

Cloud Databases

Leon Radó

December 24, 2023



**SLOVAK UNIVERSITY OF
TECHNOLOGY IN BRATISLAVA**
FACULTY OF INFORMATICS
AND INFORMATION TECHNOLOGIES

Contents

1	Introduction	4
2	Cloud Database Implementation	5
2.1	Virtual Machine Image	5
2.2	Database as a Service	5
3	Types of Cloud Databases	7
3.1	Relational Databases	7
3.2	NoSQL Databases	7
4	Database Search	8
4.1	Indexing	8
4.2	Query Language	8
5	Benefits of Cloud Databases	9
5.1	Accessibility	9
5.2	Scalability	9
5.3	Cost-Efficiency	9
5.4	Managed Services	9
5.5	Disaster Recovery	9
6	Disadvantages and Concerns of Cloud Databases	10
6.1	Data Security	10
6.2	Limited Control	10
6.3	Network Dependency	10
6.4	Vendor Lock-In	10
7	The Latest Trends in Cloud Databases	12
7.1	Serverless Databases	12
7.2	Multi-Cloud Database Solutions	12
7.3	AI in Database Management	13
8	Conclusion	14

Abstract

This article introduces cloud databases and provides a brief overview of how they operate. From the beginning, the reader will gain basic knowledge of cloud computing, especially databases. The paper explains how these databases function, what methods we have for implementation, and what benefits or drawbacks a user can experience. Later, the reader will understand what are the main types of databases as well as how the information is retrieved from these databases by the usage of indexing. Afterward, this writing describes what are the main advantages as well as drawbacks of cloud databases. In this section, essential properties of cloud databases, like cost-effectiveness, flexibility, accessibility, and security are evaluated. In the end, it mentions the latest trends and the future of cloud computing, like serverless databases and multi-cloud database solutions, which can partly eliminate all of the major concerns associated with these databases. After reading this paper, the reader will gain basic knowledge of database systems and understand the importance and rising popularity of cloud computing.

Keywords: *database, cloud computing, relational database, nosql, cloud provider*

1 Introduction

A cloud database is a particular type of database that works just like a standard on-premise database system but differs by its type of placement. While traditional databases exist on physical servers, cloud databases operate in a cloud environment. That means the database system is not located in the business infrastructure, but instead, it is accessible via the cloud by a specific cloud service provider. *Cloud computing has become the norm in today's computing world. This paradigm shift is looking for databases that are compatible in all respects.*[1] Each of these two types has an advantage over the other in terms of cost-effectiveness, accessibility, performance, security, and many other important characteristics, so it is up to the business owner to choose the right option for the purpose and requirements of his business.

In this article, I will mainly focus on the benefits and disadvantages that come with cloud databases to understand how these new types of databases work and what improvements they can offer to the world of data management.

2 Cloud Database Implementation

There are two primary methods to implement a cloud database, the thing they have in common is the presence of a third-party cloud service provider. One of these options is operating the database on the cloud independently of the provider with the virtual machine. Another option is using the database as a service (DBaaS).

2.1 Virtual Machine Image

Providers allow users to create a virtual machine image to directly access a database stored on the cloud. This virtual machine includes the database software, operating system, and necessary configurations.

The owner pays for the specific resources he needs, the cost is predictable and based on the resources and services he selects. This type of implementation requires manual and more demanding management but offers better control of the database and more options for customization.

2.2 Database as a Service

DBaaS is a more commonly used method to implement a cloud database. This service provides a database for the user through the Internet and includes hosting, setup, and management of the system, so the user doesn't have to worry about additional complex work related to maintaining a database. *Enterprises may focus on their core competencies instead of having to maintain complex IT landscapes.*[2]

DBaaS most often follows a pay-as-you-go pricing method, which means the owner pays just for the resources utilized at the moment.

