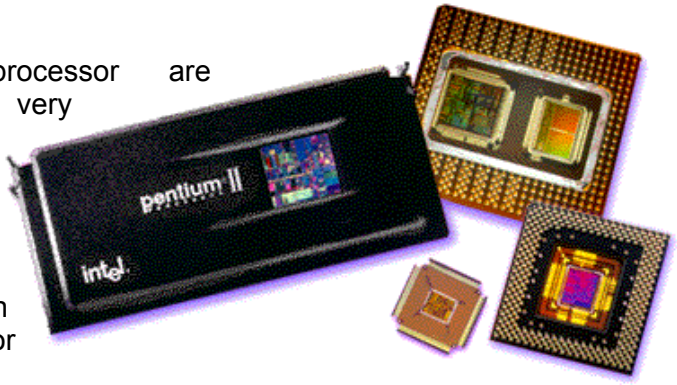
Gross Anatomy Guide

The Microprocessor

Introduction:

What is a Microprocessor?

The terms microprocessor and processor are interchangeable. Microprocessors are very complex electronic circuits. They consist of thousands or even millions of microscopic transistors packed into an extremely small chip of silicon no larger than a quarter inch square. This chip is then placed into a package like those shown above making the microprocessor easier for technicians handle. There're typically hundreds of hair-thin wires that connect microprocessor chip to the leads and pins on the package. The leads and pins are needed to connect the microprocessor to the rest of the computer's circuitry.



Microprocessors come in a variety of package types and pin configurations. *Courtesy of Intel Corporation*

The millions of transistors in the microprocessor make up circuits that determine how the device handles digital information. Sub-circuits like registers, counters, and encoders inside the microprocessor allow it make decisions and direct data according to an external program. It is theses circuits, working together to give the microprocessor the power to control all aspects of PC operation. In essence, the microprocessor is the brain of the PC.

First... A little history lesson**-** The development of the microprocessor (also known as a CPU or Central Processing Unit) is ultimately linked to the evolution of electronics technology starting early in the 1940s.

During World War II, the earliest computers were designed exclusively for military use. Military computers performed the calculations needed to direct ordinance. Early computers were not digital, but analog and produced tables that soldiers could use on the battle field and at sea to more precisely aim munitions at targets. Up to this point, trajectory calculations were performed with pencil and paper and were difficult to adjust for variations in temperature, wind speed, and direction. At sea, ships needed to have their guns trained on the horizon at all times… Difficult to do considering the constant heaving of ocean swells. Analog computers were well suited for controlling the big guns on a battleship because the ocean, by its very nature, is analogous. Keep in mind that the closest thing to a calculator in those days was a slide rule.

Going into the 1950s, vacuum tubes were the devices computers used to make decisions and perform calculations. Electron tubes were big and bulky. They required lots of current to operate and they got hot. Since each tube performed a single calculation for each element of information, hundreds were required to perform even the simplest calculations. Therefore, vacuum tube computers were the size of buildings and required air conditioning to operate, and by today's standards, were expensive and not very reliable.

In 1948, scientists at Bell Laboratories invented the transistor, the single most significant achievement to impact the world of computer technology. In the late 1950s, this invention ushered in the era of pulse circuit technology and solid state devices that made the digital

computer practical. With the advent of the transistor, computers that were once the size of a building could now be integrated into a single room or even a desk.

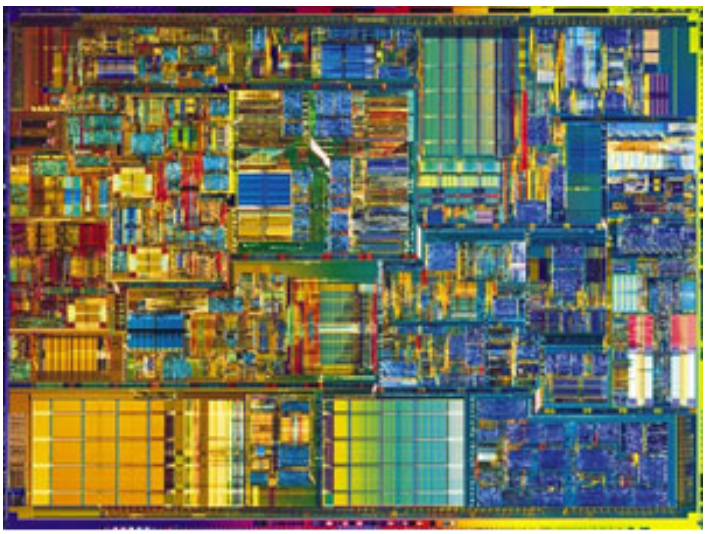
Throughout the 1960s, computers once again grew larger as they pursued more processing power and storage capacity. Big computers were still very expensive and had to operate 24 hours a day just to pay for themselves. These computers boosted processing speeds of 3000 calculations a second and memory capacities of 64 Kilobytes. Wow! Once again, computers were the size of buildings and required air conditioning to keep them cool.

Solid state technologies continued to keep pace with advances in digital computer technology. It was clear to industry that the key to processing power was based upon the number of transistors that could be integrated into the computer's circuitry. One transistor… One switch… One decision…

In the mid-1960s the trends in computer technology moved from, "Bigger is Better" to "Integration is Essential." Since Computers are really collections of sub-circuits that perform basic operations, manufacturers developed a technology of combining transistors into tiny circuits that function as a computer's:

Register	Temporarily holds data or instructions while calculations and logic operations are performed.
ALU	The ALU (Arithmetic Logic Unit) performs addition, subtraction, and logic operations on data based upon the program instructions.
Accumulator	Another type of register that holds the result of the ALU operation.
Counter	Counts the number of cycles needed to complete an operation.
Encoder	Converts information into a digital form that computers can use.
Memory	Where programs and computation results are stored.
Buffer	Controls the traffic of data and program instructions as they move from registers… to the ALU… to accumulators… to memory and back.

The process is know as integration and allowed many transistors, resistors, and capacitors to be squeezed onto a single silicon chip. The integrated circuit or IC was born. Small Scale Integration (SSI) allowed designers to integrate entire circuit boards into a single plug-in module the size of a postage stamp. Eventually, manufacturers discovered better ways of packing more and more transistors onto silicon chips using photolithography technology.



Intel Pentium 4 Die 42 million of transistors etched onto a silicon wafer no larger than a dime form this microprocessor chip. Courtesy of Intel Corporation.

In the 1970s, Large Scale Integration (LSI) combined common electronic functions into one IC. These chips were designed to be universal and led to the birth of the first microprocessors. What was once a building sized computer could now be housed on a single silicon wafer lo larger than ¼ square inch.

In 1971, Intel Corporation introduced the very first microprocessor. Although the Intel 4004 wasn't very powerful, its production was pretty significant. It could only add, subtract and handle only 4 bits at a time. What was significant was the fact that all the functions that were previously handled by SSI integrated circuits and transistors were now consolidated on one

silicon chip. The 4004 was used in the first portable electronic calculators. 4 bits, sometimes called a nibble, is the word length of this microprocessor.

The first microprocessor built into a personal computer was the Intel 8080. The 8080 hit the market in 1974 featuring an 8-bit (1 Byte) word data and address bus. As technology demands have increased, so did the processing power or microprocessors.

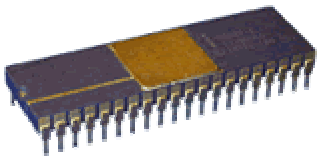
Microprocessor Evolution Timeline		
Year	**Model**	**Features**
1971	4004	2,300 transistors, 4-bit, 108KHz (60,000 operations per second)
1972	8088	3,300 transistors, 8 bit (First wide use in home computers)
1972	4040	4-bit (Like the 4004 but with larger instruction set)
1974	8080	4,500 transistors, 8-bit, 200,000 operations per second
1976	8085	8-bit/16-bit, 5 MHz
1978	8088/8086	8-bit/16-bit (the 8088 was used in the first IBM PC)
1982	286	134,000 transistors, 16 bit
1985	386	275,000 transistors, 32-bit (multitasking)
1989	486	1.2 million transistors, 32-bit (first math coprocessor), 25MHz system clock (up to 66MHz over time).
1993	Pentium	3.1 million transistors, 32-bit, 16KB L1, 512KB L2 (external), 75MHz system clock (up to 233MHz), 66MHz host bus, Socket 7.
1995	Pentium Pro	5.5 million transistors, 32-bit, 16KB L1, 256KB L2 (internal), 150MHz system clock (up to 200MHz), 66MHz host bus L2 Cache bus runs at ½ main system clock.
1997	Pentium II	7.5 million transistors, 32-bit, 32KB L1, 512KB L2, MMX, 233MHz system clock (up to 500MHz), 66MHz host bus (later 100MHz) L2 Cache bus runs at ½ main system clock.
1998	Celeron	7.5 million transistors, 32-bit, 32KB L1, 0K (later 128KB L2), MMX, 266MHz system clock (up to 400MHz), 66MHz host bus L2 Cache (when provided) bus runs at ½ main system clock.
1999	Pentium II Xeon	7.5 million transistors, 32-bit, 32KB L1, 512KB / 1MB / 2MB L2, MMX, 400MHz system clock (up to 500MHz), 100MHz host bus L2 Cache bus runs at FULL main system clock.
1999	Pentium III	7.5 million transistors, 32-bit, 32KB L1, 512KB / 1MB / 2MB L2, MMX, 400MHz system clock (up to 500MHz), 100MHz host bus L2 Cache bus runs at ½ main system clock.
1999	Pentium III Xeon	9.5 million transistors, 32-bit, 32KB L1, 512KB / 1MB / 2MB L2, MMX II, 500MHz system clock (up to 550MHz), 100MHz host bus L2 Cache bus runs at FULL main system clock.
2000	Pentium IV	42 Million Transistors, 32-bit, 8KB L1, 512KB / 256KB L2, 800MHz system clock (up to 3.2GHz), 533MHz host bus (up to 2.8GHz), L2 Cache bus runs at FULL main system clock.

Today… Very Large Scale Integration (VLSI) and Ultra Large Scale Integration (ULSI) technologies continue the trend integrating millions of transistors onto smaller and smaller chips. Computers and other consumer electronics have became smaller, more powerful, more energy efficient, and affordable enough for personal use.

Microprocessor Package Types

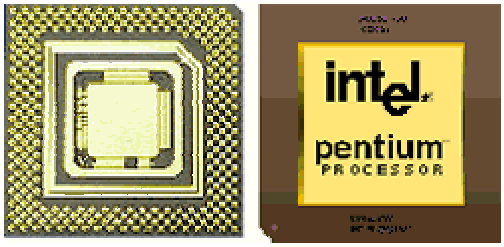
Package types are standardized designs for electronic components. The package type describes the number of pins and where the pins are located. It also describes the size and shape of the device. This is critical when you're trying to match a motherboard or heatsink to a processor. If they don't match, they won't work together.

Microprocessor Package Types



Intel 8080 Microprocessor, 40 Pin DIP

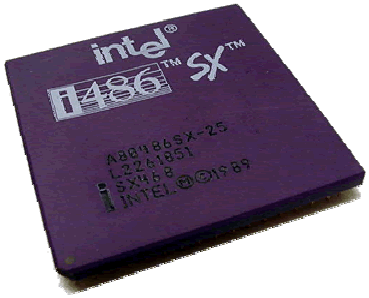
DIP Package Type: Early LSI microprocessors were packaged in 40 pin DIPs (Dual Inline Package). This device was then inserted into a socket on a printed circuit board with connections made to the circuitry through the microprocessor's pins. This type of package was used for the Intel 8085, 8086, and 8088 microprocessors



Intel Pentium Processor, 273 Pin PGA

SECC Package Type
SECC is short for Single Edge Contact Cartridge. The SECC is inserted into a Slot 1 connector on the motherboard. Instead of pins, the SECC uses goldfinger contacts on the edge connector. These contacts carry signals from the processor to the motherboard. The back of the cartridge is a thermal plate that acts as a heatsink. Inside the S.E.C.C., most processors have a printed circuit board called the substrate that links together the processor, the L2 cache and the bus termination circuits. The S.E.C.C. package was used in the Intel Pentium II processors, which have 242 contacts and the Pentium II Xeon and Pentium III Xeon processors, which have 330 contacts.

