IBM Db2
The AI Database

Powered by AI and built to
empower intelligent businesses
with multi-modal management

IBM

Highlights

- **Cloud Pak for Data:** Leverage platform integration, container-ization and data virtualization to accelerate the Journey to AI.

- **Powered by AI:** Use machine learning to tune workloads and optimize queries for significantly faster querying.

- **Built for AI:** Enjoy faster data exploration, model more complex relationships and leverage support for data science tools.

- **Graph:** Support dynamic multi-modal data management with real-time query support.

- **REST service:** Interact with Db2 data from your web, mobile, or cloud application through a set of scalable RESTful APIs.

Today's businesses run on data and the leaders that drive them must embrace forward-looking data science and artificial intelligence (AI) technologies to retain competitive differentiation. They must also reliably support increasingly complex business operations without downtime. Supporting these disparate needs once required a myriad of data platforms, but that is no longer the case.

With version 11.5.4, IBM® Db2® is extending its legacy of dependability and AI functionality by adding multi-data management designed to help optimize performance and support enterprises to find deeper insights. It is not only powered by and built for AI, but expands the utility of Db2 to include graph analytical capability.

Powered by AI

- Machine learning algorithms help to provide significantly faster query speed improvements.
- Machine learning algorithms are used to score queries and provide confidence-based results for faster insights.

Built for AI

- Support for PYTHON, GO, JSON, and Jupyter Notebooks allows data scientists to use the most innovative tools available.
- Data federation lets mission-critical data stay in place while running operations, fueling new insights with less hassle.

Highlighted 11.5.4 features for AI

- Graph: A single back-end that supports multiple data models, complex analytical modeling using graph capability and mission-critical relational database management system.
- REST: A set of scalable RESTful APIs to help users interact with Db2 data—create, discover, execute, and manage user-defined services in Db2.

Db2 adds this AI functionality on top of the capabilities enterprises have relied on for decades. Mission-critical transactional and analytics workloads are supported with hardware acceleration and in-memory technology. Db2 columnar data store functionality (the BLU Acceleration® feature) uses advanced compression techniques to deliver reliably faster response times without the limitations of in-memory-only systems. And Db2 will support almost any organization's growth requirements by providing petabyte-level, application-transparent database scalability beyond 100 nodes[1].

Db2 on Cloud Pak for Data

Db2 is also highly flexible. It can be purchased alone or as an IBM Cloud Pak® for Data cartridge, which enables access to database, data warehouse, data lake, and fast data solutions alongside solutions from every rung of the AI Ladder. Db2 also uses a common SQL engine to help ensure integration within the hybrid data management solution set, while the data virtualization capability in Cloud Pak for Data enables integration with even more sources outside of IBM products. Whether data is on-premises, in the cloud, structured, unstructured, SQL or NoSQL it can be brought together for better insight.

Better performance with AI

Db2 11.5.4's AI enhancements are designed to help companies achieve deeper insights more quickly and support AI applications with greater ease. Combined with existing Db2 performance-boosting technology, these new capabilities make Db2 the perfect solution to help your business climb the ladder to a robust AI practice.

Significantly faster querying

Query optimization is a crucial component of any data platform's performance, as even the world's fastest databases can be slowed by poorly executed queries. In fact, some studies indicate DBAs may spend up to a quarter of their time tuning queries[2], making it difficult for them to support other areas.

Db2 11.5.4 alleviates this bottleneck by using machine learning to reduce tuning requirements. It does so by monitoring SQL performance information over time and correlating it with queries using machine learning algorithms, allowing models to be created and optimized for specific SQL statements.

Thus, machine learning helps make more efficient access path cost estimates, resulting in both faster query execution and a reduction in resource consumption, yielding significant performance improvements.

In turn, faster query results mean businesses can begin putting insights to use when they will have a greater effect—either because they beat the competition to an opportunity for differentiation or because they were able to respond to consumer needs closer to the time an action took place. As an added bonus, faster query completion also allows for a higher query throughput, providing the opportunity for more insights to be gained.

