**Storage**

**08/26/2022**

**Databases**

Databases are programs that either use disk or memory to do 2 core things: **record data** and **query data**. In general, they are themselves servers that are long lived and interact with the rest of your application through network calls, with protocols on top of TCP or even HTTP.

Some databases only keep records in memory, and the users of such databases know those records may be lost forever if the machine or process dies.

For the most part though, databases need persistence of those records, and this cannot use memory. This means that you have to write your data to disk. Anything written to disk will remain through power loss or network partitions, so that's what is used to keep permanent records.

Since machines die often in a large scale system, special disk partitions or volumes are used by the database processes, and those volumes can get recovered even if the machine were to go down permanently.

**Disk**

Usually refers to either **HDD (hard-disk drive)** or **SSD (solid-state drive)**. Data written to disk will persist through power failures and general machine crashes.

Disk is also referred to as non-volatile storage.

**SSD** is far faster than HDD (see latencies of accessing data from SSD and HDD) but also far more expensive from a financial point of view.

* Because of that, HDD will typically be used for data that's rarely accessed or updated, but that's stored for a long time.
* SSD will be used for data that's frequently accessed and updated.

**Memory**

Short for **Random Access Memory (RAM).**

Data stored in memory will be lost when the process that has written that data dies.

**Persistent Storage**

Usually refers to disk, but in general it is any form of storage that persists if the process in charge of managing it dies.