SECC2 Package Type
The SECC2 package is similar to the SECC package except the SECC2 uses less casing and does not include the thermal plate. The SECC2 package was used in some later versions of the Pentium II processor and Pentium III processor which have 242 contacts.



Intel i486 SX Processor, 237 Pin PGA

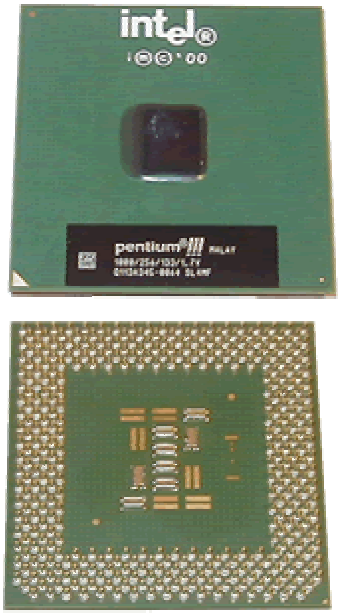
PGA Package Type:
PGA stands for Pin Grid Array. The A 132L ceramic PGA package was used for both the i386 and i486 microprocessors. The package was modified slightly with improvements in processing power to accommodate an increased number of pins (from 168 up to 237). Socket types included 486 Socket, Socket 1, Socket 2, Socket 3 and Socket 6. All supported the 486 bus.

As the Pentium processor evolved to the Pentium Pro, it supported a Socket 4, Socket 5, Socket 7 and Socket 8 connection to the motherboard. The number of pins increased from 273 to 387. The buses supported also changed as the processor design improved. This package PGA package supported the P5, P54C, P55C, and P5 bus. The Intel Pentium processor was packaged in a ceramic PGA with a heat spreader to achieve better cooling.

Intel Pentium II Processor, 242 Contact SECC

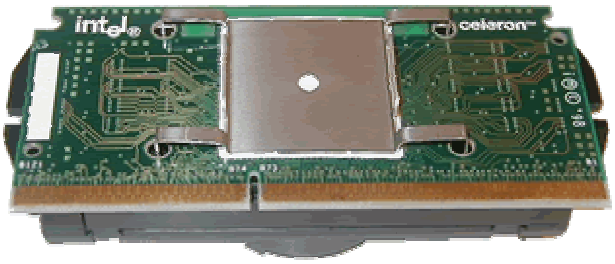
Intel Pentium III Processor, 242 Contact SECC2

SEP Package Type
SEP is short for Single Edge Processor. The SEP package is similar to a SECC or SECC2 package but it has no covering. In addition, the substrate (circuit board) is visible from the bottom side. The SEP package was used by early Intel Celeron processors, which have 242 contacts. It's connected to the motherboard through a Slot 2 connector.



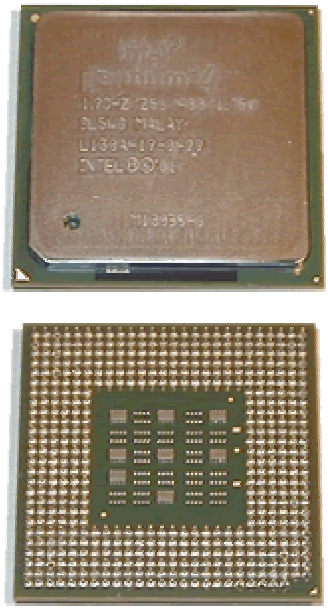
Intel Pentium III Processor, 370 Pin FC-PGA

FC-PGA2 Package Type
FC-PGA2 packages are similar to the FC-PGA package type, except these processors also have an Integrated Heat Sink (IHS). The integrated heat sink is attached directly to the die of the processor during manufacturing. Since the IHS makes a good thermal contact with the die and it offers a larger surface area for better heat dissipation, it can significantly increase thermal conductivity. The FC-PGA2 package is used in Pentium III, Intel Celeron processor (370 pins), and the Pentium 4 processor (478 pins). This processor package requires a Socket 478 connector on the motherboard and operates on a P6.8 bus

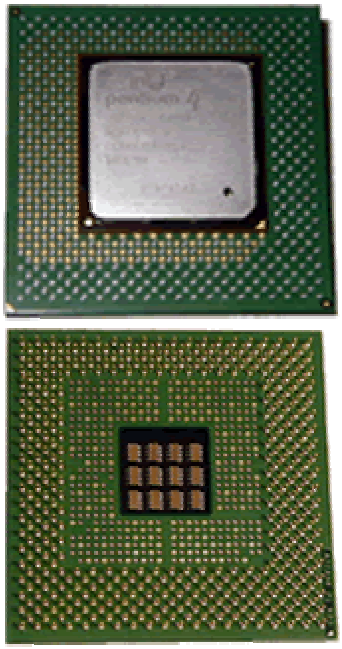


Intel Celeron Processor, 242 Contact SEP

FC-PGA Package Type
The FC-PGA package is short for Flip Chip Pin Grid Array. The chip is mounted to a small printed circuit board (substrate) where the pins have been secured. The pins are inserted into a ZIF (Zero Insertion Force) socket on the motherboard. These chips are turned upside down so that the die (the processor chip) is exposed on the top of the processor board. With the die exposed, thermal compound can be applied directly to the die which allows for more efficient cooling of the chip. FC-PGA processors also have surface-mounted capacitors and resistors soldered to the bottom of the processor board. The pins on the bottom of the chip are staggered. In addition, the pins are arranged in a way that the processor can only be inserted one way into the socket. The FC-PGA package is used with Pentium III and Intel Celeron processors, which use 370 pins. This microprocessor package type requires a Socket 370 connector on the motherboard.



Intel Pentium 4 Processor, 478 Pin FC-PGA2

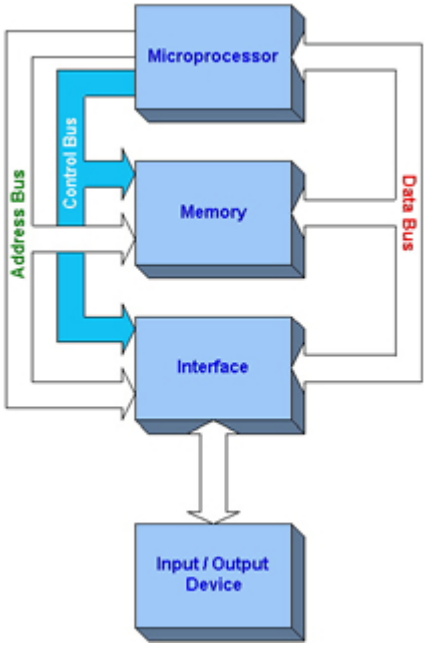


Intel Pentium 4 Processor, 423 Pin OOI

OOI or OLGA Package Type
OLGA stands for Organic Land Grid Array. The OLGA chips also use a flip chip design, where the processor is attached to the substrate face-down for better signal integrity, more efficient heat removal and lower inductance. The OOI then has an Integrated Heat Spreader (IHS) that helps heatsink dissipation to a properly attached fan heatsink. The OOI is used by the Pentium 4 processor, which has 423 pins. This processor package type requires a Socket 423 connector on the computer's motherboard and operates on P6.8 bus.

NOTE: This is not a comprehensive list of package types for all processors. These are only the most popular Intel processors. There are many other package types by Intel and other chip manufactures to consider. Further information can be found in the technical specifications for those products. There are also several web sites that list technical specs for microprocessor package types.

How Microprocessors Work in a PC
As the name implies, microprocessors are complex microscopic electronic circuits that process digital information. Technicians often use block diagrams to figure out complex operations. To the right is a block diagram of a simple microprocessor based computer. Although there's not much detail in the diagram, there's enough to get an idea of how data moves around inside a PC.



The Basic Computer – This diagram shows how the Microprocessor, Memory, and Interface read information from the Data Bus. The microprocessor sends control operations as well as addresses both the memory and interface functions. The input/output device sends and receives data to and from the interface device.

Input / output devices like the keyboard, mouse, disk drive, and printer are connected to interface circuitry that encode and decode data so it can be used by the computer and each input/output device. Notice that the microprocessor, memory, and interface circuits are all connected by control, address, and data buses. The microprocessor sends out binary addresses and control signals to enable memory locations and interface ports. Data moves back and forth between all three on the data bus.

Buses are really nothing more than paths (wires) for data that runs between devices in a circuit. They're drawn as wide arrows in the block diagram because they represent more than a single path. If this computer featured an Intel 8080 microprocessor, the address and data bus would be 8 bits wide so the wide arrow really represents 8 individual conductors and not a single wire. If this were a Pentium III microprocessor, then the buses would be 32 bits wide.

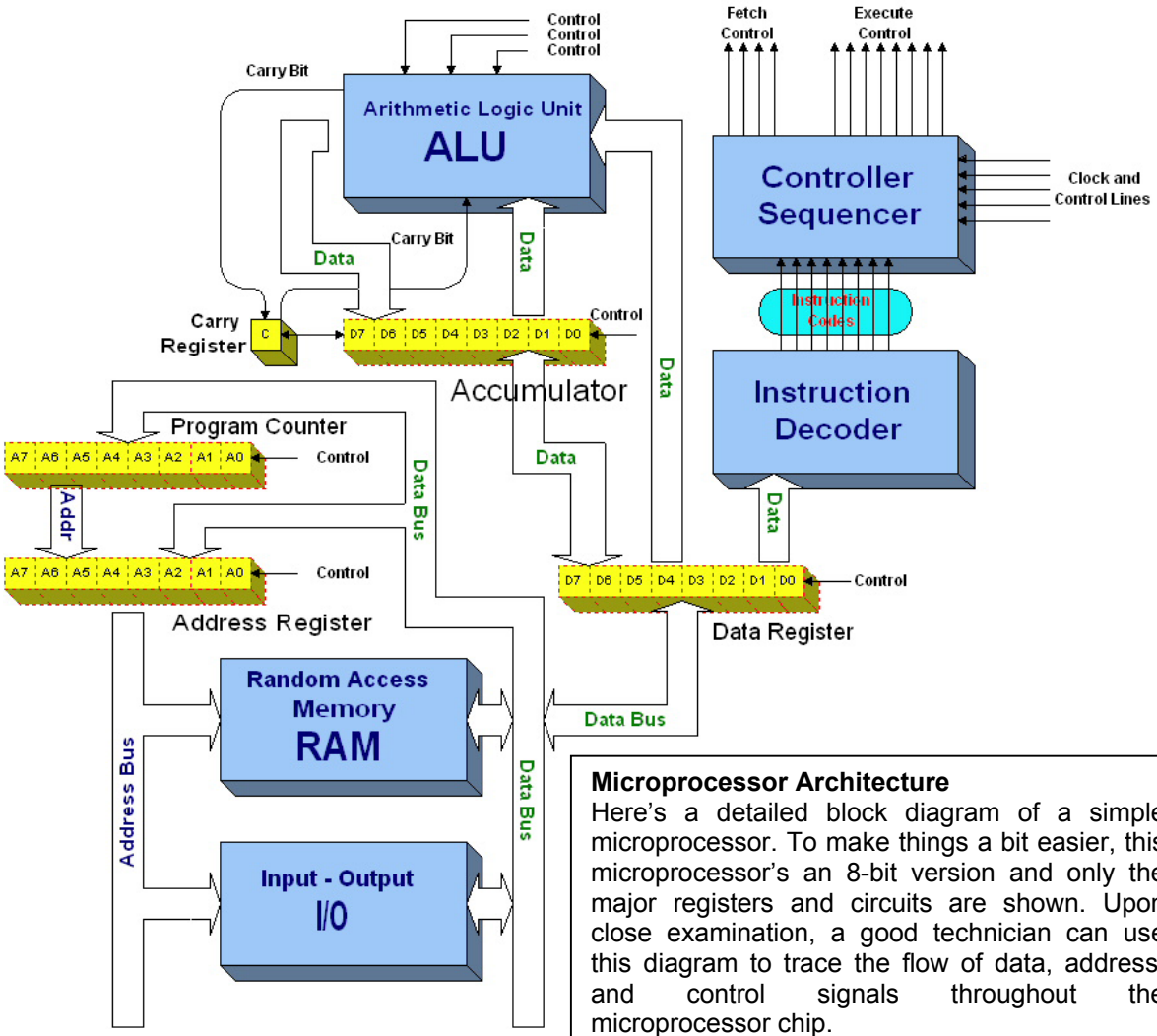
Wider buses are significant advances because they:

- allow very large numbers to be processed.
- are able to address more memory locations.
- can carry more control commands to devices.

