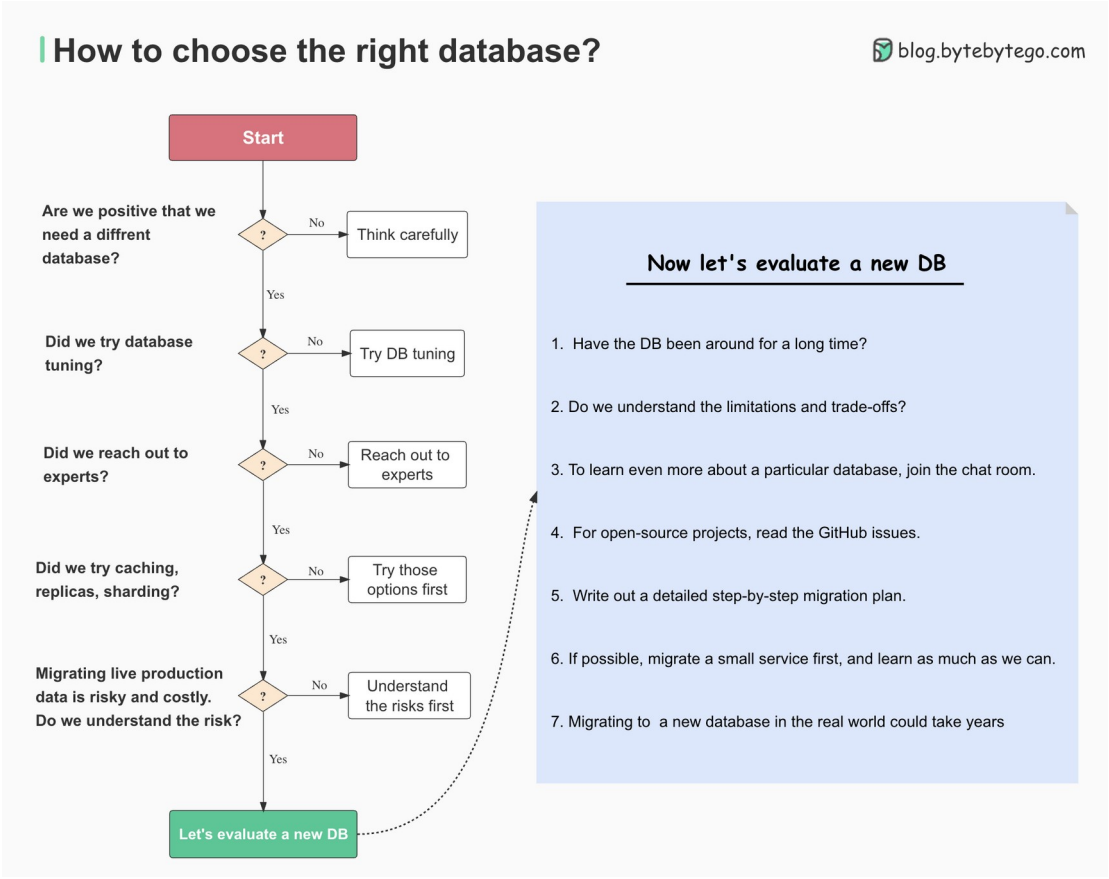
How to choose the right database

Choosing the right database is often the most important decision we'll ever make.

We are talking about a database for a real growing business, where a bad choice would lead to extended downtime, customer impact, and even data loss.

This take is probably a bit controversial.



First, are we positive that we need a different database?

Is the existing database breaking at the seams? Maybe the p95 latency is through the roof. Maybe the working set is overflowing the available memory, and even the most basic requests need to go to the disk.

Whatever the issues are, make sure they are not easily solvable.

Let’s read the database manual of our current database system. There could be a configuration knob or two that we can tweak to give us a bit more breathing room.

Can we put a cache in front of it, and give us a few more months of runway?

Can we add read replicas to shed some read load?

Can we shard the database, or partition the data in some way?

The bottom line is this: Migrating live production data is risky and costly. We better be damn sure that there is no way to keep using the current database.

We have exhausted all avenues for the current database.

How do we go about choosing the next one?

We developers are naturally drawn to the new and shiny, like moths to flame. When it comes to databases, though, boring is good.

We should prefer the ones that have been around for a long time, and have been battle tested.

Software engineering at scale is about tradeoffs. When it comes to databases, it is even more true.

Instead of reading the shiny brochures, go read the manual. There is usually a page called “Limits”. That page is a gem.

Learn as much as possible about the candidate now. The investment is relatively small at this juncture.

Once we narrow down the database options, what’s next?

Create a realistic test bench for the candidates using our data, with our real-world access patterns.

During benchmarking, pay attention to the outliers. Measure P99 of everything. The average is not meaningful.

After everything checks out, plan the migration carefully. Write out a detailed step-by-step migration plan.

Picking the right database is not glamorous, and there is a lot of hard work involved. Migrating to a new database in the real world could take years at a high scale.

Good luck.