software languages. Over the same period of time, these higher-performance microprocessors have become more widely used in the design of new electronic equipment and computers. This book presents a detailed study of the software and hardware architectures of Intel Corporation's 8088 and 8086 microprocessors. An introduction to the 80286, 80386, 80486, and Pentium processors is also included.

In this chapter we begin our study with an introduction to microprocessors and microcomputers. The following topics are discussed:

1.1 General Architecture of a Microcomputer System
1.2 Evolution of the Intel Microprocessor Architecture

## 1.1 GENERAL ARCHITECTURE OF A MICROCOMPUTER SYSTEM

The hardware of a microcomputer system can be divided into four functional sections: the input unit, microprocessing unit, memory unit, and the output unit, The block diagram in Fig. I—1 shows this general microcomputer architecture. Each of these units has a special function in terms of overall system operation. Next we will look at each of these sections in detail.

The heart of a microcomputer is its microprocessing unit (MPU). The MPU of a microcomputer is implemented with a VLSI device known as a microprocessor, or just processor for short. A microprocessor is a general-purpose processing unit built into a single integrated circuit (IC). The microprocessor used in the original IBM PC is Intel Corporation's 8088, shown in Fig. 1-2.

Earlier we indicated that the 8088 is a 16-bit microprocessor. To be more accurate, it is the 8-bit external bus version in Intel's 8086 family of 16-bit microprocessors. Even though the 8088 has an 8-bit external bus, its internal architecture is 16 bits in width and it can directly process 16-bit-wide data. For

this reason, the 8088 is considered a 16-bit microprocessor.

The 8088 MPU is the part of the microcomputer that executes instructions of the program and processes data. It is responsible for performing all arithmetic operations and making the logical decisions initiated by the computer's program. In addition to arithmetic and logic functions, the MPU controls overall system operation.

The input and output units are the means by which the MPU communicates with the outside world. Input unit such as the Keyboard on the IBM PC, allow the user to input information or commands to the MPU; for instance, a programmer could key in the lines of a BASIC program from the keyboard. Many other input devices are available for the PC; two examples include a mouse, for implementing a user-friendlier input interface, and a scanner, for reading in documents.

Figure 1-1 General architecture of a microcomputer system.

Figure 1-2 Intel Corporation's 8088 microprocessor. (Courtesy of Intel Corp.)

The most widely used output devices of a PC are the monitor and the printer. The output unit in a microcomputer is used to give the user feedback and to produce documented results. For example, key entries from the keyboard are echoed back to the monitor—that is, by looking at the information displayed on the monitor, the user can confirm that the correct entry was made. Moreover, the results produced by the MPU's processing can be either displayed or printed. For our earlier example of a BASIC program, once it is entered and corrected, a listing of the statements could be printed. Alternate output devices are also available for the microcomputer; for instance, many modern PCs are equipped with an advanced audio processing and speaker system.

The memory unit in a microcomputer is used to store information, such as number or character data. By "store" we mean that memory has the ability to hold this information for processing or for outputting at a later time. Programs that define how the computer is to operate and process data also reside in memory.

In the microcomputer system, memory can be divided into two different types: primary storage memory and secondary storage memory. Secondary storage memory is used for long-term storage of information that is not currently being used. For example, it can hold programs, files of data, and files of information. In the original IBM PC, the floppy disk drives represented the secondary storage memory subsystems. It had two

5'/,-inch drives that used double-sided, double-density floppy-diskette storage media that could each store up

to 360Kbytes (360,000 bytes) of data. This floppy diskette is an example of a removable media—that is, to

use the diskette it is inserted into the drive and locked in place. If either the diskette is full or one with a

different file or program is needed, the diskette is simply unlocked, removed, and another diskette installed.

The IBM PCXT also employed a second type of secondary storage device called a hard disk drive.

Modern hard disk drive sizes are SGbytes (5000 million bytes), 10Gbytes, 20Gbytes, 40Gbytes, 60Gbytes,

and 80Gbytes. The hard disk drive differs from the floppy disk drive in that the media is fixed, which means

that the media cannot be removed. However, being fixed is not a problem because the storage capacity of

the media is so much larger. Today, desktop PCs are equipped with a hard disk drive in the 20Gbyte to

80Gbyte range.

Both the floppy diskette and hard disk are examples of read/write media—that is, a file of data can

be read in from or written out to the storage media in the drive. Another secondary storage device that

is very popular in personal computers today is a CD drive. Here a removable compact disk (CD) is used

as the storage media. This media has very large storage capacity, more than 600Kbytes, but is read-only.

This means you cannot write information onto a CD for storage. For this reason, it is normally used to

store large programs or files of data that are not to be changed. Recently, a recordable CD (CD-R) has

become available, and CD-R drives allow the PC to both read from and write to this media.

4 The 8088 and 8086 Microprocessors

Primary storage memory is normally smaller in size and is used for the temporary storage of active information, such as the operating system of the microcomputer, the program that is currently being run, and the data that it is processing. In Fig. 1-1 we see that primary storage memory is further subdivided into program-storage memory and data-storage memory. The program section of memory is used to store instructions of the operating system and application programs. The data section normally contains data that are to be processed by the programs as they are executed (e.g., text files for a word-processor program or a database for a database-management program). However, programs can also be loaded into
data memory for execution.

Typically, primary storage memory is implemented with both read-only memory (ROM) and random-access read/write memory (RAM) integrated circuits. The original IBM PC had 48Kbytes of ROM and could be configured with 256Kbytes of RAM without adding a memory-expansion board. Modern PC/ATs made with the Pentium IV processors are typically equipped with 128Mbytes of RAM.

Data, whether they represent numbers, characters, or instructions of a program, can be stored in either ROM or RAM. In the original IBM PC, a small part of the operating system and BASIC language are resident to the computer because they are supplied in ROM. By using ROM, this information is made nonvolatile—that is, the information is not lost if power is turned off. This type of memory can only be read
from; it cannot be written into. On the other hand, data that are to be processed and information that frequently changes must be stored in a type of primary storage memory from which they can be read by the
microprocessor, modified through processing, and written back for storage. This requires a type of memory

that can be both read from and written into. For this reason, such data are stored in RAM instead of ROM.