Bus Width	Formula	Possible Bus Value
4–Bit	2^4	0 – 15
8–Bit	2^8	0 – 255
16–Bit	2^{16}	0 – 65,535
32-Bit	2^{32}	0 – 4,294,967,295

Photolithography technologies have become so sophisticated that manufacturers are now able to cram enough transistors into their microprocessors to accommodate 32–bit wide buses. The first microprocessor, the Intel 4004 featured only a 4-bit data and address bus meaning it could access a maximum of 16 memory addresses and handle numbers only as large as 15 on its data bus. The formula to figure this out is pretty easy…2^n where n represents the width of the bus. Remember that 2 in the formula represents the radix of the binary number system and since microprocessors use only binary values… 2^n will result in the maximum number of binary combinations given the bus width of n. The table above illustrates this.

To understand how a more conventional microprocessor works, it is helpful to look inside the device and learn more about its internal components and functions. Microprocessors are typically constructed from sub-circuits called registers, arithmetic logic units, counters, buffers, and decoders. Knowing how these components work together to process data gives a broader understanding of how software and hardware works… or doesn't work.



Microprocessor Architecture
Here's a detailed block diagram of a simple microprocessor. To make things a bit easier, this microprocessor's an 8-bit version and only the major registers and circuits are shown. Upon close examination, a good technician can use this diagram to trace the flow of data, address, and control signals throughout the microprocessor chip.

Microprocessor Architecture

Microprocessor architecture refers to the physical design of the IC circuit itself. The architecture defines the bus width as well as the construction of the registers, instruction decoder, controller sequencer, and arithmetic logic unit. Understanding how a microprocessor functions at this level will make you a better troubleshooter because you'll have a better understanding of how computers really work. Here are the basic functions of a basic microprocessor you should be familiar with:

Arithmetic Logic Unit (ALU): The ALU performs arithmetic and logic operations of data words that are sent to it. The Accumulator and Data Register are the two main inputs to the ALU. The control lines to the ALU determine if the input data words are added, subtracted, or logically compared. The two input words (operands) are loaded into the Accumulator and Data register and then the appropriate control signal from the Controller Sequencer is applied to the ALU to add, subtract, or compare the two binary words. The result is then loaded back into the Accumulator where it replaces the original operand.

Accumulator: Registers are digital circuits (flip-flops) that act as temporary memory circuits for single binary words. Registers hold a word until it can be processed by the ALU, loaded onto a bus, or decoded by the Instruction Decoder. The accumulator is a special type of register. It not only holds an operand for processing, it also holds the result of the ALU process. It can also be loaded directly with a value or loaded with a value from RAM with a single command from the microprocessor's instruction set.

Data Register: The Data Register is a temporary storage location for data words going to or coming from the Data Bus. Information on the Data Bus can be one of two things. Either it's a command or an operand. The Data Register holds an instruction while it's being decoded by the Instruction Decoder and also holds the operand while it's being written to RAM or sent to an I/O device like the printer or hard disk drive. It does this one binary word at a time.

Address Register: Like all registers, the address register is a temporary storage location for a single binary word. It holds the address of the memory location or I/O device where the information is going to or coming from.

Program Counter: The Program Counter points to the next location in memory where a command or operand is going to be read or written. Normally, the Program Counter counts sequentially but can be loaded with any value so data words can be retrieved from anywhere in memory.

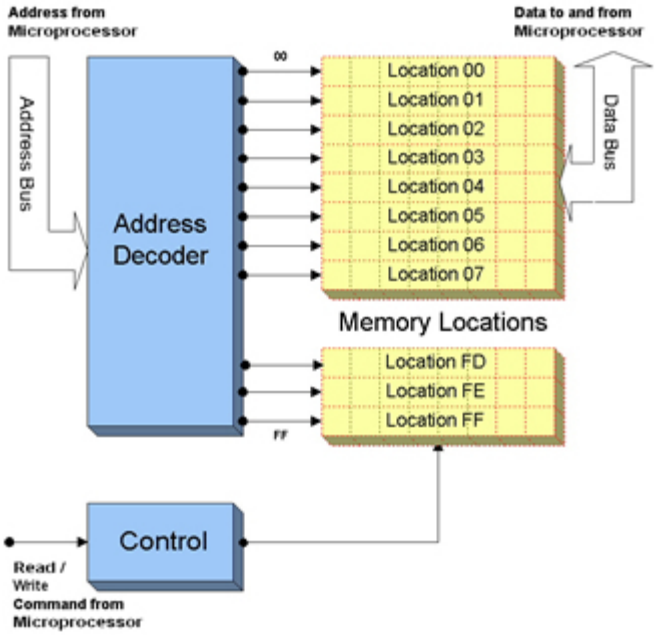
Instruction Decoder: The instruction decoder decodes data words from memory and determines the operation to be processed. Operations are individual commands that are part of the microprocessor's instruction set. Commands include operations that perform mathematical, logical, and memory manipulations. Each microprocessor has it's own instruction set.

Controller Sequencer: As each instruction is interpreted by the Instruction Decoder, the Controller Sequencer sends signals to the ALU and registers throughout the microprocessor to actually execute the instruction. The control signals from the Controller Sequencer enable registers to Read (RD) or Write (WR) data from the data and address busses. They also force registers to isolate the bus using a tri-state buffer so other registers can use it. Remember… Only one register can access the bus at any one time.

Memory

Microprocessors require memory to operate. The block diagram to the right shows the basic operation of how memory circuits typically work.

- The microprocessor loads an address onto the address bus to write/read information to/from and specific *Memory Location*.
- The *Address Decoder* reads the binary word on the *Address Bus* and enables a single control line to address the desired *Memory Location*.
- The microprocessor then sends either a *Read* or *Write* control signal to the *Control* circuit. The *Control* circuit then triggers the *Memory Location* to move data to/from the *Data Bus*.
- If the microprocessor sends a *Read* signal to the *Control* circuit, then the data word is moved from the addressed *Memory Location* to the *Data Bus*.
- If a *Write* signal is provided, then the data word is moved from the *Data Bus* into the *Memory Location*.



Memory: *Buses and control lines also connect to Cache, Random Access Memory (RAM), and Read Only Memory (ROM). In this microprocessor model, the address and data bus are 8-bits wide. That means that the microprocessor can address 256 bytes (00_{18} to FF_{16}) of memory (2^8) and can read or write 8 bits of the memory at a time.*

Random Access Memory (RAM): RAM is often referred to as main memory. The microprocessor can address any memory location in RAM depending on the data word loaded on the address bus. Early forms of computer memory used sequential logic circuits to store data. When advances in memory circuit technology allowed the microprocessor to access memory locations out of sequence, memory operations seemed to be random. However, there's really nothing random about it. The real problem with RAM is that it's volatile meaning… When the power's turned off, information in RAM is lost. That's why computers need ROM.

Read Only Memory (ROM): ROM is programmed with a permanent collection of instructions or data. ROM is addressed just like any other type of memory. The microprocessor loads a data word onto the address bus accessing the desired memory location. Since this is READ Only Memory, the microprocessor sends a read signal to the memory array moving the data word in the memory location onto the data bus.

Cache: Most microprocessors feature a small amount of really fast (and expensive) memory right on the chip. This type of memory is called cache and it really speeds up processor operations. Instead for writing and reading to and from external RAM, the microprocessor can keep some instructions close by making execution faster.

Microprocessor cache in modern chip design is identified as L1 and L2 cache and is relatively small. The Intel Pentium IV features 8KB of primary (L1) cache and either 512KB or 256KB of secondary (L2) cache… depending on the P4 model. Secondary cache is bigger than the primary cache and is positioned between L1 cache and RAM. It's faster than main memory, but slower than primary cache.

Basic PC Memory Operations: Because computers are turned off from time to time, all PCs are equipped with a ROM that contains the startup program. The most obvious ROM is the Basic Input/Output System (BIOS) IC. When the PC starts up, the microprocessor executes programs that are stored in the BIOS chip. This program tests the computer's hardware and

starts the boot process by reading the boot sector off the hard disk, floppy disk, or CDROM. The boot sector also contains a program that the BIOS stores it in RAM. The microprocessor then executes the boot sector's instructions from the RAM that fetches operating system instructions from the hard disk drive. The microprocessor continues to fetch and execute instructions from the hard disk drive until the operating system is up and running.

In early computers, this process was performed manually by entering binary commands from a switch panel. It was difficult work and took hours if not days to bring-up a large computer system. Technicians described the process as, "It's like pulling yourself up by your own bootstraps (shoelaces)." Later the process was called "bootstrap loading" and has survived to this day as "booting-up".

The Microprocessor Instruction Set

So far… All we've talked about is the architecture of the microprocessor. We've discussed what the basic microprocessor components are and what they do. You've also discovered how data is moved around inside a microprocessor as it processes digital information. Now it's time to introduce how microprocessors execute programs.

Perhaps the most powerful design feature in any microprocessor is its Instruction Set. The instruction set is the list of commands the Instruction Decoder can interpret and execute. In reality, each instruction is a binary word usually represented as a hexadecimal number as machine code. From the Intel 4004 up to the Pentium IV, each microprocessor has had it's own instruction set. As new microprocessor versions are released, the Instruction Set gets larger and includes more powerful operations to manipulate data. We're using the instruction set of the Intel 8085 microprocessor in this module because it's basic construction makes understanding assembly language a bit easier. Although it's only an 8-bit microprocessor, it's still widely used in industry as an educational tool for teaching microprocessor theory.

Chip manufacturers publish documents called data sheets that detail the architecture and application of their products. The data sheet for the Intel 8085 microprocessor includes 29 pages of technical data… Most of it beyond the scope of this course. However, pages 19 through 21 are very useful because they list the 8085's Instruction Set and command application. With this information and an 8085 simulator or emulator, you can practice writing and executing assembly language code.

