CS 35101 Computer Architecture

Section 600

Dr. Angela Guercio

Fall 2010

Structured Computer Organization

- A computer’s native language, machine language, is difficult for human’s to use to program the computer
  - Due to this difficulty, computers are often structured as a series of abstractions, each building on the one below it
  - In this way, complexity can be mastered
  - This approach is called **structured computer organization**

Languages, Levels, and Virtual Machines

- Let the machine language be called L0 (since it is at the lowest level of abstraction)
- L0 is inconvenient for human use, so let's design a new language L1 which is easier
  - A program written in L1 must be **translated** into an equivalent L0 program before it can be executed
  - Another possibility is to write a program in L0 that examines each individual instruction and executes the equivalent sequence of L0 instructions. This technique is called **interpretation** and the program is called an **interpreter**.

Languages, Levels, and Virtual Machines

  - Translation and interpretation are similar. Both methods, and a combination of the two, are widely used
- Rather than think of translation/interpretation, it is often simpler to imagine a **virtual machine** whose machine language is L1. Call this machine M1.
  - Why not implement M1 directly? It might be too expensive or complicated to construct out of electrical circuits.
  - In order to make translation practical, M0 (the real machine) and M1 must not be too different. So, L1 might still be difficult to program! Solution? Create a new VM M2. If necessary repeat until we have a useful machine. This leads to a computer consisting of a number of **layers** or **levels**, one on top of another.

Computer as Multilevel Machine

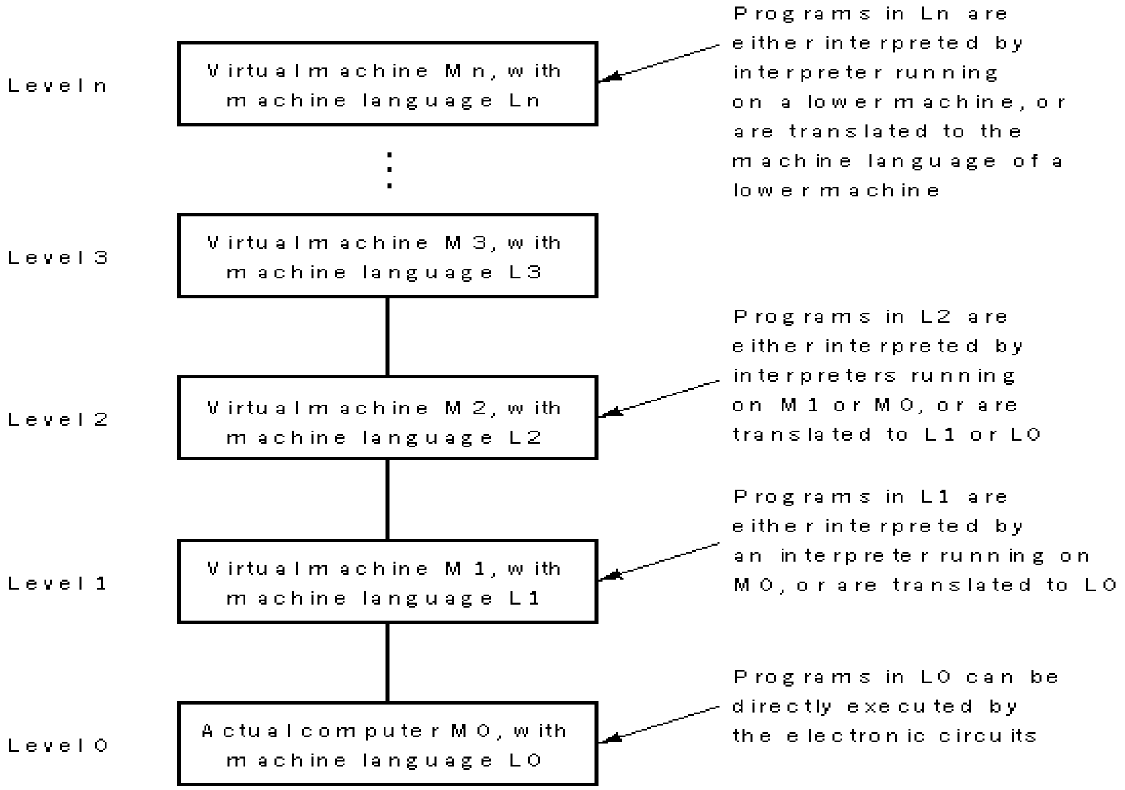


Figure 1-1. A multilevel machine.

Contemporary Multilevel Machines

- Most modern computers consist of two or more levels (as many as six)
  - The lowest level is the **digital logic level** constructed from **gates**
    - Each gate has one or more digital inputs and computes some simple function of the inputs such as AND or OR. Gates are built up from transistors.
    - A small number of gates can be combined to form a 1-bit memory. 1-bit memories can be combined to from 16, 32, or 64 bit **registers** which can hold a single binary number.

Contemporary Multilevel Machines

- The next level up is the **microarchitecture** level
  - At this level we see a collection of (typically) 8 to 32 registers that form a local memory and a circuit called an **ALU** (**Arithmetic Logic Unit**) capable of performing simple arithmetic operations.
  - The registers are connected to the ALU to form a data path over which data flow
  - On some machines the operation of the data path is controlled by a program called a **microprogram**. On other machines the data path is controlled directly by hardware.

Contemporary Multilevel Machines

- Level 2 is the **Instruction Set Architecture (ISA) level.** This level consists of the instructions that can be carried out by the computer.
- The facilities added at level 3 are carried out by an interpreter running at level 2 called an operating system. This level is called the **operating system level**.
- Levels 4 and 5 are used by application programmers (only systems programmers use the lower three levels). The languages of levels 4 and 5 are usually translated while those of levels 2 and 3 are always interpreted.

Contemporary Multilevel Machines

- Levels 4 and 5 provide symbolic languages while the machine languages of levels 1, 2, and 3 are numeric.
- Level 4 is the **assembly language level**. It provides a program called an **assembler** which translates a symbolic form of the level 1, 2, or 3 language.
- Level 5 consists of high-level languages such as BASIC, C, C++, and Java. Programs written in these languages are translated to level 3 or 4 languages by translators known as **compilers**.

