Advanced Computer Architecture
(0630561)

Lecture 1

Introduction

Prof. Kasim M. Al-Aubidy
Computer Eng. Dept.

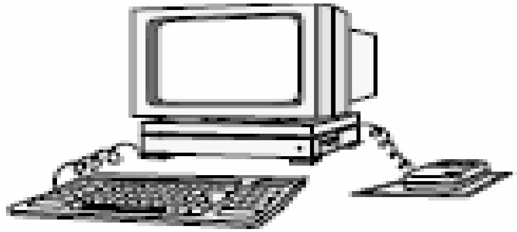
Introduction:

- Two important influences on the computer architecture are;
  - the electronic speed of the ICs.
  - the parallel activity that can be accomplished.
- Over the years, the technology has progressed s much that today we have machines operating at high speed.
- Examples:
    - CRAY 1 System: use a clock frequency of 80 MHz.
    - CRAY 2 System: use a clock frequency of 250 MHz.
    - CRAY 3 System: use a clock frequency of 1 GHz.
- CRAY in 1989 announced Gallenium Arsenide technology for their next model which employ parallel processing based on multiprocessors. The targeted peak speeds are in excess of 3 Megaflops.
- Do we really need very fast computer systems?
- Problems requiring large scale computing (memory & computing speed) are large scale problems in optimization, planning, scheduling, network flows, field problems……

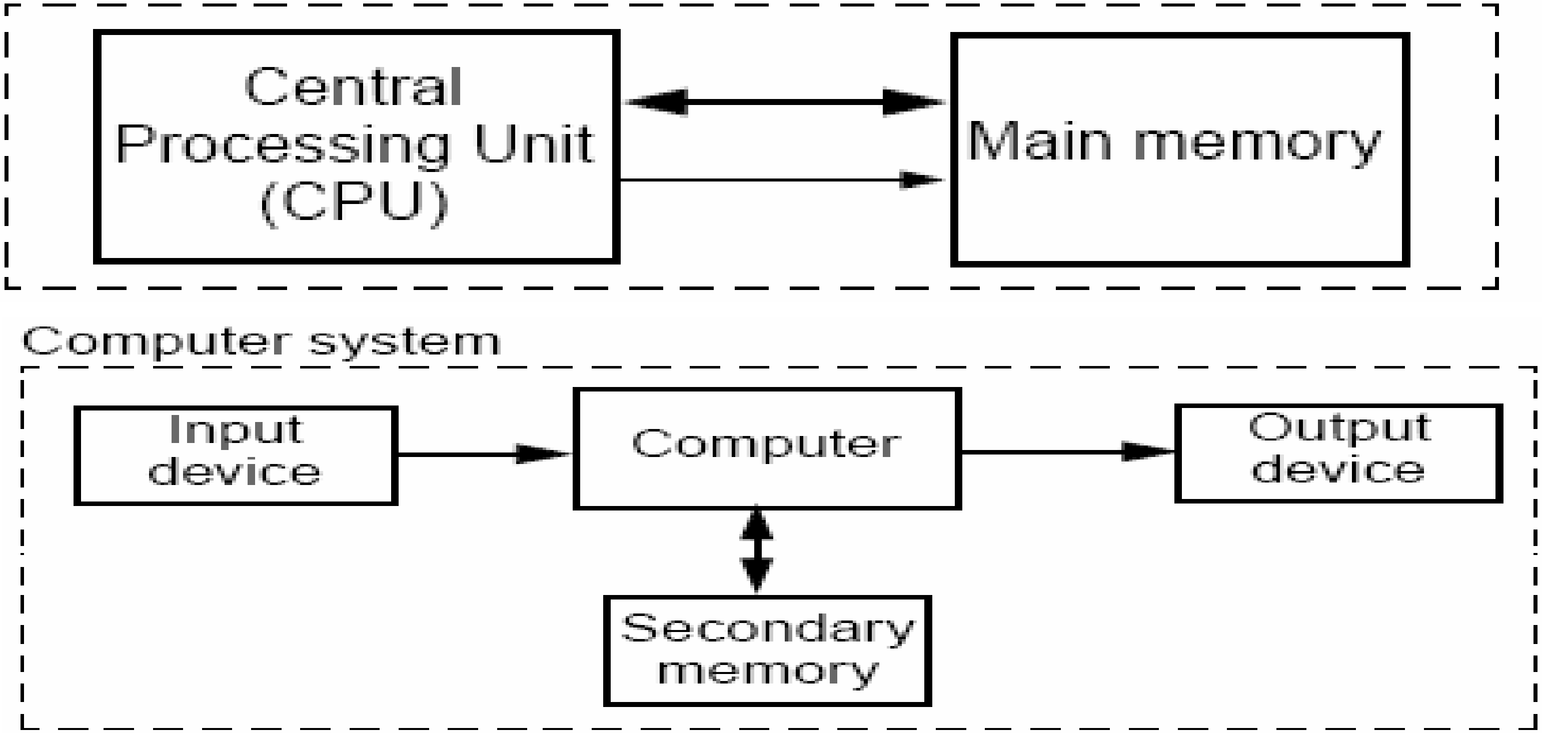
Applications that require the high speed computers are;

