DEFINITION

What is a cloud database? An in-depth cloud DBMS guide

By **Craig Stedman,** Industry Editor

A cloud database is an organized and managed collection of data in an IT system that resides on a public, private or hybrid cloud computing platform. From an overall design and functionality perspective, a cloud database is no different than an on-premises one that runs on an organization's own data center systems. The biggest difference between them lies in how the database is deployed and managed.

For example, the same database appears identical to end users and applications, whether it's on premises or in the cloud. Depending on the particular database software that's used, cloud databases can store structured, unstructured or semistructured data, just as their on-premises counterparts do.

But using a cloud database changes the responsibilities of IT and data management teams. Cloud vendors install and manage the underlying system infrastructure and, in managed services environments, the database platform. That reduces the routine management work traditionally done by IT operations workers and database administrators (DBAs). A DBA can then take on other tasks, such as optimizing databases for applications and tracking the usage and cost of cloud database systems.

Like other IT systems, database deployments are clearly shifting toward the cloud. In a report published in December 2023, Gartner said cloud databases now account for more than half of total database management system (DBMS) revenues worldwide and nearly all the revenue growth in the market. Also, in a survey of 753 cloud users conducted in late 2023 by IT management tools vendor Flexera, 65% said their organizations were using data warehouses in the public cloud, while 57% had adopted cloud-based relational database services and 44% were using NoSQL ones. All those numbers were up significantly from the previous edition of the annual survey.

This comprehensive guide to cloud databases further explains what they are, how they work and their potential IT and business benefits for organizations, as compared with on-premises databases. You'll also find information on cloud database technologies, vendors and security issues, plus more details on database administration responsibilities in the cloud. Throughout the guide, hyperlinks point to related articles that cover those topics and others in more depth.

What is a Cloud Database? Cloud vs. On-Premises Databases

◼ How cloud databases work

In businesses, databases are used to collect, organize and deliver data to executives and workers for operational and analytics applications. In general, cloud databases provide the same data processing, management and access capabilities as on-premises ones. Existing on-premises databases usually can be migrated to the cloud, along with the applications they support.

Instead of traditional software licenses, pricing is based on the use of system resources, which can be provisioned on demand as needed to meet processing workloads. Alternatively, users can reserve database instances -- typically for at least a year -- to get discounted pricing on regular workloads with consistent capacity requirements.
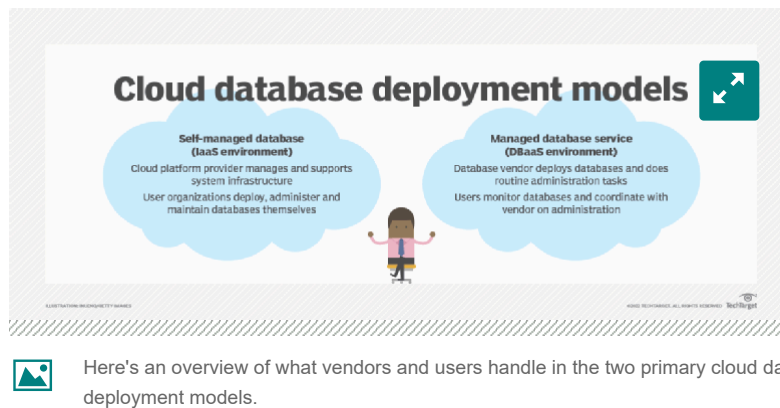
Organizations that are implementing databases in the public cloud choose between the following two deployment models:

1. **Self-managed database.** This is an infrastructure as a service (IaaS) environment, in which the database runs in a virtual machine on a system operated by a cloud provider. The provider manages and supports the cloud infrastructure, including servers, operating systems and storage devices. But the user organization is responsible for database deployment, administration and maintenance. As a result, it's akin to an on-premises deployment for the DBA, who retains full management control of the database.