A Six-Level Computer

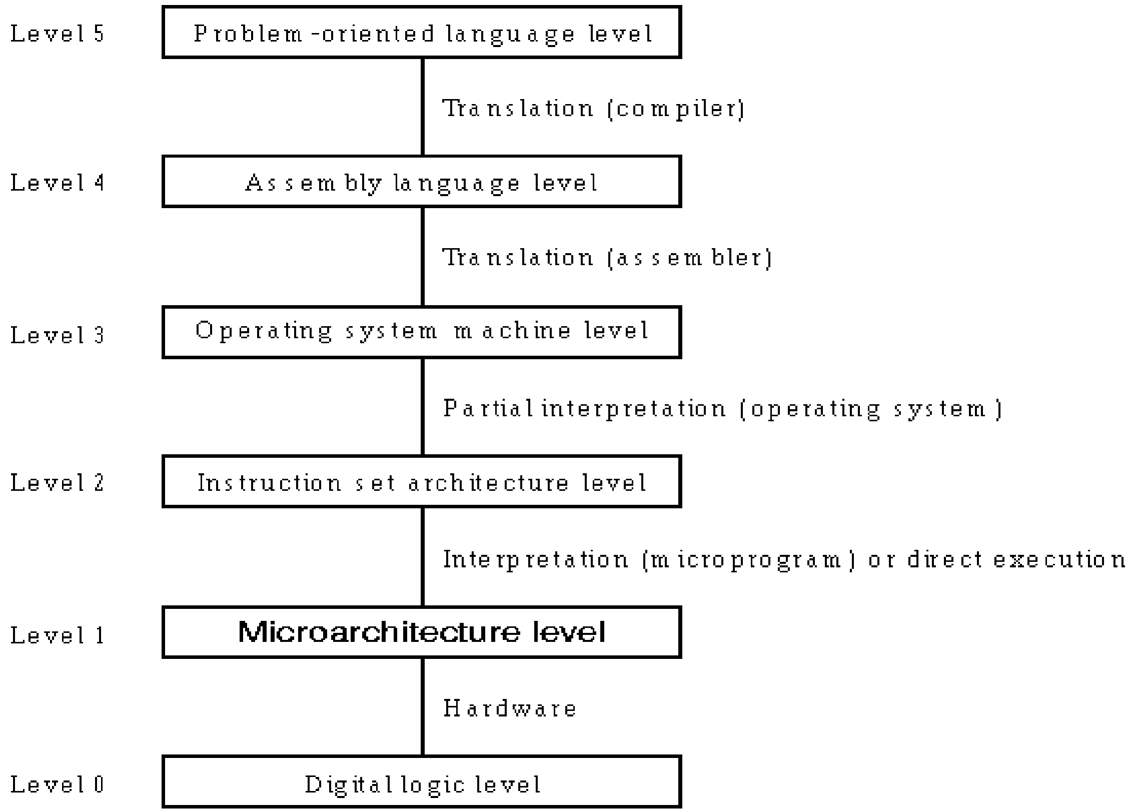


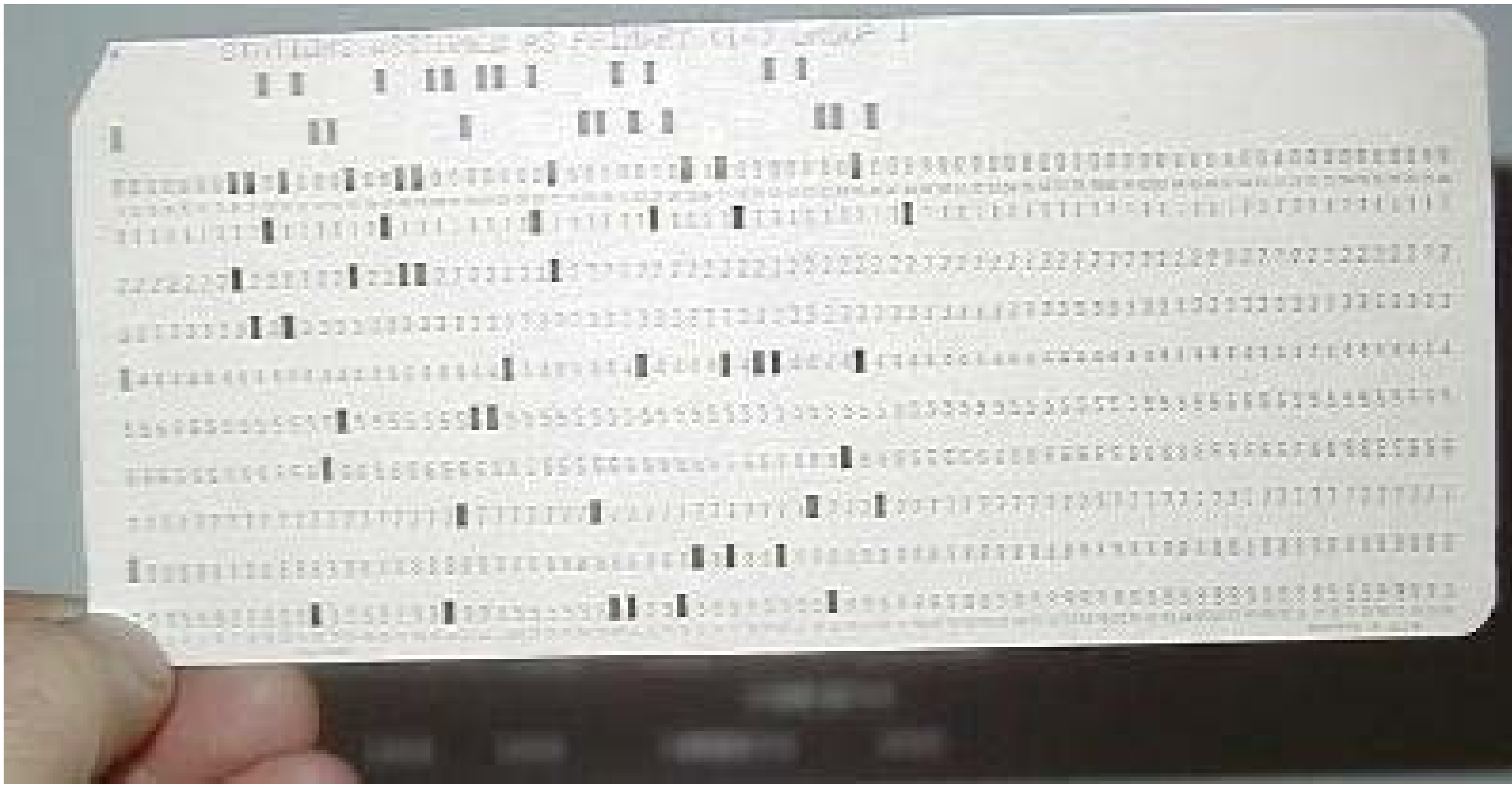
Figure 1-2. A six-level computer. The support method for each level is supported is indicated below it (along with the name of the supporting program).