- Design of VLSI circuits.
- CAD/CAM applications.
- Solving field problems.
- Weather forecasting.
- Intelligent systems.
- Modeling and simulation in economics, planning, and other areas.
- Remote sensing requiring processing of large data gathered from satellites.
- Problems in nuclear energy.

What is a computer?



A **computer** is a data processing machine which is operated automatically under the control of a list of instructions (called a program) stored in its main memory.



Technology

- Technology advances at astounding rate
  - 19th century: attempts to build mechanical computers
  - Early 20th century: mechanical counting systems (cash registers, etc.)
  - Mid 20th century: vacuum tubes as switches
  - Since: transistors, integrated circuits
- 1965: Moore's law [Gordon Moore]
  - Predicted doubling of IC capacity every 18 months
  - Has held and will continue to hold
- Drives functionality, performance, cost
  - Exponential improvement for 40+ years

Semiconductor History

Date	Event	Comments
1947	1st transistor	Bell Labs
1958	1st IC	Jack Kilby (MSEE '50) @TI Winner of 2000 Nobel prize
1971	1st microprocessor	Intel (calculator market)
1974	Intel 4004	2300 transistors
1978	Intel 8086	29K transistors
1989	Intel 80486	1M transistors
1995	Intel Pentium Pro	5.5M transistors
2006	Intel Montecito	1.7B transistors
201x	IBM	50B transistors

Technology Push

- What do these two intervals have in common?
  - 1776-1999 (224 years)
  - 2000-2001 (2 years)
- Answer: Equal progress in processor speed!
- The power of exponential growth!
- Driven by Moore's Law
  - Devices per chip doubles every 18-24 months
- Computer architects turn additional resources into
  - Speed
  - Power savings
  - Functionality

- Technology advances at varying rates
  - E.g. DRAM capacity increases at 60%/year
  - But DRAM speed only improves 10%/year
  - Creates gap with processor frequency!

Technology => dramatic change

- ❑ **Processor**
  - ➔ logic capacity: about 30% increase per year
  - ➔ clock rate: about 20% increase per year