Earlier we pointed out that the instructions of a program could also be stored in RAM. In fact, to run the PC operating system such as Windows 98®, it must be loaded into the RAM of the microcomputer. Normally the operating system, supplied on CDs, is first read from the CDs and written onto the hard disk. This is called copying of the operating system onto the hard disk. Once it is copied, the CD version of Windows 98 may not be used again. The PC is set up so that when it is turned on, Windows 98 is automatically read from the hard disk, written into the RAM, and then run.

RAM is an example of a volatile memory—that is, when power is turned off, the data that it holds are lost. This is why Windows 98 must be reloaded from the hard disk each time the PC is turned on.

## 1.2 EVOLUTION OF THE INTEL MICROPROCESSOR ARCHITECTURE

Generally, microprocessors and microcomputers are categorized in terms of the maximum number of binary bits in the data they process—that is, their word length. Over time, five standard data widths have evolved for microprocessors and microcomputers: 4-bit, 8-bit, 16-bit, 32-bit, and 64-bit.

Figure 1-3 illustrates the evolution of Intel's microprocessors since their introduction in 1972. The first microprocessor, the 4004, was designed to process data arranged as 4-bit words. This organization is also referred to as a nibble of data.

The 4004 implemented a very low performance microcomputer by today's standards. This low performance and limited system capability restricted its use to simpler, special-purpose applications. It was commonly used in electronic calculators.

Beginning in 1974, a second generation of microprocessors was introduced. These devices, the 8008, 8080, and 8085, were 8-bit microprocessors and were designed to process 8-bit (1-byte-wide) data

instead of 4-bit data. The 8080, identified in Fig. 1-3, was introduced in 1975.

These newer 8-bit microprocessors were characterized by higher-performance operation, larger system capabilities, and greater ease of programming. They were able to provide the system requirements for many applications that could not be satisfied with the earlier 4-bit microprocessors. These extended

Performance (MIPS) Pentium® processor

| Pentium® Pro processor

200

Pentium® processor

1971 1973 1975 1977 1979 1981 1983 1985 1987 1989

1991 1993 1995 1997 1999 Year

Figure 1-3 Evolution of the Intel microprocessor architecture.

capabilities led to widespread acceptance of 8-bit microcomputers for special-purpose system designs. Examples of these dedicated applications are electronic instruments, cash registers, and printers.

In the mid-1970s, many of the leading semiconductor manufacturers announced plans for development of third-generation 16-bit microprocessors. Looking at Fig. 1-3, we see that Intel's first 16-bit microprocessor, the 8086, became available in 1979 and was followed the next year by its 8-bit bus version, the 8088. This was the birth of Intel's 8086 family architecture. Other family members such as

the 80286, 80186, and 80188 were introduced in the years that followed.

These 16-bit microprocessors provided higher performance and had the ability to satisfy a broad scope of special-purpose and general-purpose microcomputer applications. They all have the ability to

handle 8-bit, 16-bit, and special-purpose data types, and their powerful instruction sets are more in line with those provided by a minicomputer.

In 1985, Intel Corporation introduced its first 32-bit microprocessor, the 80386DX, which brought true minicomputer-level performance to the microcomputer system. This device was followed by a 16-bit external bus version, the 80386SX, in 1988. Intel's second generation of 32-bit microprocessors, called the 80486DX and 80486SX, became available in 1989 and 1990, respectively. They were followed by a yet higher-performance family, the Pentium processors, in 1993. Today, its fourth generation member—the Pentium® IV processor, represents this family.

Microprocessor Performance: MIPS and iCOMP

Figure 1-3 illustrates the 8086 microprocessor families relative to their performance. Here performance is measured in what are called M/PS—that is, how many million instructions they can execute per second. Today, the number of MIPS provided by a microprocessor is the standard most frequently used to compare performance. Notice that performance has vastly increased with each new generation of microprocessor. For instance, the performance identified for the 80386 corresponds to an 80386DX device operating at 33 MHz and equals approximately 11 MIPS. With the introduction of the 80486, the level of performance capability of the architecture was raised to approximately 27 MIPS. This shows that performance of the 8086 architecture was more than doubled with the introduction of the 33-MHz 80486DX microprocessor.

The MIPS used in this chart are known as Drystone V1.1 MIPS—that is, they are measured by running a test program called the Drystone program, and the resulting performance measurements are normalized to those of a VAX 1.1 computer (VAX 1.1 was a minicomputer manufactured by Digital

Equipment Corporation). Therefore, we say that the 80486DX is capable of delivering up to 27 VAX MIPS of performance.

Intel Corporation provides another method, the iCOMP index, for comparison of the performance of its 32-bit microprocessors in a personal computer environment. In the iCOMP index chart shown in Fig. 1-4, a bar is used to represent a measure of the performance for each of Intel's MPUs. Instead of being related to the performance of a test program, such as the Dystone program, the iCOMP rating of an MPU is based on a variety of 16-bit and 32-bit MPU performance components important to the personal computer—that is, the iCOMP rating encompasses performance components that represent integer mathematics, floating-point mathematics, graphics, and video. The contribution by each of these categories is also weighted based on an estimate of their normal occurrence in widely used software applications, For this reason, iCOMP is a more broad-based rating of MPU performance for the personal computer applications.

The higher the iCOMP rating, the higher the performance offered by the MPU. Notice that the members of the 80386 family offer low performance when compared to the newer 80486 and Pentium processor families. In fact, the slowest 80386SX MPU shown in Fig. 1-4, the —20, has a performance rating of 32, whereas the fastest 80386DX, the -33, is rated at 68. Therefore, a wide range of system-performance levels can be achieved by selecting among the various members of the 80386, 80486, and Pentium processor families.

Transistor Density

The evolution of microprocessors was made possible by advances in semiconductor process technology. Semiconductor-device geometry decreased from about 5 microns in the early 1970s to

“The 80386, 80486, and Featum Processor Familie: Sofware Architecture os

mos ox-so.chU

ines0x2-s0cPU

ossea3 cu

osx cru

os seas ou

‘wos ox CPU

besos cru

ess. 25cm

esseascru

Figure 13-1 COMPO inex ring a Reprint by prison nt Corson Capea Cap 198)

execution unit the segment unit, the page unit the bus unit the prefetch unt, and the decode unit. Let tus now look more closely at each ofthe processing units of the 8O386DX.

