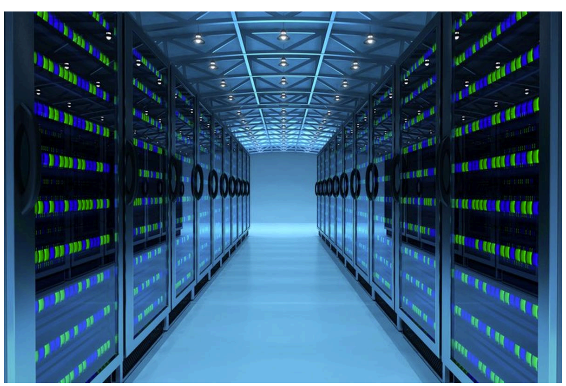
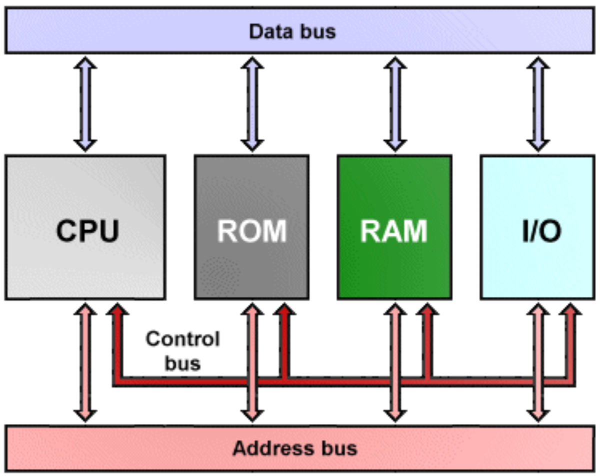
Computer Architecture: An Introduction

Computer Architecture



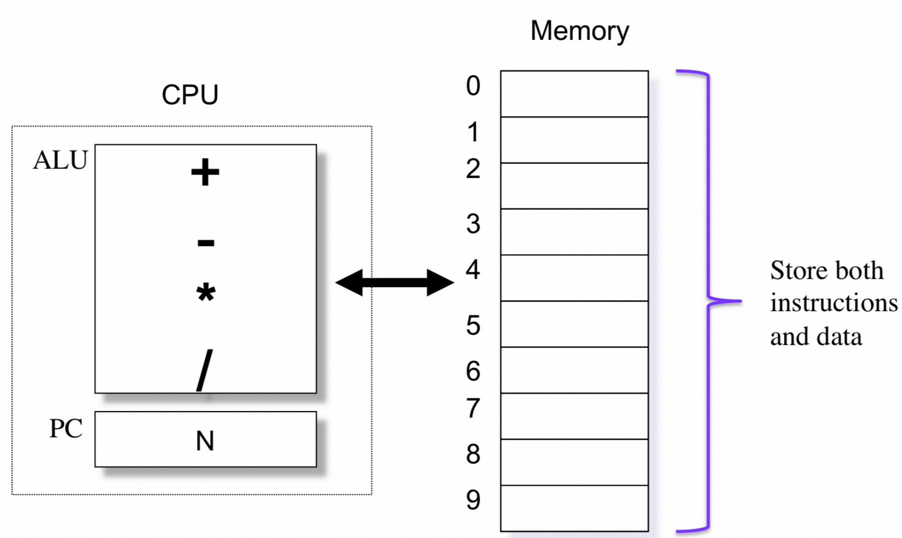
Main Components

- CPU
  - Central Processing Unit
  - Performs computation
- Memory
  - Stores data
- BUS
  - Used for carrying information
  - Communication and data transfer
- I/O Devices
  - Human Interface (mouse, keyboard, screen)
  - Networking
  - Graphics
  - Storage

