference session information
  - session\_id (Primary Key)
  - session\_code (unique), title, description
  - speaker\_name, speaker\_company
  - session\_dates, start\_time, end\_time
  - room\_track, capacity
  - created\_at, updated\_at
- session\_attendance** - Junction table tracking attendance
  - attendance\_id (Primary Key)
  - attendee\_id (Foreign Key to attendees)
  - session\_id (Foreign Key to sessions)
  - check\_in\_time, rating (1-5), feedback
  - created\_at
  - Unique constraint on (attendee\_id, session\_id)

Indexes created for optimal query performance on email, company, session dates, and foreign key relationships.

The schema is ready for data insertion. No test data has been added as requested.

Create new  
databases  
branches and add  
test data



Workload testing branch created and populated with sample data:

**Branch:** workload-testing

**Sample Data Added:**

- **10 attendees** from various companies (TechCorp, DataWorks, Innovate Solutions, etc.)
- **10 conference sessions** covering topics like:
  - Database performance tuning
  - Security best practices
  - Cloud migration strategies
  - AI/ML integration
  - High availability and disaster recovery
  - Containerization and Kubernetes
- **29 attendance records** linking attendees to sessions they attended, including:
  - Check-in timestamps
  - Session ratings (1-5 scale)
  - Feedback comments

The main branch remains clean without test data, while the workload-testing branch is ready for performance testing and workload simulations.

Use simple prompts for powerful actions

Run a workload for 10 seconds

## Task Completed

Workload test completed successfully on the workload-testing branch:

### Test Results:

- **Duration:** 10.03 seconds
- **Total queries executed:** 188
- **Queries per second:** 18.75
- **Errors:** 0

The workload simulation ran 10 different types of queries including:

- Attendee searches and filtering
- Session attendance aggregations
- Rating calculations
- Join operations across multiple tables
- Group by operations on various dimensions

All queries executed successfully with no errors, demonstrating the database schema is working correctly under load.

How many users attended the "Security best practices" session?

 Task Completed



**4 attendees** attended the "Database Security Best Practices" session in the workload-testing branch.

# Agenda

- Db2 Serverless Introduction and Overview
- Target Workloads
- Db2 Serverless and the Agentic Era
- SaaS Capabilities
- Brief Architecture Overview

- Automated patching and management of components
- Operationally Monitored and Managed 24X7X365
- Near Zero Downtime Updates and Maintenance Window Scheduling
- Automatic Backups and Point In Time Recovery
- Ongoing security monitoring and secure practices built in

## Insights when you need them ... Just ask...

You 11:51 PM

What else can you do?



11:51 PM

I'm your Db2 assistant, here to help with a wide range of database tasks:

### Performance Troubleshooting & Analysis

Diagnose and resolve performance issues including slow queries, bottlenecks, lock contention, memory pressure, transaction log problems (log full, log pressure, blocking transactions), spilling, and queuing. I identify root causes and provide actionable recommendations.

### Database State & Telemetry Retrieval

Retrieve current and historical database information including CPU usage, active queries, queued requests, application status, tablespace utilization, health checks, resilience status, configuration values, and system metrics.

### Disk Space Management & Optimization

Analyze storage usage, identify space reclamation opportunities, recommend reorganization (reorg) candidates for tables and indexes, troubleshoot tablespace growth, and provide optimization strategies.

### Db2 Documentation & Knowledge

Explain Db2 concepts, features, configuration parameters, SQL syntax, error codes, best practices, and internal behavior. Provide SQL queries and commands for you to use based on official documentation.

### IBM Db2 Support Ticket Preparation

Help you prepare comprehensive support tickets for IBM Db2 support by gathering problem details, environment information, troubleshooting history, and recommending relevant diagnostics to collect.

### Connection Information

You're currently connected to the **BLUDB** database as user **db2inst1** using the **PROD\_ANALYTICS\_SVL** profile.

What would you like help with?



PROD\_ANALYTICS\_SVL

Type something...

SQL Generate SQL

Deep Research

Add connection



**Key Insights** (Last updated: 2 seconds ago)

- 667 Queries run
- 86% Average CPU usage
- 17% Average memory usage
- 40% Peak CPU usage (20 Feb 2026 11AM)

**Select database tag** (24 database tags)

- All databases (50 databases)
- Db2 Warehouse (7 databases)
- Db2 Database (22 databases)
- pureScale (21 clusters)

Name	Tags	Alerts	CPU	Memory	Storage	Respon...
core_banking_pr...	prod us-west +1	1 ⚠ 3	32%	68%	74%	96 ms
core_banking_dr...	dr us-east +2	1 ⚠ 1	42%	73%	79%	47 ms
retail_deposits_...	uat retail +1	1 ⚠ 1	45%	77%	68%	51 ms
loan_servicing_qa	qa nonprod +5	4 ⚠	33%	70%	61%	78 ms
payments_hub_...	sepa regulated +1	1 ⚠	51%	65%	55%	45 ms

**Notifications** (Status: All notifications, Sort by: Newest)

- Rows read: core\_banking\_prd\_us\_east 01 -01/23/26, 2:48:08 PM
- Total activity time: core\_banking\_prd\_us\_east - January 23, 2026 2:15:05 PM

# Agenda

- Db2 Serverless Introduction and Overview
- Target Workloads
- Db2 Serverless and the Agentic Era
- SaaS Capabilities
- Brief Architecture Overview

