



IBM dashDB Local

Hybrid data warehousing using a software-defined environment in a private cloud

The evolution of the data warehouse

Managing a large-scale, on-premises data warehouse environments to meet today's growing analytic demands can be complex and expensive. It's time to now consider architecting a hybrid data warehouse that ultimately lowers the cost of analytics, enables unprecedented flexibility and delivers deeper insights.

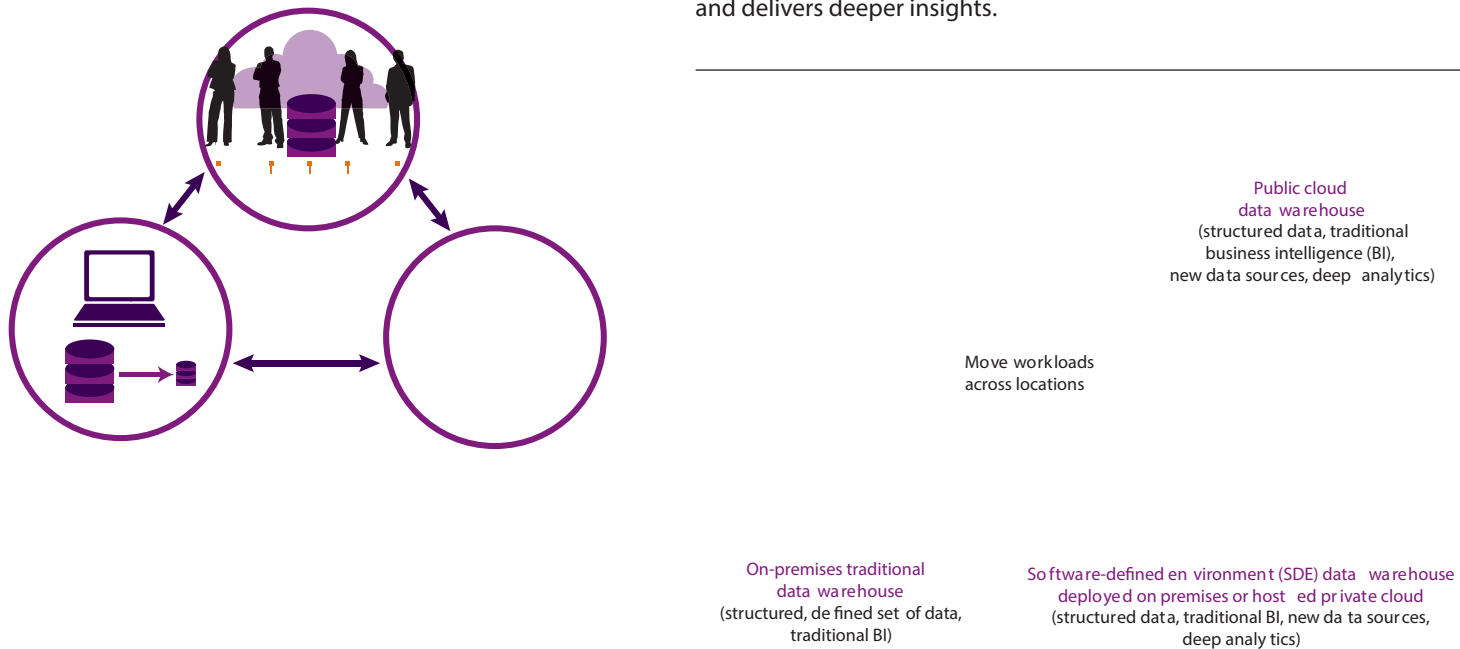


Figure 1: Hybrid data warehouse architecture



CBM

The image features the letters 'CBM' in a bold, sans-serif font. Each letter is composed of two overlapping, semi-transparent shapes. The top shape is a dark purple, and the bottom shape is a vibrant magenta. This creates a gradient effect across each letter. The entire logo is centered within a thick, solid blue square border.

A hybrid data warehouse introduces technologies that extend traditional data warehouse capabilities to provide key functionality required to meet new combinations of data, analytics and location, while addressing the following IT challenges:

- Delivering new analytic services and data sets to meet time-sensitive business initiatives
- Managing escalating costs due to massive growth in new data sources, analytic capabilities and users
- Achieving data warehouse elasticity and agility for sensitive business data

IBM dashDB Local enables hybrid data warehousing using a software-defined environment

IBM® dashDB™ Local is a client-managed, preconfigured data warehouse that runs in private clouds, virtual private clouds and other container-supported infrastructures. This data warehouse is designed to provide the ideal solution when you must maintain control of your data but want cloud-like flexibility. It includes in-memory processing to deliver fast answers to queries, as well as massively parallel processing (MPP) to help you scale out and scale up capabilities as demand grows. For analytics, you can use IBM dashDB Local to leverage familiar structured query language (SQL), integrated R and Python, or robust in-database analytics, including geospatial analytics.

