COMPUTER SYSTEMS

Sotirios G. Ziavras
Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, New Jersey 07102, U.S.A.

Keywords: Computer organization, processor, memory hierarchy, peripheral devices, bus architectures, multiprocessors, multicomputers, computation models, supercomputers.

Contents

UNESCO – EOLSS SAMPLE CHAPTERS

Summary

The fundamentals of computer systems design and organization are presented, and the conventional procedure for the execution of computer programs is described. An overview of the major features and interactions of the hardware and software components of modern computer systems is also included. Numerous computer systems have been designed and built to aid humans in information processing and numerical calculations. As a result, several models have emerged in the field of computer systems design. These models differ in the architecture of the processors, the underlying model of computation, the architecture of the main memory, or the techniques used to interconnect the basic resources within the computer. This article presents a summary of the most fundamental computer models. The performance analysis task of computer systems is touched upon to facilitate comparisons. Advances in the technology that integrates transistors on chips improve the performance of all design models by increasing the depth of the instruction execution pipelines and the number of functional units in processors, the speed of all major electronic components, the size of on-chip cache memories, etc.

1. Introduction

Modern computers are electronic and process digital information. The physical machine consists of transistors, digital circuits implemented with transistors, wires, and mechanical components in peripheral devices used for information storage. These physical entities are collectively called hardware. System and application programs are called software. A general purpose computer system is a programmable machine that can solve problems by accepting inputs and instructions on how to use these inputs. The instructions are included in computer programs (that is, software) that normally contain sequences of them. Graphical languages are rarely used to represent computer programs as collections of instructions with relationships between arbitrary pairs of them. Programs are often written in high-level languages (HLLs) that have to be translated (by appropriate software compilers) to produce machine-readable (that is, machine language) code that can be run directly by the given computer system. The machine language code contains sequences of primitive instructions for the given computer in binary representation. On the other hand, HLLs employ mnemonics of more powerful instructions, and appropriate structures to make programming easy and independent of the target computer.

From the software point of view, a computer is a six-level system consisting of the digital logic (collections of electronic gates), microarchitecture (a collection of functional units, such as ALUs - Arithmetic Logic Units, and their interconnectivity), instruction set architecture (the complete set of machine language instructions), operating system (code that monitors and controls the activities of the computer), assembly and machine language, and high-level language. The assembly language is very close to the machine language of a computer; it basically replaces the binary representation of machine instructions with mnemonics in a one-to-one fashion. From the hardware point of view, a computer is conveniently assumed to be a five-level hierarchy. The five levels correspond to network ports for connecting to the outside world (these ports may not be necessarily available, as a computer may be a standalone information processing and/or computing machine), peripheral or mass-storage devices for (applications and system) program and data storage, main memory, program and data caches (fast memories for retrieving data by content), and CPU (Central Processing Unit) or processor. Special emphasis is given in this article to the description of computer systems based on this five-level representation.

Many other components are also included in computer systems in order to enable the aforementioned basic components to function properly. For example, control and data busses are used to transmit data between any two successive levels of the hardware hierarchy and glue logic is used to implement the appropriate interfaces. The design of a computer system most often begins with the selection of a particular CPU. The other components are selected progressively based on performance requirements. Analytical techniques, software simulations, and software or hardware prototyping of the complete or partial computer system are used to make final decisions about the design. Special attention is given nowadays to hardware-software codesign, where the selection or design of components is made in unison with the development of the corresponding system software.

There exist several types of general purpose computer systems. These types are grouped together into two major computer classes, comprising sequential or conventional computers, and parallel computers, respectively.

UNESCO – EOLSS SAMPLE CHAPTERS

The class of sequential or conventional computer systems comprises:

- **Laptops and palmtops.** These are small, portable computer systems. Laptops often contain very powerful processors and have capabilities very close to those of PCs (see below); their major drawbacks are smaller screens, smaller memory, and fewer peripheral (that is, I/O - Input/Output) devices which are, however, portable. These single-user computers are implemented in the form of microcomputers. The prefix micro denotes the inclusion of a microprocessor that resides on a single chip and serves as the CPU. Other components also are included, of which the critical ones are a keyboard for data input, a screen (monitor) for information display, memory chips for temporary storage, and a hard disk for data storage.
- **PCs (personal computers) or desktops.** These computers also are of the microcomputer type. Figure 1 shows the basic components of a microcomputer. The RAM and ROM form the main memory that stores system and application programs, and data. The ROM contains only part of the operating system, and most often the part that initializes the computer. The sofware stored in the ROM is called firmware. Resource interface units form the required glue logic for the implementation of the required data exchange protocols. The control bus transfers the control signals produced by the microprocessor. To access an element in the memory or a peripheral device, the microprocessor first issues the address of that item on the address bus. The address bus is unidirectional, from the processor to the other units. While the latter value is still present on the address bus, the microprocessor issues the appropriate control signals to read or write from the corresponding location. The address issued by the microprocessor is decoded by external logic to choose the appropriate memory module or I/O device. The data is finally transferred on the bidirectional data bus. Details on how microcomputers execute programs are presented in Section 2.
- **Workstations.** They appeared in the 1980s as single-user computers with much better performance than PCs, primarily because they contain very advanced microprocessors. They often include proprietary co-processors to facilitate graphics functions because they basically target the scientific and engineering communities. They were uniprocessor computers in the early days, but multiprocessor workstations appeared for the first time in the market a few years ago. They are now often used as multi-user platforms.
- **Minicomputers.** High-performance cabinet-sized computers that can be used simultaneously by a few dozens of users. They are often used in engineering and scientific applications. They have been replaced recently by advanced workstations and networks of workstations.
- **Mainframes.** Very powerful computers that can serve many dozens or hundreds of users simultaneously. IBM has produced numerous computers of this type. They have been replaced recently in many occasions by networks of workstations.