Evolution of Multilevel Machines

- The first digital computers had only two levels: the ISA level (where programming was done) and the digital logic level. The digital logic circuits were very complicated.
  - In 1951, Maurice Wilkes suggested the idea of a three-level computer in order to simplify the hardware. The machine was to have a built-in unchangeable program (the microprogram whose function was to execute ISA-level programs. The microprogram was easier to implement in hardware then the instruction set, so the circuits needed were simpler.
  - A few three-level machines were constructed in the 1950s, more during the 1960s and by 1970 microprogrammed machines were dominant.

Evolution of Multilevel Machines

- Early computers were operated directly by a programmer (or a computer operator) who entered a deck of cards containing (e.g.) a FORTRAN program. The program was translated to machine language, output on cards and subsequently run.
- Around 1960 people tried to reduce wasted time by automating the operator’s job. A program called an **operating system** was kept in the computer at all times. The programmer provided certain control cards along with the program that were read and carried out by the operating system.



A Sample Job for FMS OS

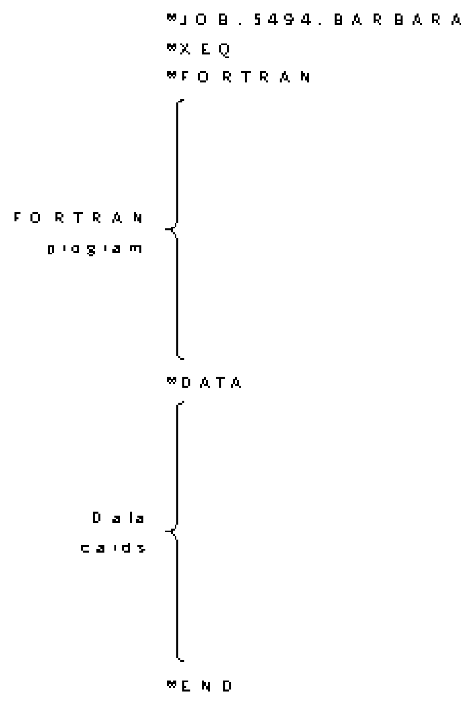


Figure 1-3. A sample job for the FMS operating system.

Evolution of Multilevel Machines

- Early operating systems read card decks and printed output on the line printer. These were known as **batch systems**. In the early 1960s **timesharing systems** in which users were connected to the CPU using terminals and the CPU was shared were introduced.
- Due to the ease of introducing new instructions in microprogrammed architectures, by the 1970s instruction sets had grown large and the microprogram large and slow. At this point researchers realized that by simplifying the instruction set and implementing it directly in hardware the computer could be much faster.

Milestones in Computer Architecture (1)

Year	Name	Made by	Comments
1834	Analytical Engine	Babbage	First attempt to build a digital computer
1936	Z1	Zuse	First working relay calculating machine
1943	COLOSSUS	British gov't	First electronic computer
1944	Mark I	Aiken	First American general-purpose computer
1946	ENIAC I	Eckert/Mauchley	Modern computer history starts here
1949	EDSAC	Wilkes	First stored-program computer
1951	Whirlwind I	M.I.T.	First real-time computer
1952	IAS	Von Neumann	Most current machines use this design
1960	PDP-1	DEC	First minicomputer (50 sold)
1961	1401	IBM	Enormously popular small business machine
1962	7094	IBM	Dominated scientific computing in the early 1960s
1963	B5000	Burroughs	First machine designed for a high-level language
1964	360	IBM	First product line designed as a family

Milestones in Computer Architecture (2)

Year	Name	Made by	Comments
1965	PDP-8	DEC	First mass-market minicomputer (50,000 sold)
1970	PDP-11	DEC	Dominated minicomputers in the 1970s
1974	8080	Intel	First general-purpose 8-bit computer on a chip
1974	CRAY-1	Cray	First vector supercomputer
1978	VAX	DEC	First 32-bit superminicomputer
1981	IBM PC	IBM	Started the modern personal computer era
1981	Osborne-1	Osborne	First portable computer
1983	Lisa	Apple	First personal computer with a GUI
1985	386	Intel	First 32-bit ancestor of the Pentium line
1985	MIPS	MIPS	First commercial RISC machine
1987	SPARC	Sun	First SPARC-based RISC workstation
1990	RS6000	IBM	First superscalar machine
1992	Alpha	DEC	First 64-bit personal computer
1993	Newton	Apple	First palmtop computer

Computer Generations

- Zeroth Generation
  Mechanical Computers (1642 – 1945)
- First Generation
  Vacuum Tubes (1945 – 1955)
- Second Generation
  Transistors (1955 – 1965)
- Third Generation
  Integrated Circuits (1965 – 1980)
- Fourth Generation
  Very Large Scale Integration (1980 – ?)

