Choosing the Right database for your application

Databases are at the heart of any application whether a mobile app or a standard application or a big enterprise application, so choosing the right one will build the basis for the success of the application. At its most basic, databases collect and store user information along with other data as per the application. There are various factors to be considered while choosing the right database for your application. Some of these factors are defined here.

Factors to be considered while choosing the right database

1. **Performance**

"That database doesn't perform well" is among the one of the most common statements we heard now a days. You must understand your ratios of reads to writes and also to check that the database engine is designed to maximize one or the other (or both). Also, it to consider that that fast database performance may be a result of a tradeoff somewhere else. For example, a database may have faster transaction velocity, but may result in weaker consistency.

2. **Read/ write ratio**

Based on database to database, some databases are faster at writes and some faster at reads and queries. The number of expected reads and writes from the application is a factor to consider while selecting the correct database.

3. **Structure of data**

The structure of the data is very important as it decides how it is stored and retrieved. As our applications deal with data present in a variety of formats, selecting the right database should include picking the right data structures for storing and retrieving the data. In correct database selection with respect to the our data, our application will take more time to retrieve

data from the database, and will also require more development efforts to work around any data issues.

4. **Volume of data to be stored**

To determine the size of the database, you must estimate how much data you will store. The quantity of data need to store and retrieve as critical application data is very important factor while choose the right database. The amount of data may vary depending on a combination of the data structure selected, the ability of the database to differentiate data across multiple file systems and servers, and even vendor-specific optimizations. So while choosing the database the factors like the overall volume of data generated by the application at a particular instance as well as the size of data to be retrieved from the database shall be taken into the consideration. If the data volume is just a less than gigabytes, then any database like in-memory databases is sufficient. But if it is in the terabytes range, or the data volume grows and touches petabytes, then the options of databases shrink is important. The larger the database means more complicated maintenance and management, and may need DBA that involves additional expenditure.

5. **Availability, latency, throughput, data consistency**

Your database engine must have the ability to be highly available. Availability is one of the most important factors for a transactional database as applications must be available 99.999% times round the clock. We may install in-house database for high availability because databases on cloud or at third party may run in multiple availability zones like offering only in weekdays or 9 to 5 availability. We may also have replication or can set up a pair of active-active servers when we cannot tolerate downtime as well as unscheduled maintenance periods. All this comes at a high price.

End-to-end response time of the application and the response time of the database is called as Latency. Ideally, the response time under 100 milliseconds is ok for a simple transaction but analytic queries can take seconds or minutes. Latency is another co-factor considered along with availability.

Number of transactions per second implies the throughput of database, and databases with high performance can support many users simultaneously.

All SQL databases have strong data consistency because all reads return the latest data. The data consistency for SQL databases can range between eventual and strong.

6. Accessibility of data or Number of simultaneous users

We must consider the load for multiple simultaneous users i.e. the number of people or users' concurrently accessing the database is another very important factor. Not only this the level of computation involved in accessing any specific data are also important factors to consider while choosing the right database. Ff the database chosen is not good enough to handle large loads it may adversely affect the processing speed of the application. It is easier to estimate the number of simultaneous users of a database used by the employees of an organization than that of a public database. For public databases, it might require scaling up to several servers for seasonal or unexpected loads.

7. Speed and scalability

Speed addresses the time taken to service all incoming reads and writes to our application. It is important to consider the speed we require for reading the data from the database and writing the data to the database. Some databases are actually designed to optimize read-heavy applications, while others are designed in a way to support write-heavy applications. A database that may handle input/output needs can actually go a long way to making a scalable architecture also.

To be continued….

Assignments

1. Explain how the number of simultaneous users affects the database selection.
2. Why Volume of data to be stored is an important factor? Explain