Machine Language Program (Hex or Machine Code)	Assembly Language Program				
	Address	Mnemonic	Operand	MC	Comment
:10 00 00 00 06 **09** 0E **05** 79 FE 00 CA 17 00 0D 78	0000	MVI B	09	06	Moves 09 into A Reg.
:16 00 80 D2 89				**09**	First Num
:08 00 10 00 13 00 14 0D C2 0E 00 76 6E	0002	MVI C	05	0E	Moves 05 into C Reg.
:00 00 00 01 FF				**05**	Second Num
Machine Code: *Programs are instructions and operands written by programmers. This code is assembled and then stored in main memory to be executed. Machine code is terse and works at the microprocessor's level. It's hard to believe, but there are highly paid machine level programmers that can write complex programs directly in hex code (Geeky!). The machine code above multiplies 9 and 5 and stores the result in the microprocessor's Accumulator.*	0004	MOV A C		79	Moves C Reg. into A
	0005	CPI	00	FE	Check if A Reg. is 0
				00	
	0007	JZ	00 17	CA	If A=0 then Jump to 0017
				17	
				00	
	000A	DCR C		0D	Decrement C Reg.
	000B	MOV A B		78	Move Reg. B into A
	000C	MVI D	00	16	Move 00 into D Reg.
				00	
Assembly Language: *The program to the right is written in assembly language and generated the machine code shown above. The program multiplies two 8-bit numbers (by repeated addition) to get a 16-bit result stored in Accumulator. Mnemonics are used to make programming commands easier to remember and edit. The code is created in a text editor and then compiled into machine code using a tool called an assembler.*	000E	ADD B		80	Add B Reg. to A
	000F	JNC	00 13	D2	Jump if no carry to 0013
				13	
				00	
	0012	INR D		14	Increment D Reg.
	0013	DCR C		0D	Decrement C Reg.
	0014	JCN	00 0E	C2	Jump if no carry to 000E
				0E	
				00	
	0017	HLT		76	Halt the program

There are 14 lines of code in the program above. Once processed by an assembler, it's compiled and reduced to 39 bytes of binary data called Machine Code. An assembler is a program that takes basic computer instructions and converts them into a pattern of bits that the computer's processor can use to perform its basic operations. This type of program is called assembler or assembly language.

Mnemonics are words or acronyms used in assembly language to represent a binary machine instruction operation code (Opcode). The opcode is the command portion of the instruction without the operand. Different processors have different instruction sets and therefore use a different set of mnemonics to represent them.

Let's look at one command in detail in this program. The instruction at memory location 000E is ADD B. The machine code (MC) for this command is 80_{16} and takes 4 clock cycles to completely execute. This instruction adds the contents of the B Register to the Accumulator. Here's what's happening inside the microprocessor when this instruction is executed:

First Clock Cycle:

- The Program Counter's Tri-State Buffer is activated loading the counter value onto the Counter Register.
- The Read (RD) line is activated.
- The Data-In Tri-State buffer is activated to read the instruction into the Instruction Register.
- Latch the instruction into the instruction register.

Second Clock Cycle:

- The ADD B instruction is decoded.
- Set the operation of the ALU to addition.
- Latch the output of the ALU into the C register.

Third Clock Cycle:

- The result is stored in the Accumulator.

Fourth Clock Cycle:

- The program counter is incremented.

It's clear to see that a lot of things happen inside the microprocessor when commands are executed. Control lines fire in a sequence clock cycles that trigger registers, controllers, and the ALU to manipulate data. Some instructions might take 2 or 3 clock cycles while more powerful operations might take 5 or 6 clock cycles.

Conclusion

An increase in the number of transistors in a microprocessor has a huge effect on its performance. The Intel 8088 microprocessor was simple in comparison to today's high end processors. A typical 8088 instruction took 15 clock cycles to execute. Because of the crude design of the chip's multiplier, it took nearly 80 clock cycles just to do one 16-bit multiplication. Adding more transistors meant more powerful multipliers capable of performing the same operation in a fewer clock cycles and a much higher processing speed.

More transistors also allow a technology called pipelining. In a pipelined architecture, instruction execution overlaps. So even though it might take 5 clock cycles to execute each instruction, there can be 5 instructions in various stages of execution simultaneously. That way it looks like one instruction completes every clock cycle.

Many modern processors have multiple instruction decoders, each with its own pipeline. This allows multiple instruction streams, which means more than one instruction can complete during each clock cycle. This technique can be quite complex to implement, so it takes lots of transistors.

The trend in processor design has been toward full 32-bit ALUs with fast floating point processors built in and pipelined execution with multiple instruction streams. There has also been a tendency toward special instructions (like the MMX instructions) that make certain operations particularly efficient. There has also been the addition of hardware virtual memory support and L1 caching on the processor chip. All of these trends push up the transistor count, leading to the multi-million transistor powerhouses available today. These processors can execute about one billion instructions per second!

Gross Anatomy

Exercise 1 – The Microprocessor

Name:	
Period:	
Date:	

Exercise Objective:

Upon completing this exercise, you'll learn what a microprocessor is and how it functions as part of a personal computer. You'll also learn how microprocessors have evolved from their early concepts to the data-handling workhorses they are today. You'll also study microprocessor architecture and construction. Finally, you'll work with pseudo computer code and understand how basic machine level programming works.

Research Resources:

Company	Web Site	Description
Webopedia	http://www.webopedia.com/	Internet technology encyclopedia.
Techweb	http://www.techweb.com/	Internet technology encyclopedia.
Search390	http://search390.techtarget.com/glossary/0,294242,sid10,00.html	Technology Glossary
Infotech Solutions	http://www.insoluz.com/	8085 Simulator Software
PC Guide	http://www.pcguide.com/ref/cpu/index.htm	The Processor
Intel Corporation	http://intel.com/education/mpworks/INDEX.HTM	How Microprocessors Work

Vocabulary:

On a separate piece of paper, define the following words using at least two sentences. Be sure your definitions are in a context regarding microprocessor technology.

- Transistor
- Register
- Assembler
- Clock Cycle
- Integrated Circuit
- ALU
- L1 and L2 Cache
- Photolithography
- Mnemonic
- Package Type
- Bus
- Opcode
- Die
- Pipelining

Questions:

1. Describe the relationship between the number of transistors in a microprocessor and its processing power.
2. Describe the Intel Pentium IV processor in terms of maximum: A) Bus Width; B) System Clock Speed; C) Host Bus Speed; D) Internal Cache; E) Package Type.
3. List the package types that support the Intel Pentium III processor.
4. What is the major advantage of the flip-chip package design?
5. The circuit in the microprocessor that performs arithmetic and logic operations is called the _____ _____ _____ .
6. Describe the differences between RAM and ROM.
7. Where are the two operands held just before they're transferred to the ALU?
8. How many memory locations can a 16-bit address bus access?
9. What is the machine code for the halt instruction in an 8085 microprocessor?
10. What does the program to the right do?
11. How many memory locations are required to store this program?
12. Write out the machine code for this program.
13. What will the contents of the accumulator be after this program is executed?
14. Download an 8085 Microprocessor Simulator or 8086 Emulator from the Internet or CSS support disk. Write and enter a program that simulates multiplying 5 and 6 by repeated addition. Show the results to your instructor.

Addr.	Instruction
0000	MVI A 00
0002	MVI B 04
0004	ADD B
0005	ADD B
0006	ADD B
0007	HLT

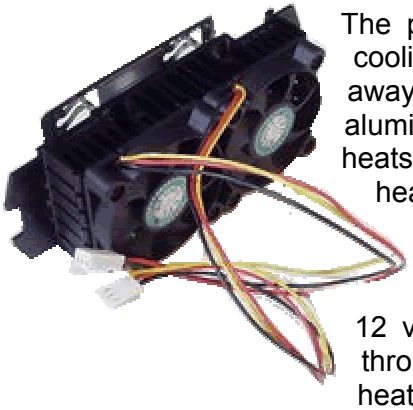
15. Upon close examination of the Pentium 4 Die photograph, it easy to visualize the silicon patterns as buildings and roadways within a city. Write an analogy comparing the buses, registers, tri-state buffers, instruction decoder, control sequencer, and memory to the traffic of a small city.

Gross Anatomy Guide

The Processor Cooler

Introduction:

Heatsink and Cooling Fan:



The processor die gets very hot while it's working. The heatsink and cooling fan keep the microprocessor running cool by drawing the heat away from the processor chip. The heat sink is a component made of an aluminum alloy. Its fins are designed to increase the surface area of the heatsink so it dissipates heat into the surrounding air more easily. The heatsink clamp insures good physical contact between the processor die and the heatsink so there's good heat transfer. Adding thermal compound or a thermal pad between the processor die and heatsink also improves thermal conduction. The fan plugs intro the 12 volt header on the system motherboard. Its job is to circulate air through the fins of the heatsink drawing the heat away from the processor chip.

This Slot 1/A Pentium II/AMD Athlon CPU cooling fan is a combination heatsink and fan designed for cooling Intel Pentium II CPUs that use the original SECC plastic cartridge (233-400 MHz), or AMD Athlon CPUs. It sports an extra-large heatsink with enough cooling capacity for any CPU.

This is a high-end custom cooler that employs a nickel-plated, copper-core, aluminum-finned heatsink and a double ball-bearing fan unit for a Socket 423 Intel Processor

Thermal Interface:

The material between the processor die and the heatsink is called the thermal interface. The thermal interface is responsible for transferring heat from one medium to another… a term known as thermal conduction. Thermal compounds have been developed and refined to improve thermal conduction allowing processors to run as cool as possible.

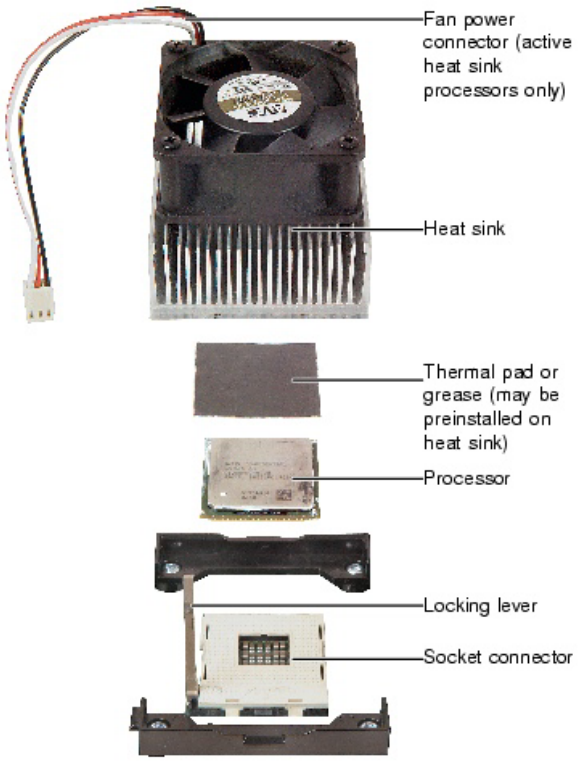
Thermal Compound

A silicone based paste loaded with some sort of thermally conductive material (silver or mercury). Do not use thermal grease and do not combine thermal grease with the thermal compound. This will cause the processor to overheat. Working with thermal compound is messy… and dangerous. Since some compounds are loaded with mercury, it's important to never touch the compound you're your skin. Use an applicator or gloves to prevent mercury being absorbed into your body through your skin. Thermal compound comes in individual blister-packs or an application syringe. Be neat! Thermal compounds can conduct electricity as well. Drips between the pins of a processor can short it out causing serious damage.

Thermal Pad

Safe and simple thermal interface appliances widely used by the computer manufacture industry. They are self-adhesive thermally conductive pieces of tape that can be applied neatly to a heatsink without taking special safety precautions. The pad is made of thermal compound sandwiched between two very thin squares of Teflon coated aluminum foil. Thermal pads are applied by peeling off the backing (one side) and then placing it on top of the processor. Then the other backing is peeled off and the heatsink installed. Once the PC is turned on, the processor's heat will melt the pad into a semi-fluid mass that will fill in the microscopic crevices in the heat sink and CPU allowing for better thermal transfer.

Thermal pads are fine for normal processor operation. However, thermal compound is better at transferring heat and is preferred for processor over-clocking or in situations requiring high processor utilization. Regardless of the type of thermal conductor you decide to use, you must use one or the processor will overheat.



Common Processor Cooler Problems

1. Misalignment of heatsink/fan over processor die.

   This can happen if both ends of the heatsink/fan clip are not properly located on the mounting tabs of the socket. If not properly aligned, the heatsink may end up resting on the backside of the socket, and not make contact with the die.

   Consequences: Burned-up processor

2. Failure to remove the Mylar plastic cover on the heatsink's thermal interface material.

   Protective Mylar film is present on heatsinks with pre-applied thermal interface material, and must be removed prior to heatsink installation.