Higher logic density gave room for instruction pipeline & cache

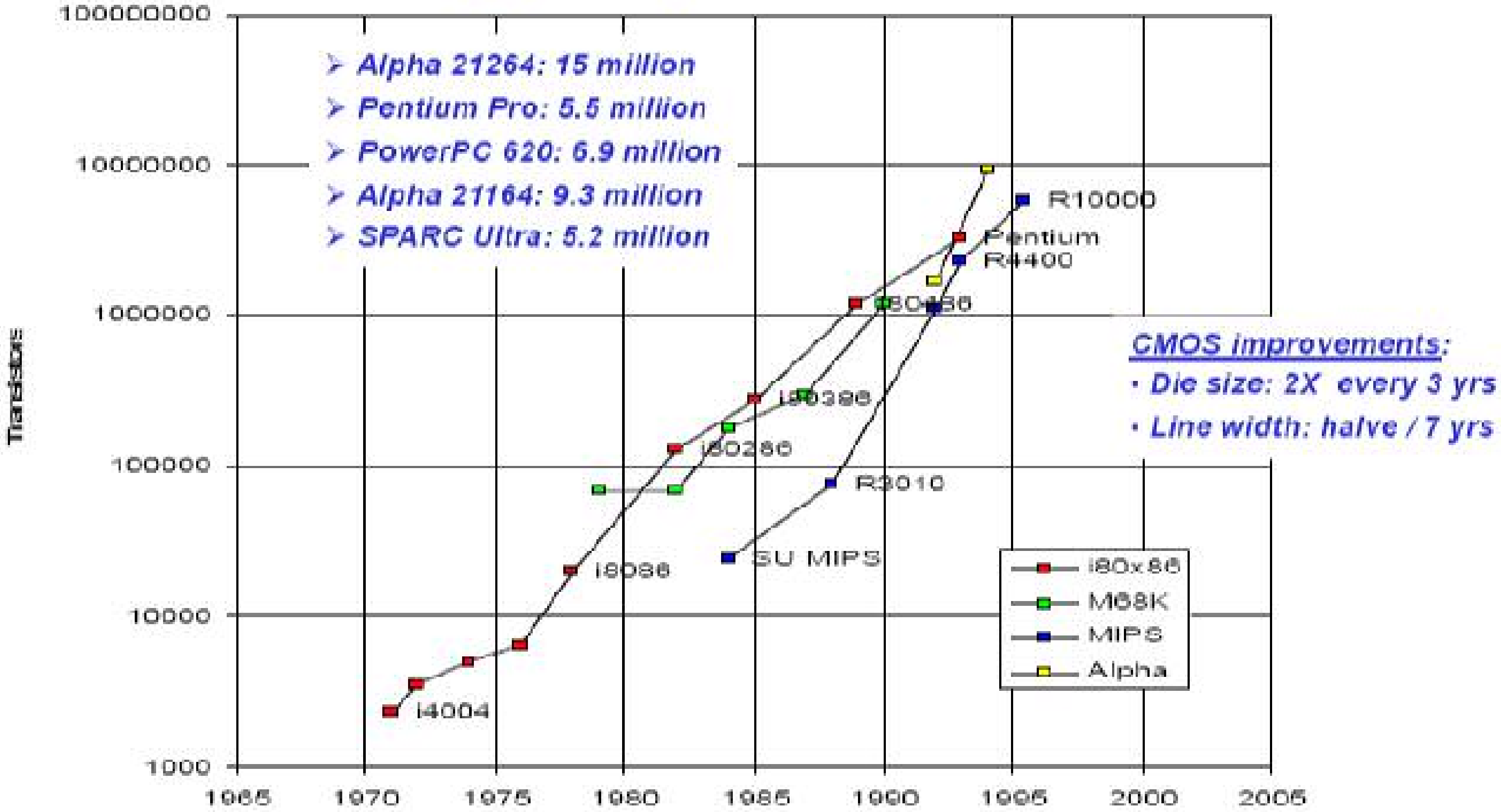
- ❑ **Memory**
  - ➔ DRAM capacity: about 40% increase per year (double every 2.5 years)
  - ➔ Memory speed: about 10% increase per year
  - ➔ Cost per bit: about 25% improvement per year