A von Neumann Machine

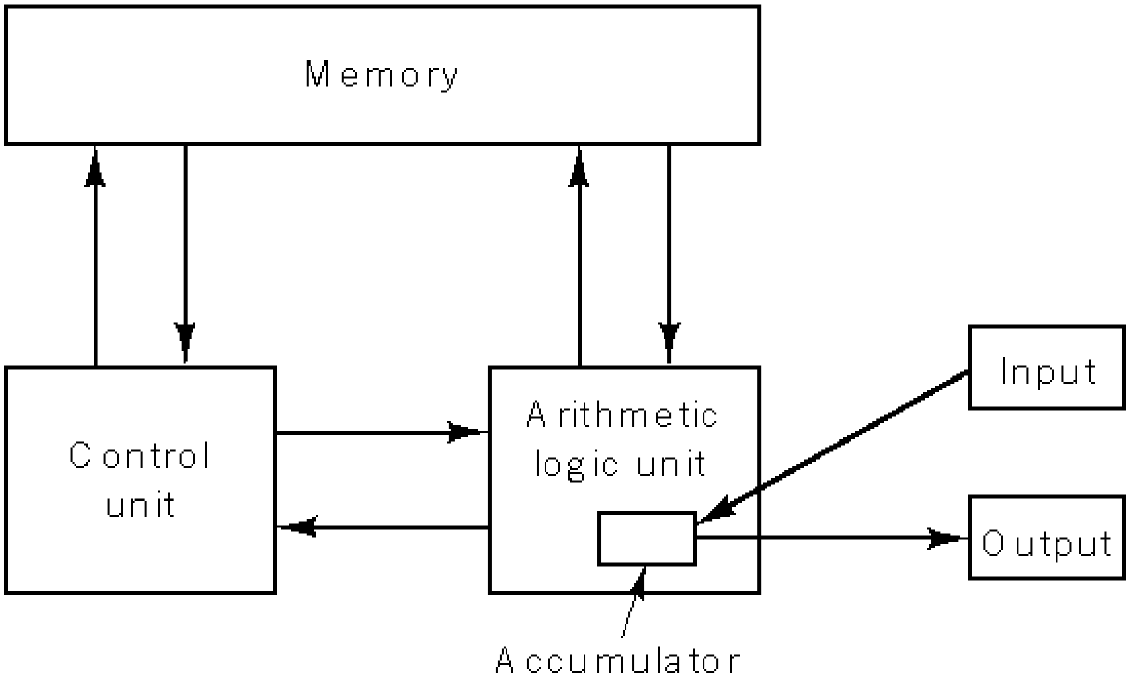


Figure 1-5. The original von Neumann machine.

First Bus-Based System

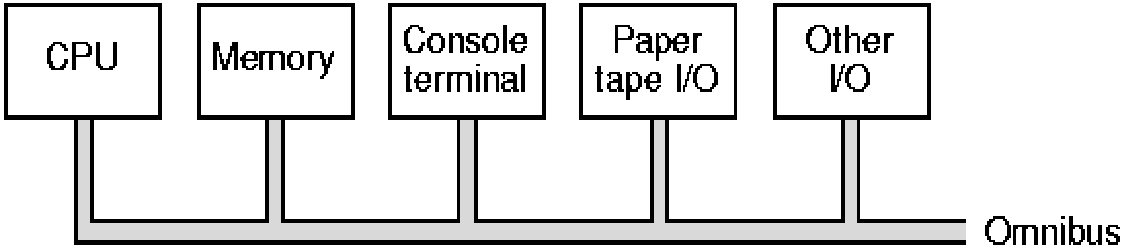


Figure 1-6. The PDP-8 omnibus.

The IBM 360 Product Line

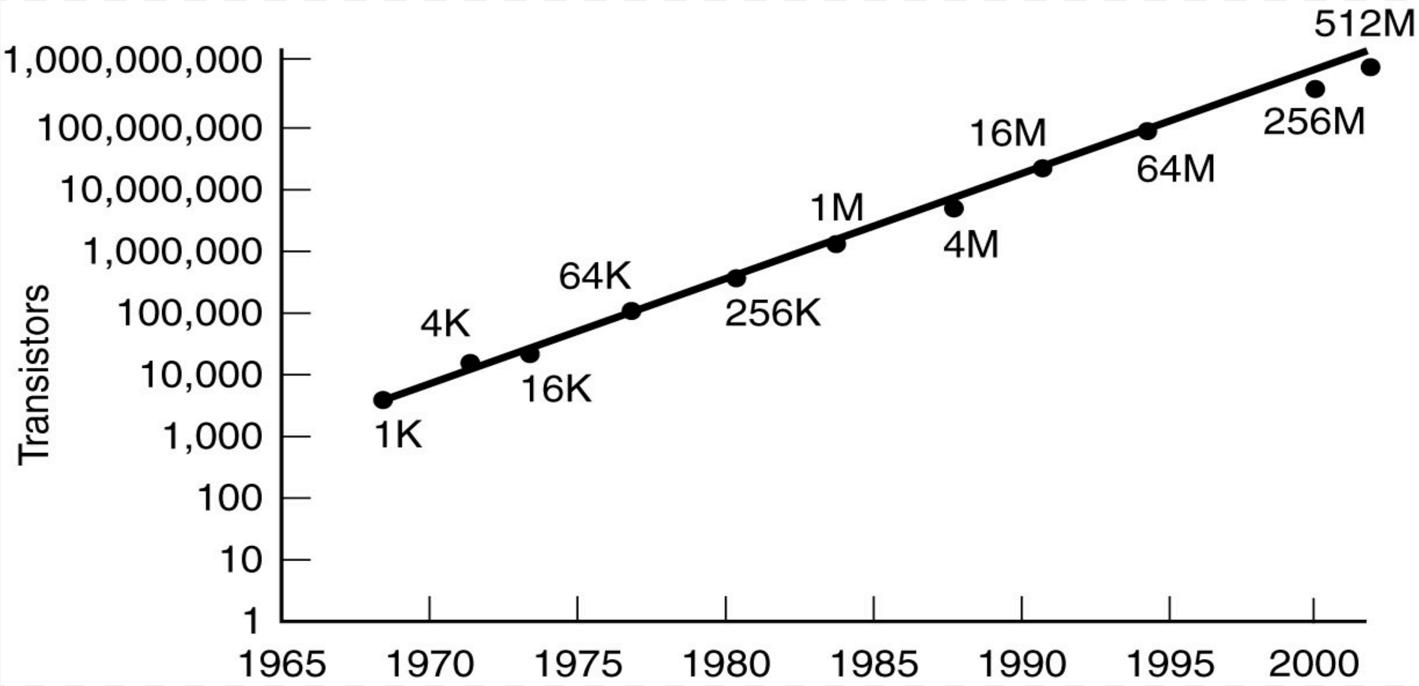
Property	Model 30	Model 40	Model 50	Model 65
Relative performance	1	3.5	10	21
Cycle time (in billionths of a sec)	1000	625	500	250
Maximum memory (bytes)	65,536	262,144	262,144	524,288
Bytes fetched per cycle	1	2	4	16
Maximum number of data channels	3	3	4	6

The initial offering of the IBM product line.