"The bus units the $0386DX°s interface to the ouside World. By interface, we mean the path by data bus, 32-bit address bos,

-bit, an 32-bit data transfers

ate supported. These buses are demultiplexed like those ofthe 80286. That is, the 80386DX has separate pins for its address and data bus lines. This demuliplexing of address and data results in higher performance and easier hardware design. Expanding the data bus width to 32 bit further improves the performance of the 8O386DX’s hardware architecture as compared 10 that of either the 8086 or 80286,

‘The bus unt is responsible for performing all external bus operations. This processing unit contains the latches and drivers for the address bus, transceivers for the data bus, and control lie for sigoaling whether a memory, inpuvoutpt, or inerupt-acknowledge bus eycle is being performed. Looking a Fig. 13-2, we find that for data accesses, the address of the storage location tobe accessed

‘input fom the pang unit, nd for code access, the prefetch unit provides the adress.

8 The 8088 and 8086 Microprocessors

Transistors per device (1000s)

7

60!

Pentium’ Pro processor,

50

Pentium® processor

301

20

80486

V+ tt

1970 1972 1974 1976 1978 1980 1982 1984 1990 1992 1994 1996 1998 Year

Figure 1-5 Device complexity.

as the 8080 were most widely used as special-purpose microcomputers—a system that has been tailored to meet the needs of a specific application. These special-purpose microcomputers were used in embedded control applications—applications in which the microcomputer performs a dedicated control function.

Embedded control applications are further divided into those that involve primarily event control and those that require data control. An example of an embedded control application that is primarily event control is a microcomputer used for industrial process control. Here the program of the micro-

processor is used to initiate a timed sequence of events. On the other hand, an application that focuses

Reprogrammable

microprocessors

VUULILLLLLLLLL LULL LULU LULL LATE Mb ddd dd ddd dd hdd bbb dd ddd ddd dd ddd LLL MLL dd hdd

microprocessors

Figure 1-6 Processors for embedded control and reprogrammable applications.

more on data control than event control is a hard disk controller interface. In this case, a block of data that is to be processed—for example, a file of data—must be quickly transferred from secondary storage memory to primary storage memory.

The spectrum of embedded control applications requires a variety of system features and performance levels. Devices developed specifically for the needs of this marketplace have stressed low cost and high integration. In Fig. 1-6, we see that highly integrated 8-bit, single-chip microcomputer devices such as the 8048 and 8051 initially replaced the earlier multichip 8080 solutions, These devices were tailored to work best as event controllers. For instance, the 8051 offers one-order-of-magnitude-higher performance than the 8080, a more powerful instruction set, and special on-chip functions such as ROM, RAM, an interval/event timer, a universal asynchronous receiver/transmitter (UART), and programmable parallel input/output ports. Today, this type of embedded control device is called a microcontroller.

Later, devices such as the 80C186XL, 80C188XL, and 80386EX were designed to meet the needs of data-control applications. They are highly integrated and have additional features such as string instructions and direct-memory access channels, which handle the movement of data better. They are known as embedded microprocessors.

The category of reprogrammable microprocessors represents the class of applications in which a microprocessor is used to implement a general-purpose microcomputer. Unlike a special-purpose microcomputer, a general-purpose microcomputer is intended to run a variety of software applications—that is, while it is in use it can be easily reprogrammed to run a different application. Two examples of reprogrammable microcomputers are the personal computer and file server. Figure 1-6 shows that the

Figure 1-7 Code and system-level compatibility.

10

fouuo>

abesoys ssew

“Nd 1p Joy Lioddns jessydusg =g—]1 aunty

po |

LZ .

aoinap a2epaqut

jeseyduag 1ud

Jou

IS ’

jonuo>

jesayduad jOs,U0> uOD joju0>

Rined Ado pur} pu Lud

Jonuo>

HSv14 uoddns

Away

/WO¥d3/WOu sng

anweUukg

el

Jouu0>

SUONeIUNWUWOD

eeg

8086, 8088, 80286, 80386, 80486, and Pentium processor are the [Intel microprocessors intended for use

in this type of application.

Architectural compatibility is a critical need of microprocessors developed for use in reprogram-

mable applications. As shown in Fig. 1—7, each new member of the 8086/8088 family provides a superset

of the earlier device's architecture—that is, the features offered by the 80386 microprocessor are a

superset of the 80286 architecture, and those of the 80286 are a superset of the original 8086/8088

architecture.

Actually, the 80286, 80386, 80486, and Pentium processors can operate in either of two modes:

the real-address mode or protected-address mode. When in real mode, they operate like a high-

performance 8086/8088. They can execute what is called the base instruction set, which is object code

compatible with the 8086/8088. For this reason, operating systems and application programs written for

the 8086 and 8088 run on the 80286, 80386, 80486, and Pentium processor architectures without

modification. Further, a number of new instructions have been added in the instruction sets of the 80286,

80386, 80486, and Pentium processors to enhance their performance and functionality. We say that

object code is upward compatible within the 8086 architecture. This means that 8086/8088 code will run

on the 80286, 80386, 80486, and Pentium processors, but the reverse is not true if any of the new

instructions are in use.

Microprocessors designed for implementing general-purpose microcomputers must offer more

advanced system features than those of a microcontroller; for example, they need to support and manage

a large memory subsystem. The 80286 is capable of managing a 1 Gbyte (gigabytes) address space, and the 80386 supports 64Tbytes (64 terabytes) of memory. Moreover, a reprogrammable microcomputer, such as a personal computer, normally runs an operating system. The architectures of the 80286, 80386, 80486, and Pentium processors have been enhanced with on-chip support for operating system functions

such as memory management, protection, and multitasking. These new features become active only when the MPU is operated in the protected mode. The 80386, 80486, and Pentium processors also have a special mode of operation known as virtual 8086 mode that permits 8086/8088 code to be run in the protected mode.

Reprogrammable microcomputers, such as those based on the 8086 family, require a variety of input/output resources. Figure 1-8 shows the kinds of interfaces that are typically implemented in a personal computer or a microcomputer system. A large family of VLSI peripheral ICs (examples are floppy disk controllers, hard disk controllers, local area network controllers, and communication controllers) is needed to support a reprogrammable microprocessor such as the 8086, 80286, 80386, 80486, and Pentium processor. For this reason, these processors are designed to implement a multi-chip microcomputer system, which can easily be configured with the appropriate set of input/output interfaces.