Performance optimization no longer implies smaller programs

- ❑ **Disk**
  - ➔ Capacity: about 60% increase per year

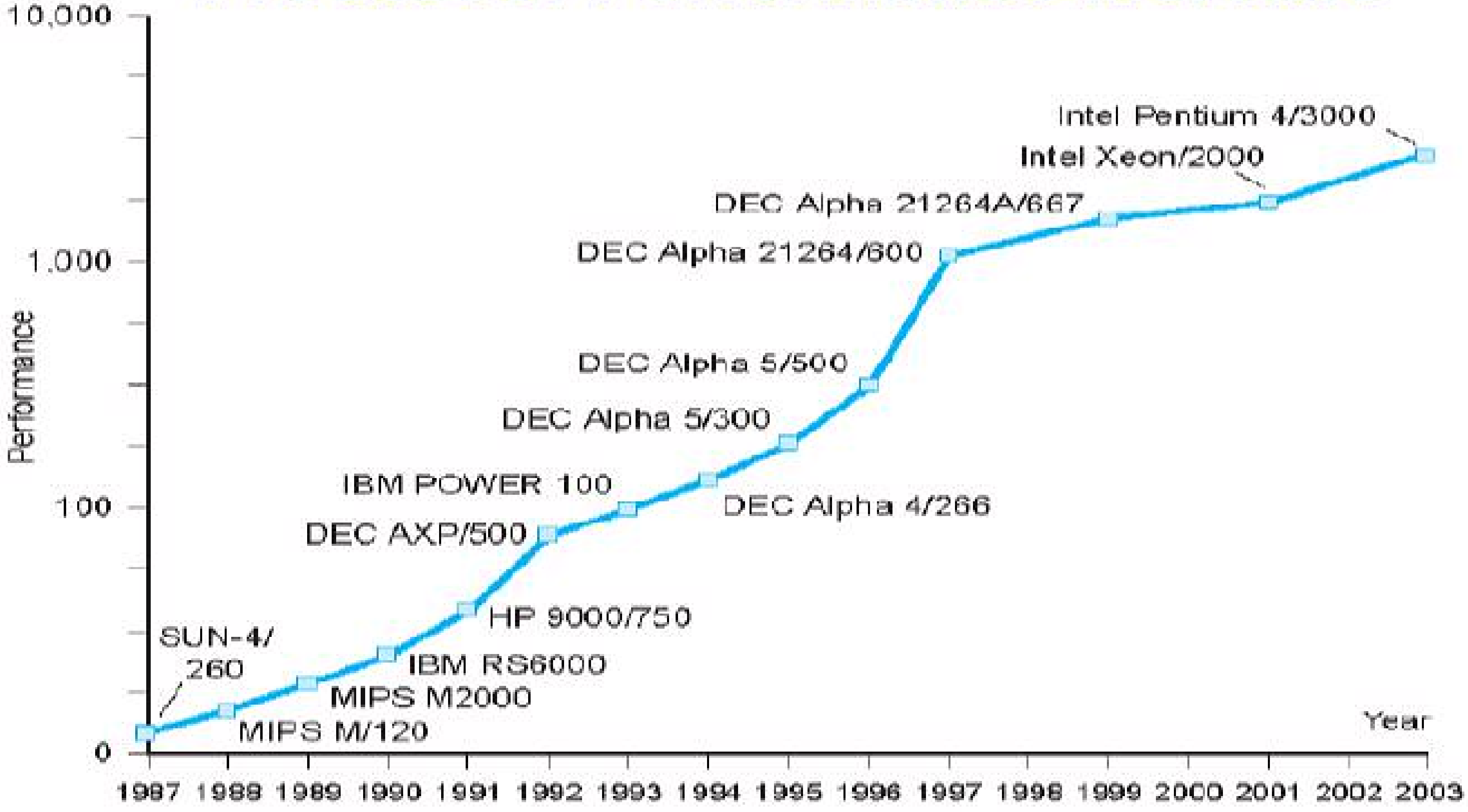
Computers became lighter and more power efficient