2. **Managed database service.** Database as a service (DBaaS) environments are fully managed by the vendor, which could be a cloud platform provider or another database vendor that runs its cloud DBMS on a platform provider's infrastructure. Under the DBaaS model, both the system infrastructure and the database platform are managed for the customer. The DBaaS vendor handles provisioning, backups, scaling, patching, upgrades and other basic database administration functions, while the DBA monitors the database and coordinates with the vendor on some administrative tasks. Similar data warehouse as a service (DWaaS) offerings are also available for deployments of cloud data warehouses.

In addition, some cloud providers -- Amazon Web Services (AWS) and Oracle, for example -- offer versions of their DBaaS technologies for installation in on-premises data centers as part of a private cloud or a hybrid cloud infrastructure that combines public and private clouds. As with a regular DBaaS environment, the provider deploys the databases on its own systems and manages them for customers. But it delivers the systems to a customer's data center to run there and then manages the databases remotely.

Many vendors now also offer serverless databases in the cloud. Like DBaaS, they're managed services, and the two terms are sometimes used interchangeably. But there are some differences. For example, serverless systems automatically provide the

processing resources required by database applications and scale up or down as workloads fluctuate, while DBaaS commonly includes a specific amount of resources with scaling options. The term *serverless* is really a misnomer -- the databases do run on a cloud provider's servers. But they're effectively serverless from a customer's standpoint.

Here's an overview of what vendors and users handle in the two primary cloud dat deployment models.

■ Types of cloud databases

A wide variety of cloud databases are available, matching the different types of database technologies that can be deployed on premises. At this point, every notable database vendor offers its software in the cloud. That includes cloud-native databases developed specifically for use in cloud environments and existing on-premises databases that now support the cloud.

The following are the key types of databases that cloud users can take advantage of:

- **Relational databases.** Relational software that's accessed and managed with the SQL programming language has dominated the database market since the 1990s and remains the most widely used DBMS technology. It's particularly well suited for transaction processing and other applications involving structured data, thanks to the relational model's support for data integrity and consistency. Relational databases organize data in tables with rows and columns and use a fixed schema to help enforce consistency rules.

- **NoSQL databases.** NoSQL systems forego the rigid schemas of relational databases, making them a better option for semistructured and unstructured data, such as text, log files, sensor data and videos. Despite the term, many do provide some SQL capabilities and can also be used to store large amounts of structured data, particularly for applications that don't require complete consistency. Because of that, NoSQL more commonly means "not only SQL" now. There are four major NoSQL product categories: document databases, key-value databases, wide-column stores and graph databases.

- **Multimodel databases.** These databases support more than one data model, enabling them to run a wider set of applications. Many relational and NoSQL databases now qualify as multimodel through add-ons -- for example, the addition of a graph module to a relational DBMS or a NoSQL document database.

- **Distributed SQL databases.** Initially labeled as NewSQL databases and still referred to by that name in some cases, these technologies distribute relational databases across multiple computing nodes to create transactional systems that can provide NoSQL-like levels of scalability.

- **Cloud data warehouses.** First developed to provide data warehousing capabilities for business intelligence and reporting applications, cloud data warehouse and DWaaS technologies typically now also support development of data lakes that contain large amounts of raw data, as well as machine learning and other advanced analytics functions.

Specialized databases are also available for particular applications. Most notably, they include time series databases that hold time-stamped data stored in sequential order; vector databases designed to support large-scale similarity searches on sets of unstructured data; more conventional database search engines; and ledger databases that create an immutable record of transactions using blockchain and other cryptographic techniques.

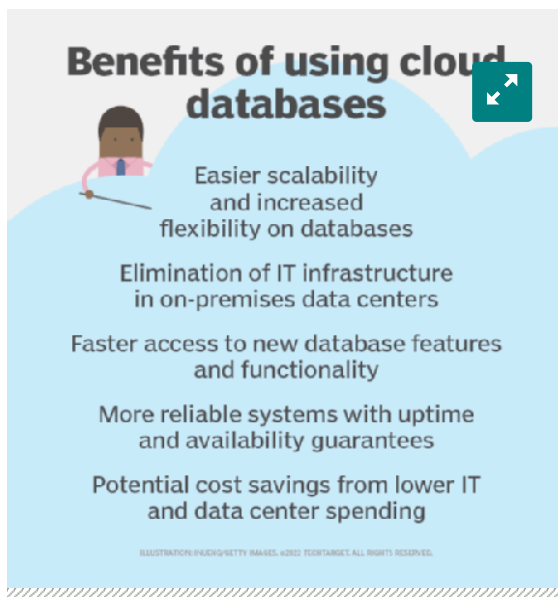## ■ Key cloud database management system components

Like other types of DBMS technologies, cloud database platforms include a set of components that work together to process and manage data. The list of key components includes the following items:

- A storage engine that manages data storage.