IBM dashDB Local complements IBM's overall hybrid data warehouse strategy, providing organizations with the highly flexible architecture that is needed in the dynamic, fast-moving world of big data and the cloud. Due to the common analytics engine between IBM dashDB Local and IBM dashDB managed service, analytic workloads can be moved across public and private clouds without application changes. IBM dashDB technology is highly compatible with IBM DB2® and IBM PureData® System for Analytics, powered by Netezza® technology, as well as Oracle SQL. This compatibility helps move analytic workloads to IBM dashDB Local or the cloud more easily, depending on the application.

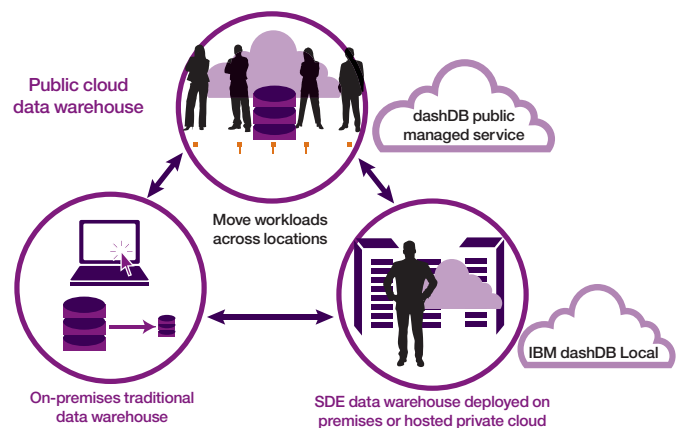


Figure 2: Introduction to IBM dashDB Local

IBM dashDB Local employs Docker container technology, which helps simplify management and reduce deployment times to minutes. It also provides elastic scaling and is designed for ease of updating and upgrading. All of these features are designed to be helpful to IT or cloud administrators. From a user standpoint, IBM dashDB Local helps provide the performance needed to quickly acquire data sets, apply analytics to solve specific business problems and operationalizing insights on the right deployment option.

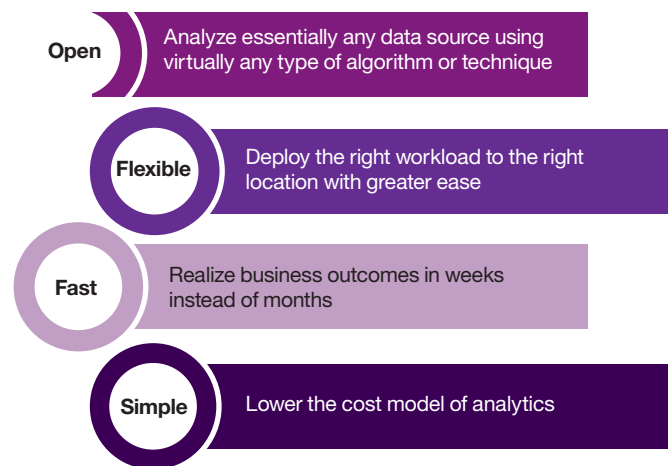


Figure 3: IBM dashDB Local key differentiators

Open

Users demand that IT departments deliver new analytic services, such as R, Python and Spark, as well as support a variety of data types. Not being able to serve these requests leads to the proliferation of analytic silos and less control of data. IBM dashDB Local helps IT departments gain control by delivering an open solution that's designed to make it easier to process data using virtually any type of algorithm or technique, across a range of data sources. Additionally, you can run IBM dashDB Local on standard hardware supporting Docker installations on Linux®, cloud, and on Apple OSX and Microsoft® Windows® platforms with minimal prerequisites.

The openness of IBM dashDB Local allows you to:

- Load a wide range of structured and unstructured data sources with greater ease as data is requested, including geospatial data, object storage and more.
- Use familiar business intelligence tools, as well as open source R, Python and Spark for in-database processing.
- Connect Esri ArcGIS to perform geospatial analytics.
- Run on standard hardware you already have in your data center.