Moore’s Law

- **Moore’s law** is named after Gordon Moore, co-founder and Chairman of Intel, who discovered it in 1965
  - The law states that the number of transistors that can be put on a chip doubles every 18 months
  - Many observers expect Moore’s law to continue to hold into the 21st Century, possibly around 2020

Technological and Economic Forces



Moore's law predicts a 60-percent annual increase in the number of transistors that can be put on a chip.
The data points given in this figure are memory sizes, in bits.

The Computer Spectrum

Although PCs are the most common type of computer, there are others as well

Type	Price ($)	Example application
Disposable computer	0.5	Greeting cards
Microcontroller	5	Watches, cars, appliances
Game computer	50	Home video games
Personal computer	500	Desktop or notebook computer
Server	5K	Network server
Collection of Workstations	50–500K	Departmental minisupercomputer
Mainframe	5M	Batch data processing in a bank

The current spectrum of computers available. The prices should be taken with a grain (or better yet, a metric ton) of salt.

The Computer Spectrum

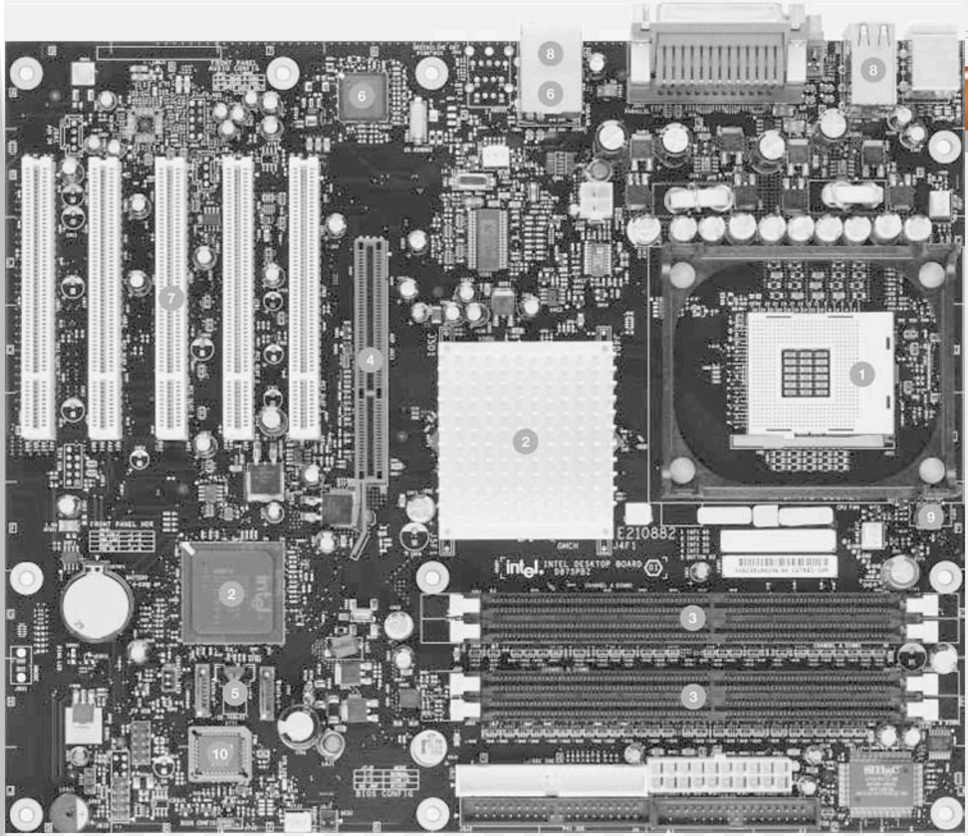
- **NOW** (**Network of Workstations**) or **COW** (**Cluster of Workstations**) consist of standard PCs or workstations connected by gigabit/sec networks, and running specialized software that allow all the machines to work together on a single problem
- Mainframes are large computers which have high I/O capacity and vast storage capacities
- Supercomputers are typically highly parallel machines made from high-end components

Typical Prices of Computers

Type	Price ($)	Example application
Disposable computer	1	Greeting cards
Embedded computer	10	Watches, cars, appliances
Game computer	100	Home video games
Personal computer	1K	Desktop or portable computer
Server	10K	Network server
Collection of Workstations	100K	Departmental minisupercomputer
Mainframe	1M	Batch data processing in a bank
Supercomputer	10M	Long range weather prediction

Figure 1-9. The current spectrum of computers available. The prices should be taken with a grain (or better yet, a metric ton) of salt.

Personal Computer



1. Pentium 4 socket
2. 875P Support chip
3. Memory sockets
4. AGP connector
5. Disk interface
6. Gigabit Ethernet
7. Five PCI slots
8. USB 2.0 ports
9. Cooling technology
10. BIOS

A printed circuit board is at the heart of every personal computer. This figure is a photograph of the Intel D875PBZ board. The photograph is copyrighted by the Intel Corporation, 2003 and is used by permission.