REVIEW PROBLEMS

_ Section 1.1

1, What are the four building blocks of a microcomputer system?

2. What is the heart of the microcomputer system called?

3. Is the 8088 an 8-bit or a 16-bit microprocessor? Explain.

4, What is the primary input unit of the PC? Give two other examples of input units available for the PC.

5. What are the primary output devices of the PC?

6. Into what two sections is the memory of a PC partitioned?

7. What is the storage capacity of the standard 5'/,-inch floppy diskette of the original PC? What is the storage capacity of the standard hard disk drive of the original PCXT?

12 The 8088 and 8086 Microprocessors

8. What do ROM and RAM stand for?

9, How much ROM was provided in the original PC's processor board? What was the maximum amount of RAM that could be implemented on this processor board?

10. Why must Windows 98 be reloaded from the hard disk each time power is turned on?

Section 1.2

11. What are the standard data word lengths for which microprocessors have been developed?

12. What was the first 4-bit microprocessor introduced by Intel Corporation? Also name 8-bit,16-bit and 32-bit microprocessors introduced.

13. Name five 16-bit members of the 8086 family architecture.

14, What does MIPS stand for?

15. Approximately how many MIPS are delivered by the 33 MHz 80486DX?

16. What is the name of the program that is used to run the MIPS measurement test for the data in Fig. 1-3?

17. What is the iCOMP rating of an 80386SX-25 MPU? An 80386DX-25 MPU?

18. Approximately how many transistors are used to implement the 8088 microprocessor, the 80286 microprocessor, the 80386DX microprocessor, the 80486DX microprocessor and the Pentium processor?

19. What is an embedded microcontroller?

20. Name the two groups into which embedded processors are categorized based on applications.

21. What is the difference between a multichip microcomputer and a single-chip microcomputer?

22. Name six 8086 family microprocessors intended for use in reprogrammable microcomputer applications.

23. List the two modes of operation for 80386.

24. What is meant by upward software compatibility relative to 8086 architecture microprocessors?

25. List three advanced architectural features provided by the 80386DX microprocessor.

26. Give three types of VLSI peripheral support devices needed in a reprogrammable microcomputer system.

# Software Architecture of the 8088 and 8086 Microprocessors

## A INTRODUCTION

In this chapter we shall study in detail the 8088 and 8086 microprocessors and their assembly language programming. To program either the 8088 or 8086 using assembly language, we must understand how the microprocessor and its memory and input/output subsystems operate from a software point of view. For this reason, in this chapter, we will examine the software architecture of the 8088 and 8086 microprocessors. The description that follows frequently refers only to the 8088 microprocessor, but everything that is described for the 8088 also applies to the 8086. This is because the software architecture of the 8086 is identical to that of the 8088. The following topics are covered here.

2.1

2.2

2.3

2.4

2.5

2.6

zt

2.8

2.9

Microarchitecture of the 8088/8086 Microprocessor

Software Model of the 8088/8086 Microprocessor

Memory Address Space and Data Organization

Data Types

Segment Registers and Memory Segmentation

Dedicated, Reserved, and General-Use Memory

Instruction Pointer

Data Registers

Pointer and Index Registers

2.10 Status Register

14 The 8088 and 8086 Microprocessors

2.11 Generating a Memory Address

2.12 The Stack

2.13 Input/Output Address Space

## 2.1 MICROARCHITECTURE OF THE 8088/8086 MICROPROCESSOR

The microarchitecture of a processor is its internal architecture—that is, the circuit building blocks that implement the software and hardware architectures of the 8088/8086 microprocessors. Due to the requirement of additional features and higher performance, the microarchitecture of a microprocessor family evolves over time. In fact, a new microarchitecture is introduced for Intel's 8086 family every few years. Each new generation of processors (the 8088/8086, 80286, 80386, 80846, and Pentium processors) represents significant changes in the microarchitecture of the 8086.

The microarchitectures of the 8088 and 8086 microprocessors are similar. They both employ parallel processing—that is, they are implemented with several simultaneously operating processing units. Figure 2—1(a) illustrates the internal architecture of the 8088 and 8086 microprocessors. They contain two processing units: the bus interface unit (BIU) and the execution unit (EU). Each unit has dedicated functions and both operate at the same time. In essence, this parallel processing effectively makes the fetch and execution of instructions independent operations. This results in efficient use of the system bus and higher performance for 8088/8086 microcomputer systems.

The bus interface unit is the 8088/8086's connection to the outside world. By interface, we mean the path by which it connects to external devices. The BIU is responsible for performing all external bus operations, such as instruction fetching, reading and writing of data operands for memory, and inputting or outputting data for input/output peripherals. These information transfers take

place over the system bus. This bus includes an 8-bit bidirectional data bus for the 8088 (16 bits for the 8086), a 20-bit address bus, and the signals needed to control transfers over the bus. The BIU is not only responsible for performing bus operations, it also performs other functions related to instruction and data acquisition. For instance, it is responsible for instruction queuing and address generation.

To implement these functions, the BIU contains the segment registers, the instruction pointer, the address generation adder, bus control logic, and an instruction queue. Figure 2-1(b) shows the bus interface unit of the 8088/8086 in detail. The BIU uses a mechanism known as an instruction queue to implement a pipelined architecture. This queue permits the 8088 to prefetch up to 4 bytes (6 bytes for the 8086)
of instruction code. Whenever the queue is not full—that is, it has room for atleast 2 more bytes, and, at the same time, the execution unit is not asking it to read or write data from memory—the BIU is free to look ahead in the program by prefetching the next sequential instructions. Prefetched instructions are held
in the first-in first-out (FIFO) queue. Whenever a byte is loaded at the input end of the queue, it is automatically shifted up through the FIFO to the empty location nearest the output. Here the code is held until
the execution unit is ready to accept it. Since instructions are normally waiting in the queue, the time needed to fetch many instructions of the microcomputer's program is eliminated. If the queue is full and the EU is not requesting access to data in memory, the BIU does not need to perform any bus operations.
These intervals of no bus activity, which occur between bus operations are known as idle states.