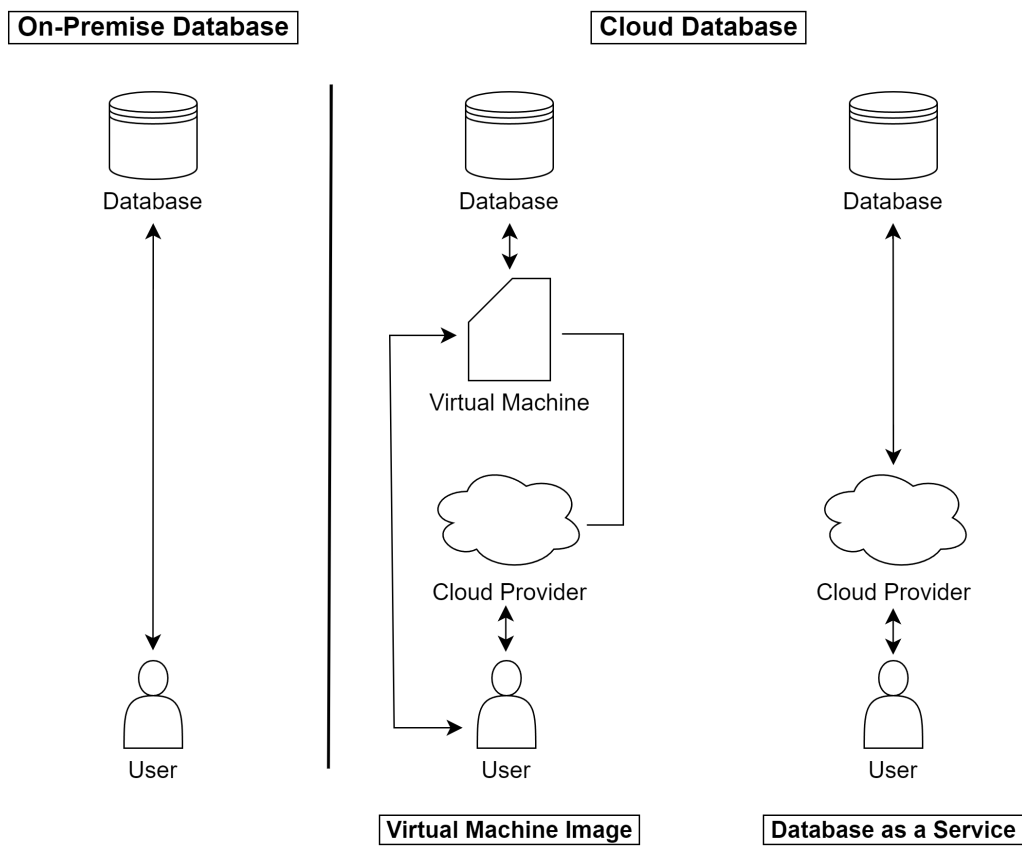


Figure 1: Comparison of Database Systems

3 Types of Cloud Databases

Just like standard on-premise databases, cloud databases can be divided into relational databases and non-relational, usually referred to as NoSQL databases. When choosing the right database it's important to consider the strengths, weaknesses, and use cases of both types.

3.1 Relational Databases

These databases store structured information organized into rows and columns, the structure of the data has to be defined in advance. Relation databases are typically used for applications with structured data, for example, transactional systems, and offer consistency and reliability, however, they are usually scaled vertically, which means they can be demanding on server resources like CPU and RAM.

3.2 NoSQL Databases

Non-relational databases are designed to work with a large amount of unstructured data, that is most often stored in key-value pairs without predefined format. This type of database is adequate for example for web applications that store a variety of information. *NoSQL databases were born to meet performance needs, leaving other details like atomicity in the background.*[3]

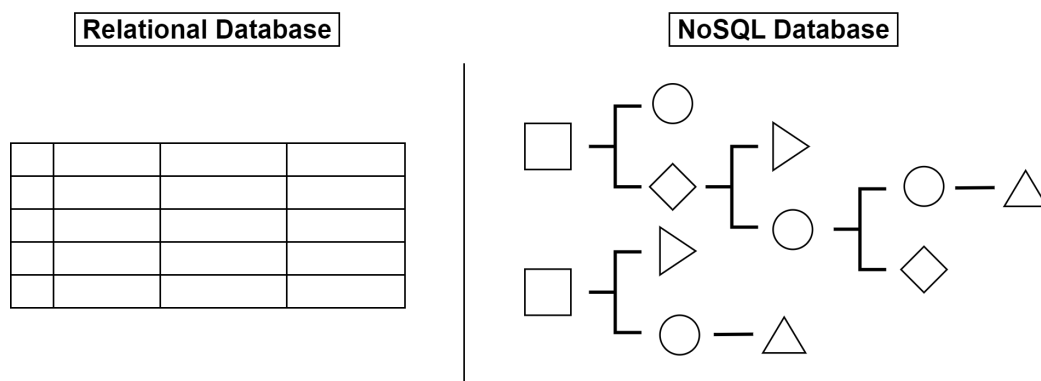


Figure 2: Structure of Relational and NoSQL Database[4]

4 Database Search

In order to retrieve information from a database, the corresponding database has to go through the process of database search. After the user provides a search query related to the information he wants to extract, the database starts searching for the relevant data.

4.1 Indexing

Comparing every record stored in the database to the specific search query can be really time-consuming, especially for larger database systems. However, this problem can be effectively resolved by using indexes. *To gain the speed benefits of indexing at retrieval time, the indexing algorithm is proposed and an index table is built by this algorithm in advance.*[5]

Indexing ensures that each record in the database obtains a new key-value pair data structure, where the key is a column with the content that might be potentially queried, and the value is a reference to the exact location in the database.

4.2 Query Language

When the user needs to extract information from a database, it is necessary for him to use language compatible with that database. This language will mostly depend on whether the database is relational or non-relational but also on the respective database provider.