   Consequences: Melted Mylar and a burned-up processor

3. Leaving the thermal interface material exposed to the environment for too long.

   Dust particles may stick to the material. This causes gaps to form between the heatsink and die resulting in poor heat conduction between the two surfaces.

   Consequences: Shortened functional life and burned up processor

4. Re-using the thermal interface material

   If a heatsink/fan is removed from a processor, the heatsink/fan, and die must be completely and appropriately cleaned, i.e., all residual thermal interface material must be removed from the surface of the heatsink/fan. A new pad of thermal interface material must be applied to the heatsink/fan. The same holds true even if the heatsink/fan is removed before power is applied. Once the thermal pad is compressed between the heatsink and processor die, removal can cause the pad to be distorted making it unusable.

   Consequences: Burned-up processor

5. “Pressing” on the heatsink/fan while attaching the heatsink

Assemblers do this to facilitate the hooking of the clip on the Socket A mounting tab. Pressing down on the heatsink/fan during installation is not necessary and may damage the processor or motherboard

Consequences: Cracked die or chipped corners

6. Plugging the heatsink/fan into the wrong connector

Motherboards typically have two fan connectors in the vicinity of the processor socket: (1) CPU FAN PLUG, and (2) SYSTEM FAN PLUG. SYSTEM FAN PLUG has a delay at start-up and, as a result, should not be utilized for plugging the heatsink/fan. The CPU FAN PLUG should be utilized because it applies power immediately.

Consequences: Burned-up processor

7. “Scraping” the motherboard with the heatsink clip during installation

As the clip is fastened onto the processor socket, it may scrape the surface of the motherboard. This can result in damaged traces on some motherboards. Avoid problems by placing a thin insulating material under the socket mounting tabs during heatsink installation.

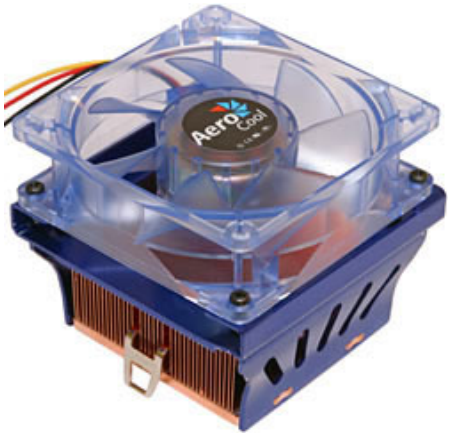
Consequences: Burned-up processor or damaged motherboard

Conclusion:

There’s an entire industry dedicated to keeping PC processors cool. CPU coolers range from stock to stupendous featuring accessories such as air, water, and even liquid nitrogen circulation systems.

It’s important to remember that there are three component parts to the processor cooler. The heatsink, cooling fan, and thermal interface all play a vital role in keeping a PC running.

- Heatsinks are designed for a specific processor package type.
- Thermal compound fills in the gap between the die and heatsink facilitating a better transfer of heat. It must be applied safely, neatly, and correctly to avoid personal injury or damage to the processor. The key to the cooler’s success is a complete and proper interface between the processor die and the heat sink.
- The fins on the heatsink increase its surface area permitting heat to be distributed further away from the processor die. The fan circulates cool air through the fins dissipating even more heat.



Gross Anatomy Guide

Exercise 2 – The Processor Cooler

Name:	
Period:	
Date:	

Exercise Objective:

At the completion of this exercise, you will demonstrate your knowledge of processor heatsinks. You will correctly answer a variety of questions regarding the identification, operation, and selection of heatsinks.

Research Resources:

Company	Web Site	Description
Webopedia	http://www.webopedia.com/	Internet technology encyclopedia.
Techweb	http://www.techweb.com/	Internet technology encyclopedia.
Gateway Computers	http://support.gateway.com/s/Manuals/Desktops/8508817/8508817a.htm	Removing Old Processors
Dan's Data	http://www.dansdata.com/coolercomp.htm	CPU Coolers Compared
Tom's Hardware	http://www6.tomshardware.com/cpu/20010521/	Can't Touch This!

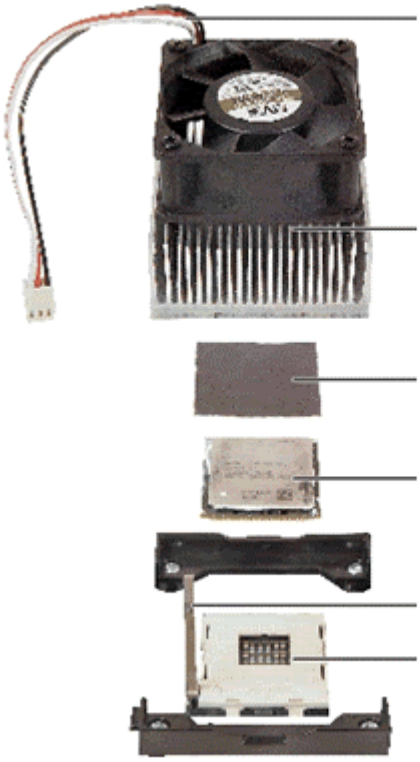
Vocabulary:

On a separate piece of paper, define the following words using at least two sentences. Be sure your definitions are in a context regarding heatsink technology.

- Heatsink
- Thermal Pad
- Thermal Interface
- Thermal Conduction
- Thermal Compound
- Cooling Fan

Questions:

1. You should never allow thermal compound to touch your skin because it contains _____.
2. What material is used to fill in the gaps between the heatsink and processor die to allow a better transfer of heat?
3. What's a good substitute for Thermal Compound?
4. During a repair operation you needed to remove the heatsink from the processor. What must you do before reinstalling the heatsink?
5. Label the parts in the pictorial to the right.
6. Which connector should the fan be plugged into on the motherboard?



Gross Anatomy Guide

Computer Memory

Introduction:

To recap what we know so far… We know that the microprocessor's job is to execute instructions that are written into programs. Since microprocessors essentially act on only one command at a time, the instructions and operands must be held in queue someplace until it's their turn to be executed. This is where computer memory comes into play. Computer memory is where computer programs and operands are stored until they're acted upon.

RAM and ROM

Consider two basic families of computer memory… They are Serial Access Memory (SAM) and Random Access Memory (RAM).

Paper Tape: *A long strip of heavy paper, usually an inch wide, in which holes could be punched, 5 to 9 per row. For computer use, usually 8 holes were used: 7 data bits and 1 parity bit. Paper tape was also used in telecommunications (telex) and in the printing industry as the input medium for hot-metal typesetting machines and is still used for numerical control of milling and drilling machines.*

Regarding Serial Access Memory, data is read and written sequentially from or to the data storage media. Storage media like Hollerith (punch) cards, paper tape, or magnetic tape are all examples of SAM. The entire program is processed one byte at a time as it's read or written from or to the storage medium. There were big drawbacks to SAM technology with the major one being its excruciatingly slow access speed. If the needed data or instruction is at the end of a tape, then the entire program must be processed first until it could be found. There was no way to go directly to the needed information. Additionally, the only types of programs that could be run were *straight–line* programs. There was no ability for a program to jump forward to an instruction nor could they loop back if certain results were met.

SAM is old technology, but there are still some examples of it being used today. Magnetic backup tapes still store information sequentially as do milling and drilling machines. Even computer video memory works by emulating serial access to process the texture buffer memory.

If SAM technology has an opposite, it's Random Access Memory or RAM. Random Access Memory is a considerable improvement over SAM because data can be retrieved from any memory location, in any order, at any time. Today's microprocessors use only Random Access Memory to process data. In fact, the term RAM is synonymous with computer memory.

Computer memory is either *volatile* or *nonvolatile*. Simply put, nonvolatile memory holds its data even when power is removed from the memory system. Good examples of nonvolatile memory are punch cards, paper tape, magnetic tape, and magnetic disk storage. Later, we'll discuss ROM, PROM, EPROM, and flash memory which are also nonvolatile forms of memory. On the other hand, volatile memory loses its data anytime the computer looses power. Computer memory (RAM) is volatile memory. If you have ever tried to recover a word processing document after a power failure you know what I mean. If it wasn't saved to disk it was gone with no hope of recovery. Some computer users even purchase battery backup systems called Uninterruptible Power Supplies (UPS) to protect themselves in case of problems with electrical power.

Memory Speed

Memory Access Time / Memory Cycle Time – Besides capacity, the important characteristics of computer memory are memory access time and memory cycle time. These two qualities go hand-in-hand when determining a memory system's performance. The faster the memory access time the faster the memory cycle time. Manufacturers and chip retailers often refer to this characteristic as *memory speed*. Memory speeds range from the (1980) 200 nS (nanoseconds) to today's scorching speeds of up to 4 nS. That's 1 memory operation every 4 billionths of a second – That's Zippy! As semiconductor technologies advance, memory speeds are bound to improve. The following metrics combine to make-up the memory system's memory access time:

- **Read Access Time –** Sometimes called *read time*, it's the time it takes for the memory system to place information on the data bus after the desired memory location is addressed.

- **Write Access Time –** Sometimes called write time, it's the time it takes for the memory system to take information from the data bus and store it in a memory location.

- **Average Access Time –** The write time and read time are seldom the same because of the different circuitry that's required to perform each operation. The average of the read and write time is used by industry to indicate the speed of memory chips or modules.

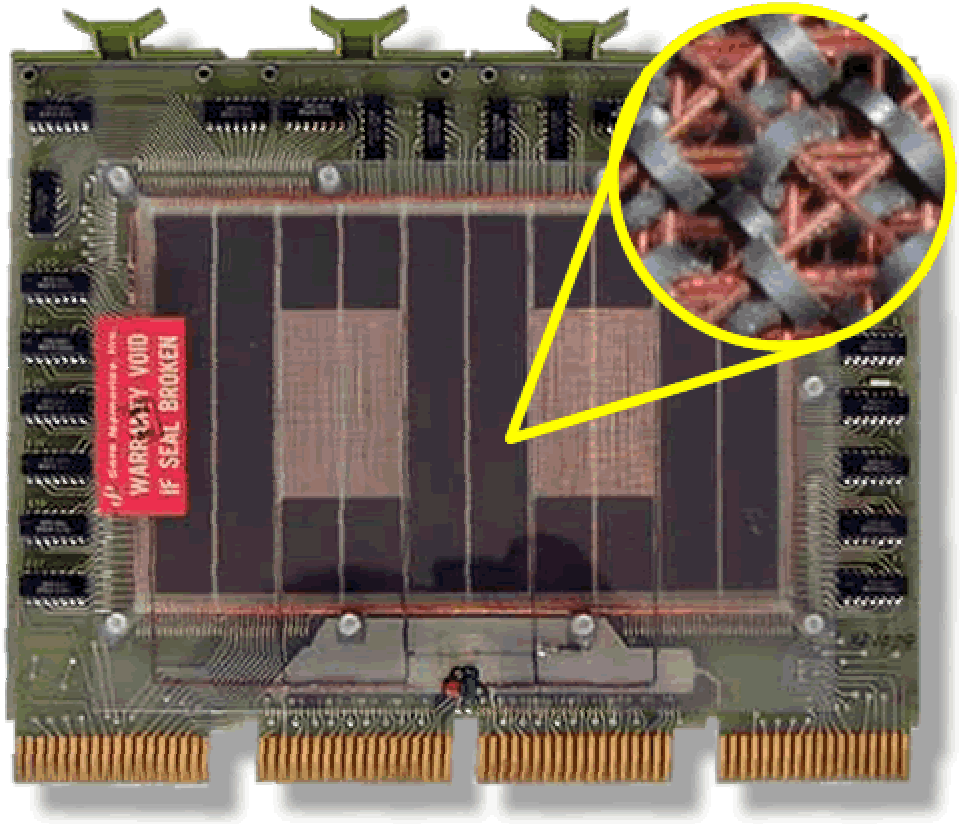
- **Memory System Cycle Time –** Shortest possible time between two memory operations. The memory system's cycle time depends on the computers clock speed and microprocessor architecture. The microprocessor's design may prevent access to memory during specific instructions therefore affecting memory system cycle time.