The execution unit is responsible for decoding and executing instructions. Notice in Fig. 2—1(b)
that it consists of the arithmetic logic unit (ALU), status and control flags, general-purpose registers, and temporary-operand registers. The EU accesses instructions from the output end of the instruction queue and data from the general-purpose registers or memory. It reads one instruction byte after the other from
the output of the queue, decodes them, generates data addresses if necessary, passes them to the BIU and

requests it to perform the read or write operations to memory or I/O, and performs the operation

Software Architecture of the 8088 and 8086 Microprocessors 5

BUS

INSTRUCTION | |NTERFACE

PIPELINE UNIT

SYSTEM BUS

(a)

EXECUTION UNIT BUS INTERFACE UNIT

(EU) (BIU)

— 16 ———> —it—>

MULTIPLEXED BUS

ADDRESS

GENERATION

AND BUS

ONTROL

INSTRUCTION

QUEUE

6 BYTES (8086)

4 BYTES (8088)

(b)

Figure 2-1 (a) Pipelined architecture of the 8088/8086 microprocessors. (Reprinted with permission of Intel Corporation, Copyright/Intel

Corp. 1981) (b) Execution and bus interface units. (Reprinted with permission of Intel Corp., Copyright/Intel Corp. 1981)

specified by the instruction. The ALU performs the arithmetic, logic, and shift operations required by an instruction. During execution of the instruction, the EU may test the status and control flags, and updates these flags based on the results of executing the instruction. If the queue is empty, the EU waits for the next instruction byte to be fetched and shifted to the top of the queue.

## 2.2 SOFTWARE MODEL OF THE 8088/8086 MICROPROCESSOR

The purpose of developing a software model is to aid the programmer in understanding the operation of the microcomputer system from a software point of view. To be able to program a microprocessor, one does not need to know all of its hardware architectural features. For instance, we do not necessarily need to know the function of the signals at its various pins, their electrical connections, or their electrical switching characteristics. Also the function, interconnection, and operation of the internal circuits of the microprocessor need not be considered. What is important to the programmer is to know the various registers within the device and to understand their purpose, functions, operating capabilities, and limitations. Furthermore, it is essential that the programmer know show external memory and input/output peripherals are organized, how information is arranged in registers, memory, and input/output, and how memory and I/O are addressed to obtain instructions and data. This information represents the software architecture of the processor. Unlike the microarchitecture, the software architecture changes only slightly from generation to generation of processor.

The software model in Fig. 2-2 illustrates the software architecture of the 8088 microprocessor.

Looking at this diagram, we see that it includes 13 16-bit internal registers: the instruction pointer (IP), four

data registers (AX, BX, CX, and DX), wo pointer registers (BP and SP), two index registers (SI and D1),

and four segment registers (CS, DS, SS, and ES), In addition, there is another register called the status reg-

ister (SR), with nine of its bits implemented as status and control flags.

Figure 2-2 shows that the 8088 architecture implements independent memory and input/output

address spaces. Notice that the memory address space is 1,048,576 bytes (1 Mbyte) in length and the I/O

address space is 65,536 bytes (64 Kbytes) in length. Our concern here is what can be done with this software

architecture and how to do it through software. For this purpose, we will now begin a detailed study of the

elements of the model and their relationship to software.

## 2.3 MEMORY ADDRESS SPACE AND DATA ORGANIZATION

Now that we have introduced the idea of a software model, let us look at how information such as numbers,

characters, and instructions are stored in memory. As shown in Fig. 2—3, the 8088 microcomputer supports

1 Mbyte of external memory. This memory space is organized from a software point of view as individual

bytes of data stored at consecutive addresses over the address range 00000,,, to FFFFF ,,. Therefore, memory

in an 8088-based microcomputer is actually organized as 8-bit bytes, not as 16-bit words.

The 8088 can access any two consecutive bytes as a word of data. In this case, the lower-addressed

byte is the least significant byte of the word, and the higher-addressed byte is its most significant byte. Figure

2-4(a) shows how a word of data is stored in memory. Notice that the storage location at the lower address,

00724 ,,, contains the value 00000010, = 02,,.The contents of the next-higher-addressed storage location, 00725 ,,, are 01010101, = 55,,. These two bytes represent the word 0101010100000010, = 5502,,.

To permit efficient use of memory, words of data can be stored at what are called even- or odd-addressed word boundaries. The least significant bit of the address determines the type of word boundary. If this bit is 0, the word is at an even-address boundary—that is, a word at an even-address boundary corresponds to two consecutive bytes, with the least significant byte located at an even

Memory space

External memory

address space

Code segment

(64 Kbytes)

000046

Data segment

(64 Kbytes)

input/output

address space

Stack segment

(64 Kbytes)

Extra segment

(64 Kbytes)

FFFFFi6

Figure 2-2 Software model of the 8088/8086 microprocessor.

|__Frrer |

|__FFFFE |

FFFFD

FFFFC

Figure 2-3 Memory address space of the 8088/8086 microprocessor.

address. For example, the word in Fig. 2—4(a) has its least significant byte at address 00724,,. Therefore, it is stored at an even-address boundary.

A word of data stored at an even-address boundary, such as 00000, , 00002,,, 00004, ,, and so on, is said to be an aligned word—that is, all aligned words are located at an address that is a multiple of 2. On the other hand, a word of data stored at an odd-address boundary, such as 00001 ,,, 00003,,, or 00005,, and so on, is called a misaligned word. Figure 2-5 shows some aligned and misaligned words of data. Here words 0, 2, 4, and 6 are examples of aligned-data words, while words | and 5 are misaligned words. Notice that misaligned word | consists of byte 1 from aligned word 0 and byte 2 from aligned word 2.

When expressing addresses and data in hexadecimal form, it is common to use the letter H to specify the base. For instance, the number 00AB;, can also be written as OOABH.

EXAMPLE 2.1

What is the data word shown in Fig. 2—4(b)? Express the result in hexadecimal form. Is it stored at an even- or odd-addressed word boundary? Is it an aligned or misaligned word of data?

Solution

The most significant byte of the word is stored at address 0072C,, and equals

11111101, = FD,,= FDH

Its least significant byte is stored at address 0072B ,, and is

10101010, = AAy, = AAH

Together the two bytes give the word

1111110110101010, = FDAA,, = FDAAH

Expressing the address of the least significant byte in binary form gives

0072BH = 0072B,, — 00000000011100101011,

Because the rightmost bit (LSB) is logic 1, the word is stored at an odd-address boundary in memory; therefore, it is a misaligned word of data.

