

(10)

Called - many integrated wire (MIC)
• The multiple and MIC strategy involves a homogeneous collection of general purpose, processors, on a single chip.

GPU, GPGPU

A GPU is a core designed to perform parallel operation on graphics data. It is found on a plug in graphics card. It is used to encode and render 2D and 3D graphics as well as process video.

• GPUs perform parallel operation on multiple sets of data.

They are increasingly being used as vector process for a variety of application. That require repetitive computations.

1) Quick Path interconnect (QPI) Protocol Layer)

In this layer the Packet is defined as the unit of transfer.

(Ans D)

Cortex - A

The Cortex - A and Cortex - A50 are application processor intended for mobile devices such as smart phones and e-book readers as well as consumer device such as digital TV and cable internet modems. These processor run at higher clock frequency and support a memory management unit (MMU).

←—————→

Cortex - R is

The Cortex - R is designed to support real-time applications in which the timing of events needs to be controlled with rapid response to events. They can run at a fairly high clock frequency and have very low response latency.

Cortex - M is

Cortex - M series processor have been developed primarily for the micro controller domain where the need for fast, highly deterministic interrupt

- is stored in AC.
- (a) The value in the PC (30) is loaded into the MAR
- b) The value in Location 302 (which is the instruction with the value 294) is loaded into the MBR and the PC is incremented
- c) The value in the ~~MBR~~ MBR is loaded into the IR.
- b) a) The address (portion of the IR (94)) is loaded into the MAR.
- b) The value in the AC is loaded into the MAR
- c) The value in the MBR is stored in Location 941

Q. 4 (Ans F)

(a) $2^{24} = 16 \text{ Mbytes}$

(b) (i) If the local address bus is 32 bits the whole address can be transferred at once and decoded in memory. However, because the data bus

(3)

- RU (Recovery unit)
The RU keeps a copy of the complete state of the system that includes all registers, collects hardware fault signals etc.
- COP
The COP is responsible for the data compression and encryption function for each etc.
- I-cache
This is a 64-kB L1 instruction cache allowing IFU to prefetch instructions before they are needed.
- L2 Control
This is a control logic that manages the traffic through the two ~~manages~~ L2 caches.

Data L2

A 1-MB data cache for all memory traffic other than instruction.

Inst - L2:

A 1-MB L2 instruction cache.

(8)

- memory latency and throughput.
memory ~~increase~~ access speed
(latency) and transfer speed.
throughput lag processor speeds.

(6) Consider a program running on a single processor such that a fraction $(1-f)$ of the execution is a function code that is inherently sequential and a fraction f of that involves code is inherently parallelizable with no scheduling overhead.

Let T be the total execution time of the program using a single processor. Then the speed up using a parallel processor with N processors that fully exploits the parallel portion of the program is as follows:

Speed up. $\frac{\text{Time to execute Program on a Single Processor}}{\text{Time to execute Program on } N \text{ Parallel processors}}$

Time to execute Program on N Parallel processors.

(9)

$$\text{Speed up} = \frac{T(1-f) + Tf}{\frac{T(1-f) + Tf}{N}} = \frac{1}{1-f + \frac{f}{N}}$$

(10) Multicore:

The use of multiple processor on the same chip, also referred to as multiple cores, or multicore also referred to the potential to increase performance without increasing the clock rate.

If the software can support

the effective use of multiple processor then doubling the number of processor almost double performance.

Two core chips were quickly followed by four-core chips then 8, then 16, and so on.

(MTC)

The leap in performance as well as the challenge in developing software to exploit such a large number of cores has led to the introduction of new terms

(b)

2.2) The techniques used in Contemporary processors to increase speed are following.

- Pipelining This enables a processor to work simultaneously on multiple instruction by performing a different phase each of the multiple instruction at the same time.

- Branch Prediction

Branch Prediction potentially increase the amount of work available for the processor to execute.

- Super scalar Execution:

This is the ability to issue more than one instruction in every processor clock cycle. In effect multiple parallel pipelines are used.

- Data flow analysis:

The processor analyze which instruction are dependent on each other's result or data to create an optimized schedule of instruction.

(7)

• Speculative execution :-

This enables the processor to keep its execution engines as busy as possible by executing instructions

that are likely to be needed.

←—————→
aim f) The problems created due to increase in clock speed and logic density of the processor are.
Power. As the density of logic and the clock speed on a chip increase, so does the Power density the difficulty of dissipating the heat generated on high density high speed chips is becoming a serious design issue.