# A short description of the Architecture

COMPONENT	PRIMARY ROLE	KEY CHARACTERISTICS
Control Plane	Provides Service Interfaces	User Interfaces, API, MCP Servers, Provisioning, Getting Started Tutorials
Db2 Service Proxy	Connection handling and routing to Db2 Database Instances	Stable endpoint, helps scale-from-zero, routes traffic to data plane, improves maintenance, maintains connections
Db2 Serverless Instances	Execute database workloads	Elastic compute, auto scale up/down every second, ephemeral, provides isolation and performs the work of the Db2 Engine.
Db2 Serverless Storage	Persist database data	Durable, physically independent of compute

# Agenda

- Db2 Serverless Introduction and Overview
- Target Workloads
- Db2 Serverless and the Agentic Era
- SaaS Capabilities
- Brief Architecture Overview

IDUG

2026

Sydney | March 16 - 18

# AU Db2 TECH CONFERENCE

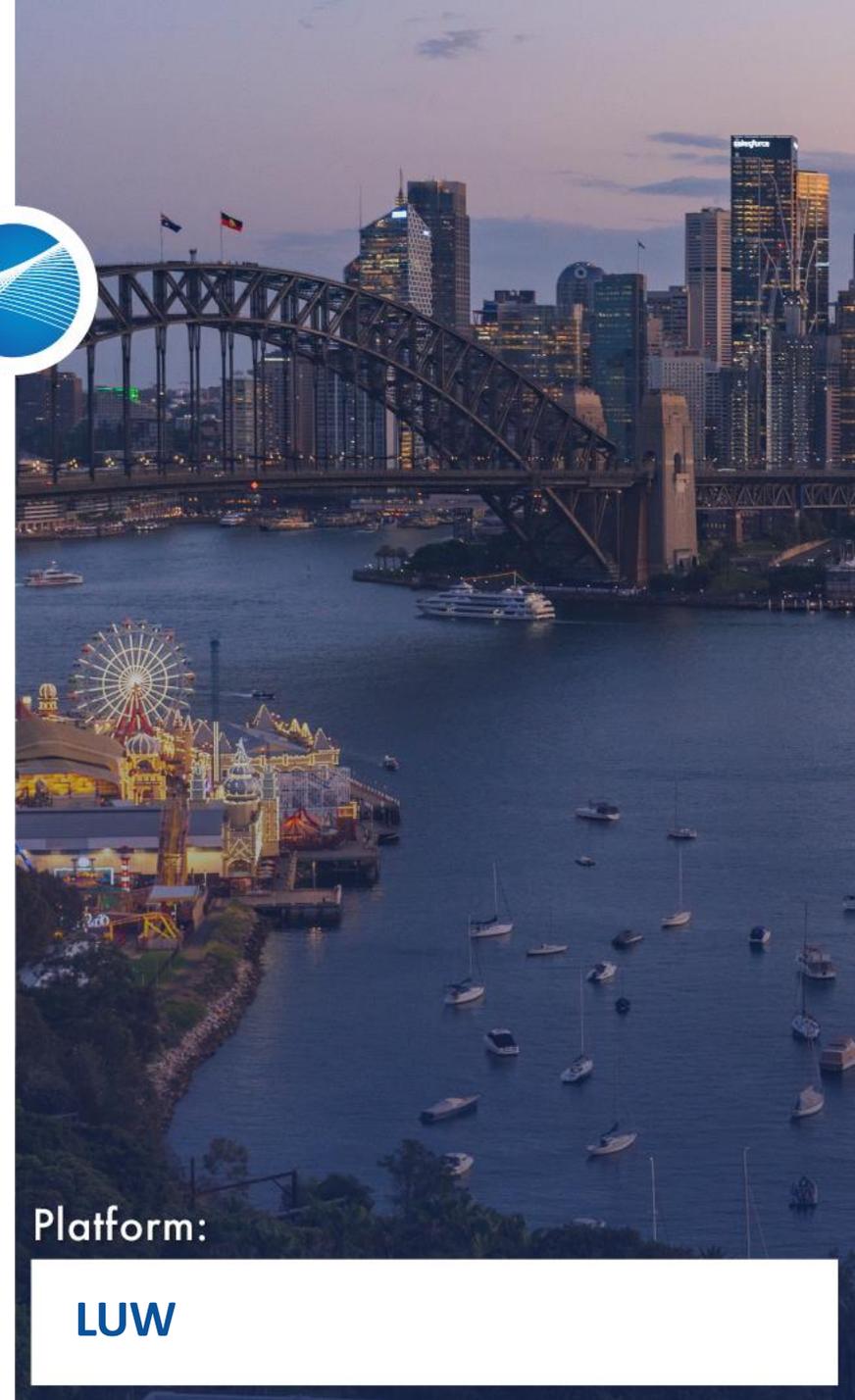
An Introduction to Db2 Serverless

Andrew Hilden, *IBM*

Contact: [ahilden@ca.ibm.com](mailto:ahilden@ca.ibm.com)

---

Session Code: B06



Platform:

Luw

An aerial view of the Sydney Harbour Bridge and the city skyline at dusk. The bridge is illuminated, and the city lights are visible against the twilight sky. The water in the harbour is dark, with many sailboats and yachts anchored. The overall scene is a mix of urban architecture and natural beauty.

# IDUG

2026 Australia **Db2** Tech Conference