- A metadata catalog that contains data about database objects.

- A database access language, such as SQL, for querying and modifying data.

- A query optimization engine and a separate query processor.

- A lock manager to control concurrent access to data.

- A log manager to record changes made to the data.

- A set of database management utilities.

■ Cloud database benefits

Compared with running databases on premises, cloud databases offer the following potential IT and business advantages to user organizations:

- **Increased scalability and flexibility.** Cloud database systems can be easily scaled up by adding more processing and storage capacity when workloads increase. Some vendors offer autoscaling features that do so dynamically, without users even needing to submit a request -- a capability that's particularly prevalent with serverless databases. In addition, an organization can quickly deploy new databases and shut down ones it no longer needs, matching its database strategy to the speed of business.

Benefits of using cloud databases

Easier scalability and increased flexibility on databases

Elimination of IT infrastructure in on-premises data centers

Faster access to new database features and functionality

More reliable systems with uptime and availability guarantees

Potential cost savings from lower IT and data center spending

ILLUSTRATION:RHJDIKQ/GETTY IMAGES. ©2022 TECHTARGET. ALL RIGHTS RESERVED.

These are some of the IT and business benefits that organizations can get from using cloud databases.

- **Elimination of IT infrastructure.** Because the cloud provider is responsible for the system infrastructure in a cloud database environment, an organization might be able to reduce its own IT footprint by decommissioning systems, especially if it's moving on-premises databases to the cloud. At the very least, it can avoid the need to add more systems when it deploys new databases.

- **Faster access to new features.** With on-premises databases, users need to wait for and then install a software upgrade to get new features and functionality. DBaaS and serverless database vendors can update cloud databases on an ongoing basis, which enables organizations to take advantage of new features as soon as they're available.

- **More reliable systems with guaranteed uptime.** Cloud vendors provide high availability, automated backup and disaster recovery capabilities that often are more advanced than what user organizations have implemented. The vendors also guarantee uptime percentages as part of their cloud service-level agreement (SLA) with customers, giving them an incentive to keep cloud database platforms running smoothly.

- **Cost savings.** Reduced capital expenditures, data center operating costs and space needs in IT facilities, as well as possible IT staff cuts, can result in lower spending overall. But that isn't a sure thing: Pay-as-you-go cloud services can cost more than planned if resource utilization exceeds expectations or, conversely, if excess capacity goes unnoticed. A cloud database environment needs to be monitored closely to keep cloud costs under control.

On the other hand, on-premises databases might still be best for some organizations, particularly if they want to retain full control of the database environment or need to for regulatory compliance purposes. Other factors to consider when deciding between cloud and on-premises databases include the amount of data that would be transferred into and out of a cloud-based system and the choice of database administration and performance monitoring tools.

■ Migrating databases to the cloud

As mentioned above, migrating on-premises databases to a cloud environment can enable an organization to retire in-house IT systems and gain the other benefits of using cloud databases. Relocating a database to the cloud can also be an effective way to boost data processing efficiency and application performance as part of a broader cloud deployment.

But database migration can be a complex process. Before starting one, organizations need to consider various factors and plan a database migration strategy. For example, whether to migrate to a self-managed IaaS environment or a vendor-managed DBaaS one is a fundamental decision. Another is whether to migrate to the cloud version of the current DBMS or a different database technology. Changing databases can have financial or functional benefits, but it could also cause compatibility issues.

Even some related on-premises and cloud database technologies don't fully match up on features. For example, Microsoft's Azure SQL Database relational cloud service shares a common codebase with its SQL Server on-premises database, but there are differences between the two products that could require some reengineering of SQL Server databases before they can be migrated to Azure SQL Database. Azure SQL Managed Instance, a version of the cloud software that Microsoft developed to make database migration easier, still isn't 100% compatible with SQL Server.