The double word is another data form that can be processed by the 8088 microcomputer. A double word corresponds to four consecutive bytes of data stored in memory; an example of double-word data is a pointer. A pointer is a two-word address element that is used to access data or code in memory. The

Address | Memory Memory Address | Memory

(binary) (hexadecimal) (binary)

O0r2816 Ps § | — 0072Cy9) 11111101 _|

(a) (b)

Figure 2-4 (a) Storing a word of data in memory, (b) An example.

address. For example, the word in Fig. 2—4(a) has its least significant byte at address 00724,,.

Therefore, it is stored at an even-address boundary.

A word of data stored at an even-address boundary, such as 00000, , 00002,,, 00004, ,, and so on, is

said to be an aligned word—that is, all aligned words are located at an address that is a multiple of 2. On the

other hand, a word of data stored at an odd-address boundary, such as 00001 ,,, 00003,,, or 00005,, and so

on, is called a misaligned word. Figure 2-5 shows some aligned and misaligned words of data. Here words

0, 2, 4, and 6 are examples of aligned-data words, while words | and 5 are misaligned words. Notice that

misaligned word | consists of byte 1 from aligned word 0 and byte 2 from aligned word 2.

When expressing addresses and data in hexadecimal form, it is common to use the letter H to

specify the base. For instance, the number 00AB;, can also be written as OOABH.

EXAMPLE 2.1

What is the data word shown in Fig. 2—4(b)? Express the result in hexadecimal form. Is it stored at an

even- or odd-addressed word boundary? Is it an aligned or misaligned word of data?

Solution

The most significant byte of the word is stored at address 0072C,, and equals

11111101, = FD,,= FDH

Its least significant byte is stored at address 0072B ,, and is

10101010, = AAy, = AAH

Together the two bytes give the word

1111110110101010, = FDAA,, = FDAAH

Expressing the address of the least significant byte in binary form gives

0072BH = 0072B,, — 00000000011100101011,

Because the rightmost bit (LSB) is logic 1, the word is stored at an odd-address boundary in memory; therefore, it is a misaligned word of data.

The double word is another data form that can be processed by the 8088 microcomputer. A double word corresponds to four consecutive bytes of data stored in memory; an example of double-word data is a pointer. A pointer is a two-word address element that is used to access data or code in memory. The

Address | Memory Memory Address | Memory

(binary) (hexadecimal) (binary)

O0r2816 Ps § | — 0072Cy9) 11111101 _|

(a) (b)

Figure 2-4 (a) Storing a word of data in memory, (b) An example.

Figure 2-5 Examples of aligned and misaligned data words.

word of this pointer that is stored at the higher address is called the segment base address, and the word

at the lower address is called the offset.

Just like for words, a double word of data can be aligned or misaligned. An aligned double word is located at an address that is a multiple of 4 (e.g., 00000,,, 00004, ,, and 00008,,). A number of aligned and misaligned double words of data are shown in Fig. 2-6. Of these six examples, only double words

0 and 4 are aligned double words.

Physical

Address memory Aligned

double

00008H words |

00007H Double

word

o0006H Double 5 |

word

0000SH 4 | uk

word

00004H Double 3

Double moe

00003H

00002H

oooo1H | eytet |

00000H

double words

Figure 2-6 Examples of aligned and misaligned double words of data.

An example showing the storage of a pointer in memory is given in Fig. 2~7(a). Here the higher-addressed word, which represents the segment base address, is stored starting at even-address boundary

00006 ,,,. The most significant byte of this word is at address 00007), and equals 00111011, = 3B,¢. Its least significant byte is at address 00006,, and equals 01001100, = 4C,,. Combining these two values, we get the segment base address, which equals 0011101101001100, = 3B4C,«.

The offset part of the pointer is the lower-addressed word. Its least significant byte is stored at address 00004,,; this location contains 01100101, = 65,4. The most significant byte is at address 00005,,, which contains 00000000, = 00,,. The resulting offset is 0000000001100101, = 0065,,. The complete double word is 3B4C\ 0065 4. Since this double word starts at address 00004, . it is an example of an aligned double word of data.

Address Memory Memory Address | Memory
(binary) (hexadecimal) (hexadecimal)

3B | 000Big} AO
| 4 c | c000Ayg] 00 —C*
f 0 0 | oo0aig|_ 5
r& 5 | —gooo8yg |__F |

(a) (b)

Figure 2-7 (a) Storing a 32-bit pointer in memory. (b) An example.

EXAMPLE 2.2

How should the pointer with segment base address equal to A000,, and offset address 55FF,, be stored at an even-address boundary starting at 00008 ,,,? Is the double word aligned or misaligned?

Solution

Storage of the two-word pointer requires four consecutive byte locations in memory, starting at address 00008 ,,. The least-significant byte of the offset is stored at address 00008,,, and is shown as FF,,, in Fig. 2-7(b). The most significant byte of the offset, 55), is stored at address 00009, ,. These two bytes are followed by the least significant byte of the segment base address, 00,,, at address 0000A,,, and its most significant byte, AO,,, at address 0000B,,,. Since the double word is stored in memory starting at address 00008, ,, it is aligned.

## 2.4 DATA TYPES

The preceding section identified the fundamental data formats of the 8088 as the byte (8 bits), word (16 bits),

and double word (32 bits). It also showed how each of these elements is stored in memory. The next step is

to examine the types of data that can be coded into these formats for processing.

The 8088 microprocessor directly processes data expressed in a number of different data types. Let us

begin with the integer data type. The 8088 can process data as either unsigned or signed integer numbers;

cach type of integer can be either byte-wide or word-wide. Figure 2—8(a) represents an unsigned byte integer;

this data type can be used to represent decimal numbers in the range 0 through 255. The unsigned word inte-

ger is shown in Fig. 2—8(b); it can be used to represent decimal numbers in the range 0 through 65,535.

MSB LSB

(a)

MSB

LSB

Pst |] Tt ttt tT tT tt tT JP)

(b)

Figure 2-8 (a) Unsigned byte integer. (b) Unsigned word integer.

EXAMPLE 2.3

What value does the unsigned word integer 1000), represent?

