Bridging to the cloud: The right data management platform can make the difference
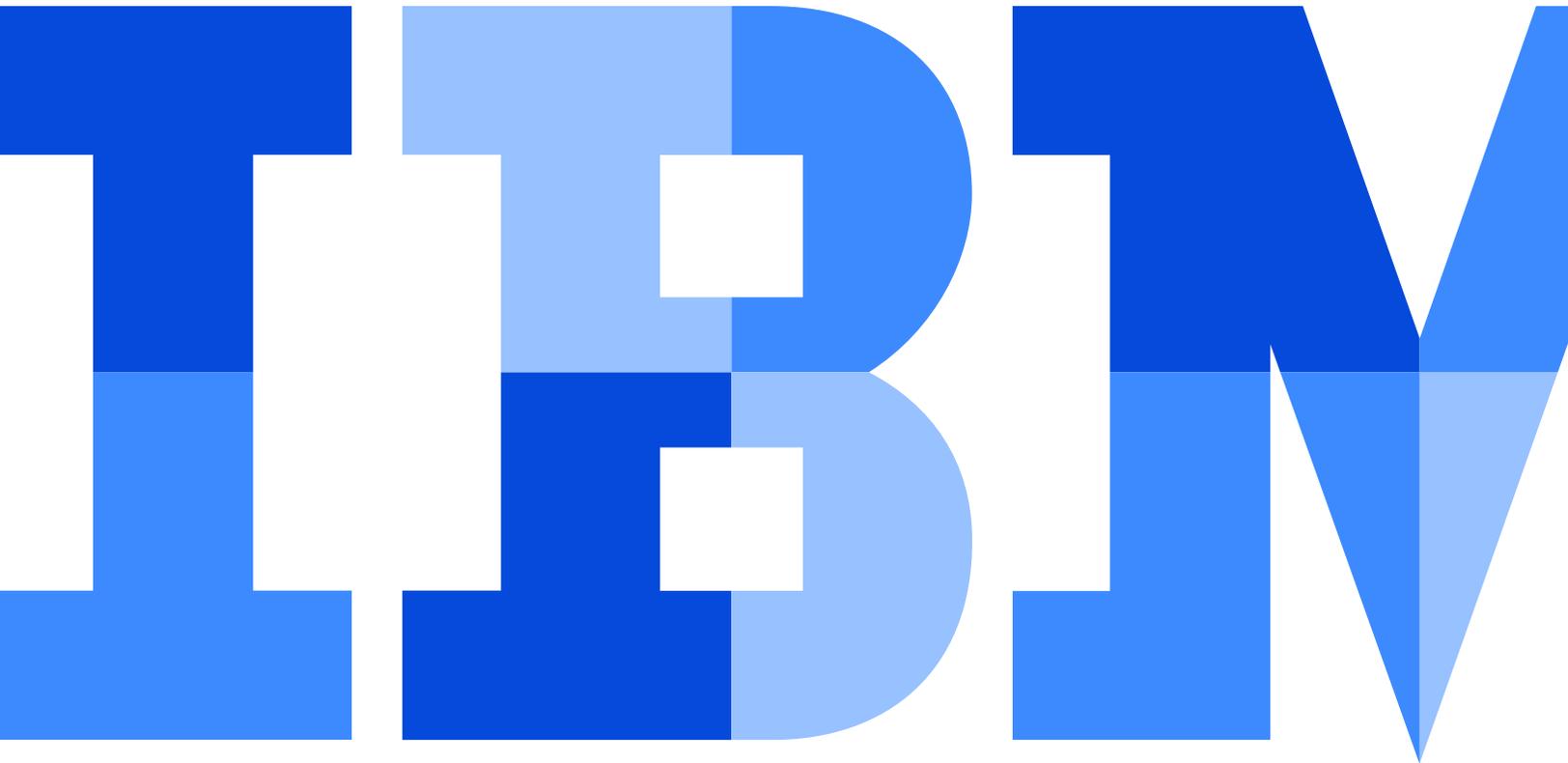
IBM

Contents

For years, enterprises have been encouraged to make the jump to the cloud from their on-premises systems. However, as noted by Aberdeen in their research report, "Hybrid IT is the new normal for IT Infrastructure," companies are finding that the advice to transition 100 percent to the cloud has been somewhat extreme and misguided. While the cloud does have advantages for certain data management workloads, on-premises deployments are often still necessary to address much of a business's data requirements. Given this fact, companies should focus less on full migrations to the cloud and more on building bridges to the cloud that extends and fuels digital business transformation on their own terms and at their own pace.