R C Delay

The speed at which electrons can flow on a chip between transistor is limited by the resistance and capacitance of the metal wires connecting them specifically delay increase as the RC product increase.

(5)

Ques 1) Embedded system is

The term embedded system refers to the use of electronics and software within a product as opposed to a general purpose computer such as laptop or desktop systems. Today many devices that use electronic power have an embedded computing system.

Different embedded system used in everyday life are cell phones, digital cameras, video cameras, calculator, microwave ovens, home security system

← →
Ques 2) Different desktop application that require the great power of contemporary microprocessor based system.

- image processing
- three-dimensional rendering
- speech recognition
- video conferencing
- multimedia authoring
- voice and video annotation of field
- Simulation modeling

1) IAS operation

The IAS operation by repeatedly performing an instruction cycle. Each instruction cycle consists of two subcycles.

1: Fetch cycle:

During fetch cycle the opcode of the next instruction is loaded into the IR and the address portion is loaded into the IR and the MAR. This instruction may be taken from the IBR or it can be obtained from memory by loading a word into the MBR and then down to the IBR and MBR.

Execute cycle:

Control circuitry interpret the opcode and encodes the instruction by sending out the appropriate control signal to cause data to

be moved or an operation to be performed by the ALU.

(2)

- LSU: (Load store unit)
It is responsible for handling all types of operand access of all length modes and formulas as defined in the architecture.

- XU (Translation Unit)
This unit translates logical address from instruction into physical address in main memory.

FXU (Fixed Point Unit)

The FXU executes fixed point arithmetic operation.

BFU (Binary Floating Point Unit)

The BFU handles all binary and hexadecimal floating point operation as well as fixed point multiplication operation.

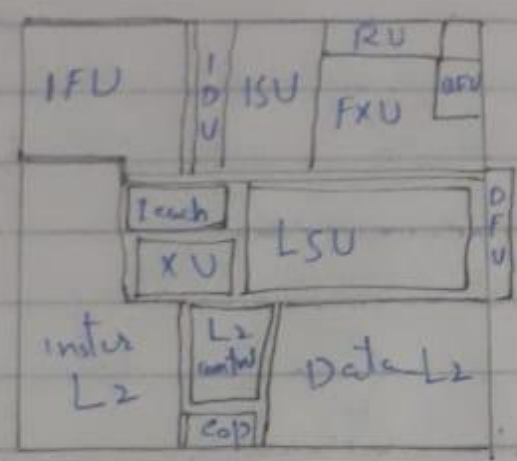
DFU Decimal Floating Unit.

The DFU handles both fixed point and floating operation on numbers that are stored as decimal digits.

①

Name: 15269 Anwaris Ghafar
 ID #: 15269
 Teacher #: M. Amin
 Date: 27-04-2020

(Q no. 1)



Function of each Sub area:

• ISU (Instruction Sequence Unit)

Determine the sequence in which instructions are executed in what is referred to as a superscalar architecture.

• IFU (Instruction Fetch Unit)

Logic for fetching unit.

• IDU: (Instruction Decode Unit)

The IDU is fed from the IFU buffers and is responsible for the parsing and is responsible for the architecture operation codes.

Name : Anvita Shafiq

ID : 15269

Paper : Computer architecture.

38)

$$T_s = 1.8 \text{ ms}$$
$$\text{speed up} = 11/1.8.$$

(d)

By depending on the information given it is not obvious how to quantify this effect in Amdahl's equation, therefore if it is supposed that the fraction of code which is parallelizable is $f = 1$ then Amdahl's decreases to $\text{speed up} = N = 8$.
Therefore, the actual speed up is only 75% of the theoretical speed up.