Confidence-based query results

One of AI's core benefits is that it augments people's decision-making ability in the course of their daily activities. In this spirit, an available capability for Db2 leverages machine learning to score the accuracy of historic SQL query results and use those scores to prioritize and reorder future results, prioritizing ones with higher degrees of confidence. In this way, users can quickly identify and deliver the most accurate insights for the business without having to rely solely on experience or instinct. Businesses can then pursue these opportunities confident in the knowledge that they have a greater chance of succeeding.

More intuitive insights

AI has an incredible potential to democratize access to insights throughout an organization. Db2 realizes this with the inclusion of IBM Db2 Augmented Data Explorer (ADE). ADE provides an intuitive self-service analytics portal, allowing users of every skill set to explore their data and generate insights. It brings up interesting statistical insights about their data without the need for complex search queries or extensive training. Users can discover insights through automatically generated graphs and visualizations, including the graphing of complex patterns and multivariate clusters. And results are displayed with natural language summaries, simplifying the interpretation of results even further. This multiplies the organization's ability to derive insight by de-centralizing the task and including more individuals while simultaneously allowing those insights to be acted upon more quickly than they could be if they needed to be communicated through a corporate structure.

Data science support

Data scientists are among the most valuable resources a company has. Consequently, every effort should be made to ensure their time isn't wasted. Db2 supports languages like PYTHON and GO, architectures like JSON and collaborative development environments like Jupyter Notebooks—helping ensure developers and DBAs have access to innovative data science tools. Ultimately, this serves two functions: Db2 developers can now access and utilize data science features to drive deeper insights, and data scientists can rely upon the performance, dependability, and general enterprise-readiness of Db2.

Multi-modal data management with Graph

Graph connects nodes to create relationships that can be retrieved using queries. The node represents an object, and by using an edge, graph databases are able to represent the relationship between nodes. Graph databases store complex sets of data and support dynamic multidimensional data management. The availability of Graph in Db2 11.5.4 enables a multi-modal database management approach, whereby data duplication and migration is avoided. This results in significant cost savings since an organization does not need to maintain multiple databases.

Db2 REST Service

From your web, mobile, or cloud application, interact with Db2 data through a set of scalable APIs to create, discover, execute, and manage user-defined services REST, which stands for Representational State Transfer, can be used with recurring tasks to run queries on the database based on pre-defined criteria; the REST service will provide notifications whenever changes in the data meet the predetermined criteria.

Db2 BLU Acceleration

The effectiveness and speed of analytics are often hindered by infrastructures that are unable to keep pace with the rate of data growth and change. IBM BLU Acceleration seeks to eliminate that concern by increasing processing speed with several advanced technologies, including:

– **In-memory computing:** This technology dynamically optimizes movement of data from storage to system memory to CPU memory. Db2 supports in-memory computing on existing infrastructures along with deeper analytics. This in-memory capability is optimized for SAP applications, transactional and analytics workloads. It's also ideal for workloads migrated from Oracle Databases, where Db2 provides an average of 98 percent compatibility.
– **Massively Parallel Processing (MPP):** The MPP-based cluster architecture in Db2 enhances query response time to better provide insight from real-time operational and historical data. Multi-core and single-instruction multiple-data (SIMD) processing are also available.
– **Actionable compression**: Reduce the size of data that needs to be stored while preserving the order of the data. This allows analytics to occur without the need to decompress the data, saving time and storage costs.

– **Data skipping**: By assessing which data is most relevant to a query and bypassing irrelevant data, Db2 is able to save time by foregoing unnecessary processing.
– **Column-based Db2 shadow tables**: BLU Acceleration utilizes an extra storage engine and integrated runtime directly in the core Db2 system to support the storage and analysis of column-organized tables in parallel to traditional row-based tables. This enables analytics on operational data directly within a transactional environment without compromising transactional performance and therefore allows investments in existing enterprise resource planning (ERP) environments and skills to be maintained, avoiding business disruptions.

Collectively, these technologies provide faster processing by simplifying or eliminating steps capable of slowing analysis. Potential benefits include more timely answers to business questions so users can take quicker action, massive storage savings thanks to actionable compression, and ease of implementation and management for both transactional and analytic data workloads.

The flexibility to use all data easily

The rate of change within most industries and the increase in data of all types, speeds and locations demands a database that provides extreme flexibility with considerable simplicity. The integration of Db2 with other hybrid data management solutions—including those from IBM, open-source solutions, and solutions from other vendors—as well as its various deployment options help achieve that goal.