The following example shows a search query for data from the 'books' category by the author 'Joel Lord' from the relational and NoSQL databases, respectively.

```
SELECT * FROM books WHERE author = "Joel Lord"
```

Figure 3: Search Query in SQL (Structured Query Language)[6]

```
db.books.find({"author": "Joel Lord"})
```

Figure 4: Search Query in MongoDB Query Language[6]

5 Benefits of Cloud Databases

5.1 Accessibility

Accessing a cloud database isn't limited by geographical location and can be done from anywhere with a stable Internet connection. This feature is especially useful for organizations that need access to the database from different locations.

5.2 Scalability

Cloud databases are designed to be highly scalable, which means they can process vast amounts of data with great performance. They also support a large amount of users and applications without the need for additional hardware installation.

5.3 Cost-Efficiency

Because of the missing need for investment into hardware installation, cloud databases are more cost-efficient. *Cloud computing can provide almost immediate access to hardware resources, with no upfront capital investments for users, leading to a faster time to market in many businesses.*[7]

The pay-as-you-go model also gives assurance that enterprises pay for the resources, they actually utilize. However, it's always important to regularly analyze and optimize the resources to avoid unnecessary overspending.

5.4 Managed Services

Cloud databases are in most cases managed by cloud providers, which results in enhanced reliability because of their high level of expertise in this field. Owners don't have to worry about tasks like database administration, patching, and maintenance.

5.5 Disaster Recovery

Cloud providers offer services to automatically backup and recover the data in need, which results in a decreased risk of data loss in case of natural disaster or hardware malfunction. *Since all the data is stored in the cloud, backing it up and restoring the same is relatively much easier than storing the same on a physical device.*[8]

6 Disadvantages and Concerns of Cloud Databases

6.1 Data Security

Despite the fact cloud databases have high security standards, there is always a risk of information theft and unauthorized access. *The chances that the host company may violate the privacy of its customers and access data without permission, make some potential customers worried.*[9] Since the data aren't stored locally, the owners need to reconsider their choice of storing sensitive information.

6.2 Limited Control

These types of databases are almost fully managed by a third-party provider, which can be unsuitable for businesses with special requirements because they aren't aware of how their data are accessed and managed. Options for customization are also very limited.

6.3 Network Dependency

Cloud databases rely on the Internet connection. This can be really problematic during network outages or other associated issues, especially for applications requiring low-latency real-time response.

Furthermore, this type of service is always prone to serious malfunctions and interruptions. *Even the best cloud service providers run into this kind of trouble, in spite of keeping up high standards of maintenance.*[10]

6.4 Vendor Lock-In

Organizations that began to use a cloud database are heavily dependent on the particular cloud provider. Migrating to another cloud database or an on-premise database in the future can be very difficult and expensive.

Cloud Databases			
	Benefits	Disadvantages	Concerns
	Accessibility	Limited Control	Data Security
	Scalability	Network Dependency	Vendor Lock-In
	Cost Efficiency		
	Managed Services		
	Disaster Recovery		

Table 1: Benefits, Disadvantages, and Concerns of Cloud Databases

7 The Latest Trends in Cloud Databases

In the last few years, it has been possible to observe a high increase in the number of organizations transferring their data to the cloud. This transition led to the latest innovations and further development in this area. *Advances in cloud computing technologies have changed the scene now and computing is being offered as a service over the internet.*[11]

7.1 Serverless Databases

A serverless database is a type of cloud database that doesn't require any additional management on the server side from the user. These databases are automatically scaled, meaning the storage resources are automatically allocated and scaled up and down based on the current demand. Users can see information about the number of reads and writes and the size of utilized resources.

7.2 Multi-Cloud Database Solutions

Many enterprises utilize cloud databases by at least two different providers. This strategy is referred to as a multi-cloud database solution, and although it has its challenges, it comes with significant security improvements. *The same tasks are performed on multiple clouds and their results are compared thereby the user can be assured of the integrity of the data.*[12]

With the implementation of this architecture, users are able to choose what data they want to store in each database as a result of not being dependent on one provider. This method also lowers the risk of potential vendor lock-in, and in case of natural disasters or service outages, businesses can rely on other providers.