Solution

First, the hexadecimal integer is converted to binary form:

1000, , = 0001000000000000,

Next, we find the value for the binary number: :

000 1000000000000, = 2!2 = 4096

The signed byte integer and signed word integer in Figs. 2~9(a) and (b) are similar to the unsigned

integer data types just introduced; however, here the most significant bit is a sign bit. A zero in this bit position

identifies a positive number. For this reason, the signed integer byte can represent decimal numbers in the range

+127 to—128, and the signed integer word permits numbers in the range +32,767 to —32,768, respectively. For

example, the number +3 expressed as a signed integer byte is 00000011, (03,,). On the other hand, the 8088

always expresses negative numbers in 2’s-complement notation. Therefore, —3 is coded as 11111101, (FD,,).

Sign bit

(a)

MSB

LSB

Pa TE PE Eh ELT ee

Sign bit

(b)

Figure 2-9 (a) Signed byte integer. (b) Signed word integer.

EXAMPLE 2.4

A signed word integer equals FEFF;,. What decimal number does it represent?

Solution

Expressing the hexadecimal number in binary form gives

FEFF), = 1111111011111111,

Since the most significant bit is 1, the number is negative and is in 2's complement form. Converting to its binary equivalent by subtracting 1 from the least significant bit and then complementing all bits gives

FEFF , , = —0000000100000001,

=-257

The 8088 can also process data that is coded as binary-coded decimal (BCD) numbers. Figure 2—10(a) lists the BCD values for decimal numbers 0 through 9. BCD data can be stored in either unpacked or packed form. For instance, the unpacked BCD byte in Fig. 2-10(b) shows that a single BCD digit is stored in the four least significant bits, and the upper four bits are set to 0. Figure 2—-10(c) shows a byte with packed BCD digits. Here two BCD numbers are stored in a byte. The upper four bits represent the most significant digit of a two-digit BCD number.

0

1

2

3

4

5

6

7

8

9

(a)

MSB LSB

EL TY fot | foo |

ecb Digit

(b)

MSB LSB

BCD Digit 1 BCD Digit 0

(c)

Figure 2-10 (a) BCD numbers. (b) An Unpacked BCD digit. (c) Packed BCD digits.

EXAMPLE 2.5

The packed BCD data stored at byte address 01000,, equal 10010001,. What is the two-digit decimal number?

Solution

Writing the value 10010001, as separate BCD digits gives

10010001, = 1001 5¢)0001 gop = 91 9

Information expressed in ASCII (American Standard Code for Information Interchange) can also be directly processed by the 8088 microprocessor. The chart in Fig. 2-11 (a) shows how numbers, letters, and control characters are coded in ASCII. For instance, the number 5 is coded as

H,Hy = 0110101, = 35H

where H denotes that the ASCII-coded number is in hexadecimal form. As shown in Fig. 2—11(b), ASCII data are stored as one character per byte.

by 1

bg 0

bs 1

ee Pelee

o

mio =

FREER

o

m

rc

<|5|3 8)

PLE

rile

=F isis

Vv

peetrs | [<fe tri t i |

Por} o fortes t= pw pita oy

jt itoleljsoirs| [>in iri"]- |

east} e Le poet de jo ff 6 pom

(a)

Figure 2-11 (a) ASCII table.

}o 001)

foot

foot

fo 100)

jo 1014)

rer

arte

000)

root,

Pore)

Prom

Tr +o0]

aiel=l Ls bepetel*lel+|=17%

°o

4

BEEEEEBEE

24 The 8088 and 8086 Microprocessors

MSB LSB

or} TT Tt YT IPS)

——_$_———]———

ASCIl Digit

(b)

Figure 2-11 (b) ASCII digit.

EXAMPLE 2.6

Byte addresses 01100,, through 01104,, contain the ASCII data 01000001, 01010011, 01000011, 01001001, and 01001001, respectively. What do the data stand for?

Solution

Using the chart in Fig. 2—11(a), the data are converted to ASCII as follows:

(01100H) = 01000001 , scr =A

(O1101H) = 01010011 ,scq = S

(01102H) = 01000011 ,scy = C

(01103H) = 01001001 ascy = 1

(01104H) = 01001001 seq = 1

A 2.5 SEGMENT REGISTERS AND MEMORY SEGMENTATION

Even though the 8088 has a | Mbyte address space, not all this memory is active at one time. Actually, the | Mbytes of memory are partitioned into 64 Kbyte (65,536) segments. A segment represents an independently addressable unit of memory consisting of 64 K consecutive byte-wide storage locations. Each segment is assigned a base address that identifies its starting point—that is, its lowest address byte-
storage location.

Only four of these 64 Kbyte segments are active at a time: the code segment, stack segment, data segment, and extra segment. The segments of memory that are active, as shown in Fig. 2-12, are identified by the values of addresses held in the 8088's four internal segment registers: CS (code segment), SS (stack segment), DS (data segment), and ES (extra segment). Each of these registers contains a 16-bit base address that points to the lowest addressed byte of the segment in memory. Four segments give a maximum of 256 Kbytes of active memory. Of this, 64 Kbytes are for program storage (code), 64 Kbytes are for a stack, and 128 Kbytes are for data storage.

The values held in these registers are referred to as the current-segment register values; for example, the value in CS points to the first word-wide storage location in the current code segment. Code is always fetched from memory as words, not as bytes.

FFFFFH

00000H

Figure 2-12 Active segments of memory.

Figure 2-13 illustrates the segmentation of memory. In this diagram, the 64 Kbyte segments are identified with letters such as A, B, and C. The data segment (DS) register contains the value B. Therefore, the second 64 Kbyte segment of memory from the top, labeled B, acts as the current data storage segment This is one of the segments in which data that are to be processed by the microcomputer are stored. For this reason, this part of the microcomputer's memory address space must contain read/write storage locations that can be accessed by instructions as storage locations for source and destination operands. CS selects segment E as the code segment. It is this segment of memory from which instructions of the program are currently being fetched for execution. The stack segment (SS) register contains H, thereby selecting the 64 Kbyte segment labeled as H for use as a stack. Finally, the extra segment (ES) register is loaded with value J such that segment J of memory functions as a second 64 Kbyte data storage segment.

The segment registers are said to be user accessible. This means that the programmer can change their contents through software. Therefore, for a program to gain access to another part of memory, one simply
has to change the value of the appropriate register or registers. For instance, a new data space, with up to 128 Kbytes, is brought in simply by changing the values in DS and ES.