Part of IBM Cloud Pak for Data

Db2 can be purchased as part of a platform that brings together solutions across the entire AI Ladder. This includes databases, data warehouses, data lakes, and fast data solutions as well as capabilities from other rungs of the ladder like Watson™ Studio OpenScale™ and many more. Businesses can start with just Db2 and add additional capabilities as needed by purchasing additional VPCs. In addition, governance is baked in from the beginning for better organization along with data virtualization. This enables access to data no matter where it resides, resulting in cost reductions since ETL processes are eliminated. And because Cloud Pak for Data is containerized it can work on any hardware that supports Linux® containers and any cloud that supports Red Hat®; this ensures that organizations are not locked in to a particular vendor, but rather have the choice to select a cloud vendor aligned to their AI strategy.

Better data access with a common SQL engine

Even with the right data management solutions at your disposal, a database will lose its effectiveness if it cannot reliably connect to the full range of data sources in an efficient, timely way. Db2 overcomes this challenge with its common SQL engine. The common SQL engine's built-in data federation allows Db2 users to access data from Db2 family offerings like IBM Db2 Warehouse, IBM Db2 Big SQL, IBM Db2 Event Store, IBM Integrated Analytics System and existing Netezza® offerings. Its data federation capabilities also extend to Oracle, Teradata, Microsoft SQL Server, cloud sources such as Amazon Redshift, and open-source solutions such as Hive.

This makes running queries across multiple data sources fast and easy, because they can be processed where the data resides, eliminating the need for data movement. In addition, you can write a query once and have it work across any offering in the Db2 family, such as the ones noted above. Access to more data allows more complete insights to be generated and shrewder actions to be taken. The time saved by no longer needing to move data or rewrite queries also means those actions can happen faster and more effort can be put towards uncovering other insights.

Cloud and multicloud flexibility

Many businesses have embraced a hybrid data management model that leverages both cloud and on-premises deployments and, more recently, have turned to multicloud strategies. In fact, in the Institute for Business Value's recent survey, 98% of companies predicted they will use multiple hybrid clouds within the next three years. The reasons for cloud and multicloud use include corporate mandates to reduce capital expenditures, spinning up short-term projects, and avoiding vendor lock-in. Yet, no matter why the decision is made, databases must be able to flexibly accommodate the business's cloud needs. Db2 provides the same experience and features across on-premises, hosted, and cloud options so developers can leverage existing skills and streamline transitions between ground and cloud. Moreover, Db2 can be deployed on a variety of public cloud platforms including IBM Cloud™ and Amazon Web Services so that you can continue (or begin) your multicloud strategy.

The IBM Db2 Hosted offering features the same functionality as its on-premises equivalent—along with rapid self-service cloud deployment and pay-as-you-go licensing.

Very Large Database (VLDB) improvements

Db2 11.5.4 brings big data to the online transaction processing (OLTP) system—previously a domain of warehousing only. With the increasing number of data sources and the extreme growth of data volumes in today's organizations, the OLTP system may contain hundreds of terabytes of data. Some of the enhancements aimed at very large database sizes and user populations include:

– Concurrency and scalability for recently and commonly referenced pages
– Features for higher throughput
– Performing online table reorganization tasks at the partition level
– IBM Db2 Database Partitioning Feature (DPF)
– Db2 Workload Management

Straightforward deployment options

Db2 recently simplified its editions structure to make the path from prototype to production deployment easier. All editions of Db2 11.5.4 now share a common install image, from the freely downloaded version to Db2 Standard and Advanced Editions. A simple, optional packaging structure is also offered including Db2 Advanced Recovery and Db2 Performance Management offerings.