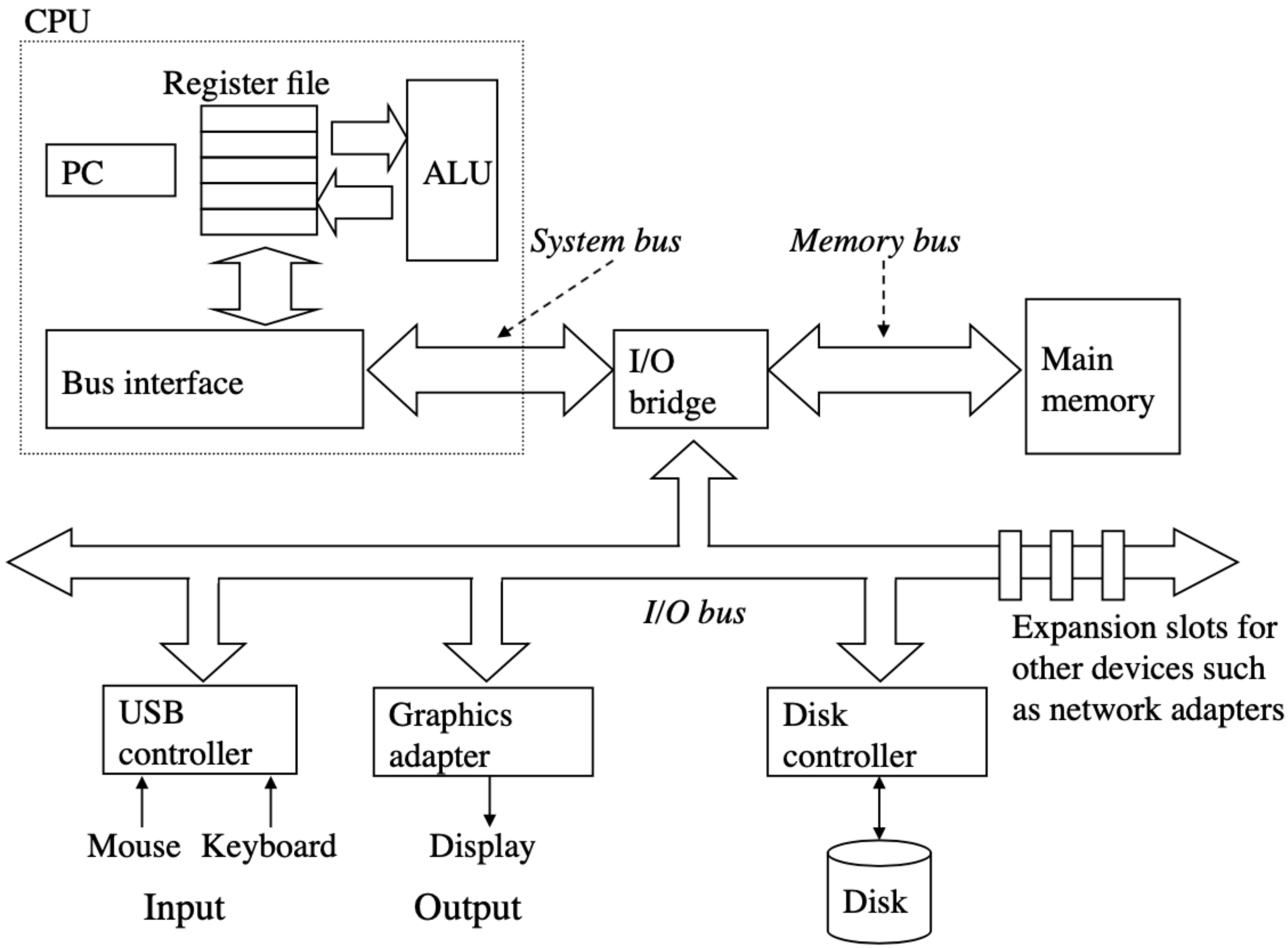


Von Neuman Model

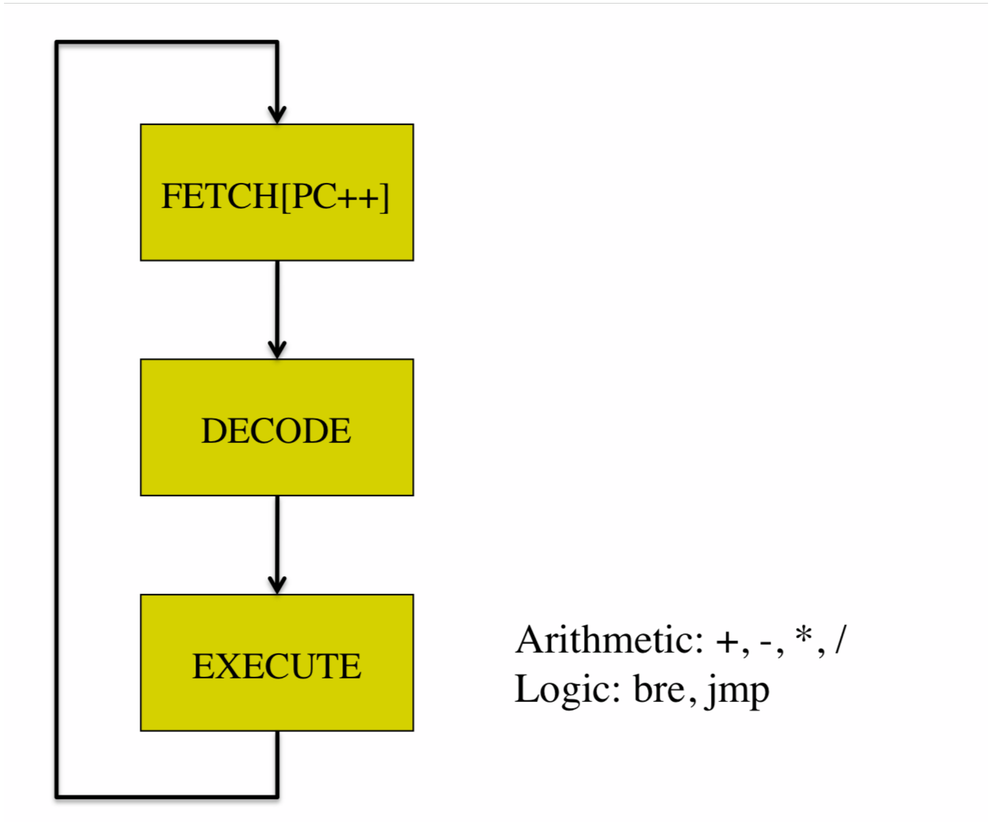
- Computer Architecture Model
- Central processing unit
  - Arithmetic Processing Unit (ALU)
  - Control Unit (instruction register and program Counter
- Memory Component that stores both data and instructions