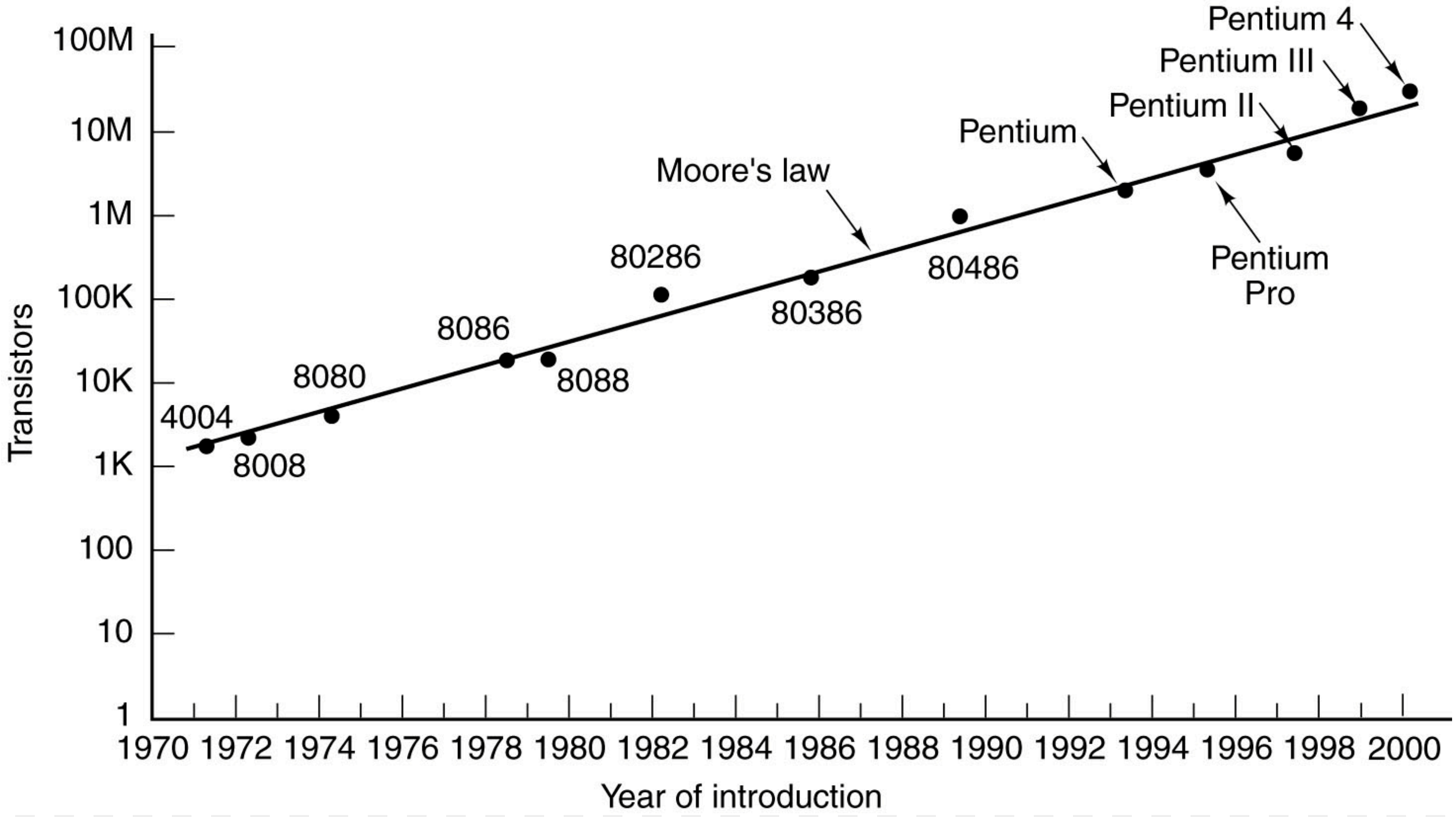
Pentium II

- The Intel Corporation was formed in 1968.
- In 1970, Intel manufactured the first single-chip CPU, the 4-bit 4004 for a Japanese company to use in an electronic calculator.
- The 8088, a 16-bit CPU was chosen as the CPU for the original IBM PC.
- A series of backward compatible chips (80286, 386, 486, Pentium, Pentium Pro and Pentium II) followed.

The Intel CPU Family

Chip	Date	MHz	Transistors	Memory	Notes
4004	4/1971	0.108	2300	640	First microprocessor on a chip
8008	4/1972	0.108	3500	16 KB	First 8-bit microprocessor
8080	4/1974	2	6000	64 KB	First general-purpose CPU on a chip
8086	6/1978	5–10	29,000	1 MB	First 16-bit CPU on a chip
8088	6/1979	5–8	29,000	1 MB	Used in IBM PC
80286	2/1982	8–12	134,000	16 MB	Memory protection present
80386	10/1985	16–33	275,000	4 GB	First 32-bit CPU
80486	4/1989	25–100	1.2M	4 GB	Built-in 8-KB cache memory
Pentium	3/1993	60–233	3.1M	4 GB	Two pipelines; later models had MMX
Pentium Pro	3/1995	150–200	5.5M	4 GB	Two levels of cache built in
Pentium II	5/1997	233–450	7.5M	4 GB	Pentium Pro plus MMX instructions
Pentium III	2/1999	650–1400	9.5M	4 GB	SSE Instructions for 3D graphics
Pentium 4	11/2000	1300–3800	42M	4 GB	Hyperthreading; more SSE instructions