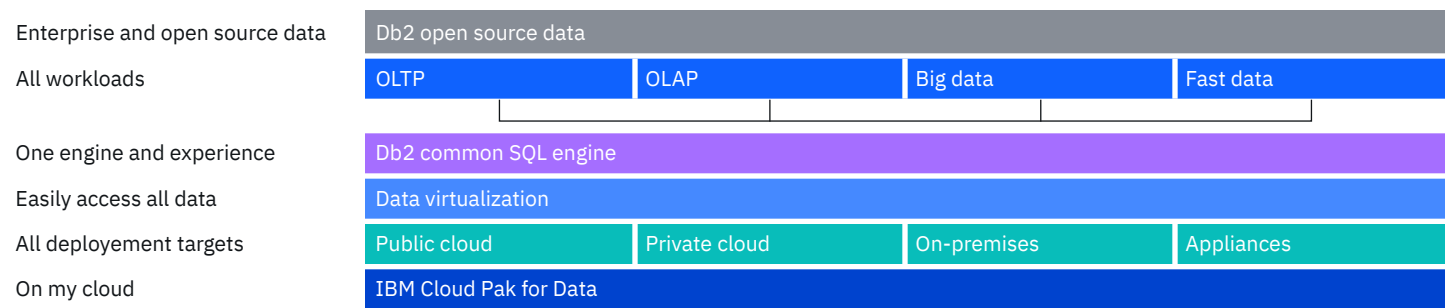| Enterprise and open source data	Db2 open source data			
All workloads	OLTP	OLAP	Big data	Fast data
One engine and experience	Db2 common SQL engine			
Easily access all data	Data virtualization			
All deployement targets	Public cloud	Private cloud	On-premises	Appliances
On my cloud	IBM Cloud Pak for Data			

Figure 1: IBM Hybrid Data Management provides one SQL engine across a myriad of workloads and deployments.

The four editions build upon one another for increasingly complex use cases:

– **IBM Db2 Community Edition:** This is an unsupported edition, intended for a single application developer to design, build, test and prototype applications for deployment on any of the Db2 client or server platforms. It includes all of the capabilities of Db2 Standard and Advanced Editions, but is limited to 4 cores and 16 GB of RAM and cannot be used in production systems.
– **IBM Db2 Base Edition:** This edition is designed for small to medium projects. It includes all the features of the Advanced editions, but is restricted to 8 cores and 64 GB of RAM and no storage limit and includes support. This version provides an ideal transition for users requiring support from the Community edition, at an optimal price point.
– **IBM Db2 Standard Edition:** This edition is ideal for medium-sized businesses and departmental deployments. It includes all of the same capabilities as the Advanced Edition, but is limited to 16 cores and 128 GB of RAM. It is provided as a standalone offering and delivered natively in IBM Cloud Pak for Data.
– **IBM Db2 Advanced Edition:** This edition is designed to meet the needs of medium-to-large businesses and is ideal for transactional and operational analytic workloads. It has no memory, storage, socket or core limits and can be deployed on physical and virtual servers. It is delivered as an extension for IBM Cloud Pak for Data, enabling administrators to access capabilities beyond the transactional database with greater ease.
– **IBM Db2 Advanced Recovery Solution:** This optional software bundle may be purchased separately. It helps improve data availability, mitigate risks and accelerate crucial administrative tasks. It can be used with all Db2 editions previously mentioned and includes IBM Db2 Merge Backup for Linux, UNIX and Windows V3.1; IBM Db2 Recovery Expert for Linux, UNIX and Windows V5.1; and IBM InfoSphere® Optim™ High Performance Unload for Db2 for Linux, UNIX and Windows V6.

The reliability enterprises depend on

The reliability of a database is dependent on several factors. The database must be secure enough that you can trust it with sensitive data and stay compliant with government regulations. It must have backup and disaster recovery so that it remains available even when unforeseen circumstances occur. And it must have tools capable of meeting specific business needs in a straightforward, simple way.

Robust security and encryption

Db2 11.5.4 builds upon its rich legacy of security by providing support for Key Management Interoperability Protocol 1.1 (KMIP 1.1). This allows integration with centralized enterprise key managers such as the IBM Security™ Key Lifecycle Manager and other products that support this industry-standard protocol. This centralized key management for many databases and file systems across the enterprise not only enhances security, but reduces the complexity for users as well, saving time they can put toward more valuable tasks. In addition, Db2 can be hosted in IBM data centers around the world to help meet regulatory requirements for keeping data in specific regions or countries.

Backups and recovery

Database availability is a paramount concern for most organizations, whether during day-to-day activities or in the event of a disaster. Db2 helps provide this availability in a number of ways. Foremost, IBM Db2 pureScale® clustering technology is designed to avoid both planned and unplanned outages with Geographically Dispersed Db2 pureScale Clusters (GDPCs). It helps support disaster recovery over multiple sites that are far enough apart to be on separate power grids. This feature can mean virtually no costly downtime, even during maintenance.