Von Neumann Model: In Practice

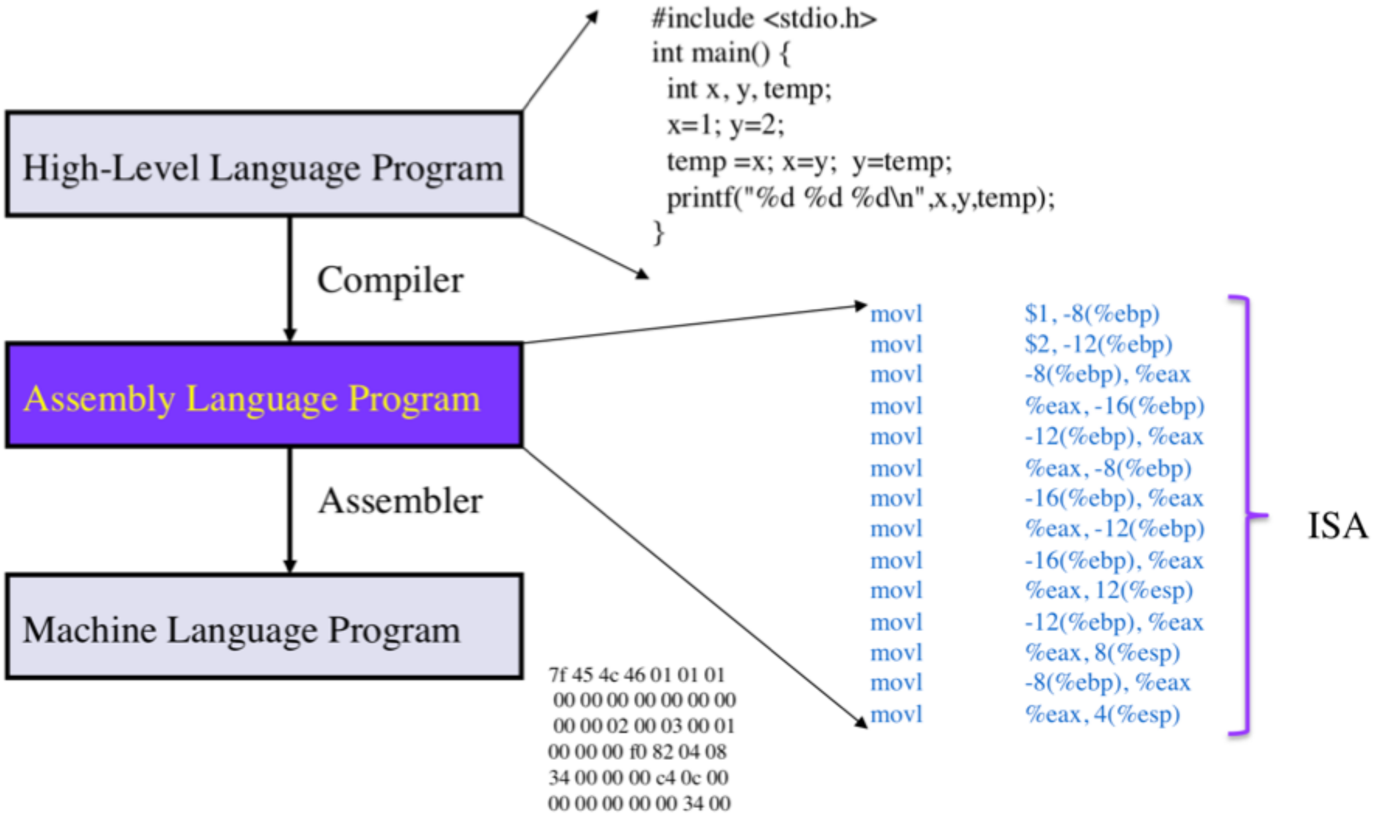


Basic CPU Function



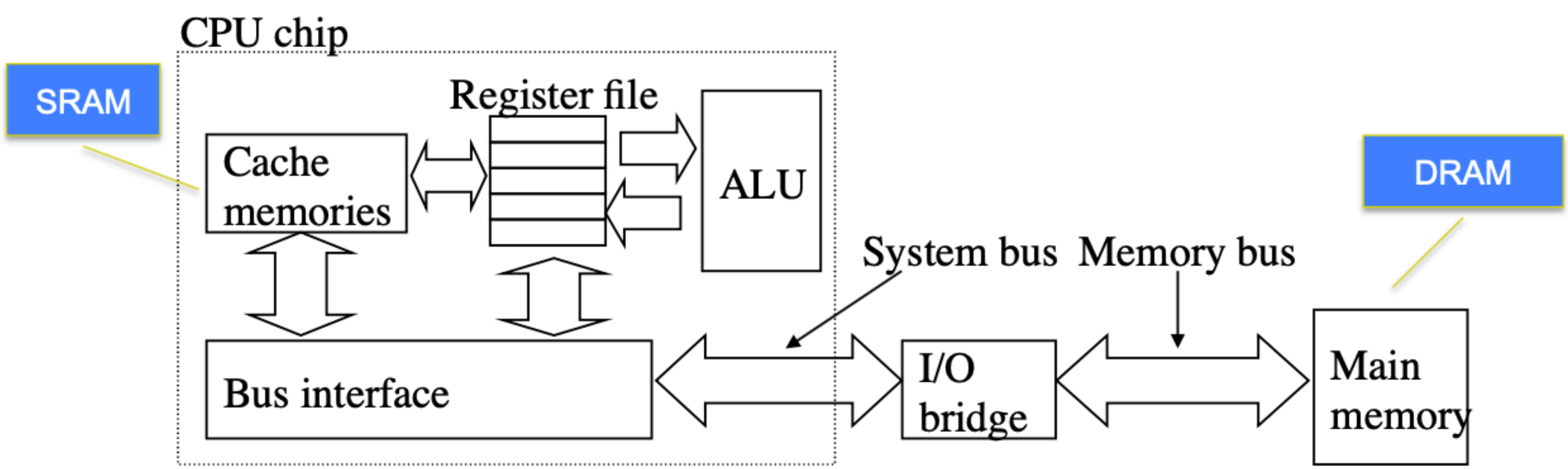
Program to Hardware

- How logic turns into something machines can understand



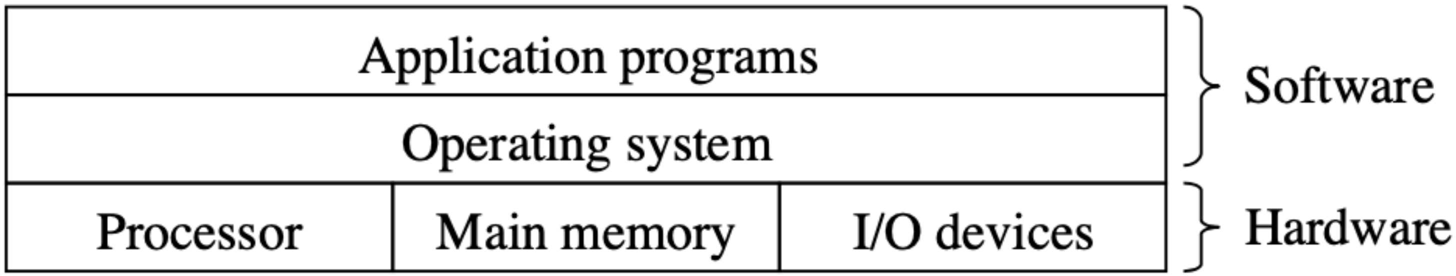
Memory Hierarchy

- A program that resides on disk must be loaded into memory for execution
- Why do we have levels of memory?
  - Larger devices are slower
  - Faster devices are more expensive
    - Typically less storage per dollar

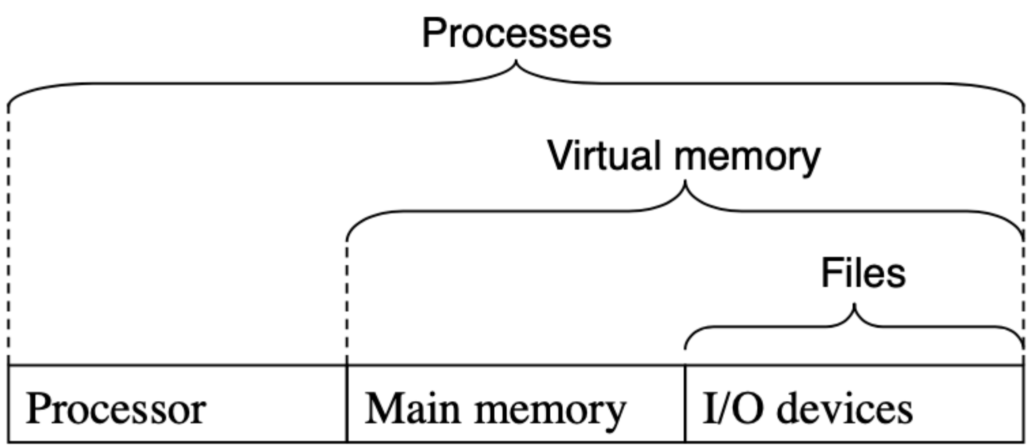


The Operating System

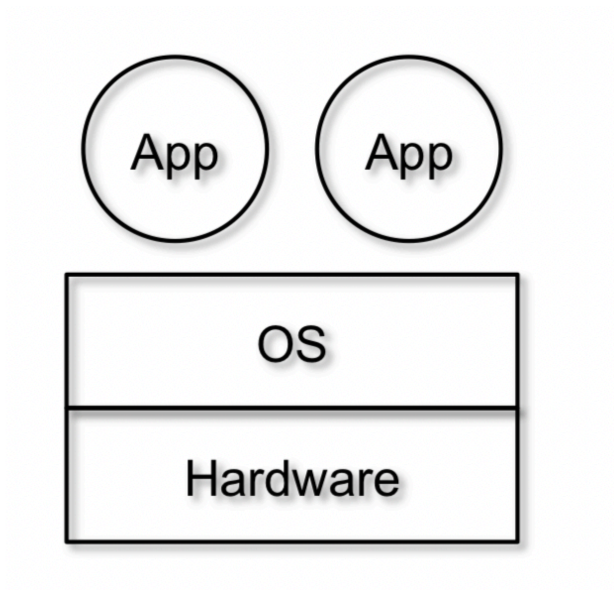
- An operating system manages the hardware

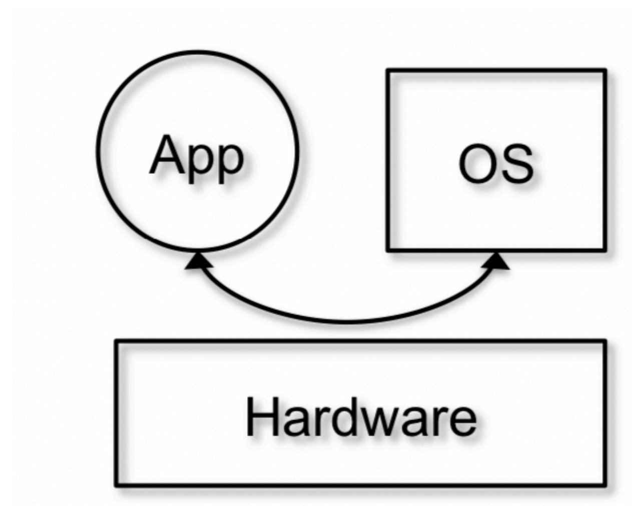


- Manages hardware through abstractions



Operating System Reality





- Main Idea: OS sits on top of hardware facilitate applications applications
- Reality: Apps typically run directly on hardware. Will switch over when OS when needed

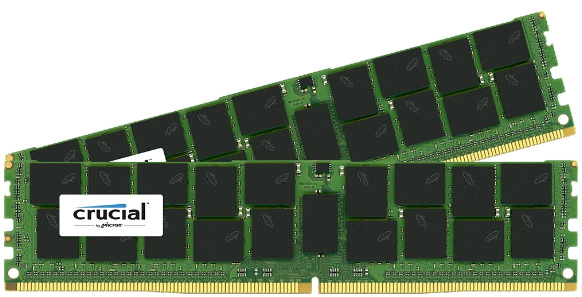
Computer Architecture

- How to design and build components
- In this class
  - Understand how programs run on current system
  - Understand how current architecture affect programming languages
- Additionally you'll be able to think about
  - How can I make programs faster
  - How can I make programs more power efficient

Architecture Trends

- How have Computer Architecture changed over time?

