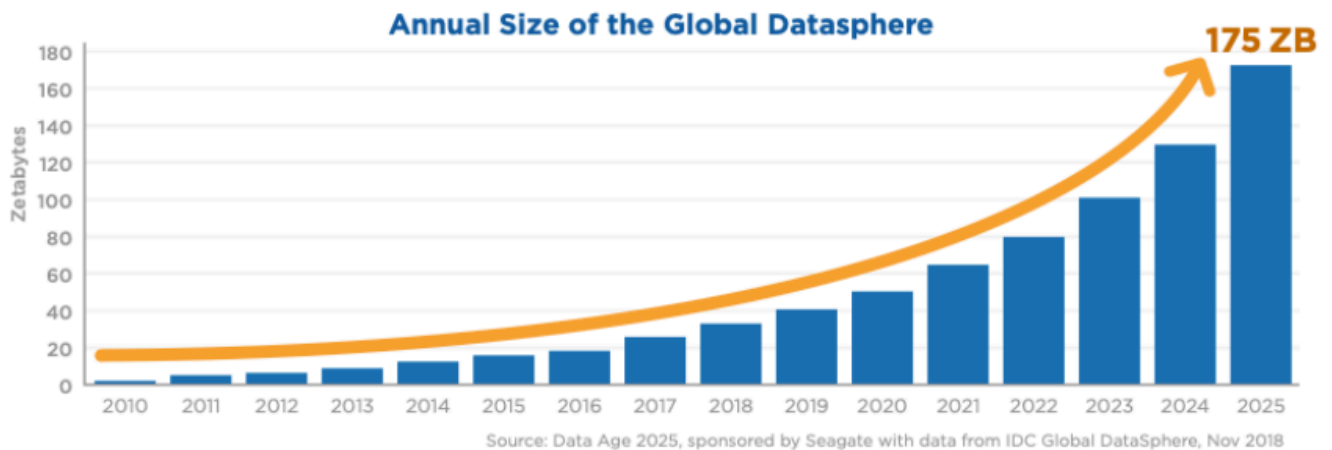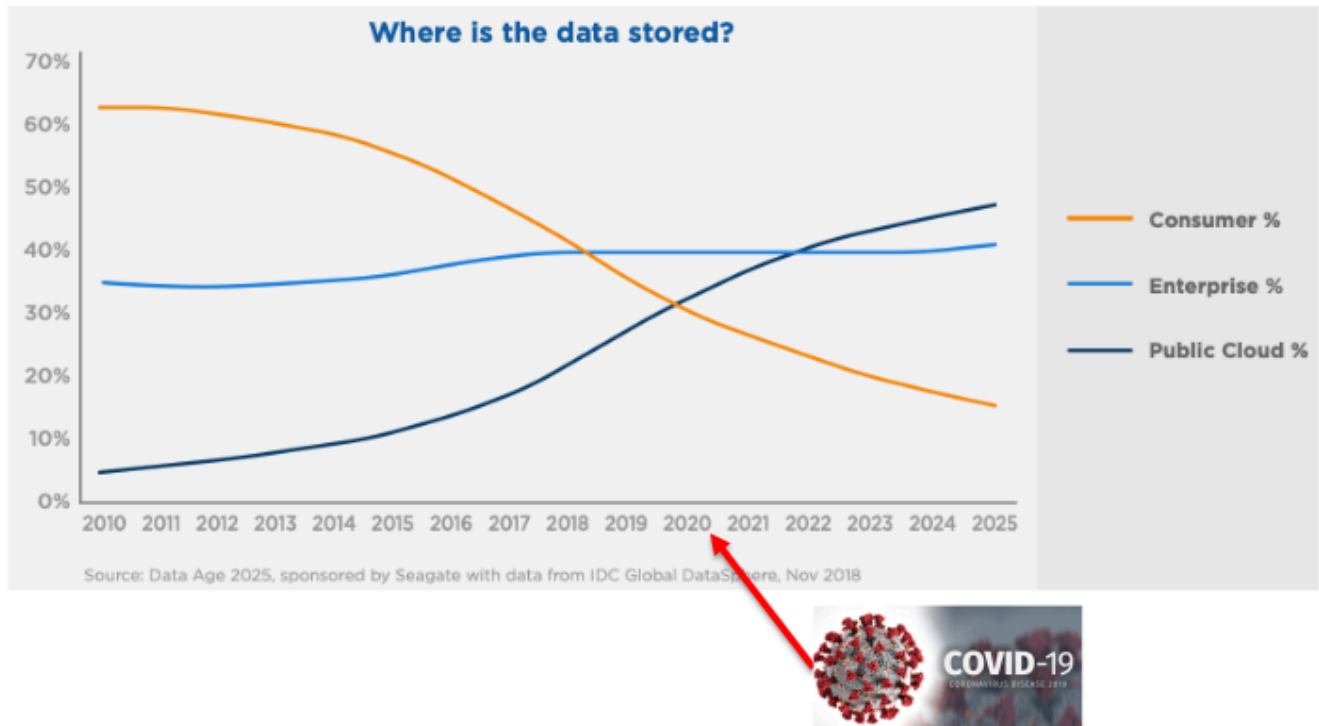# Storage