◾ Cloud DBMS vendor landscape

Not surprisingly, the top cloud platform providers -- AWS, Google Cloud, Microsoft and Oracle -- are also the leading database vendors in the cloud, according to Gartner. They all support both IaaS and DBaaS environments on their own platforms and offer different types of cloud databases, including relational, NoSQL, data warehouse and special-purpose ones. Each makes more than 10 separate database services available to users.

The following are some other prominent cloud database vendors, based on vendor rankings by consulting firms such as Gartner and Forrester Research, DBMS popularity rankings on the DB-Engines website and additional research by TechTarget editors:

- IBM and SAP, two other major IT vendors that have transitioned from on-premises databases and now offer broad sets of cloud DBMS services.

- NoSQL database vendors Couchbase, DataStax, MongoDB, Neo4j and Redis, among others.

- Cloud data warehouse vendors Snowflake and Yellowbrick Data.

- Analytics database vendors Cloudera, Databricks and Teradata, which support data warehouses, data lakes and data lakehouses that combine aspects of the other two technologies.

- Multimodel database vendors InterSystems and Progress Software, which acquired the former MarkLogic in 2023.

- Distributed SQL database vendors Cockroach Labs and Yugabyte.

- Database search engine vendor Elastic.

- Alibaba Cloud and Tencent Cloud, two cloud platform providers that primarily operate in China and have extensive database portfolios.

Open source database options

Organizations can also use various open source databases in the cloud. Like other open source software, the databases are developed through a community process and their source code is openly available, although database vendors lead the development work in many cases. Popular open source relational databases include MySQL, PostgreSQL, MariaDB, Firebird and SQLite. Many NoSQL databases are also available under open source licenses.

Considerations to take into account in weighing open source vs. proprietary databases include cost, technical support needs and requirements for specific features and functionality. Open source databases can also help organizations avoid vendor lock-in because they're available from multiple providers. In addition, compatibility with technologies such as MySQL and PostgreSQL is built into some proprietary databases -- Amazon Aurora from AWS and Google's AlloyDB for PostgreSQL being two examples. As a result, users can often switch from one database service to another compatible one.

The open source and proprietary categories aren't mutually exclusive, though. While the community editions of open source databases can be deployed for free, vendors commonly offer commercial support or versions with proprietary features. For example, Oracle owns MySQL and sells several editions of the database, which is also offered commercially by AWS, Google, Microsoft and many other vendors. Similarly, PostgreSQL and MariaDB are available from a variety of vendors, including EDB, which focuses on PostgreSQL, and MariaDB PLC, which leads that database's development.

Some vendors that created open source databases have now switched to software licenses that aren't fully open source. Such licenses, often referred to as *source available* ones, align with most open source tenets. But they require other cloud providers looking to offer DBaaS implementations of a database to purchase a commercial license or make modified and related source code publicly available for others to use. Vendors that use these kinds of licenses include Couchbase, MongoDB, Redis, Cockroach Labs and Elastic.

■ What to evaluate when choosing a cloud database

The database is one of the most important technologies in any IT environment. Here are some of the features and issues organizations should examine when they evaluate cloud databases for planned deployments:

- **Performance.** As with any type of IT system, this is probably the top factor to consider, especially if the database will be supporting high-performance workloads. Scalability is a critical part of that -- for example, to make sure that real-time processing jobs don't bog down. Performance monitoring and tuning capabilities are another key aspect to look at.

- **Cost.** The major cloud providers offer free online cost calculators that can be used to check different scenarios on pricing models, service configurations, processing regions and other parameters to help balance expected resource needs and the available budget.

- **Availability.** High availability, disaster recovery and [data backup and recovery](#) capabilities should all be assessed, too, along with the cloud vendor's uptime SLA.

- **Security.** Securing a DBaaS environment isn't solely the vendor's responsibility, but it's crucial to know what it will handle and what security tools and measures it will apply.