Development, deployment and delivery of applications and services depend and thrive on this hybrid approach. Workloads should be able to move between on-premises, public cloud and private cloud deployments as needed with efficient, reliable and speedy data access. As discussed in the recent ebook, "8 reasons to upgrade your data management," if this agility isn't possible with a company's current data environment, the organization should modernize its architecture. Having an on-premises system that easily integrates with the cloud better allows for use cases, such as:

- Deploying apps on-premises after cloud dev/test
- Adding capacity to on-premises systems with cloud bursting
- Combining sensitive data within your firewall and cloud data for analysis
- Improving data accessibility and self-service within an organization
- Managing workloads that accelerate data science and development initiatives

Only an enterprise-grade data management platform can deliver the solutions businesses require in a flexible, holistic and secure manner that allows data to reside where it fits best. This paper outlines what readers should consider when making a strategic commitment to a platform that will act as a bridge connecting new and legacy on-premises environments to the public and private cloud.

Data management requirements for optimal hybrid cloud use

To address the growing list of potential opportunities and challenges, businesses must have access to a wide breadth of data management capabilities that can be added or swapped as necessary to meet specific on-demand needs without procurement delays. These capabilities must be able to evolve with the market to address forward-looking needs. As enterprises build upon on-premises offerings with interconnected cloud solutions that create broad, forward-looking, and seamless hybrid environment, certain capabilities are critical, such as:

- **Support for all data types.** The amount of unstructured data available has seen astronomical growth recently due to factors like Internet of Things (IoT), social media, chat bots, image recognition, audio and speech applications, and expansive written content, such as email and text documents. Hybrid data management platforms must have solutions available that can support this unstructured data, traditional structured transactional data sets and semi-structured data—ideally using the same codebase. Without the ability to support each data type with an option best suited to its specific requirements, companies risk overlooking information needed to provide deeper insights or encountering delays.

- **The ability to grow AI capabilities as needs change.** Hybrid environments must be fluid—allowing organizations to pick the best technology for the job across all components of the AI Ladder and begin using it with little to no wait. One of the best ways to accomplish this is through IBM Cloud Pak for Data which has many capabilities that help across the AI Ladder built in or available as extensions. To introduce additional capabilities, simply add VPCs. Even if you only start with IBM Db2 on Cloud Pak for Data and the inherent governance capabilities of the platform, you can easily start with the base offering to build out IBM Db2 Warehouse, Watson Studio Open Scale and many other solutions spanning the entire ladder.

- **Interconnectivity and a single view of data.** Having a hybrid cloud environment necessitates a high level of interconnectivity; multiple data management options are of little value if data will remain siloed. Data virtualization and the ability to run analytics on data where it resides across multiple locations is essential to form the most complete insights capable of guiding your business.

  Similar codebases within the platform also helps avoid rewrites to queries and applications. By leveraging data federation, a hybrid cloud environment provides a single view of all data, simplifying the process of querying data across multiple sources. A good example of a technology that includes these capabilities is the IBM Db2 family's common SQL engine.

- **Simple scalability.** As demand spikes and users increase, data management platforms must be able to accommodate requirements for additional storage and computing power without unnecessary burdens on IT staff. One of the best ways to do this is to have a set amount of storage and power available on premises and then cloud bursting to meet peak needs. For those organizations maintaining a long-term cloud deployment, the ability to independently scale computing power and storage means they aren't stuck paying for extra resources.

- **Multicloud options.** A recent IBM Institute for Business Value (IBV) survey indicated that 98 percent of organizations "plan to use multiple hybrid clouds within three years". Often it's done to avoid vendor lock-in. Hybrid data management platforms should offer solutions on popular cloud services in ways that can easily communicate with existing on-premises deployments to ensure clients have a wide range of choices to build their data environments.

- **AI functionality.** Platforms leveraging artificial intelligence (AI) not only help build more innovative applications, they also help connect on-premises and cloud data management solutions better. AI-based query optimization helps determine the most efficient path to data over time, further eliminating inefficiencies that might otherwise separate data sitting in disparate deployments. Thus, on-premises and cloud data can be used together to drive insight much more easily. AI functionality can also return queries based on confidence in their accuracy, delivering the insights most likely to be correct first. Data scientist and developers will appreciate the support for a broad range of languages and frameworks in AI-ready data management technology.