Setup is also simple. Db2 11.5.4 provides the ability to deploy pureScale and be up and running in a matter of hours with the simplicity of a push-button installation for pureScale clusters. According to IBM internal testing, the streamlined setup can improve the install process by at least 40 percent (sockets) or 25 percent (RDMA); plus, it reduces the number of steps from 30 to four for a native IBM General Parallel File System (IBM GPFS) replication setup process. This type of setup also introduces smarter defaults with intuitive options and quick pre-deployment validation across hosts, and can increase resiliency for aborted and partial installations with clean rollback for re-execution. Db2 11.5.4 supports all sync modes (SYNC, NEARSYNC, ASYNC and SASYNC) between the clusters for high-availability/disaster recovery (HADR) locally, across long distances or in the cloud.

Additional HADR support is available through change-queue-based replication and change data capture (CDC) replication, for the widest range of options. On-premises Db2 instances can also replicate data to other Db2 instances or Db2 Hosted, which can be your fail-safe in the event of a primary data center disaster. And Db2 users can also take advantage of backup and log compression acceleration. Db2 11.5.4 enables businesses to offload compression to hardware on POWER9™ processors and use IBM Active Memory Expansion to significantly reduce CPU consumption and elapsed time while maintaining most compression storage benefits.

Db2 Tools

Db2 provides access to a comprehensive set of database management solutions, covering a complete range of capabilities for the enterprise. This enables developers, architects and database administrators to design, develop, manage and deploy both transactional and warehouse databases with greater efficiency, scalability, performance and availability. To help you accelerate adoption and enhance the value of key Db2 features, all tools have been updated to support BLU Acceleration, compression and pureScale capabilities. A few of the tools available for Db2 include:

Tools	Description or function
IBM Data Server Manager	IBM Data Server Manager helps users administer, monitor, manage and optimize the performance of their IBM data management platforms across the hybrid data enterprise. It also enables DBAs and other IT staff to proactively manage performance and prevent problems before they impact the business.
IBM Advanced Recovery Feature	IBM Db2 Advanced Recovery Feature combines three Db2 tools for advanced database backup, recovery and data extraction. These tools help improve data availability, accelerate crucial administrative tasks and mitigate the risk of downtime, which can be very costly.
IBM Db2 Augmented Data Explorer[3]	IBM Db2 Augmented Data Explorer is an easy-to-use, web-based platform that connects to Db2 databases, whether on premises or on cloud, to instantly pull key insights through automatically generated visualizations and natural language summaries.
IBM Db2 Connect	IBM Db2 Connect helps manage access to your enterprise information, whether on premises or on cloud. For agile enterprises, it can deliver improved application enablement and a robust, highly scalable communications infrastructure for connecting data to web, Windows, UNIX and Linux applications.
IBM Data Studio	IBM Data Studio provides an integrated, modular environment to facilitate database development and IBM Db2 administration. It also offers improved collaboration through an open source integrated environment and database development tools for Db2 for z/OS®, Db2 for i, IBM Informix® and Db2 Big SQL.
IBM InfoSphere Data Architect	IBM InfoSphere Data Architect is a collaborative enterprise data modeling and design solution that can simplify and accelerate integration design for initiatives related to business intelligence, master data management and service-oriented architecture.

For more information about IBM database management solutions for Db2, visit: ibm.com/analytics/db2/tools.

Advanced Db2
11.5.4 capabilities

Below is a summary of key capabilities and technologies included in Db2 11.5.4. Each is designed to help businesses operate more efficiently and effectively.