CPU



Memory

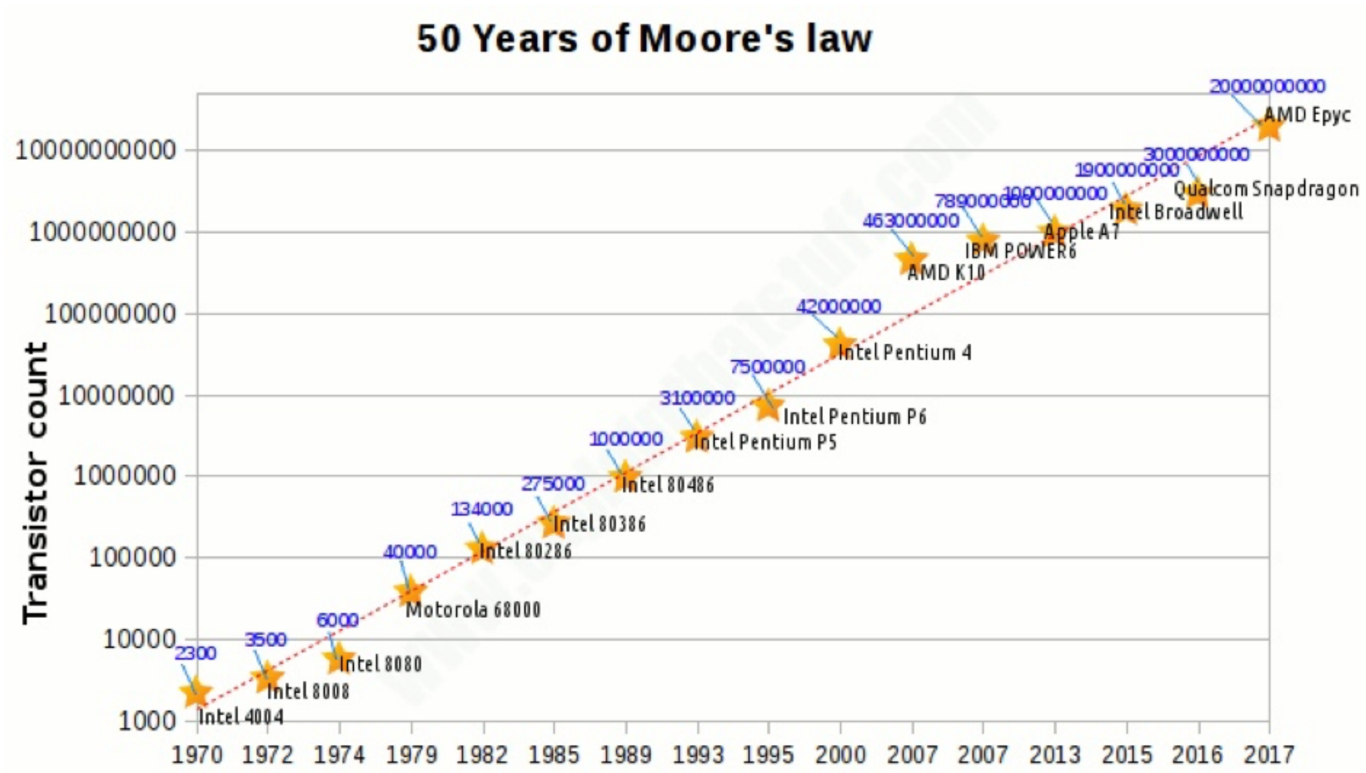
Storage

Moore’s Law

- Gordon Moore, an Intel Engineer
- Observed architectural improvement trend
    - # of transistors on a chip doubles every 18 months
- General exponential growth in hardware:
    - Process speed x2 every 18 months
    - Memory capacity x2 every 2 years
    - Disk capacity x2 every year