Contrary to sequential computers that use a single CPU to solve a problem, parallel computer systems employ many CPUs in appropriately connected structures. This new class of computers comprises multiprocessors, multicomputers, and vector supercomputers. These types of computer systems are discussed in detail in Section 3.

UNESCO – EOLSS SAMPLE CHAPTERS

Also, distributed computer systems can be developed, where several complete computer systems are connected together in a networked fashion to solve a problem in parallel. For the sake of brevity, we do not discuss distributed computer systems any further. Section 2 presents in detail the class of sequential computers. Without loss of generality, for a better description emphasis is given in Section 2 to microcomputers.

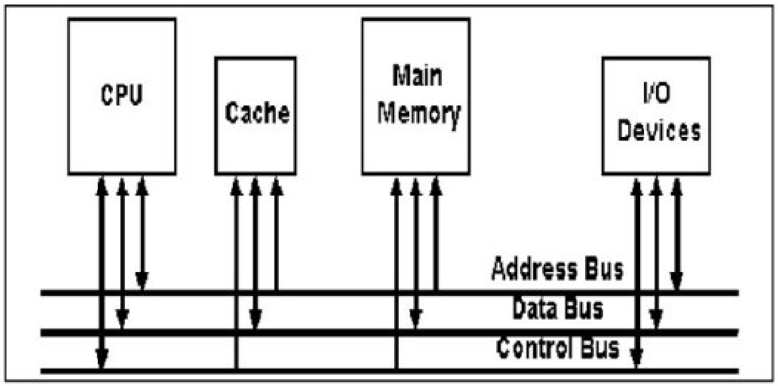

Figure 1: The architecture and basic components of a microcomputer.

2. Sequential/Conventional Computers

2.1. Basic Resources

Let us present briefly the basic components of a sequential computer and discuss their interoperability. The basic resources of complete computer systems are the CPU, instruction and data caches, main memory, peripheral devices, busses for the interconnection of these components, and some interfaces that implement the data exchange protocols. They are studied in the following subsections.

-
-
-

TO ACCESS ALL THE **20 PAGES** OF THIS CHAPTER,
Visit: http://www.eolss.net/Eolss-sampleAllChapter.aspx

Bibliography

Amdahl G.M. (1967). Validity of the Single Processor Approach to Achieving Large Scale Computing Capabilities. *Proceedings of the AFIPS (American Federation of Information Processing Societies) Conference*, pp. 483-485. [This paper introduces Amdahl's law that determines an upper bound on the speedup for the execution of programs with fixed workload on parallel computers.]

Amdahl G.M., Blaauw G.A., and Brooks F.P., Jr. (1964). Architecture of the IBM System/360. *IBM Journal of Research and Development*.

Russell R.M. (1978). The CRAY-1 Computer System. *Communicatons of the ACM* **21**(1), pp. 63-72.

Dennis J.B. and Misunas D.P. (1974). A Preliminary Architecture for a Basic Data-Flow Processor. *Computer Architecture News* **3**(4), pp. 126-132. [The first paper to describe the architecture of an entire processor based on the dataflow execution model. It differs from the conventional model which is control (or program-counter) driven.]

Moore G.E. (1965). Cramming More Components Onto Integrated Circuits. *Electronics*, pp. 114-117. [It introduces Moore's Law that governs advances in processor performance based on the expected degree of integration for silicon transistors.]

Goodman J.R. (1983). Using Cache Memory to Reduce Processor-Memory Traffic. *Proceedings of the 10th International Symposium on Computer Architecture*, pp. 124-131. [It introduces the concept of including cache in computer system design.]

Seitz C.L. (1985). The Cosmic Cube. *Communications of the ACM*, pp. 22-33. [Describes a processor interconnection technology that became very popular in the 1980s.]

Biographical Sketch

Dr. Sotirios G. Ziavras received in 1984 the Diploma in Electrical Engineering from the National Technical University, Athens, Greece. He received in 1985 the M.Sc. in Electrical and Computer Engineering from Ohio University, Athens, Ohio, USA. He was awarded in 1990 the Ph.D. in Computer Science from George Washington University, Washington, D.C., USA, where he was a Distinguished Graduate Teaching Assistant and a Graduate Research Assistant. From 1988 to 1989, he was also with the Center for Automation Research at the University of Maryland, College Park, Maryland, USA. He was a Visiting Assistant Professor in the Electrical and Computer Engineering Department at George Mason University, Fairfax, Virginia, USA, in Spring 1990. He is currently a Professor of Electrical and Computer Engineering at New Jersey Institute of Technology, Newark, New Jersey, USA. His major research interests are processor and computer systems designs, and parallel computing systems and algorithms.

UNESCO – EOLSS
SAMPLE CHAPTERS