■ Cloud database architecture considerations

The most straightforward approach for deploying cloud databases is to use a single public cloud platform. That ensures consistency on the underlying cloud infrastructure and a single cloud provider to work with, even if multiple DBaaS vendors are involved. But it might not always be feasible or meet an organization's IT and business needs. As a result, IT and data management teams might need to consider the following architectural strategies.

Hybrid cloud architecture

One option is [deploying databases across a hybrid cloud](#), putting some of them in a public cloud and others in a private cloud that's set up in an on-premises data center. Alexander Wurm, a senior analyst at advisory services firm Nucleus Research, said using a hybrid cloud enables organizations to "reap the benefits of the modern cloud, such as regular updates and elastic scalability, without interfering with the security and reliability of existing on-premises infrastructure in support of mission-critical workloads."

Some of the items to consider when planning a hybrid cloud database strategy include the following:

- Data migration requirements.

- Data security.

- Consistency and compatibility across cloud and on-premises platforms.

- Potential data latency issues.

- How to group applications and databases together into logical units to make the deployment process more manageable.

Multi-cloud architecture

A multi-cloud database architecture involves the use of multiple public cloud platforms. It can help avoid cloud provider lock-in and enable organizations to deploy different databases and applications in the cloud platform that best suits them. A multi-cloud strategy can also be incorporated into a hybrid cloud environment for an even more expansive approach to database deployment.

For organizations looking to take advantage of more than one public cloud, multi-cloud database management best practices include the following steps:

1. Start with a comprehensive plan and a governance framework.

2. Run the right database in the right cloud.

3. Use data services that support multi-cloud environments.

4. Exploit managed database services, or DBaaS.

5. Consider database portability across multiple clouds.

6. Reduce the number of different databases.

7. Reduce the number of the same databases.

8. Optimize data access for applications and end users.

9. Keep data local in one cloud platform when possible.

10. Connect cloud networks to reduce data latency.
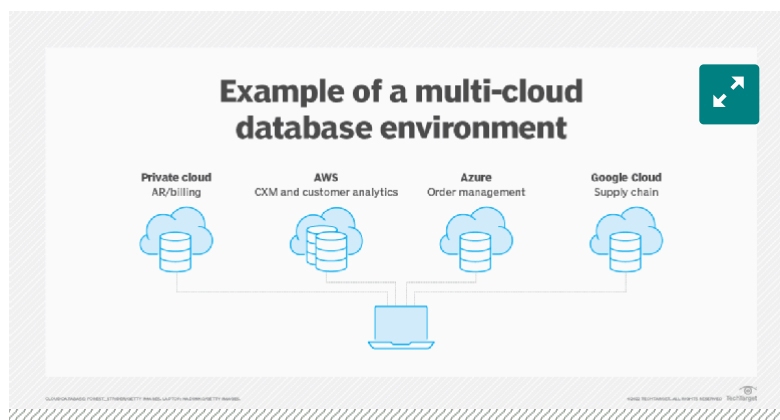
This sample diagram of a multi-cloud database architecture shows databases runr different applications in multiple clouds.

■ Cloud database security

As mentioned above, cloud database security isn't all on the vendor. What it handles can vary from vendor to vendor. But under the shared responsibility model for cloud security, users need to fully manage database security in IaaS environments where they deploy and manage the DBMS themselves. DBaaS vendors take on more responsibility for securing the database platform, but DBAs or security teams in organizations are usually still on the hook for things such as identity and access management, endpoint security, application security and some aspects of data security.

The following are some common challenges in securing cloud databases:

- Configuring and maintaining access controls.

- Managing database encryption.

- Enforcing user privileges and permissions.

To help avoid data breaches and exposures, database security best practices for user organizations include changing default logins and user credentials, using self-managed cryptographic keys and enabling full security logging capabilities, among other steps.

Cloud database management roles and responsibilities

Even in a DBaaS or DWaaS environment, DBAs play the lead role in managing an organization's cloud databases. The difference is that the cloud vendor takes over most of the regular, ongoing administration of a database platform. Instead of handling those basic tasks directly, the DBA can step in when necessary -- for example, to adjust data backup or system maintenance schedules because of application needs.