Transistor Count Trend

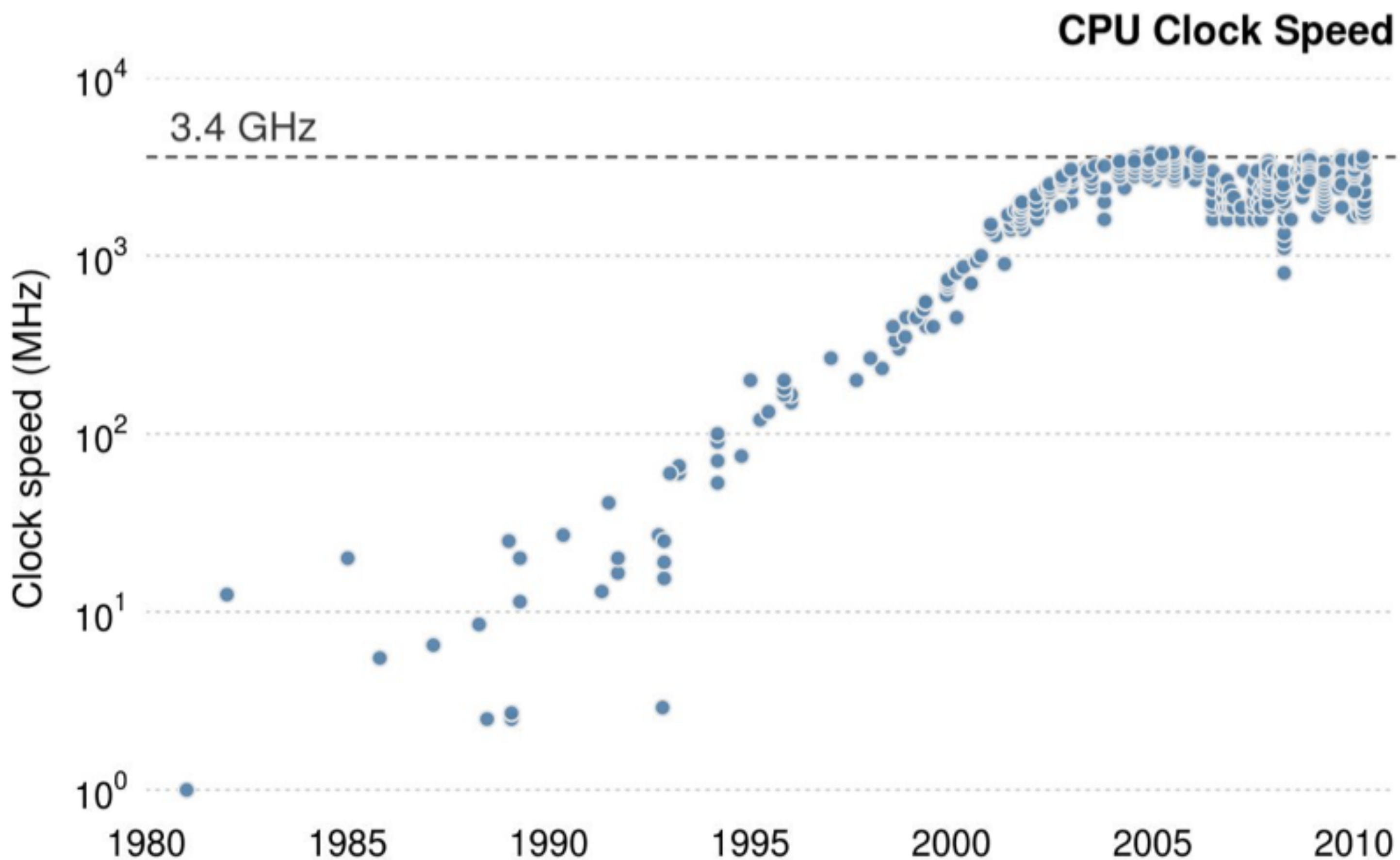
- Transistors are needed to implement logical gates within a CPU



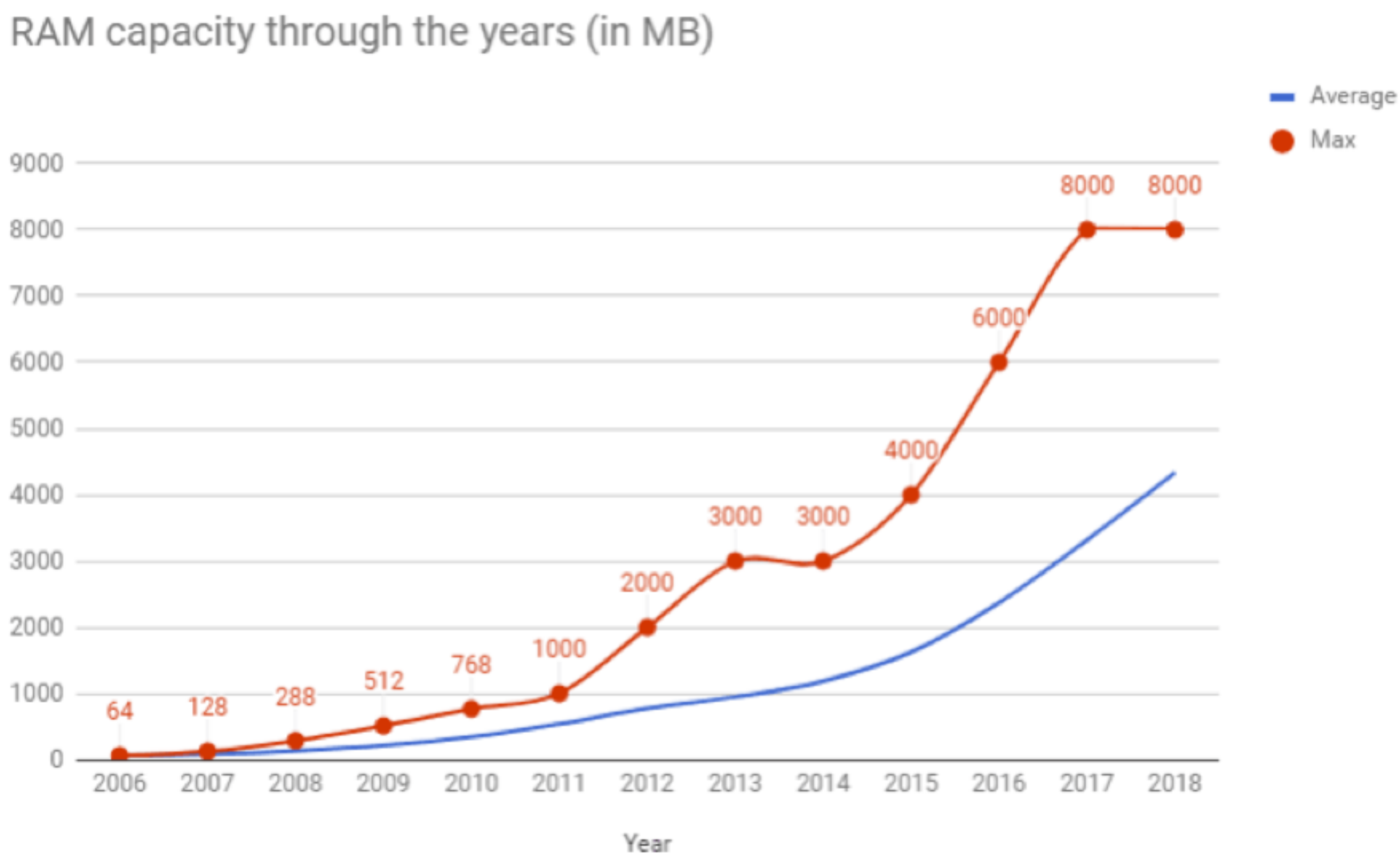
- 32-core AMD Epyc has 19,200,000,000 transistors