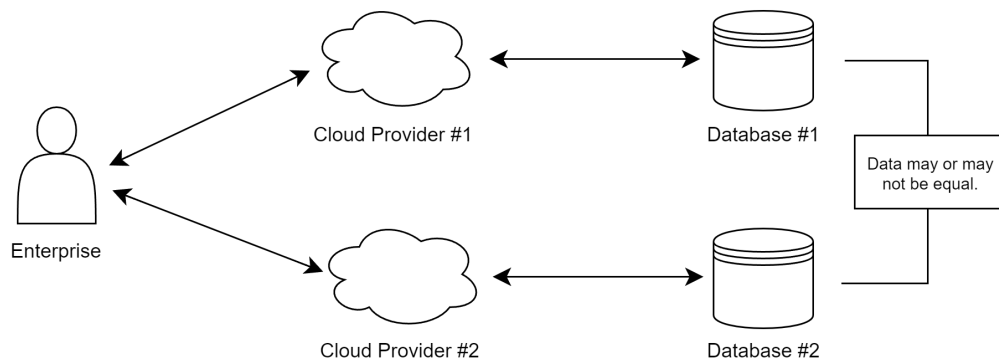


Figure 5: Multi-Cloud Architecture[13]

7.3 AI in Database Management

With the expansion of artificial intelligence in recent years, apart from everything else, it has also become a tool to enhance the database management process. Machine learning algorithms help to boost performance, predict analytics, and perform automated maintenance tasks. Last but not least, embedding natural language processing (NLP) for querying allows users to query for data by using human language.

8 Conclusion

To sum up, in these times, we are experiencing a significant shift in movement towards cloud computing and databases. *The rise of cloud computing opens a whole new future for the information technology area. However, open systems and shared resources give place to many security challenges, making security one of the main obstacles to the deployment of cloud computing technologies.*[14]

This paper introduced various cloud database systems and their implementation process and later compared their benefits and drawbacks. We can identify cloud databases as cost-effective, fast, scalable, and easier to utilize but on the other hand also less secure, and dependent on the network connection and particular cloud provider.

Although these disadvantages can be partially minimized by multi-cloud architecture, a cloud database still might not be the optimal solution for every occasion. It's always on the organization to choose the best option for the type and specifications of the stored data. However, cloud databases are becoming more and more widespread every moment and this rise doesn't seem to stop anytime soon.

References

- [1] Niraja Jain, A V Deshpande, B Raghu, and V Khanaa. The effectiveness of cloud databases. In *2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET)*, pages 1–3, 2017.
- [2] Wolfgang Lehner and Kai-Uwe Sattler. Database as a service (dbaas). In *2010 IEEE 26th International Conference on Data Engineering (ICDE 2010)*, pages 1216–1217, 2010.
- [3] Ticiana Capris, Pedro Melo, Nuno M. Garcia, Ivan Miguel Pires, and Eftim Zdravevski. Comparison of sql and nosql databases with different workloads: MongoDB vs mysql evaluation. In *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, pages 214–218, 2022.
- [4] M. Sluga. *AWS Certified Developer - Associate (DVA-C01) Cert Guide*. Cert Guide. Pearson Higher Education & Professional Group, 2020.
- [5] Phyo Thu Thu Khine, Htwe Pa Pa Win, and Khin Nwe Ni Tun. Keyword searching and browsing system over relational databases. In *2011 Sixth International Conference on Digital Information Management*, pages 121–126, 2011.
- [6] Inc. MongoDB. Database search introduction. *MongoDB - Basics*, November 2023. <https://www.mongodb.com/basics/database-search>.
- [7] M.G. Avram. Advantages and challenges of adopting cloud computing from an enterprise perspective. *Procedia Technology*, 12:529–534, 2014. The 7th International Conference Interdisciplinarity in Engineering, INTER-ENG 2013, 10-11 October 2013, Petru Maior University of Tirgu Mures, Romania.
- [8] Anca Apostu, Florina Puican, Geanina Ularu, George Suciu, and Gyorgy Todoran. New classes of applications in the cloud. evaluating advantages and disadvantages of cloud computing for telemetry applications. *Database Systems Journal*, 5(1), 2014.
- [9] Manjeet Singh. Study on cloud computing and cloud database. In *International Conference on Computing, Communication & Automation*, pages 708–713, 2015.
- [10] Anca Apostu, Florina Puican, Geanina Ularu, George Suciu, Gyorgy Todoran, et al. Study on advantages and disadvantages of cloud

computing—the advantages of telemetry applications in the cloud. *Recent advances in applied computer science and digital services*, 2103, 2013.

- [11] Niraja Jain. The future of database services: Cloud database. In *2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET)*, pages 1–5, 2017.
- [12] Kiran Baby and Anupriya Vysala. Multicloud architecture for augmenting security in clouds. In *2015 Global Conference on Communication Technologies (GCCT)*, pages 474–478, 2015.
- [13] Google Cloud. Multi-cloud database management: Architectures, use cases, and best practices. *Cloud Architecture Center*, October 2022. <https://cloud.google.com/architecture/multi-cloud-database-management>.
- [14] Sara Ibn El Ahrache, Parisa Ghodous, Hassan Badir, and Abderrahmane Sbihi. Secure processing of large scale databases in a cloud computing environment. In *2013 ACS International Conference on Computer Systems and Applications (AICCSA)*, pages 1–4, 2013.