The END

(Q: 4 Part d)

Instruction type	CPI	Instruction. mix
Arithmetic & logic	1	60%
Load/store with cache hit	2	18%
Branch.	4	12%
memory reference with cache miss	12	10%

The average CPI = $(1 \times 0.6) + (2 \times 0.18) + (4 \times 0.12) + (12 \times 0.1) = 2.64$

Therefore CPI has been increased since the time for memory access is also increased

(b) $\text{mips} = 400 / 2.64 = 152$

There is corresponding drop in the mips rate

(c) $T = T_c (\text{mips} \times 10^6)$

for the One processor

$$T_1 = (2 \times 10^6) / (178 \times 10^6) = 11 \text{ ms}$$

for the processor each processor execute $1/8$ if the 2 million instruction plus the 25,000

36

(Part c)

Effective CPI's

$$CPI = (1 \times 46000) + (2 \times 33000) + (2 \times 16000) \\ + (2 \times 9000) / 100$$

$$CPI = 162000 / 100$$

$$CPI = 1620$$

MIPS

$$\text{MIPS rate} = 60 \text{ MHz} / 1620 \times 10^6$$

$$\text{MIPS rate} = 60 \times 10^6 \text{ Hz} / 1620 \times 10^6 \\ = 60 \text{ Hz} / 1620$$

$$\text{MIPS rate} = 0.037$$

Execution Time ∴

$$I = Ic / (\text{MIPS} \times 10^6)$$

$$T = 104000 / (0.037 \times 10^6)$$

$$\bar{T} = 104000 / 37 \times 10^6$$

$$\bar{T} = 2811 \times 10^{-3}$$

$$\bar{T} = 2.811 \text{ Sec}$$

(35)

(Q4 Part B)

(a) opcode 000000
operand 00000000010

(b) At the beginning the CPU have to fetch the instruction from the memory, then the instruction will include the address of the data which is required to load. Through the execution time, the memory will be accessed in that time to load the data contents which is located at that address for a total of two trips to memory.

(34)

in only 16 bits. It will require 2 cycle to fetch a 32-bit instruction or operand.

2) The 16 bit of the address placed on the address bus can't access complex memory. Thus, a more complex memory interface control is needed to latch the first part of the address and then the second part.

c) The Program Counter must be at least 24 bits. Typically a 32-bit microprocessor will have a 32-bit external address bus, and a 32-bit program counter. Unless on-chip segment registers are used that may work the instruction register is to contain the whole instruction. It will have to be 32-bit long. It will contain only the Op code (called the op code register then it will have to be 8 bits long.

(32)

is loaded into the IR.

2) a) The address portion of the IR (940) is loaded into the MAR

b) The value in location 940 is loaded into the MBR.

c) The value in the MBR is loaded into the AC

3) a) The value in the PC (301) is loaded into the MAR.

b) The value in location 301 (which is the instruction with the value 5941) PC is incremented.

c) The value of in the MBR is loaded into the IR.

4) a) The address portion of the IR (941) is loaded into the MAR

b) The value in location 941 is loaded into the MBR.

c) The old value of the AC and the value of location MBR are added and the result

(31)

(Q. No. 4) (G)

A bus cycle takes 0.25 μ s so a memory cycle takes 2 μ s. If both operands are even aligned, it takes 2 μ s to fetch the two operands. If one is odd aligned, the time required is ~~3~~ 3 μ s. If both are odd aligned, the time required is 4 μ s.

←—————→
(Ans E)

Six steps

- ① The PC contains 300, the address of the first instruction. This value is loaded into the MAR.
- ② The value in location 300 (which) is the instruction with the value of 1940 in hexadecimal is loaded into the MAR and the PC is incremented. These two steps can be done parallel.
- ③ The value in MBR is

RH instruction:

21 = 00100001 = STOR (M(n))

$m(x)$ = refers to the memory address
Location OFB

The second 5 bits of o8c
should read - STOR (M(OFB))

Finally the assembly language
code for o8c OLO FAH OFB is
LOAD -M(OFA)
STOR (M(OFB))

1) In o8A address the $M(OFA)$
transfer to the accumulator and
transfer contents of accumulator to
memory location OFB.

In o8B address the $M(OFA)$ transfer
to the accumulator and take
next instruction from left half of
 $M(O8D)$.

In o8c address the $M(OFA)$ transfer
to the accumulator and transfer
contents of accumulator to memory

Location OFB

(29)

Here is simple way to understand this problem

Contents are divided up into two bit 5 bit instruction LH and RH

LH instruction = 010FA

opcode = 01

address = 0FA

RH instruction = 0F08D

Contents are divided up to two 5 bits instruction LH and RH

LH instruction 020FA

opcode = 02

address = 0FA

RH instruction = 210FB

opcode = 21

address = 0FB

LH instruction;

01 = 0000010 = LOAD - M(x)

m(x) refers to the memory address

location 0FA.

The first 5 bits of 08c should read Load M(0FA).

