IIT Madras
BSc Degree

Backend Systems

Memory Hierarchy

Types of storage elements

- On-chip registers: 10s-100s of bytes

Types of storage elements

- On-chip registers: 10s-100s of bytes
- SRAM (cache): 0.1 - 1 MB

Types of storage elements

- On-chip registers: 10s-100s of bytes
- SRAM (cache): 0.1 - 1 MB
- DRAM: 0.1 - 10 GB

Types of storage elements

- On-chip registers: 10s-100s of bytes
- SRAM (cache): 0.1 - 1 MB
- DRAM: 0.1 - 10 GB
- Solid-state disk (SSD) - Flash: 1-100 GB

Types of storage elements

- On-chip registers: 10s-100s of bytes
- SRAM (cache): 0.1 - 1 MB
- DRAM: 0.1 - 10 GB
- Solid-state disk (SSD) - Flash: 1-100 GB
- Magnetic disk (HDD - hard disk drive?): 0.1 - 10 TB

Types of storage elements

- On-chip registers: 10s-100s of bytes
- SRAM (cache): 0.1 - 1 MB
- DRAM: 0.1 - 10 GB
- Solid-state disk (SSD) - Flash: 1-100 GB
- Magnetic disk (HDD - hard disk drive?): 0.1 - 10 TB
- Optical, magnetic, holographic, . . .

Storage Parameters

- Latency: time to read first value from a storage location (lower is better)
  - Register < SRAM < DRAM < SSD < HDD

Storage Parameters

- Latency: time to read first value from a storage location (lower is better)
  - Register < SRAM < DRAM < SSD < HDD
- Throughput: number of bytes/second that can be read (higher is better)
  - DRAM > SSD > HDD (regs, SRAM limited capacity)

Storage Parameters

- Latency: time to read first value from a storage location (lower is better)
  - Register < SRAM < DRAM < SSD < HDD
- Throughput: number of bytes/second that can be read (higher is better)
  - DRAM > SSD > HDD (regs, SRAM limited capacity)
- Density: number of bits stored per unit area / cost (higher is better)
  - Volume manufacture important
  - HDD > SSD > DRAM > SRAM > Regs

Computer Organization

- CPU has as many registers as possible

Computer Organization

- CPU has as many registers as possible
- Backed by L1, L2, L3 cache (SRAM)

Computer Organization

- CPU has as many registers as possible
- Backed by L1, L2, L3 cache (SRAM)
- Backed by several GB of DRAM working memory

Computer Organization

- CPU has as many registers as possible
- Backed by L1, L2, L3 cache (SRAM)
- Backed by several GB of DRAM working memory
- Backed by SSD for high throughput

Computer Organization

- CPU has as many registers as possible
- Backed by L1, L2, L3 cache (SRAM)
- Backed by several GB of DRAM working memory
- Backed by SSD for high throughput
- Backed by HDD for high capacity

Computer Organization

- CPU has as many registers as possible
- Backed by L1, L2, L3 cache (SRAM)
- Backed by several GB of DRAM working memory
- Backed by SSD for high throughput
- Backed by HDD for high capacity
- Backed by long-term storage, backup

Cold storage (?)

- Backups and archives:
  - Huge amounts of data
  - Not read very often
  - Can tolerate high read latency
- Amazon Glacier, Google, Azure Cold/Archive storage classes
- High latency of retrieval: up to 48 hours
- Very high durability
- Very low cost

Impact on application development

- Plan the storage needs based on application growth
- Speed of app determined by types of data stored, how stored
- Some data stores are more efficient for some types of read/write operations

Developer must be aware of choices and what kind of database to choose for a given application