Trend: Clock Speed

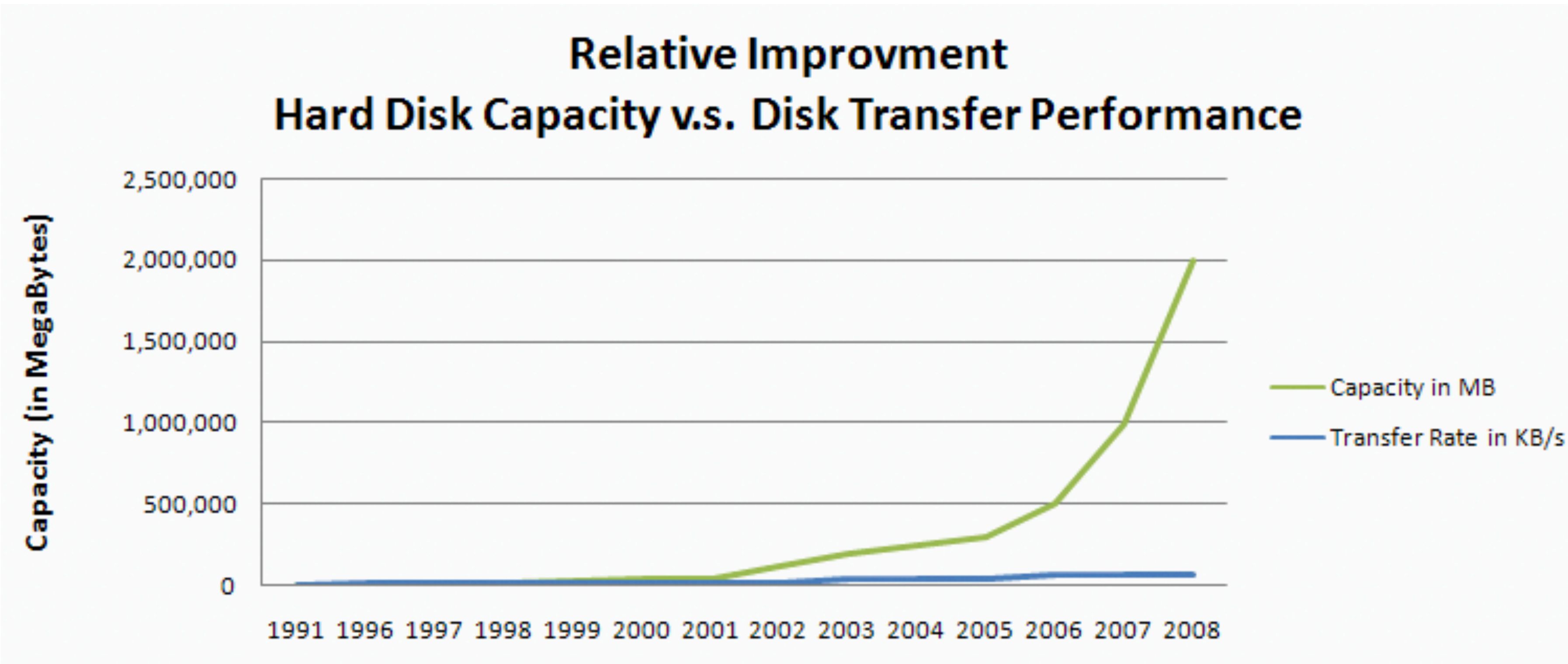
- Clock speed is the rate at which a processer completes a processing cycle