History of Computer Memory

Each significant breakthrough in computer technology has always been preceded by an advance in semiconductor technology. Improvements upon transistor and circuit integration are the key to making computers smaller, faster, and more efficient. As a technician, it's important to maintain perspective regarding advances in technology. By their very nature, technical advances are built upon the success of previous technologies. Memory systems are a good example. If you understand how early memory systems worked, you'll have a better understanding and appreciation of the technology designed into modern computer memory.

Core Memory – In the 1950's, vacuum tubes were replaced with transistors and stacks of memory cards called core memory became the hot new technology. In core memory, each bit was stored in a tiny, doughnut shaped permanent magnet. The permanent magnet could be magnetized in one of two directions. The direction of the magnetism depended on the direction

of the currents flowing through the memory wires. Since each memory location could be independently addressed and the magnets maintained their magnetism even after the current was removed, core memory was considered nonvolatile RAM memory. Core memories are now obsolete because they were big, expensive, and used lots of current to operate. The term core memory is still sometimes used to describe a computer's main memory even though modern memory systems are now comprised of volatile semiconductor memory circuits.



Core Memory – *Each bit is stored in a small, doughnut-shaped permanent magnet. The magnification inset shows the magnets and the hair-thin memory wires. The direction of the magnetic field determined if the core's value was 1 or 0. This circuit board, built in 1960, held 4096 magnets with enough core memory to access 1 x 4096 bits of data. Seven more cards would provide an 8 x 4095 bit memory system or 4 kilobytes. Cost - $7,500 per card.*

Semiconductor Memory

As with other IC construction techniques, there are two major semiconductor technologies used to solid-state memories.

Bipolar – Although bipolar memory is very fast, this type of memory is seldom used because they are expensive to construct, require a lot of power to operate, and are too large to easily integrate into an IC circuit. For that reason, bipolar memory is used for applications that require the greatest speed. Primary (L1) cache would be a good application for bipolar memory construction techniques.

MOS (Metal-Oxide-Semiconductor) - MOS memory is the most popular type of computer memory and comes in two flavors… Static and Dynamic.

> **Static Memory –** Static memory circuits are easier to build because they don't require much in the way of additional support circuitry to operate. They're a little pricey though. Each memory cell requires 6 integrated MOS transistors to store one data bit. Static memory is faster than Dynamic memory.

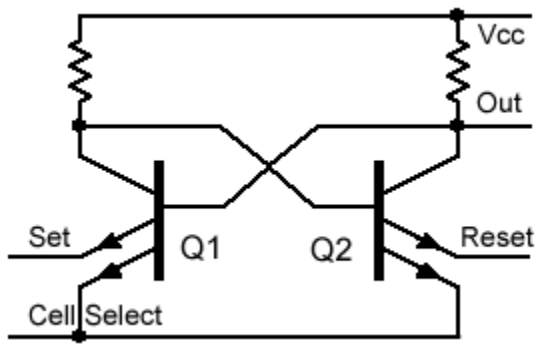
Dynamic Memory – Dynamic memory systems are cheaper to construct, but require additional support circuits to implement. They also require an additional refresh signal to maintain the contents of the chip's memory. Dynamic RAM is used for larger memory system applications.

Both static and dynamic memory technologies are steadily improving. As IC construction techniques advance, larger and larger memory circuits will be integrated onto smaller and smaller silicon wafers. As memory capacities increase, their cost will decrease providing computer users with the opportunity to increase computing power by increasing main memory. Software publishers can then write more powerful programs making computers more useful to the people that use them.

How RAM Works

The integrated circuits that make up a memory system contain thousands upon thousands of transistors, resistors, and capacitors. These components are arranged into circuits called flip-flops …the fundamental design of any computer memory cell. Each flip-flop cell must be selected before it can be forced into one of two states… set (1) or resent (0).

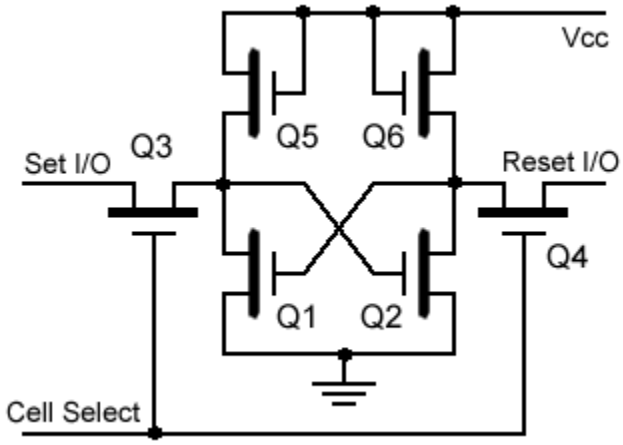
The Memory Cell



Bipolar Memory Cell – This is a simplified schematic of a single bipolar memory cell or flip-flop. Either a 1 or a 0 is stored at the collector of the multiple emitter NPN transistors Q1 and Q2 depending on the state of the flip-flop. The flip-flop can be either Set (Out =1) or Reset (Out =0) with a LOW logic signal on those inputs.

- If Reset receives a LOW signal, Q2 is turned ON which turns Q1 OFF. Q2 sinks the collector logic level so Out = 0.
- If Set receives a LOW signal, Q1 is turned ON which turns Q2 OFF. Q2 opens making Vcc available through R2 so Out = 1.

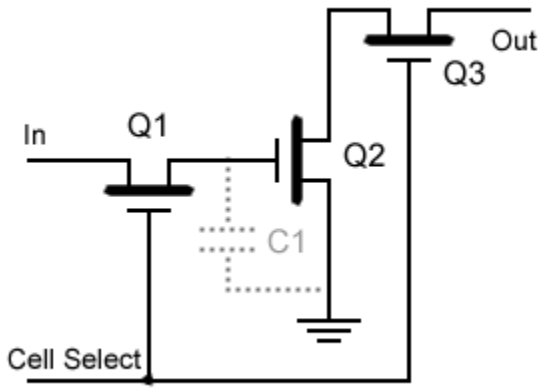
A LOW Cell Select line enables the cell to be read from or written to. Bipolar memory is fast because its transistors change states very quickly.



Static Memory Cell – Here's a concept drawing of a MOS static-memory cell. This is also a flip-flop circuit, but this one's made up of **MOS transistors**. Notice that the gates of transistors Q1 and Q2 are cross-coupled so the output of one transistor changes the state of the other.

The Cell Select signal (HI) biases the gates of the transistors so current flows in and out of Q3 and Q4. The Set I/O and Reset I/O lines not only set and reset the flip-flop but also act as the cell's output. Transistors Q5 and Q6 act as bias resistors for Q1 and Q2. Just like the bipolar memory cell, the static memory cell will retain the set or reset state until power is removed from the circuit.

Notice that current paths do not cross PN junctions in MOS circuitry like they do in bipolar transistor memories. Instead, signals must travel through channels of doped silicon. This and the high capacitance between the gate and the channel make MOS memory slower than bipolar memory



Dynamic Memory Cell – Dynamic memory circuits are much simpler in design. There are fewer transistors per memory cell and no need for support devices like resistors or capacitors to make them work. Since dynamic memory circuits use half the components of static memory, dynamic memory can hold twice as many memory cells than the same size chip of static memory.

Dynamic memory in not built around the flip-flop and is not set or reset to change logic states. Instead, this type of MOS memory cell uses the internal capacitance of Q2 (C1). If charged logic 1 is output. If the capacitor is not charged, then the output is 0.

To keep the capacitor charged, the data is refreshed every few milliseconds using external circuitry called refresh logic. The logic circuits access each memory cell every few tenths of a millisecond to keep the memory contents valid.

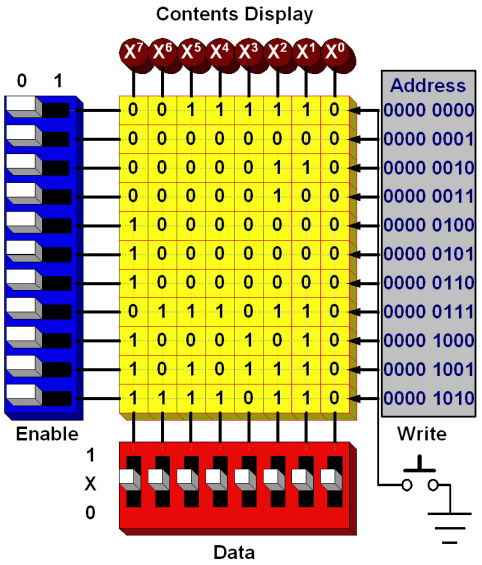
The Memory Matrix

Here's a very simple model of a section of a computer's main memory. The yellow block represents a matrix of 11 addresses each with the storage capacity of 8 data-bits.

The blue Enable Switch selects only one address at a time while the red Data switch is program bit patterns to be loaded into a memory location. The Write Switch is a momentary contact switch that initiates the loading of the 8-bit word from the Data switch into the enabled memory location. The red LED's (X^0 through X^7) indicate the contents of any enabled memory location.

In this model, there's already a program loaded that that will multiply the numbers 3 and 4 by adding 4 three times.

Addr.	Instruction	Operand	MC (Hex)	Binary
0000	MVI A	00	3E 00	0011 0000 1110 0000
0002	MVI B 04	04	06 04	0000 0000 0110 0100
0004	ADD B		80	1000 0000
0005	ADD B		80	1000 0000
0006	ADD B		80	1000 0000
0007	HLT		76	0111 0110



Word Lines – Include the Enable and Write signals. These control lines are made of metal and carry information that control information to the bits.

Data Lines – Data lines are made of polysilicon provide the logic levels to the bits. Data lines can operate in three states.

1. Logic 1
2. Logic 0
3. High Impedance

Bit – Each bit is a group of transistors at the intersection of a word and bit line.

Reading Information From Memory

1. Set the data lines to high impedance (X)

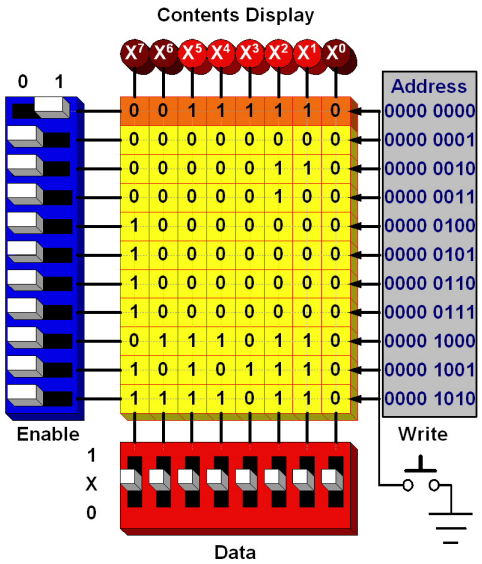
   The data lines must be set to high impedance so the enabled bits are read and not the data lines.

2. Enable the word lines

   Turning on the word line allows the microprocessor to read the contents of all cells on the word line.

3. Contents Display

   The contents of the word line are displayed.