- Data-driven world
  - 80s-90s data was primarily generated by humans
  - Nowadays machines generate data at an unprecedented rate
    - Multiple media (image/video/audio/socialmedia) as big-data source
    - Sensors, surveillance cameras, digital medical imaging devices...
    - Industry4.0 and AI

Annual Size of the Global Datasphere

175 ZB

Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018

Nowadays internet connected devices are double than humans. In 1990 people generated Megabytes of data a day. The trend show a growth where there is a centralization of data, 80% of data stored is a replica(Need of backups and the increase in data production lead to this situation).
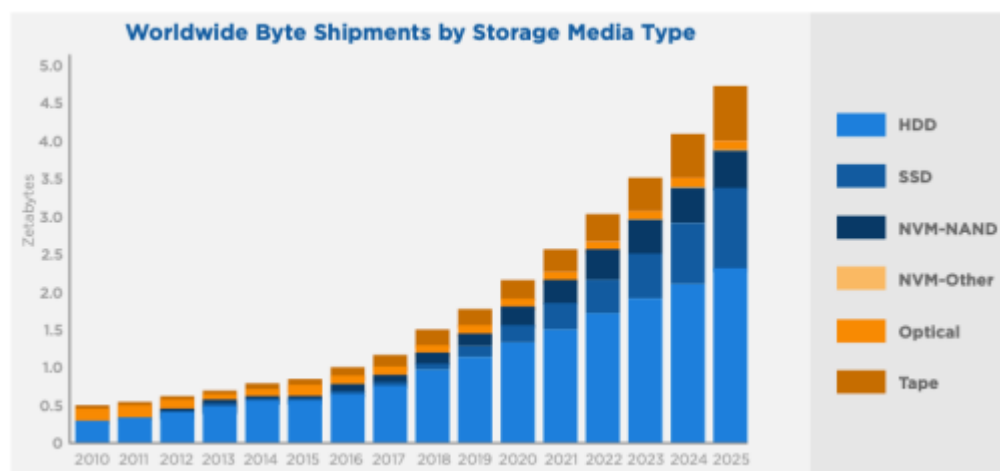
The growt favors the centralized storage strategy

- Limiting redundant data
- Automatizing replication & backup
- Reducing management costs

Data in the cloud as it is convenient to start using it.

Storage technologies is dominated by HDD. New technologies appeared as SSD, NVMe, Tapes(Amazon S3 Glacier, needs at least 12 hour to retrieve data, but it is more cheaper)

There are also hybrids as SSD are expensive so it was created a new system where a small SSD is combined with a large HDD upgrading the speed of 5 times over traditional HDD but SSD are used as a cache so you have to warm up this cache as it is treated as the cache of the operative system. It can also arrive to cost more than SSD in base at how much cache it uses.