Trend: Memory Capacity

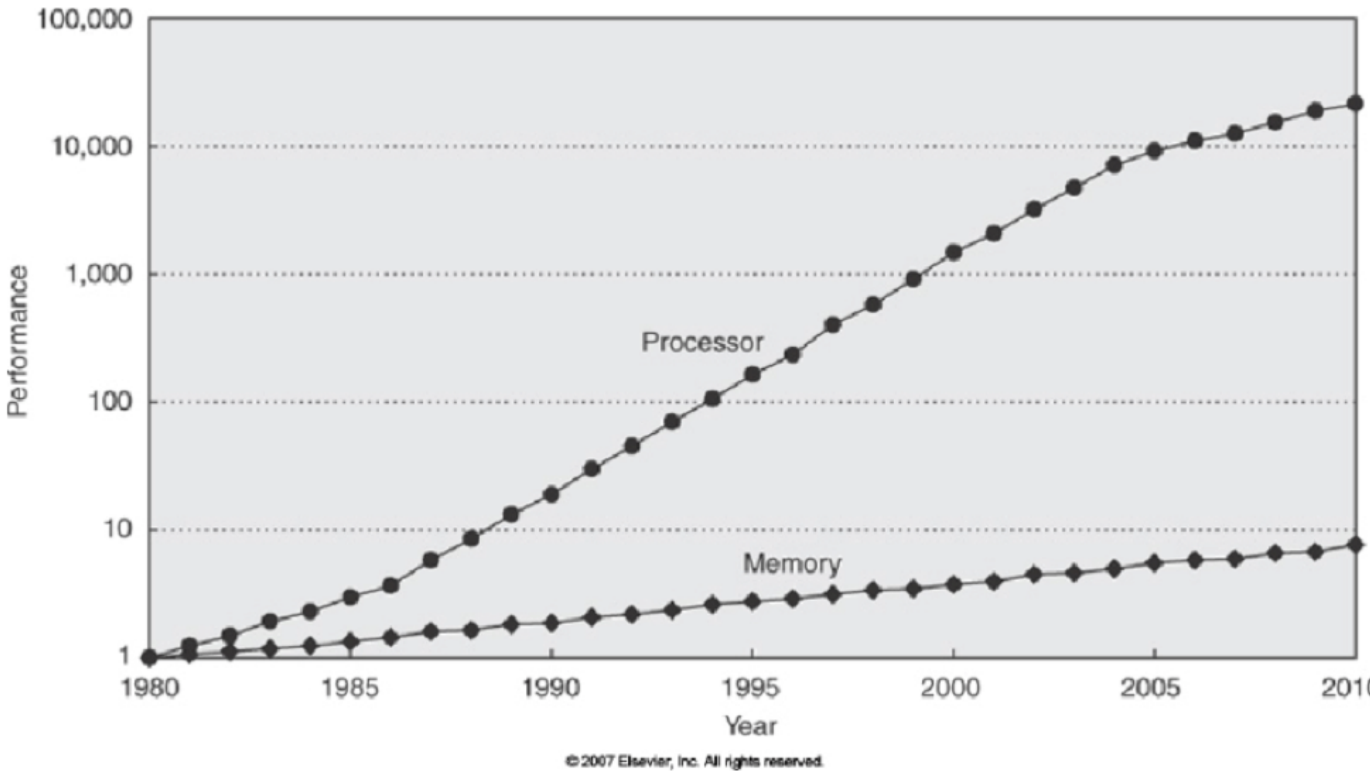


Trend: Disk Capacity



Why are these trends important when it comes to designing system components?

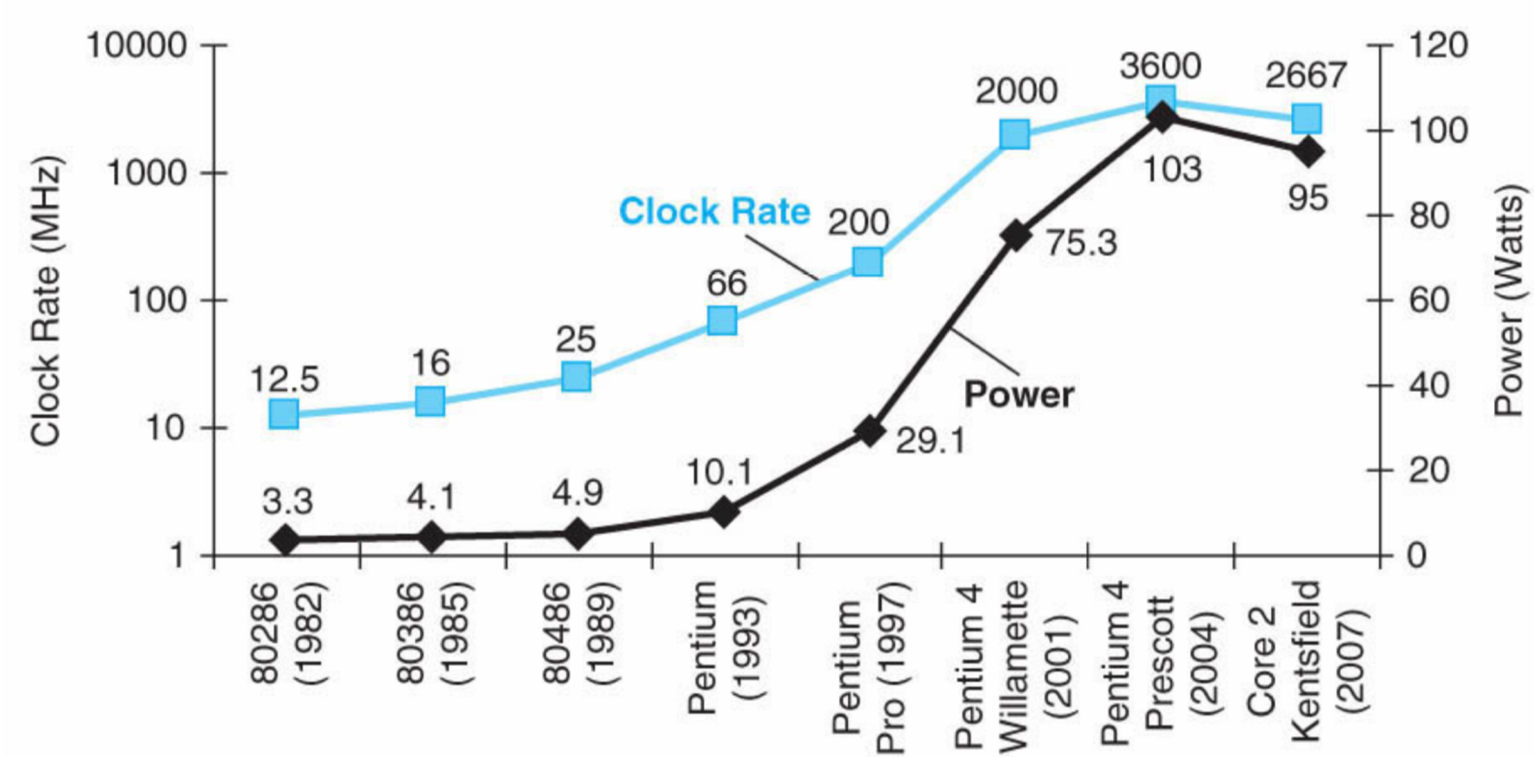
CPU vs Memory Performance Gap



- CPU performance grows faster than Memory
- Implication:
  - Need to make efficient use of the memory hierarchy
  - Registers/Caches

Power Wall

- More transistors -> higher power densities -> higher temperature
- Scaling performance limited by fundamental constraints
  - Power delivery
  - Cooling
- Goal is to maximize performance per watt



In Summary

- Todays systems are complex
- Faster processor and systems enable new domains of applications
- Understanding the system is crucial to
  - Develop efficient applications
  - Ensure good tradeoffs between performance/power etc.
  - Address various constraints (memory, power, etc.)