- **Performance, availability and security.** Enterprises using the cloud within their hybrid environment need performance, availability and security that match its on-premises counterpart. A 99.99% uptime service-level agreement (SLA) and enterprise-grade support are key to providing a highly performant and highly available solution, mitigating the effects of unplanned outages and ensuring data integrity with offsite data center failover options.

   Security patching and maintenance should be fully managed by cloud providers, with the option for high-availability plans that incorporate rolling security updates. Daily backups, at-rest encryption, and SSL connections are mandatory for building the confidence enterprises need to continue moving workstreams to the cloud.

Meeting these requirements provides organizations with the flexibility to enter at any point in the cloud adoption curve, retain existing investments and leverage the combined capabilities across on-premises and cloud environments as part of a single platform.

How the IBM Db2 family builds a bridge across on premises and cloud

The IBM family of Db2 products is designed to meet these exacting demands with multiple deployment options and flexibility. Database, data warehouse, data lake and fast data solutions are all available to address the needs of varying types, amounts and speeds of data, as well as different workload requirements. These solutions include on-premises, hosted, and cloud deployment options along with multicloud availability on popular cloud providers, such as Amazon Web Services (AWS), Red Hat and IBM Cloud™ for specific products.

Each Db2 offering also employs a common SQL engine, which uses its built-in data virtualization capabilities helps ensure you can access data across the entire family, allowing users to write a query once and have it work across the Db2 family of offerings. The data virtualization also extends to appliances like the IBM Integrated Analytics System and IBM PureData® for Analytics, other vendors like Oracle, Teradata, and Microsoft SQL Server, cloud sources like Amazon Redshift, and open source solutions like Apache Hive.

Each Db2 offering employs a common SQL engine and can be used as a part of IBM Cloud Pak for Data which enables data virtualization capabilities. This helps ensure you can access data across the entire family of Db2 products and use them seamlessly with solutions spanning the entire AI Ladder. Data Virtualization in particular means that you can access data wherever it happens to reside without moving it.

Using Db2 as part of Cloud Pak for Data is beneficial in several ways. Foremost, expanding AI capabilities as the business grows becomes much easier. Simply adding additional VPCs opens access to a vast range of solutions. In addition, governance is standard as part of Cloud Pak for Data, so you can feel more secure in the idea that clean data will be organized and ready for use as part of AI. Data virtualization improves efficiency which would be lost if data needed to be moved or transferred. Finally, the ability to deploy on any cloud ensures that there is no vendor lock in.
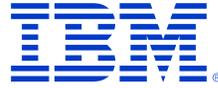
IBM Hybrid Data Management solutions focus on more than delivering an architecture that links on-premises and the cloud to meet current needs. Many Db2 family offerings are being infused with AI functionality to help businesses drive competitive advantage now and in the future. As described in Db2 V11.5 data sheet, these solutions are powered by AI and built for AI. This technology allows machine learning (ML) to provide up to 10 times the query speed improvements[1], ML-SQL to provide confidence-based results, and natural language querying to allow business users more easily find insights. The added support of popular data science tools and complex relationship modeling helps ensure that AI applications can be built more easily and deeper insights can now be found.

These AI capabilities Complement existing technologies that help boost performance, save time or control costs. For example, IBM Db2 BLU Acceleration® technology helps deliver in-memory speeds at scale with a combination of massively parallel processing (MPP), in-memory computing, actionable compression, and data skipping to improve query speed and reduce storage costs. High-availability disaster recovery (HADR) options and security features like encryption are also available to help improve resilience, business continuity and regulatory compliance.

Why IBM Db2 products make a difference

Embracing a joint on-premises and cloud strategy for data management is quickly becoming a reality for almost all business. Let the IBM Db2 set of solutions along with IBM Cloud Pak for Data help you bridge the gap with a wide breadth and depth of offerings and deployment options across the AI Ladder. The strong integration and flexibility will help you meet current needs while preparing for the future.

For more information

Learn more about the AI capabilities available with Db2 in the eBook, "Db2: The AI Database"

Start a free trial of IBM Cloud Pak for Data

Schedule a free one-on-one consultation with a data management expert

1 www.ibmbigdatahub.com/blog/why-analysts-are-hailing-db2-115-database-ai

Please Recycle