(28)

opcode = 01

address = 0FB

Since this is hexadecimal from you have to convert this number to binary form.

LH instruction;

01 = 00000001 = LOAD M(x)

m(x) refers to the memory address Location of A

The first 5 bits of 08A should read - LOAD M(0FA)

RH instruction;

21 = 00100001 = STOR M(x)

m(x) refers to the memory address Location = 0FB

The second 5 bits of 08A should read - STOR M(0FB)

Finally the assembly language code for 08A 010FA 210FB is

LOAD (M(0FA))

STOR M(0FB)

(27)

Programming in Software is

The new method of programming which is a sequence of codes or instructions is called Software Programming.

In this method programming is much easier instead of rewiring the hardware for each new programme all we need to do is provide a new sequence of codes each code is in effect an instruction and part of the hardware interprets each instruction and generate control signal.

Q: NO: 4

(A) (a)

1. Here is simple way to understand this problem contents are divided up into two 8 bits instruction LH and RH
LH instruction: 010FA.

(26)

checked by the processor after the processor has enabled interrupts when a user program is executing and an interrupt occurs interrupts are disabled immediately.

Nested interrupt:

Nested interrupt is to allow interrupt of higher priority to cause a lower priority interrupt handler to be itself interrupted.

A user program begins at $t=0$. At $t=t_0$, a printer interrupt occurs. User information is placed on the system stack and execution continues at the printer interrupt service routine (ISR) while this routine services is still executing at $t=t_1$, a communication interrupt occurs.

(25)

management is coupled with the desire for externally low gate count and lowest possible power consumption.

(Ans E)

In the interrupt cycle, the processor checks to see if any interrupts have occurred indicated by the presence of an interrupt signal. If no interrupts are pending the processor proceeds to the fetch cycle and fetches the next instruction of the current program.

(Ans F)

Disabled interrupt :-

A disabled interrupt simply means that the processor can and will ignore that interrupt request signal if an interrupt occurs during this time. It generally remains the pending and will be

(Ans C)

Microprocessor :-

A microprocessor chips included register on ALU and some sort of control unit or instruction processing logic. As transistor density increased it became possible to increase the complexity of the instruction set architecture and ~~ultimately~~ ultimately set architecture and ultimately to add memory and more than one processor chips include multiple cores and substantial amount of cache memory.

Microcontroller :-

A microcontroller is a single chip that contains the processor non-volatile memory for the program (ROM) volatile memory for input and output (RAM) a clock and an I/O control unit. The processor portion of the microcontroller has a much lower silicon area than other microprocessor and much higher energy efficiency.

of design effort on
Complex instruction Set:

CISC The x86 incorporates the
Sophisticated design principles one
found only on main frames and
super computer and serves as an
excellent example of CISC design.

RISC:

An alternative approach to processor
design is the reduced instruction
set computer (RISC)

The ARM variety of embedded
is used in a wide variety
of embedded systems and is one
of the most powerful and
best designed RISC.

Based system on the market
In this section and the next
we provide a brief overview
of these two systems.

(21)

Q. no. 3

(Ans A)

Computer architecture:

Computer architecture refers to those attributes of a system visible to programs or, put another way, those attributes that have a direct impact on the logical execution of a program. A term that is often used interchangeably with Computer architecture is instruction set architecture (ISA).

Computer organization:

Computer organization refers to the operational unit and their interconnections that realize the architectural specification.

(Example of Architectural attributes include the instruction set the number of bits used to represent various data types (e.g., numbers).

I/O mechanism and techniques for addressing memory.

1) (ISC) The current x86 flag represents the result of decades

Data Lines:

The data lines provide a path for moving data among system modules. These lines collectively are called the data bus. The data bus may consist of:

- Address Lines:

The address lines are used to designate the source or destination of the data on the data bus. The width of the address bus determines the maximum possible memory capacity of the system.

- Control Lines:

The control lines are used to control the access and the use of the data and address lines. Because the data and address lines must be a means of controlling their use, typically control lines include:

memory write, memory read,
I/O write I/O read reset etc.

Hardware failure:

It is generated by a failure such as power failure or memory parity error.

(Ans 17)

Bus interconnection scheme:

The most common computer interconnection structures are based on the use of one or more system buses.

A system bus exists typically of about fifty to hundreds of separate lines. The lines can be classified into three functional groups: data, address and control lines.

Data Lines:

The data lines provide a path for moving data among system modules. These lines, collectively, are called the data bus.