Feature	Description or function
Common SQL engine	Allows businesses to write SQL queries once and deploy them anywhere against any data form factor—used throughout the Db2 family of Hybrid Data Management solutions.
Graph	Using nodes and relationships, graph collects and aggregates information-relating entities to enable complex modeling and analytics as well as provide intelligent visual insights.
Machine Learning Query Optimizer	Improves the performance and efficiency of queries using machine learning algorithms for significantly faster query performance.
Confidence-based Query results	Leverages machine learning to score SQL query results, prioritizing results with higher degrees of confidence so businesses can act on insights with a better likelihood of accuracy.
Data virtualization	Eliminates data movement and provides a single view of all data, making queries across multiple data sources fast and easy.
IBM Db2 Augmented Data Explorer[3]	Makes it easy for business users to pull and act upon key insights through automatically generated visualizations and natural language summaries. This capability is available as an add-on tool, free of charge.
Db2 REST service	As a REST (Representational State Transfer) service provider, Db2 enables your web, mobile, and cloud applications to interact with Db2 data through a set of scalable RESTful APIs. You can use the APIs to create, discover, execute, and manage user-defined services in Db2.
BLU Acceleration	Delivers breakthrough performance of in-memory columnar processing without the cost or limitations of in-memory-only systems, dramatically simplifying and speeding the delivery of business insights.
Compression	Helps reduce storage needs and increase performance using multiple techniques including table and index compression with page-level compression as well as Db2 BLU Acceleration with advanced encoding to maximize compression of columnar tables.
Continuous data ingest	Loads data continuously from multiple sources throughout the organization to support faster decision-making.
IBM Database Partitioning Feature (DPF)	Enables massively parallel processing by transparently splitting the database across multiple partitions and using the power of multiple servers to satisfy requests for large amounts of information.
Db2 pureScale	Delivers high availability and exceptional scalability to applications transparently, utilizing shared-disk, cluster technology to free them from the complexities of the underlying database architecture.
Db2 Workload Management	Enables fine-grain resource allocation, monitoring and management of workloads based on service classes, workload characteristics, elapsed time, time of day and more.
Federation Server	Supports federation between Db2 and other databases. It includes federation between Db2 and Oracle Database and Microsoft SQL Server for staged migration or long-term coexistence strategies.

Feature	Description or function
Python UDF support	Enables the ability to call Python code directly and supports the use of machine learning algorithms as database operations.
Materialized query tables (MQTs)	Improves the performance of complex queries with precomputed results of the whole query or parts of queries.
MQ replication / Change data capture (CDC)	Replicates large volumes of data at very low levels of latency.
Multi-temperature data management	Helps maximize performance and reduce overall media costs with storage tiering and the ability to transfer data in real time between different types of storage media.
Column store	Improves performance and reduces consumption of processor, memory and I/O resources for analytics workloads by directing scans to values in a specific column or columns, avoiding the need to process all data in a table.
Data skipping	Reduces processor, memory and I/O resource consumption by automatically avoiding the processing of data that is not needed for a query.
Shadow tables	Provides the performance benefits of BLU Acceleration to analytic queries that must be executed in OLTP environments with column-organized copies of row-organized tables implemented as MQTs that are maintained by replication.

Let Db2 help you climb the AI ladder

The foundation of a solid AI practice starts with hybrid data management solutions capable of providing the widest range of data to deliver the deepest insights. Db2 addresses these needs by infusing AI into the database itself. It is both powered by AI for greater optimization and query speed as well as built for AI with data science tools, data virtualization, and the ability to model more complex relationships.

AI infusion improves a Db2 database's performance significantly when paired with the processing advancements, compression, and data skipping already available in BLU Acceleration. The flexibility of the common SQL engine and IBM Cloud Pak for Data also help to ensure data of all types can be placed where it fits best and still be leveraged for more complete insights. And Db2 comes with the security, backup, and disaster recovery features an enterprise needs to maintain high levels of availability.

Experience Db2: The AI database

There has never been a better time to embrace AI data management. Begin experiencing what Db2 has to offer today with the first of its four editions—the free download for trials and developers.

Learn more about the remaining editions and optional packages by contacting your IBM representative or business partner, scheduling a free one-on-one consultation with a Db2 expert or visiting ibm.com/analytics/database-management

Additionally, IBM Global Financing provides numerous payment options to help you acquire the technology needed to grow your business. IBM provides full lifecycle management of IT products and services, from acquisition to disposition. For more information, visit: ibm.com/financing

1   Based on IBM design for normal operation under a typical workload using HADR and pureScale clusters. Individual results will vary depending on individual workloads, configurations and conditions, network availability and bandwidth.

2   Based on internal tests and reported client experience from 28 September 2011 to 07 March 2012.

3   Denotes a beta feature.