Technology Impact on Processors



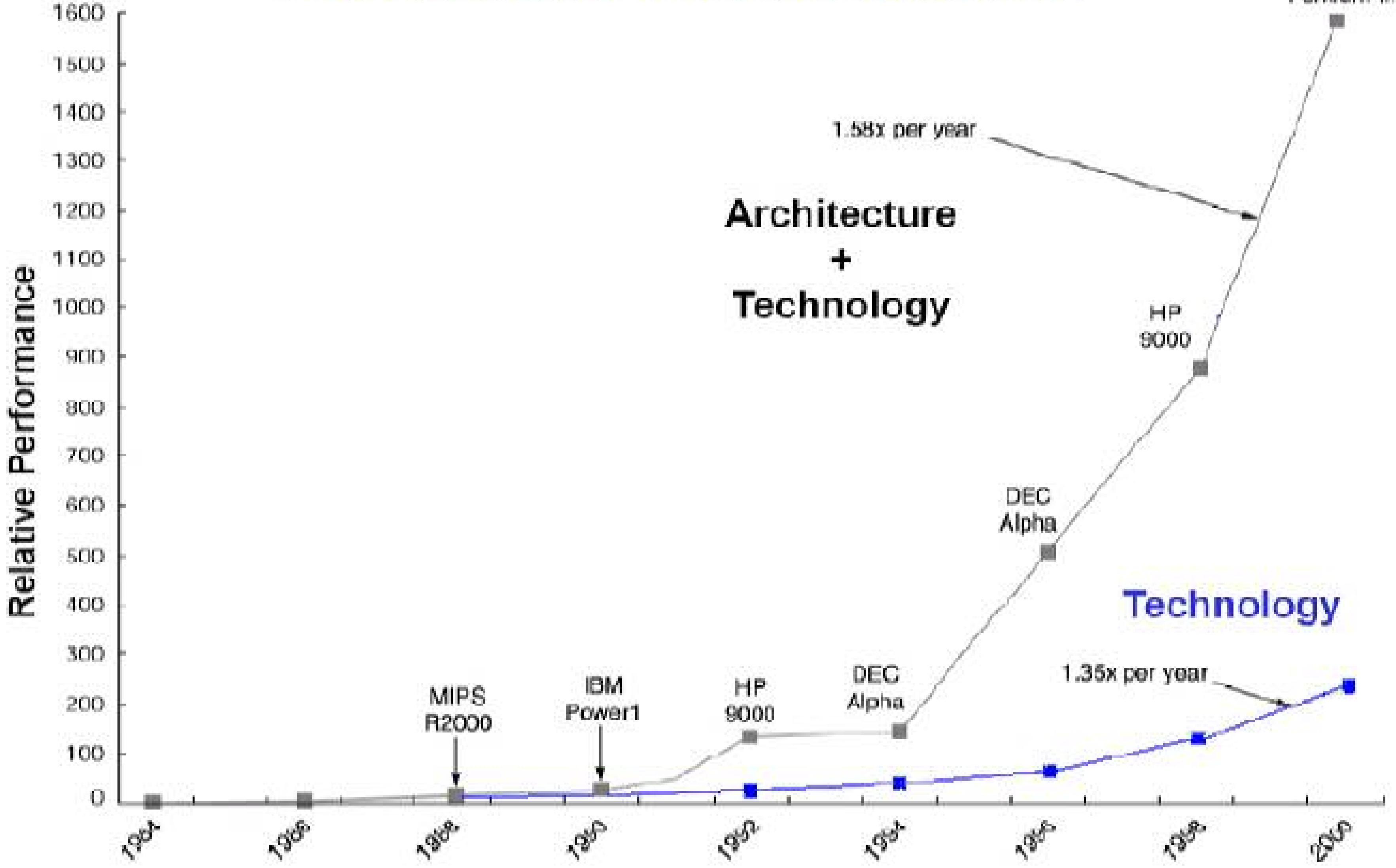
- In ~1985 the single-chip processor and the single-board computer emerged
- In the 2009+ timeframe, today's mainframes may be a single-chip computer