The data bus may exist of ~~of~~ 32, 64, 128, or even more separate wires referred to as width of data bus.

(18)

Data operand (do):

Perform the operation indicated in the instruction.

• Operand store (os):

write the result into memory or out to I/O.

← →

(Ans F)

Classes of interrupts:

• Program:

It is generated by some condition that occurs as a result of an instruction execution such as arithmetic overflow, division by zero, attempt to execute an illegal machine instruction or reference outside a user's allowed memory space.

Timer:

It is generated by a timer within the processor. This allows the operating system to request service from the processor, or to signal a variety of error conditions.

Ans (E).

Instruction cycle state Diagram:

The states in instruction cycle state diagram are follows:

- **Instruction address calculation (Iac)**
Determine the address of the next instruction to be executed usually this involves adding a fixed number to the address of the previous instruction.
- **Instruction fetch (If):**
Read instruction from its memory location onto the processor.
- **Instruction operation decoding (Iod)**
Analyze instruction to determine type of operation to be performed and operand to be used.
- **Operational address calculation (Oac)**
If the operation involve reference to an operand in memory or available via I/O then determine the address of the operand.
- **Operation fetch (Of):**
Fetch the operand from memory or read it in the from I/O

Ans (c)

Stored Program Computer:

A fundamental design approach first implemented in the IAS Computer is known as the "stored program concept".

This idea is usually attributed to the mathematician John von Neumann.

The first publication of the idea was in a 1945 proposal by von Neumann for a new computer.

The EDVAC (Electronic Discrete Variable Computer.)

In 1946, von Neumann and his colleagues began the design of a new stored program computer referred to as the IAS Computer at the Princeton Institute for Advanced Studies, it consists of:

- ⇒ A main memory which stores both data and instruction.
 - ⇒ An arithmetic and logic unit (ALU) capable of operating binary data.
-

(1)

Ans B The characteristics of computer family are as follows.

- Similar or identical instruction sets:
In some cases the lower end of that is a subset of that of the top end of the family. This means that programs can move up but not down.
- Similar or identical operating systems:
The same basic operating system is available for all family members.
- Increasing Speed:
The rate of instruction execution increases in going from lower to higher family members.
- Increasing number of I/O ports:
The number of I/O ports increases in going from lower to higher family members.
- Increasing memory sizes:
At a given time the cost of a system increases in going from lower to higher family members.

(Ans D)

Moore's Law is

The famous Moore's law which was proposed by Gordon Moore's co-founder of Intel, in 1965, Moore observed that the number of transistor that could be put on year and correctly predicted that this pace would continue into the near future.

The consequences of Moore's Law are

The cost of computer logic and memory circuitry has fallen at a dramatic rate.

The computer become smaller making it more convenient place in a variety of environment. There is a reduction in power requirement.

The interconnection on the integrated circuit are much more variable than solder connections.

(13)

2. main memory
It stores data.

Q no 2
Ans (A)

Structural components of computer:
These are four main structural components of computer.

1) Central processing unit (CPU)
It controls the operation of the computer and performs its data processing function often simply referred to as processor.

2. Main memory:
It stores data.

3. I/O: It moves data between the computer and its central equipment.

4. System interconnection:
Some mechanism that provides for communication among CPU, main memory and I/O. A common example of system interconnection is by means of a system bus.

PCIe links from the chipset may attach to the following kinds of devices that implements PCIe.

* Switch:

The switch manages multiple PCIe streams.

* PCIe endpoint

An I/O device or Controller that implements PCIe.

Such as a gigabit ethernet switch or graphics, disk, interface, etc

Legacy endpoint

Legacy endpoint category is intended for existing designs that have been migrated to PCI express and it allows legacy behaviours. Such as use of I/O space and locked transaction.

PCIe / PCI bridge

Allow older PCI devices to be connected to the PCIe-based system.

(17)

The Packet contents definition is standardized with some flexibility allowed to meet differing market segment requirements.

One key function performed at this layer is a cache coherence protocol, which deals with making sure that main memory values held in multiple caches are consistent.



1) Physical and logical Architecture of PCIe:

Root Complex:

It is also called chipset or a host bridge which connects the processor and memory subsystem to the PCI express switch fabric comprising one or more PCIe and PCI switch devices.

The root complex acts as a buffering device to deal with difference in data rates between I/O controllers and memory and processor components.