Storage Limitations

by Sophia

WHAT'S COVERED

This lesson explores the storage limitations of various databases, in four parts. Specifically, this lesson will cover:

1. Introduction

When creating a new database, choosing the right DBMS is important. This decision affects not only the database's performance but also user experience, cost, reliability, and storage options. This lesson looks at the question from the perspective of that last item: storage.

The choices you make will affect your database's three V's: velocity, volume, and variety. You must consider storage type, capacity, speed, failure rate, cost, and data input and output.

BIG IDEA

Your storage capacity needs will feature heavily in your DBMS choice, as not all database systems have the same storage offerings. Storage limitations may originate within the DBMS software or can be vendor-imposed.

Most commercial databases have different storage limitations (in other words, volume), and different versions of specific databases may also have different storage limitations. In PostgreSQL, MariaDB, and MySQL, the database itself can be unlimited in size, so your storage space is limited only by your hardware. Other databases, like Oracle, Microsoft SQL Server, and SQLite, have a maximum cap in terms of database size.

These caps can affect the database's security, compliance, cost, access methods, and integration with infrastructure. They can also affect your options in terms of redundancy, scaling, and availability. Data size and data management are multifaceted, and there are many factors to consider.

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.

Some of these factors include the following:

- Compliance and Security: Identify any specific compliance requirements (e.g., GDPR, HIPAA) that impact data storage, encryption, and access controls. Implement security measures to protect sensitive data.
- Cost: Evaluate the cost of various storage options (e.g., local storage, cloud storage) based on the database size, performance requirements, and budget constraints.
- Data Access Patterns: Understand data access and query usage patterns. Frequent or complex queries may necessitate additional indexing or caching strategies to improve database performance.
- Data Volume: Determine the expected amount of data the database will handle over time. You should consider both the current volume of data and possible future growth to make sure the storage will be able to handle the data without running out of space.
- Integration With Infrastructure: Consider how the database storage integrates with the overall IT infrastructure and the compatibility with the chosen database management system.
- Redundancy and High Availability: Consider redundancy and high-availability solutions to prevent data loss in case of hardware failures or disasters. Replication, clustering, or data mirroring can be used to enhance availability.
- Scalability: Plan for future growth and ensure the database storage can scale easily to accommodate increasing data volumes and performance demands.

2. Overall Size

Two of the biggest database vendors include Microsoft SQL Server and Oracle.

SQL Server has an Express edition that limits the database size to 10 GB. The database itself is free and is scalable and completely compatible with the paid editions, allowing seamless swapping. However, this limitation does force organizations to make the switch to the paid version at the point where the database is heavily used.

Oracle also has a free version, with a current limit of 12 GB for the database, which also is scalable and can be swapped to a paid version. In comparison, the enterprise edition of SQL Server has a maximum size of 524 petabytes (PB), while Oracle has over 2047 PB. As a comparison, 1 PB is equivalent to 1,000,000 GB.

3. Data Type Size Limits

There may be different size limitations even for data types within a database. This can create some challenges with data migration, especially if you have data that goes beyond a limit. Take a variable character (VARCHAR) column as an example. Different databases have different max sizes: In ANSI SQL, there is no limit defined, but in MySQL there is a 65,535-character limit, and in SQL Server, there is a 2 GB limit. In Oracle, there is a 32,767-character limit, and in PostgreSQL, there is a 1 GB limit.

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.

Another example is the character large object (CLOB) data type. This data type might hold XML data, file data, or a temp table. In MySQL, there is a 216-byte limit; in SQL Server, there is a 231 – 1 byte limit; in PostgreSQL, there is a 1 GB limit; and in Oracle, there is a 4 GB – 1 byte limit. As such, if we need to transfer data between databases, we must recognize that this could be an issue and carefully plan how data migration would work. It is usually best to store CLOBs separately from the database due to their unstructured nature and their size compared to conventional database entries. A reference to the object's storage location is then created by a link.

KEY CONCEPT

The physical space available on a system can also be a concern. A number of issues can arise when the database itself runs out of disk space.

Symptoms of being low on storage space include performance issues, a sharp increase in failed transactions, application errors, and otherwise unpredictable behavior and data corruption. Ideally, there would be a monitor on the system to alert when the data storage capacity is low.

4. Table Size Limits

Tables also have size limitations. Different databases have different criteria for these limits. Some of them are based on the number of rows, while others are based on specific storage size. Some databases, like Teradata, can be unlimited, while others, like MySQL, can have a max size of 256 TB. In SQL Server, a table can be as large as the database itself, with 524,272 TB. Others, like SQLite, are based on the file size.

SUMMARY

In this lesson, you learned in the **introduction** that in order for a database to store and manage data effectively, it is critical to have sufficient **overall size** of the storage space. Having sufficient storage space prevents the database from losing data and interrupting service due to growing data volume. You learned about **data type size limits** and that the amount of space needed to handle transactional data, log files, indexes, and other database components is necessary to prevent data corruption, transaction failures, and slow performance. You also learned about **table size limits**. Planning and monitoring storage space correctly helps ensure data integrity, efficient query processing, and overall database reliability and availability.

Source: THIS TUTORIAL WAS AUTHORED BY DR. VINCENT TRAN, PHD (2020) AND Faithe Wempen (2024) FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.

© 2024 SOPHIA Learning, LLC. SOPHIA is a registered trademark of SOPHIA Learning, LLC.