Perfect storage medium

Tom Rochette <tom.rochette@coreteks.org>

August 30, 2025 — 861fb9d0

0.1 Context

The perfect storage medium is an important aspect of building an agent capable of storing and retrieving information as fast as possible. Thus, it is critical to define what would be the optimal ways to fetch and put data into storage as well as the current state of the art in that matter.

0.2 Learned in this study

0.3 Things to explore

- How to store an immense volume of data
  - While taking less space
  - While being easy to work with
  - While being as fast as possible
  - While still potentially allowing operations such as lookup or search

1 Overview

2 References

File systems and databases already implement various methods which may interest us. Thus, it is appropriate to look into existing tools to find how specific problems have been solved.

- Databases
  - RDBMS
    - Unique identifier per schema
    - Foreign key relations
  - NoSQL
    - Graph based
    - Document based
      - UUID to uniquely identify documents and prevent collisions
- Filesystem
  - Hierarchical data structure
- Compression algorithms
  - Lossless
    - Compress repeating patterns using dictionaries
  - Lossy
    - Compress by removing data that has the least impact on the final result

3 Properties

- Prevent duplication of data (compression)

- Fast/Instant retrieval (lookup)
- Fast/Instant storage (hashing)

4 Data storage

Due to current hardware limitations, it is currently important to separate data storage into two categories:

- **Short term / volatile storage** which concerns itself with speed and efficiency of data retrieval.
- **Long term / persistent storage** which concerns itself with compression and conversion to short term / volatile storage.

4.1 Short term / Volatile storage (STS)

4.2 Long term / Persistent storage (LTS)

5 See also

6 References