Processor Performance Increase

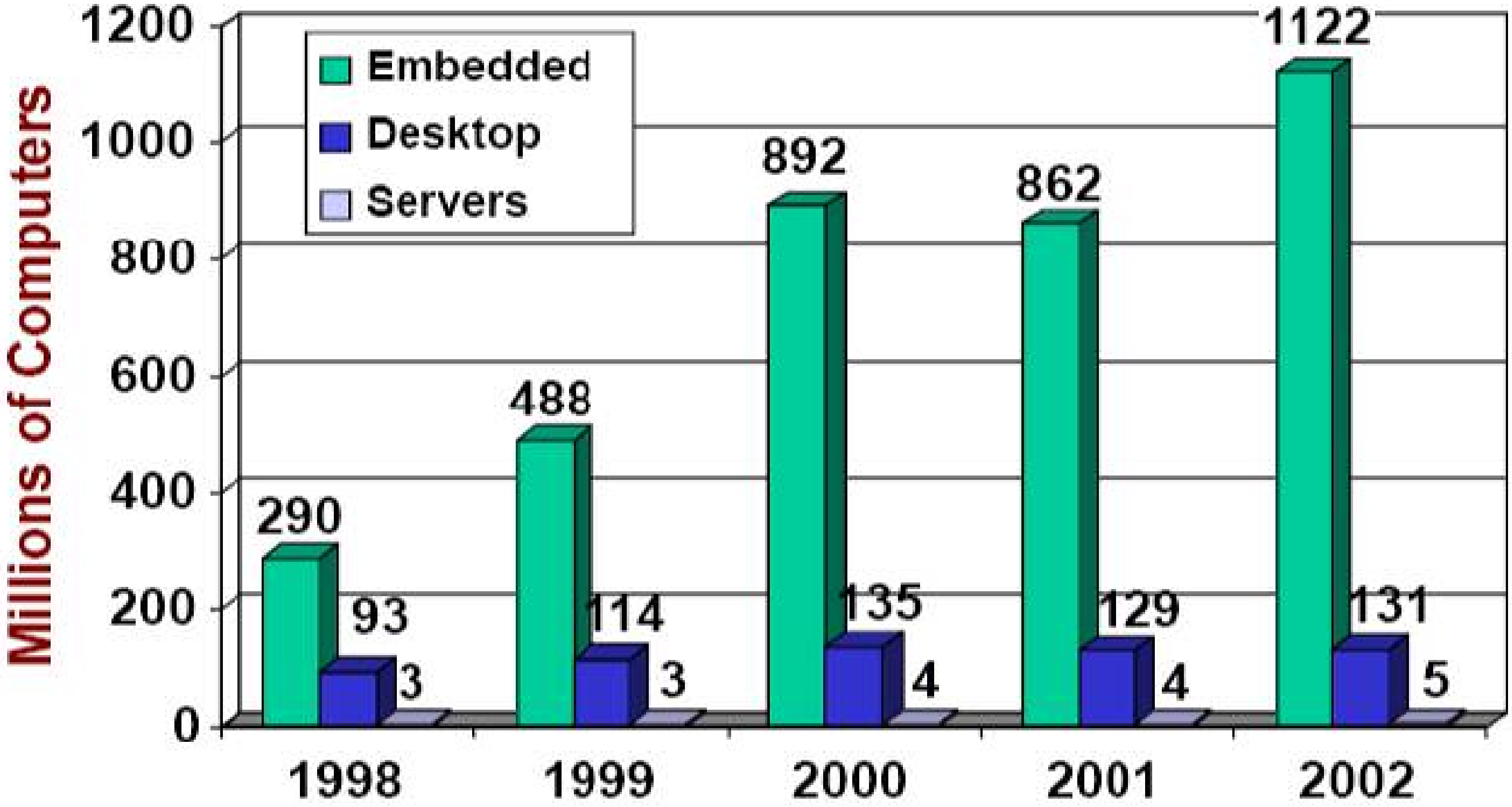


Performance now improves - 50% per year (2x every 1.5 years)