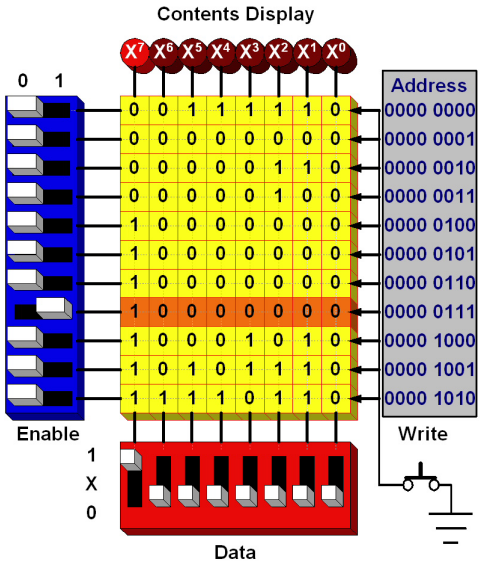
Writing Information to Memory

1. Set the data lines to 1s and 0s to represent the data word you want to write to memory.

2. Enable the word line at the desired address

3. Send a Write signal to the memory matrix.

   Only the enabled word line will receive the data word. To change the contents of a word line, simply write new data over the old.



Common Memory Types

Although memory performance improves with advances in semiconductor technology, computer memory can be categorized according to its architecture and performance. Detailed information regarding common memory types is available from various memory manufacturers and distributors.

There are two memory configurations to consider as well - Parity and Non-Parity. Parity memory includes built-in error detection/correction circuitry. Non-parity, sometimes called Non-ECC (Error Correction Code), memory does not. Most desktop and laptop PCs use non-ECC memory whereas servers typically use parity memory modules.

SRAM

Static Random Access Memory is an integrated circuit similar to a DRAM with the exception that the memory does not need to be refreshed. Unlike volatile memory like DRAM, SRAM retains its contents even when memory address and enable signals are turned off. Since SRAM is fast (and pricey), it's primarily used as cache memory.

DRAM
Dynamic Random Access Memory is the most popular type of memory component used in personal computer systems. "Dynamic" means the DRAMs need a constant "refresh" signal (pulses of current through every memory cell) to keep the stored information alive. There are several varieties of DRAM to consider:

- **FPM**
  Fast Page Mode – Considered to be the first type of PC DRAM, FPM featured faster sequential access to DRAM. Each memory location is addressed only once. After which it can be accessed multiple times without having to readdress it. Maximum transfer rate to L2 cache is approximately 176 MBps.

- **EDO**
  Extended Data Out. DRAM can be addressed without timing constraints. Asynchronous operation improves memory access time permitting memory speeds comparable to FPM DRAM at a lower cost. EDO DRAM is about 5% faster than FPM DRAM. Maximum transfer rate to L2 cache is approximately 264 MBps.

- **BEDO**
  Burst EDO is a variant on EDO DRAM in which read or write cycles are batched in bursts of four. Burst EDO bus speeds range from 40MHz to 66MHz, as opposed to the 33MHz bus speeds that can be accomplished using fast page mode or EDO DRAM.

SDRAM
Synchronous Dynamic Random Access Memory delivers bursts of data at high speeds using a synchronous interface. Its is actually SDR SDRAM (Single Data Rate SDRAM) but is usually used to referred to as just "SDRAM." SDRAM is about five percent faster than EDO RAM and is the most common form in desktops today. Maximum transfer rate to L2 cache is approximately 528 MBps.

DDR
Double Data Rate is a type of SDRAM in which data is sent on both the rising and falling edges of clock cycles in a data burst. It is usually referred to as DDR as opposed to DDR SDRAM. DDR is just like SDRAM except it features greater speed. Maximum transfer rate to L2 cache is approximately 1,064 MBps (for DDR SDRAM 133 MHZ).

Since DDR reads data on both the rising and falling edges of the clock signal and SDRAM only carries information on the rising edge of a signal, the DDR module transfers data twice as fast as SDR SDRAM. For example, instead of a data rate of 133MHz, DDR memory transfers data at 266MHz. DDR is not backward compatible with SDRAM-designed motherboards.

SLDRAM
Synchronous Link Dynamic Random Access Memory or SLDRAM is a type of SDRAM that uses a multiplexed command bus allowing fewer pins to increase bandwidth and allow higher Front Side Bus (FSB) speeds.

RDRAM
Rambus DRAM is a revolutionary type of DRAM that uses a 16 - 18 bit data path and is designed to operate with FSB speeds of 800MHz, producing a burst transfer rate of 1.6 GHz. RDRAM uses a Rambus in-line memory module (RIMM), which is similar in size and pin configuration to a standard DIMM. What makes RDRAM different than other types of memory is its use of a special high-speed data bus called the Rambus channel.

Since RDRAM uses an FSB of 800MHz, it would be easy to assume that RDRAM is faster than DDR SDRAM. That's not necessary true. Although RDRAM has an FSB speed nearly twice as fast as the highest rated DDR, RDRAM transmits in bursts with some latency between bursts.

Using trains as an analogy… The DDR train moves along at a top speed of 50 miles per hour and has 10 full passenger cards. The RDRAM moves twice the speed, but only has half the passenger cars. Although the RDRAM train moves faster, in the end both trains move the same amount of data… I mean people.

VRAM
VideoRAM is sometimes called Multiport Dynamic Random Access Memory (MPDRAM). VRAM is a type of RAM specifically used for video adapters and 3D graphic accelerators. VRAM features two independent access ports instead of one that's found in the other types of RAM. This allows the CPU and graphics processor to access the RAM at the same time. VRAM is typically located on the graphics card or the main board if the motherboard has an onboard graphic interface. The amount of VRAM determines the resolution and color depth of the display. True MPDRAM is pretty expensive so most graphics cards use SGRAM (synchronous graphics RAM) instead. Performance is nearly the same, but SGRAM is cheaper.

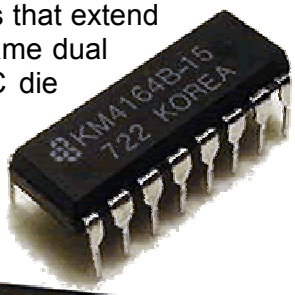
Memory Specifications
Microprocessor manufacturers such and Intel and AMD develop specifications so memory manufacturers build products that work with their motherboards and processor chip sets. The most common specification used to describe memory standards is the *Front Side Bus (FSB)*. Specified in terms of bit-width and its speed in MHz, the FSB is the communication bus that runs between the processor, L2 cache, chip set, DRAM and the AGP socket. Below is a list of common memory specifications found in PCs that use Intel processors and chip-sets.

- **PC-100** The PC-100 specification defines the requirements for SDRAM modules used on 100MHz FSB motherboards.
- **PC-133** The PC-133 specification details the requirements for SDRAM modules used on 133MHz FSB motherboards. PC133 SDRAM can be used on 100MHz FSB motherboards but will not yield a performance advantage over PC100 memory at 100MHz.
- **PC-2100** The PC-2100 specification details the requirements for 184 Pin DIMM DDR modules used on 133MHz (266MHz) FSB motherboards. Capable of up to 2.1GbS of bandwidth.
- **PC-2700** The PC-2700 (DDR333) specification details the requirements for 184 Pin DIMM DDR modules used on 166MHz (333MHz) FSB motherboards. Capable of up to 2.7GbS of bandwidth… an increase of 25% above the top operating speed of the PC-2100 specification.
- **PC-3200** The PC-3200 (DDR400) specification details the requirements for 184 Pin DIMM DDR modules used on 200Hz (400MHz) FSB motherboards. Capable of up to 3.2GbS of bandwidth.
- **PC-3500** The PC-3500 (DDR433) specification details the requirements for 184 Pin DIMM DDR modules used on 166MHz (433MHz) FSB motherboards. Capable of up to 3.5GbS of bandwidth.
- **PC-4000** The PC-4000 specification details the requirements for 184 Pin DIMM DDR modules used on 250MHz (500MHz) FSB motherboards. Capable of up to 4.0GbS of bandwidth.
- **PC-4200** The PC-4200 specification details the requirements for 184 Pin DIMM DDR modules used on 266MHz (433MHz) FSB motherboards. Capable of up to 4.2GbS of bandwidth.

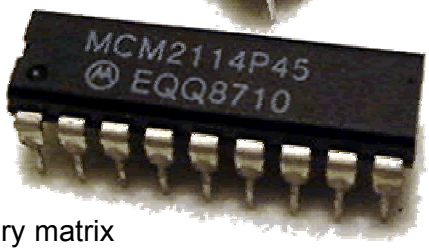
- **PC-800** The PC-800 specification details the requirements for 184 Pin RIMM PC-800 modules used on 800MHz FSB motherboards. Data bus width multiplexed down to 16 to 18 bit widths instead of the 64 to 72 bit widths as used with DDR modules. Another difference with Rambus memory is that all memory slots in the motherboard must be populated. Even if all the memory is contained in a single module, the "unused" sockets must be populated with a PCB, known as a continuity module, to complete the circuit. . Capable of up to 1.6GbS of bandwidth.

Memory Module Package Types

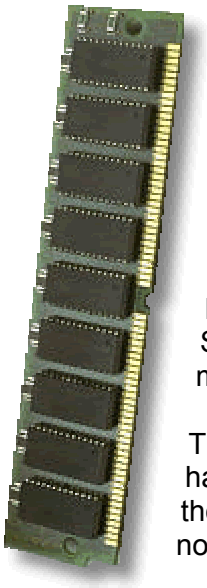
16, 18, 24 Pin DIP Modules Vary in Size Approx .75” x 5/16” (19mm x 8mm) is shown
A dual inline package (DIP) is a discrete electronic component. Built into the package is the integrated circuitry that makes up the matrix of memory cells. The tin pins that extend from the side of the package form two rows that are inline, hence the name dual inline package. The pins form the communication paths between the IC die inside the module to the system.



DIP memory chips are commonly found in first generation desktop computers, such as the 8088, Z80, 6800, 286, 386, and 486 processor based computers. By today’s standards, DIP memory holds miniscule amounts of data. Depending on their design, DIP chips provide various storage capacities and bit-widths. The 2114 Static RAM is a 1K by 4 bit memory chip operating at 200ns. Meaning… You would need two of them to provide 1KB of storage for an 8 Bit microprocessor. The 4164 Dynamic RAM is 64K x 1 bit. Therefore, you would need 8 of these chips to form a memory matrix 64KB in size. The notch on the left and the small indentation in the lower left corner of the chip indicates pin 1 of the IC.



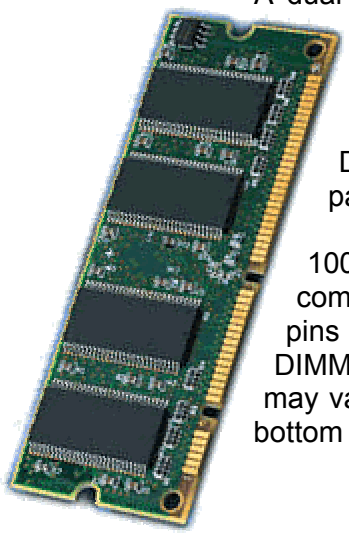
72-pin SIMM Module Size Approx. 4.25" x 1" (10.8cm x 2.54cm)
A single inline memory module (SIMM) consists of a number of memory components (usually black) that are attached to a printed circuit board (usually green). The gold or tin pins on the bottom of the SIMM provide a connection between the module and a socket on a larger printed circuit board. The pins on the front and back of a SIMM are connected to each other, providing a single line of communication paths between the module and the system.



72-pin SIMMs are commonly found in older desktop computers, such as the 486 and early Pentium models. Each 72-pin SIMM provides a 32-bit data path, so they can be installed singly in 32-bit systems (486 models) but must be installed in pairs in 64-bit systems (Pentium and Athlon models). 72-pin SIMMs are available in FPM or EDO. When upgrading, be sure to match the memory technology that is already in your system.

The number of black components on a 72-pin SIMM may vary, but they always have exactly 72 pins. 72-pin SIMMs are approximately 4.25" long and 1" high, though the heights may vary. They have one notch on the bottom left and one notch in the center of the module.