Cloud databases also add some new responsibilities to the DBA's role. In particular, monitoring the usage and cost of cloud database systems is a critical task for a DBA. That helps organizations avoid budget overruns and identify required changes in configurations or selected performance levels.
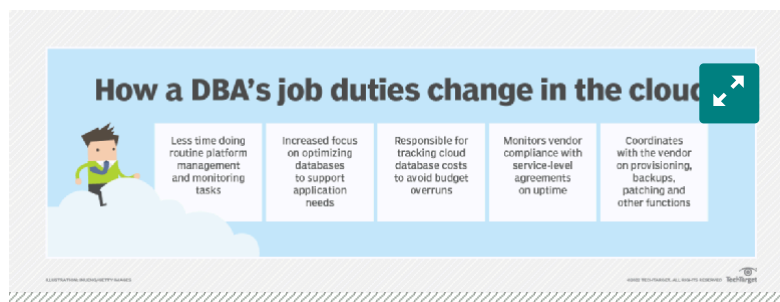
These are the key ways that cloud systems change the responsibilities of a databa administrator.

Cloud database trends to watch

The following are some current and emerging trends involving cloud databases:

- **Relational software's market dominance waning as users -- and vendors -- adopt alternatives.** While relational databases are still far and away the most-used DBMS technology, other types are increasingly being deployed by users and added to vendor product portfolios, independent analysts Merv Adrian and Sanjeev Mohan wrote in a February 2024 blog post. Relational software once accounted for more than 90% of worldwide DBMS revenue, but its share of the market has dipped below 80%, according to Gartner. Partly in response to that drop-off, relational DBMS vendors are building more nonrelational capabilities into their products. Gartner predicted that by 2027, relational systems will include 80% of the practical functionality of NoSQL databases, up from 60% in 2022.

- **Surging interest in vector databases to support generative AI development.** One of the nonrelational technologies that's seeing broader adoption is vector database software. Vector databases provide numerical representations of unstructured data in a multidimensional space to help users find similar data in, for example, large amounts of text. They've been a niche technology since the early 2000s, but the rise of generative AI has significantly expanded their use. Vector databases are well suited to storing, managing and retrieving the data used in the large language models that underpin ChatGPT and other GenAI tools. As a result, new vector database use cases are emerging in areas such as customer support, fraud detection and natural language processing.

- **Addition of GenAI capabilities to other types of databases.** In addition to pushing vector databases for generative AI uses, DBMS vendors are adding GenAI tools to help users develop and manage other databases. For example, GenAI assistants can be used to write database application code and generate SQL queries.

- **Incorporation of cloud databases into broader data ecosystems.** Database vendors increasingly are moving to integrate their software more tightly with other data management technologies, according to Gartner. It describes unified frameworks of that sort as *data ecosystems*, while others refer to them as *modern data stacks*. One aspect of the ongoing work involves tying cloud databases to data fabrics, an architecture for automating data integration processes and making them reusable.

Craig Stedman is an industry editor who creates in-depth packages of content on analytics, data management, cybersecurity and other technology areas for TechTarget Editorial.

Freelance technology writer Robert Sheldon and former TechTarget news writer Joel Shore contributed to this article.

This was last updated in April 2024

⤵ **Continue Reading About What is a cloud database? An in-depth cloud DBMS guide**

- ▪ Vector vs. graph vs. relational database: Which to choose?

- ▪ Vector search now a critical component of GenAI development

- ▪ Comparing DBMS vs. RDBMS: Key differences

- ▪ Enterprise database security best practices

⤵ **Dig Deeper on Cloud app development and management**

Cloud database comparison: AWS, Microsoft, Google and Oracle

By: Chris Foot

database management system (DBMS)

By: Kinza Yasar

database (DB)

By: Alexander Gillis

Compare NoSQL database types in the cloud

By: Kurt Marko

Search Data Center

📄 **Object storage, VM offerings emerge for HPE GreenLake**

HPE aims to attract new customers to its GreenLake ecosystem by launching standalone virtualization software. It also eyes AI ...

Understand the power usage effectiveness metric