Processor Performance



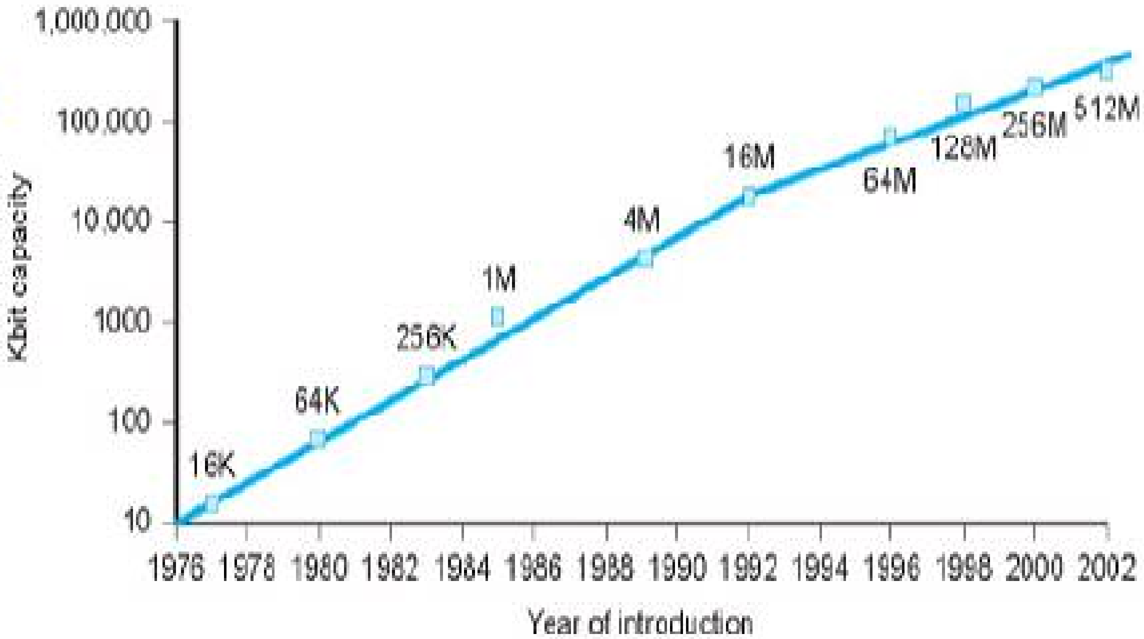
Where is the Market?



Technology Impact on Design

- DRAM capacity has been consistently quadrupled every 3 years, a 60% increase per year, resulting over 16,000 times in 20 years (recently slowed down doubling every 2 years or 4 times every 4 years)
- Processor organization is becoming a main focus of performance optimization
- Technology advances got H/W designer to focus not only on performance but also on functional integration and power consumption (e.g. system on a chip)
- Programming is more concerned with cache and no longer constrained by the RAM size

Year	Size(Mb)	Cyc time
1980	0.0625	250 ns
1983	0.25	220 ns
1986	1	190 ns
1989	4	165 ns
1992	16	145 ns
1996	64	120 ns
2000	256	100 ns