Moore’s Law

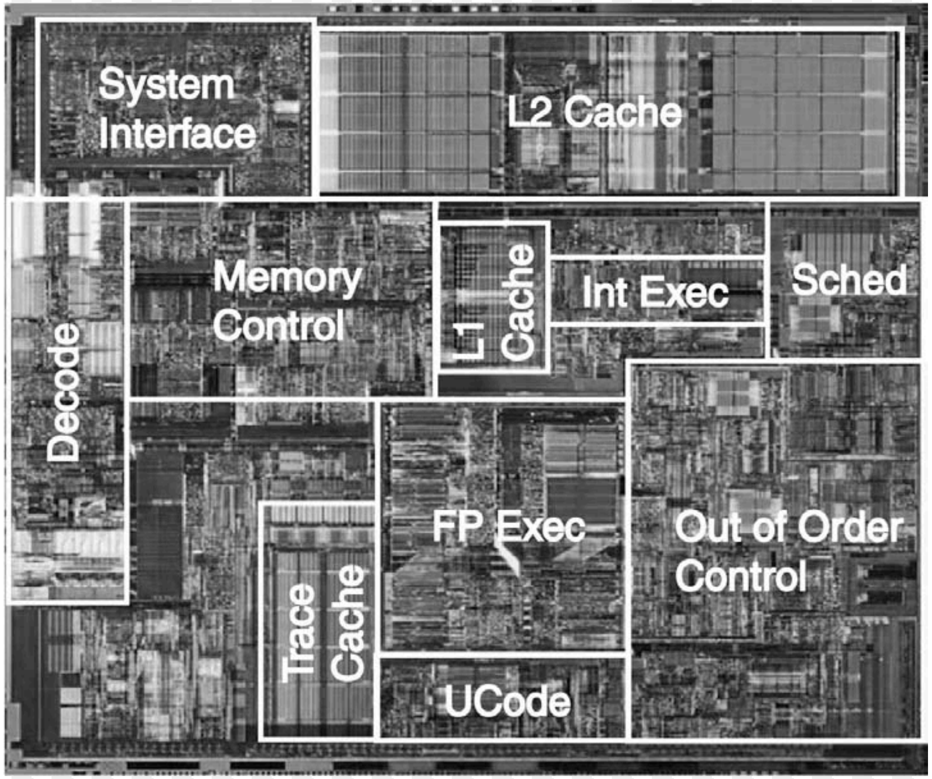


Transistors
100M
10M
1M
100K
10K
1K
100
10
1
1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000
Year of introduction
Moore's law
4004
8008
8080
8086
8088
80286
80386
80486
Pentium
Pentium Pro
Pentium II
Pentium III
Pentium 4


intel
entium II
COMPAQ


pentium II
intel

The Pentium 4 Chip



The Pentium 4 chip. The photograph is copyrighted by the Intel Corporation, 2003 and is used by permission.

Example Computer Families

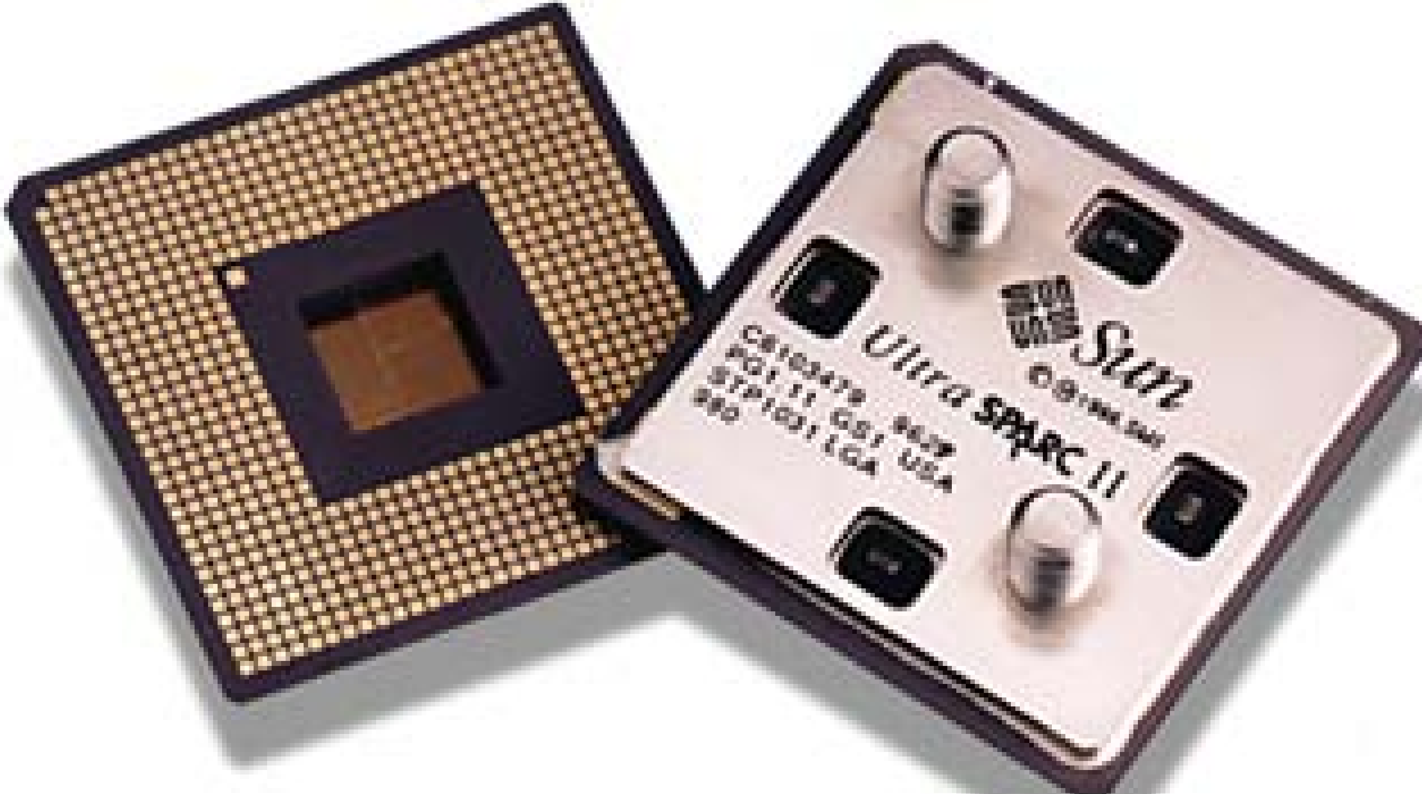
- Pentium 4 by Intel
- UltraSPARC III by Sun Microsystems
- The 8051 chip by Intel, used for embedded systems

UltraSPARC II

- In the 1970s, UNIX was popular at universities, but it ran only on timeshared minicomputers such as the VAX and PDP-11
- In 1981, a Stanford graduate student built a personal UNIX workstation using off-the-shelf parts. It was called the SUN-1.
- Early Sun workstations used Motorola CPUs.
- In 1987, Sun decided to design its own CPU based on a Cal Berkeley design called the RISC II.

UltraSPARC II

- The new CPU was called the **SPARC (Scalable Processor ARChitecture)** and was used in the Sun-4.
- The SPARC was licensed to several semiconductor manufacturers who developed binary compatible versions.
- The first SPARC was a 32-bit machine with only 55 instructions (an FPU added 14 additional instructions).
- A 64-bit version, the UltraSPARC I was developed in 1995. This machine was aimed at high-end applications (e.g. web and database servers).