There is one restriction on the value assigned to a segment as a base address: it must reside on a 16-byte address boundary. This is because increasing the 16-bit value in a segment register by

FFFFFH 7

ewes CE]

|

STACK: SS: wenn

emacs [a }, ||

OH

Figure 2-13 Contiguous, adjacent, disjointed, and overlapping segments. (Reprinted by permission of Intel Corp., Copyright/Intel

Corp. 1979)

1 actually increases the corresponding memory address by 16; examples of valid base addresses are 00000,,, 00010,,, and 00020,,. Other than this restriction, segments can be set up to be contiguous, adjacent, disjointed, or even overlapping; for example, in Fig. 2-13, segments A and B are contiguous, whereas segments B and C are overlapping.

## 2.6 DEDICATED, RESERVED, AND GENERAL-USE MEMORY

Any part of the 8088 microcomputer's 1 Mbyte address space can be implemented for the user's access; however, some address locations have dedicated functions and should not be used as general memory for storage of data or instructions of a program. Let us now look at these reserved, dedicated use, and general-use parts of memory.

Figure 2-14 shows the reserved, dedicated-use, and general-use parts of the 8088/8086's address space. Notice that storage locations from address 00000), to 00013,, are dedicated, and those from address 00014), to 0007F |, are reserved. These 128 bytes of memory are used for storage of pointers to interrupt service routines. The dedicated part is used to store the pointers for the 8088's internal interrupts and exceptions. On the other hand, the reserved locations are saved to store pointers that are used by the user-defined interrupts. As indicated earlier, a pointer is a two-word address element and requires 4 bytes of memory. The word of this pointer at the higher address is called the segment base address and the word at the lower address is the offset. Therefore, this section of memory contains up to 32 pointers.

The part of the address space labeled open in Fig. 2-14 is general-use memory and is where data or instructions of the program are stored. Notice that the general- use area of memory is the range from addresses 80,, through FFFEF |«.

FFFFFH

[reserve | FFFFCH

FFFFBH

FFFEFH

OPEN

80H

7FH

DEDICATED

0H

Figure 2-14 Dedicated-use,reserved, and general-use memory. (Reprinted by permission of Intel Corp., Copyright/Intel Corp. 1979)

At the high end of the memory address space is another reserved pointer area, located from address FFFFC,, through FFFFF},. These four memory locations are reserved for use with future products and should not be used. Intel Corporation, the original manufacturer of the 8088, has identified the 12 storage

locations from address FFFFO,, through FFFFB,, as dedicated for functions such as storage of the hard-ware reset jump instruction. For instance, address FFFF0,, is where the 8088/8086 begins execution after receiving a reset.

## 2.7 INSTRUCTION POINTER

The register that we will consider next in the 8088's software model shown in Fig. 2-2 is the instruction pointer (IP). IP is also 16 bits in length and identifies the location of the next word of instruction code to be fetched from the current code segment of memory. The IP is similar to a program counter; however, it contains the offset of the next word of instruction code instead of its actual address. This is because IP and CS are both 16 bits in length, but a 20-bit address is needed to access memory. Internal to the 8088, the offset in IP is combined with the current value in CS to generate the address of the instruction code. Therefore, the value of the address for the next code access is often denoted as CS:IP.

During normal operation, the 8088 fetches instructions from the code segment of memory, stores them in its instruction queue, and executes them one after the other. Every time a word of code is fetched from memory, the 8088 updates the value in IP such that it points to the first byte of the next sequential word of code—that is, IP is incremented by 2. Actually, the 8088 prefetches up to four bytes of instruction code into its internal code queue and holds them there waiting for execution,

After an instruction is read from the output of the instruction queue, it is decoded; if necessary, operands are read from either the data segment of memory or internal registers. Next, the operation specified in the instruction is performed on the operands and the result is written back either in an internal register or a storage location in memory. The 8088 is now ready to execute the next instruction in the code queue.

Executing an instruction that loads a new value into the CS register changes the active code segment; thus, any 64 Kbyte segment of memory can be used to store the instruction code.

A 2.8 DATA REGISTERS

As Fig. 2-2 shows, the 8088 has four general-purpose data registers. During program execution, they hold temporary values of frequently used intermediate results. Software can read, load, or modify their contents. Any of the general-purpose data registers can be used as the source or destination of an operand during an arithmetic operation such as ADD or a logic operation such as AND. For instance, the values of two pieces of data, A and B, could be moved from memory into separate data registers and operations such as addition, subtraction, and multiplication performed on them. The advantage of storing these data in internal registers instead of memory during processing is that they can be accessed much faster.

The four registers, known as the data registers, are shown in detail in Fig. 2~15(a). Notice that they are referred to as the accumulator register (A), the base register (B), the count register (C), and the data register (D). These names imply special functions they are meant to perform for the 8088 microprocessor. Figure 2—15(b) summarizes these operations. Notice that string and loop operations

Byte multiply, byte divide, byte
\/O, translate, decimal arithmetic

Byte multiply, byte divide
Translate

String operations, loops

Variable shift and rotate

Word multiply, word divide,

indirect VO

(b)

Figure 2-15 (a) General-purpose data registers. (Reprinted by permission of Intel Corp., Copyright/Intel Corp. 1979)

(b) Dedicated register functions. (Reprinted by permission of Intel Corp., Copyright/Intel Corp. 1979)

use the C register. For example, the value in the C register is the number of bytes to be processed in a string operation. This is the reason it is given the name count register. Another example of the dedicated use of data registers is that all input/output operations must use accumulator register AL or AX for data.

Each of these registers can be accessed either as a whole (16 bits) for word data operations or as two 8-bit registers for byte-wide daa operations. An X after the register letter identifies the reference of a register as a word; for instance the 16-bit accumulator is referenced as AX. Similarly, the other three word registers are referred to as BX, CX, and DX.

On the other hand, when referencing one of these registers on a byte-wide basis, following the register name with the letter H or L, respectively, identifies the high byte and low byte. For the A register, the most significant byte is referred to as AH and the least significant byte as AL; the other byte-wide register pairs are BH and BL, CH and CL, and DH and DL. When software places a new value in one byte of a register, for instance AL, the value in the other byte (AH) does not change. This ability to