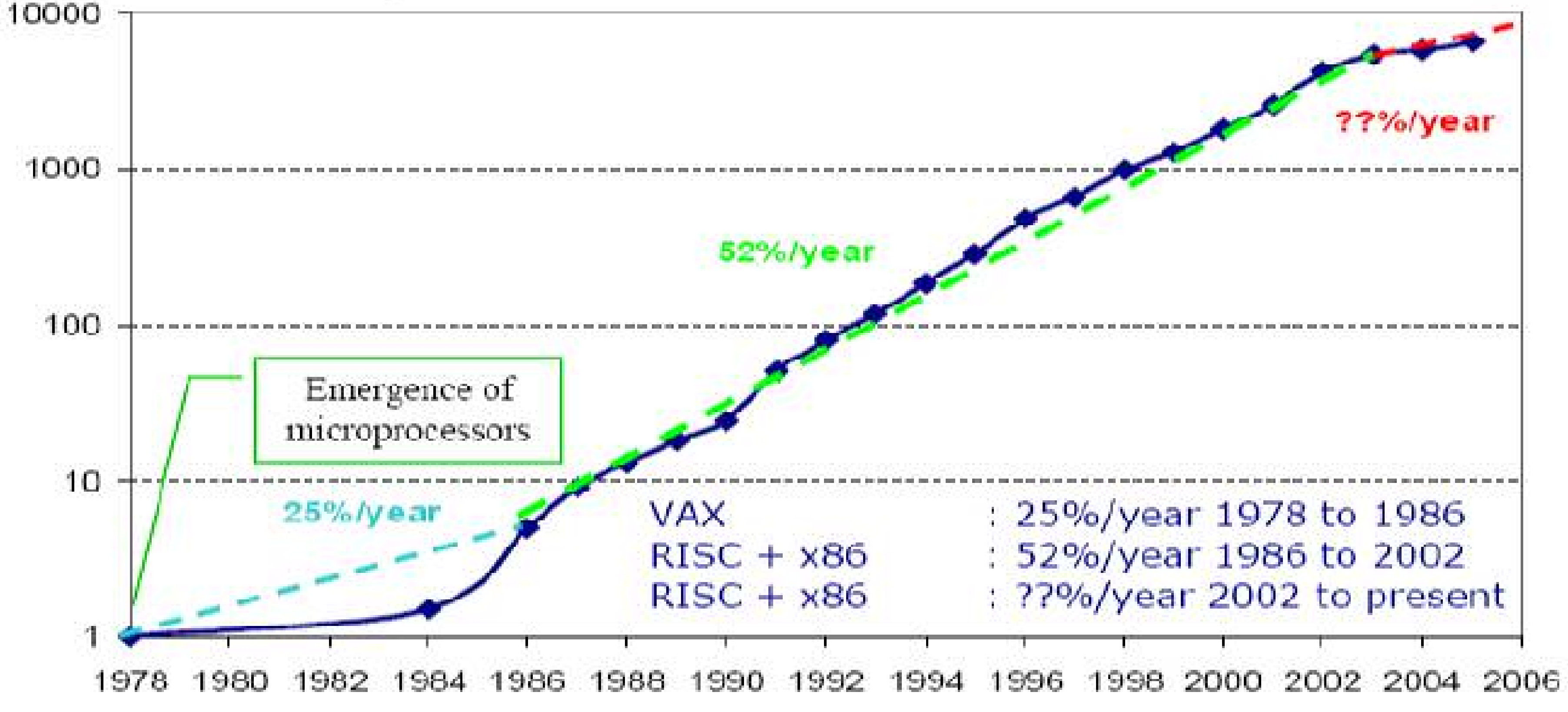


Historical Perspective

Year	Name	Size (Ft.^{3})	Power (Watt)	Perform. (adds/sec)	Mem. (KB)	Price	Price/ Perform. vs. UNIVAC
1951	UNIVAC 1	1000	124K	1.9K	48	$1M	1
1964	IBM S/360 model 50	60	10K	500K	64	$1M	263
1965	PDP-8	8	500	330K	4	$16K	10,855
1976	Cray-1	58	60K	166M	32,768	$4M	21,842
1981	IBM PC	1	150	240K	256	$3K	42,105
1991	HP 9000/ model 750	2	500	50M	16,384	$7.4K	3,556,188
1996	Intel PPro PC 200 Mhz	2	500	400M	16,384	$4.4K	47,846,890

After adjusting for inflation, price/performance has improved by about 240 million in 45 years (about 54% per year)

Uniprocessor Performance



Since 2002, performance improvement has dropped to 20% per yr due to:

- maximum power dissipation of air-cooled chips
- little ILP left for exploitation
- almost unchanged memory latency

A Short History

- 1960s, large mainframes

- 1970s, mini-computers

PDP 7

- 1980s, personal computers, work stations

- 1990s, high performance computers, PDA, etc.
- 2000s, embedded computers

Design Issues of Classes of Computers

- Desktop computing
  - $500 to $5,000
  - price-performance, graphics performance
- Server
  - $5,000 to $5,000,000
  - throughput, availability/dependability, scalability
- Embedded system
  - $10 to $100,000
  - price, power consumption, application-specific performance, the need to minimize memory

Topic Coverage:

- Introduction.
- CISC & RISC Architecture.
- Evolution of computer architecture.
  - ✓ Multiprocessors & multicomputers
  - ✓ Vector supercomputers & SIMD supercomputers
  - ✓ VLSI models, Dataflow machines.
- Pipeline Processing.
- Program behavior & network properties.
- Hardware & software parallelism.
- Principles of scalable performance.
- Performance metrics & measures.
- Parallel Computer Architectures: