1. A storage device is a device for recording (storing) information (data). It must be non-volatile in order to be a permanent home for data. Its speed is very slow compared with DRAM.

2. Dynamic random-access memory (DRAM) is a type of random-access memory that stores each bit of data in a separate capacitor within an integrated circuit. DRAM’s purpose is holding executable for programs and accessed pages.

3. Computer architecture: one or multiple cores CPUs with SRAM cache for computation; DRAM or NVRAM managed by memory controller to store data to be executed by CPUs; SSD or HDD managed by IO controller to store data permanently.

4. Storage system contains different types of storage devices: HDD, Flash-based SSD, NVRAM, SWD, OSD, Kinetic drive, Tape, Holographic storage device

5. In order to get some specific data from HDD, its arm has to seek to the right track and wait the track rotates to the right position. Normally, seek time is much longer than rotation time. To have a higher throughput, it’s better to put data into outer track. S.M.A.R.T. (Self-Monitoring, Analysis and Reporting Technology) is a monitoring system included in computer hard disk drives (HDDs) and solid-state drives (SSDs) that detects and reports on various indicators of drive reliability, with the intent of enabling the anticipation of hardware failures.

6. Flash memory is an electronic non-volatile computer storage medium that can be electrically erased and reprogrammed. There are two main types of flash memory, which are named after the NAND and NOR logic gates. NAND based flash is page addressable in size of 4KB to 16KB. Access time of a page read is around 25us, a page write is around 200us. Each page must be erased before written again. Erase operation performs on blocks, which consist 32 pages to 64 pages and take around 2.5ms. Flash memory has limited erase count. For SLC, it has around 100K erase count. For MLC, it has around 10K erase count. NOR based flash is page addressable.

7. Dynamic random-access memory (DRAM) is the most common technology used for the main memory. Despite DRAM’s advantages of high endurance and fast read/write access speed, DRAM suffers from data loss in the event of power failure or a system crash. To solve this problem, non-volatile DIMMs combining DRAM’s fast access speed and Flash’s persistence.
together provide a reliable main memory system. In addition, new types of non-volatile memory (NVM) such as phase change memory (PCM), Memristor and SST-RAM, have rapidly developed into possible candidates for the main memory in future computer systems. These emerging NVM technologies may offer other advantages in addition to their non-volatile nature. For examples, Memristor and PCM can achieve higher density, Memristor and SST-RAM can have faster read accesses and lower energy consumption.

8. A tape drive is a data storage device that reads and writes data on a magnetic tape. Magnetic tape data storage is typically used for offline, archival data storage. Tape media generally has a favorable unit cost and a long archival stability.

- Recoding order

- Zone concept

9. Storage servers contain multiple storage devices, e.g. HDD, SSD, SWD to hold data shared by multiple clients. These servers also have DRAM or NVRAM as a caching layer to accelerate IO speed. Each client has their own CPU and main memory or even local storage.

10. Tiered Storage contains different types of storage devices with different access speed:

- NVRAM > SSD > HDD >= SWD > Tape

Cloud storage

Must common tiered storage management scheme is in a hierarchical fashion. So data usually dump from a faster device to a slower device when they become unpopular. An alternative fashion is side-by-side, in which data will be classified and allocated to storage devices directly instead of demotion or promotion.
One important issue in tiered storage management is data allocation and data relocation. We need to find out which data are hot and which data are cold in order to fully utilize storage devices’ feature.

11. Software stack

12. A backup refers to the copying and archiving of computer data so it may be used to restore the original after a data loss event. There are two types of backups: incremental backup and full backup. Incremental backup happens more frequently (e.g. every day) to just store the changes from last time. Full backup happens rarely (e.g. every week) to make a complete copy of existing data.

13. A full backup of a large data set may take a long time to complete. To avoid downtime, high-availability systems may instead perform the backup on a snapshot—a read-only copy of the data set frozen at a point in time—and allow applications to continue writing to their data. A snapshot is the state of a system at a particular point in time.

14. Data archiving is the process of moving data that is no longer actively used to a separate data storage device for long-term retention.

15. Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data. This technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. In the deduplication process, unique chunks of data are identified and stored during a process of analysis. As the analysis continues, other chunks are compared to the stored copy and whenever a match occurs, the redundant chunk is replaced with a small reference that points to the stored chunk.