Sun
UltraSPARC II

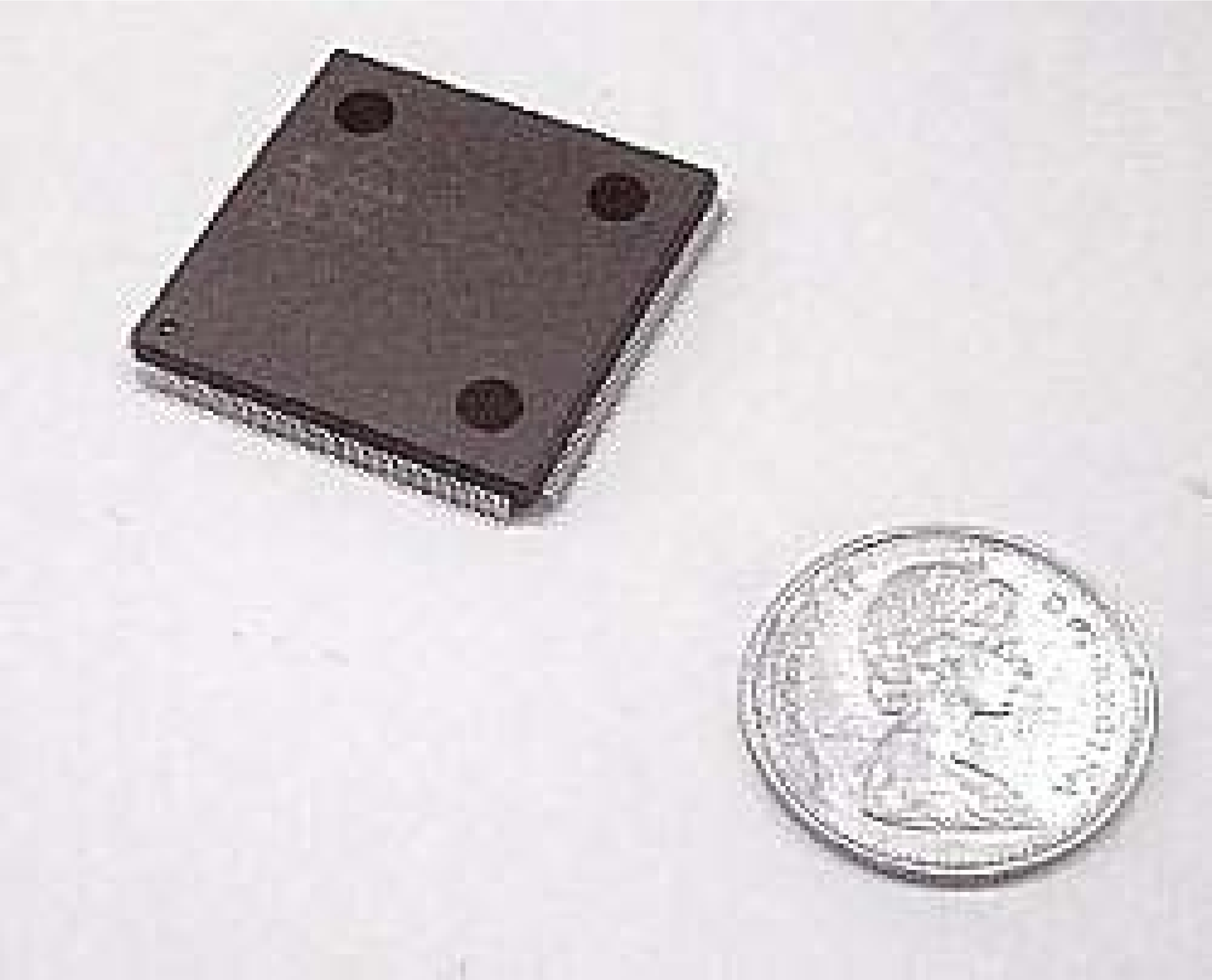
picoJava II

- The Java programming language defines a **Java Virtual Machine (JVM)** in order to allow portability across many architectures
- Usually Java programs are interpreted by a JVM interpreter written in C, but interpretation is slow.
  - One solution is to have a **JIT (Just In Time)** compiler for the machine running the JVM
  - Another alternative is to design hardware JVM chips, thus avoiding a level of software interpretation or JIT compilation.

picoJava II

- Such Java chips are especially useful in embedded systems
- The picoJava II is not a concrete chip, but a chip design which is the basis for a number of chips such as the Sun microJava 701 CPU.
- The picoJava has two optional units: a cache and a floating-point unit which can be included or removed as the manufacturer sees fit.

0227
EER1
DE180227
192
129
FUJITSU
193
128
JAPAN
picoJava II
256
65
64
CPU1
DBCN2
DBCN1
40
30
20
10
1



MCS-51 Family

Chip	Program memory	Mem. type	RAM	Timers	Interrupts
8031	0 KB		128	2	5
8051	4 KB	ROM	128	2	5
8751	8 KB	EPROM	128	2	5
8032	0 KB		256	3	6
8052	8 KB	ROM	256	3	6
8752	8 KB	EPROM	256	3	6

Members of the MCS-51 family.

Metric Units

Exp.	Explicit	Prefix	Exp.	Explicit	Prefix
10^{-3}	0.001	milli	10^{3}	1,000	Kilo
10^{-6}	0.000001	micro	10^{6}	1,000,000	Mega
10^{-9}	0.000000001	nano	10^{9}	1,000,000,000	Giga
10^{-12}	0.000000000001	pico	10^{12}	1,000,000,000,000	Tera
10^{-15}	0.000000000000001	femto	10^{15}	1,000,000,000,000,000	Peta
10^{-18}	0.000000000000000001	atto	10^{18}	1,000,000,000,000,000,000	Exa
10^{-21}	0.000000000000000000001	zepto	10^{21}	1,000,000,000,000,000,000,000	Zetta
10^{-24}	0.000000000000000000000001	yocto	10^{24}	1,000,000,000,000,000,000,000,000	Yotta

The principal metric prefixes.