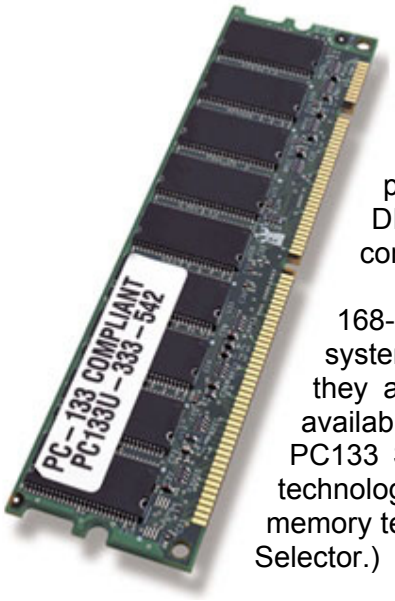
100-pin DIMM Module Size Approx. 3.5" x 1.25" (8.89cm x 3.175cm)



A dual inline memory module (DIMM) consists of a number of memory components (usually black) that are attached to a printed circuit board (usually green). The gold or tin pins on the bottom of the DIMM provide a connection between the module and a socket on a larger printed circuit board. The pins on the front and back of a DIMM are not connected, providing two lines of communication paths between the module and the system.

100-pin DIMMs are commonly found in printers. The number of black components on a 100-pin DIMM may vary, but they always have 50 pins on the front and 50 pins on the back for a total of 100. 100-pin DIMMs are approximately 3.5" long and 1.25" high, though the heights may vary. They have two small notches within the row of pins along the bottom of the module.

168-pin DIMM Module Size Approx. 5.375" x 1.375" (13.65cm x 2.54cm)

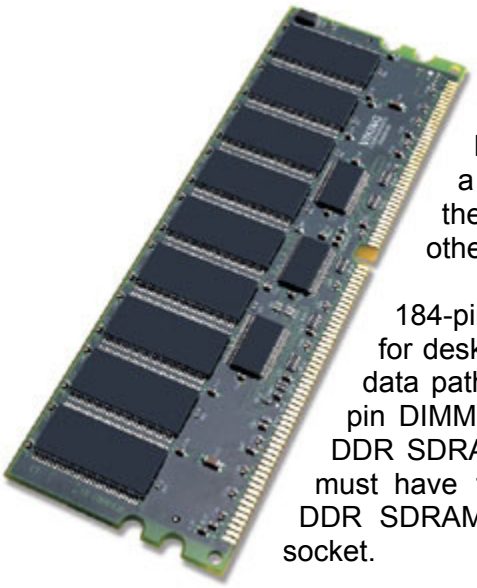


A dual inline memory module (DIMM) consists of a number of memory components (usually black) that are attached to a printed circuit board (usually green). The gold or tin pins on the bottom of the DIMM provide a connection between the module and a socket on a larger printed circuit board. The pins on the front and back of a DIMM are not connected, providing two lines of communication paths between the module and the system.

168-pin DIMMs are commonly found in Pentium and Athlon systems. Each 168-pin DIMM provides a 64-bit data path, so they are installed singly in 64-bit systems. 168-pin DIMMs are available in FPM, EDO, 66MHz SDRAM, PC100 SDRAM, and PC133 SDRAM. When upgrading, be sure to match the memory technology that is already in your system. (Information on which memory technology is used by your system is included in the Memory Selector.)

The number of black components on a 168-pin DIMM may vary, but they always have 84 pins on the front and 84 pins on the back for a total of 168. 168-pin DIMMs are approximately 5.375" long and 1.375" high, though the heights may vary. They have two small notches within the row of pins along the bottom of the module.

184-pin DIMM Module Size Approx. 5.375" x 1.25" (13.65cm x 2.54cm)



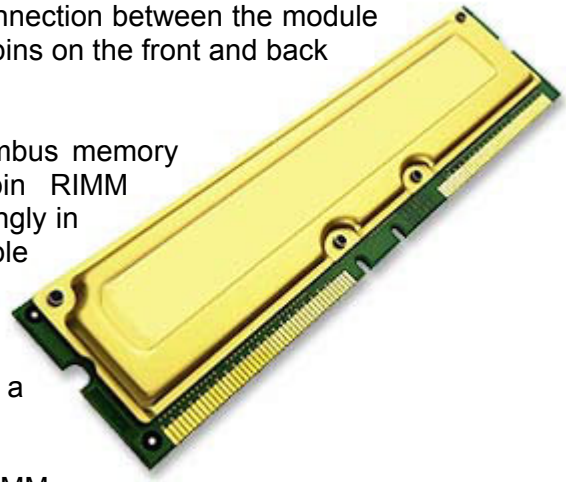
A dual inline memory module (DIMM) consists of a number of memory components (usually black) that are attached to a printed circuit board (usually green). The gold or tin pins on the bottom of the DIMM provide a connection between the module and a socket on a larger printed circuit board. The pins on the front and back of a DIMM are not connected to each other.

184-pin DIMMs are used to provide DDR SDRAM memory for desktop computers. Each 184-pin DIMM provides a 64-bit data path, so they are installed singly in 64-bit systems. 184-pin DIMMs are available in PC2100 DDR SDRAM or PC2700 DDR SDRAM. To use DDR memory, your system motherboard must have 184-pin DIMM slots and a DDR-enabled chipset. A DDR SDRAM DIMM will not fit into a standard SDRAM DIMM socket.

The number of black components on a 184-pin DIMM may vary, but they always have 92 pins on the front and 92 pins on the back for a total of 184. 184-pin DIMMs are approximately 5.375" long and 1.25" high, though the heights may vary. While 184-pin DIMMs and 168-pin DIMMs are approximately the same size, 184-pin DIMMs have only one notch within the row of pins.

184-pin RIMM Module Size Approx. 5.375" x 1.25" (13.65cm x 2.54cm)

A Rambus inline memory module (RIMM) consists of a number of memory components (usually black) that are attached to a printed circuit board (usually green). The gold or tin pins on the bottom of the RIMM provide a connection between the module and a socket on a larger printed circuit board. The pins on the front and back of a RIMM are not connected to each other.



184-pin RDRAM RIMMs are used to provide Rambus memory (RDRAM) for desktop computers. Each 184-pin RIMM provides a 64-bit data path, so they are installed singly in 64-bit systems. 184-pin RDRAM RIMMs are available for Rambus memory only. To use RDRAM memory, your system motherboard must have 184-pin RDRAM RIMM slots and a RDRAM-enabled chipset. A RDRAM RIMM will not fit into a standard DDR SDRAM DIMM socket.

The number of black components on a 184-pin RIMM may vary but are typically covered by a metal heat spreader. The modules always have 92 pins on the front and 92 pins on the back for a total of 184. 184-pin RDRAM RIMMs are approximately 5.375" long and 1.25" high, though the heights may vary. While 184-pin RDRAM RIMMs and 184-pin DDR RIMMs are approximately the same size, 184-pin RDRAM RIMMs have two notches at the center of the card's edge connector.

PC Memory Capacities

There are two limitations that will determine the maximum amount of memory a PC can handle:

1. The operating system's maximum memory rating.
2. The system maximum is the amount of memory that's supported by the motherboard's chipset and design. This information is contained in the mainboard's user manual. In most cases, the system maximum is lower than the OS maximum.

Exceeding the either the OS or the system maximum will slow system performance and cause other processing errors. Here are the OS maximums for popular versions of Microsoft Windows:

Windows Version	**Maximum Memory**
Windows 95	2GB
Windows 98	4GB
Windows 98SE	4GB
Windows ME	4GB
Windows NT	4GB
Windows 2000 Professional	4GB
Windows 2000 Advanced Server	4GB 8GB with PAE enabled
Windows 2000 Datacenter Server	4GB 64GB with PAE enabled
Windows XP Home:	4GB
Windows XP Professional	4GB

Conclusion

Motherboard and processor designers develop memory specifications so manufacturers can build memory products that will work across a broad range of PC mainboards. It's important to realize that mainboards are designed to support specific types of memory and memory modules cannot be mixed and matched on the same mainboard in any system. Not only will they not function, the may not even fit into the sockets. It's vital to match the memory specifications against the requirements of the mainboard before installing any memory.

Whether it's FPM, EDO, SDRAM, DDR, RDRAM, parity, or Non-ECC, the right type of memory is the one that's designed to work with the mainboard. That's why it's so important to be knowledgeable about memory types, specifications, and package types. One thing's for sure though… The more memory installed in a PC, the better it will operate.

Gross Anatomy Guide

Exercise 3 – Computer Memory

Exercise Objective:
At the completion of this exercise, you will demonstrate your knowledge of computer memory operation, memory types, memory specifications, and package types. You will correctly answer a variety of questions regarding the identification, operation, and selection of computer memory.

Name:	
Period:	
Date:	

Research Resources:

Company	Web Site	Description
Webopedia	http://www.webopedia.com/	Internet technology encyclopedia.
Techweb	http://www.techweb.com/	Internet technology encyclopedia.
Crucial	http://www.crucial.com	Crucial Memory Sales
Rose-Hulman College	http://www.rose-hulman.edu/Class/ee/yoder/ece332/Papers/RAM Technologies.pdf	RAM Technologies Description
Intel Corporation	http://www.intel.com/intel/intelis/museum/exhibits/memory_tech/index.htm	Memory Technology

Vocabulary:
On a separate piece of paper, decipher the following acronyms. Be sure the terminology you use is in the same context as computer memory technology.

- SAM
- FPM
- MOS
- SIMM
- FSB
- RAM
- EDO
- SDR SDRAM
- DIP
- SRAM
- BEDO
- DDR
- DIMM
- DRAM
- ECC
- RDRAM
- RIMM

Questions:

1. What are some examples of SAM technology and why was it widely abandoned in favor of RAM technologies?

2. Core memory was the very first type of random access memory. How did it work and why did it become obsolete?

3. Complete the following table of memory specifications, types, and package types:

Specification	FSB	Bandwidth	Memory Type(s)	Package Type(s)
PC-100				
PC-133				
PC-2100				
PC-2700				
PC-3200				
PC-3500				
PC-4000				
PC-4200				
PC-800				

4. Reorder the following memory types from the highest to the lowest speed.

Memory Types	In Order By Speed
SDRAM	1.
RDRAM	2.
SRAM	3.
DDR	4.
ED0 DRAM	5.
FPM DRAM	6.
SLDRAM	7.

5. What is the basic difference between Video RAM and main memory?

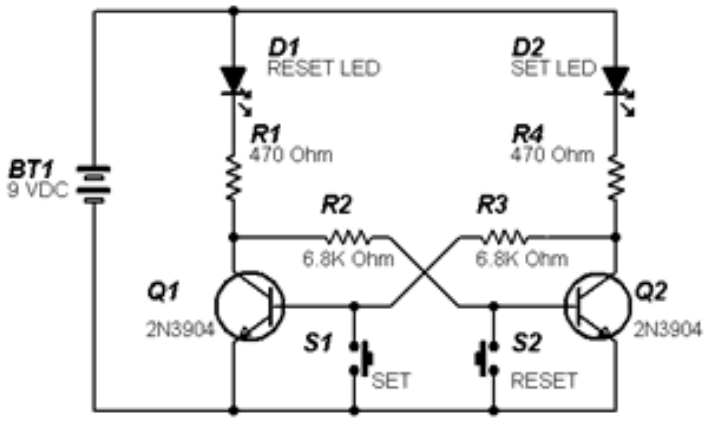
6. Consider the memory specification PC-4000. What does the numeric portion (4000) of the specification code mean?

7. What two limitations determine the maximum memory capacity of a PC?

8. What's the maximum memory allowed when using Windows 2000 Professional?

9. Which is faster… DDR SDRAM or RDRAM?

10. Use an electronics part kit to build and test this transistor Set/Reset flip-flop memory. Circle the correct LED state based upon the switch settings.



	Switch	LED	
Set:	Closed	ON	OFF
Reset:	Open	ON	OFF
Set:	Open	ON	OFF
Reset:	Closed	ON	OFF