Flexible

Various teams throughout your organization create analytic solutions using different data sources and tools. To bring all of these together into a single application can require several runtimes, as well as an optimal data flow that may span on premises and in the cloud. Additionally, IT departments must efficiently manage workloads to address the latest business needs, such as business-sensitive data and unpredictable demand. IBM dashDB technology addresses these challenges with a hybrid data warehouse architecture. IBM dashDB Local provides cloud-like agility and elasticity, while delivering advanced analytics to support the latest programming model and data source. Because IBM dashDB Local is part of a family of common database technologies, you can write your application once against it, and move that workload to the right location. Locations include public cloud, private cloud, or on premises—with minimal or no application changes required.

IBM dashDB Local flexibility allows you to:

- Choose to run Spark or SQL for analytical processing.
- Move workloads between locations –public, private, and on-premises data warehouse.
- Leave data where it resides using build-in IBM Fluid Query for federated queries.
- Employ elastic scaling across a wide range of infrastructure resources.

Fast

Even organizations with established data warehouses can benefit from a faster method for gaining business outcomes through analytics. In under approximately 30 minutes, users can deliver an optimized, private cloud data warehouse that complements and extends the core on-premises data warehouse. This hybrid architecture jumpstarts new analytic projects with minimal risk. The SDE, private cloud deployment takes advantage of underused resources, with self-service provisioning of the right combinations of data and analytic services. By leveraging an MPP architecture for IBM BLU Acceleration® in-memory processing and Netezza in-database analytics, users can quickly build and test analytic models against higher volumes of data.

IBM dashDB Local is fast because of:

- Push-button deployment in less than 30 minutes
- Self-service provisioning of the right combinations of data and analytic services or solutions
- In-memory IBM BLU Acceleration®, in-database Netezza analytics and Spark
- Availability in single node and MPP

Simple

An SDE is designed to optimize the entire computing infrastructure, including compute, storage and network resources. Additionally, an SDE automatically tailors itself to meet the needs of the required workload. IBM dashDB Local is delivered through Docker container technology and takes advantage of an SDE. For example, it auto-provisions resources to handle changing workload needs. It also makes deployment and management more efficient, with elastic scaling and ease of updates and upgrades. IBM dashDB Local provisions a full data warehouse stack, including Spark, in minutes to help you manage the service in your own public or private cloud, while maintaining existing operational and security processes.

IBM dashDB Local is designed for simplicity because of:

- Container technology that eases deployment and management
- System resources that dynamically adjust to satisfy variable workload demands
- Built-in Spark, which means you don't need to install and configure Spark separately
- Embedded high availability (HA) and disaster recovery (DR)

Is IBM dashDB Local right for you?

IBM dashDB Local can help you address needs, such as:

- More data warehouse capacity by leveraging an SDE to achieve elasticity to continually meet service levels, and maximize the use of existing resources, such as commodity hardware
- Dynamically provisioned resources to quickly gain access to the right combinations of analytic and data services
- A warehouse or data mart designed to deploy quickly and easily, with little tuning or management
- A cloud strategy built to keep data more directly under business control or on premises due to internal requirements and other mandates
- A cost-effective, high-performance processing engine to gain deeper insights from massive amounts of data being generated by mobile, web and the Internet of Things (IoT) applications
- A cost-effective alternative to rewriting applications to work with Hadoop, especially when working with structured data and commodity hardware

Get started: Example use cases

The following use cases are meant to inspire you to get started with IBM dashDB Local.

- **Prototyping, development or test ecosystem:** Quickly and more easily test new applications and data sources before production implementation.
- **Departmental or accelerated analytic projects:** Quickly start an analytic service that can meet requirements for a range of data sources, advanced analytics and application development.
- **Data warehousing as a service (DWaaS) or hybrid data warehouse:** Have the option to partially or fully migrate a subset of applications, data or both to the cloud from an on-premise warehouse.

“For a long time we have been providing traditional, on-premises database services and for approximately five years we have also offered IaaS cloud services. We recently stepped into IBM cloud data and analytic services for the first time and are looking forward to the general availability of IBM dashDB Local to utilize in additional customer projects and applications.”

— T-Systems

For more information

To learn more about IBM dashDB Local, please contact your IBM representative or IBM Business Partner, or visit: ibm.biz/dashDBLocal

To try IBM dashDB Local, [install](#) the Docker engine on the host server and configure POSIX-compliant Clustered File System Storage, for example, Global File System 2 (GFS2), GPFS™.



© Copyright IBM Corporation 2016

IBM Analytics
Route 100
Somers, NY 10589

Produced in the United States of America
December 2016

IBM, the IBM logo, ibm.com, BLU Acceleration, dashDB, DB2, GPFS, and PureData are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

Netezza is a registered trademark of IBM International Group B.V., an IBM Company.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

It is the user's responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.



Please Recycle