Pre-mid exam exercises – CE332 – Computer Architecture & Organization

Sub-CLO 1: Basic Computer Functions & Structures

1. Give an explanation about Moore's law.
   Answer : Moore's Law menyatakan bahwa jumlah transistor pada sebuah chip semikonduktor akan sekitar dua kali lipat setiap 18 hingga 24 bulan. **(Week 1 page 30)**

   Konsekuensi Moore's Law :
   - Peningkatan kinerja karena jumlah transistor yang terus bertambah
   - Penghematan konsumsi ruang karena transistor yang lebih kecil dan padat
   - Harga jual yang lebih murah karena produksi yang lebih efisien
   - Hubungan antar interchip yang lebih sedikit menyebabkan efisiensi kinerja
   - Jarak antar sinyal menjadi lebih pendek karena ukuran transistor yang lebih kecil menyebabkan konsumsi daya yang lebih sedikit. Daya yang lebih sedikit juga menghasilkan panas yang lebih rendah sehingga kebutuhan pendinginan berkurang.

2. What is the key distinguishing feature of a microprocessor? Alvin
   Answer : The key distinguishing feature of a microprocessor is that the microprocessor has the ability to integrate all of the essential components of the CPU onto a single chip.

3. Give explanation about Von Neumann computer concepts in modern computer systems today. Aditya
   Answer : Konsep Arsitektur Von Neumann adalah arsitektur komputer yang menggunakan memori yang sama untuk menyimpan data program dan data pengguna. Hal ini memungkinkan komputer untuk menjalankan berbagai aplikasi dan tugas dengan memuat dan menjalankan program-program berbeda tanpa perlu mengubah perangkat keras secara fisik. Arsitektur Von Neumann, yang dicetuskan oleh John von Neumann pada tahun 1940-an, merupakan dasar dari sistem komputer modern. Konsep ini terdiri dari beberapa komponen utama:
   1. Central Processing Unit (CPU): Otak komputer yang menangani operasi aritmatika dan logika. CPU modern memiliki beberapa inti (cores) untuk melakukan banyak tugas sekaligus.
   2. Memori: Data dan instruksi program disimpan dalam Random Access Memory (RAM) untuk diakses dengan cepat oleh CPU.
   3. Perangkat Input/Output (I/O): Perangkat modern seperti keyboard, monitor, dan antarmuka jaringan memungkinkan komunikasi dengan dunia luar.

   Komputer modern telah meningkatkan konsep-konsep ini dengan mengintegrasikan beberapa inti, meningkatkan kapasitas memori, dan mengoptimalkan I/O, menghasilkan sistem komputer yang kuat dan efisien yang digunakan dalam berbagai aplikasi saat ini. Ketiga modul tersebut dihubungkan dengan jalur komunikasi yang disebut system bus

Sub-CLO 2: Processor performance analysis and benchmark

1. List and briefly define some of the techniques used in contemporary processors to increase speed. Dimas Takeda **(Week 2 page 6-7)**

- pipelining : instructions are processed simultaneously as in a production line
- branch prediction : the processor guesses what will be used next into the cache memory
- superscalar execution : the processor is able to execute multiple instructions in a single cycle
- data flow analysis : optimizing the instructions set by reorganizing the execution order
- speculative execution : temporary locations are made to contain the execution which was done based from the branch prediction to minimize idle time in the execution engine

2. Memory always lags in speed compared to a processor. Describe the efforts need to overcome the imbalance between memory and processor speeds Sifa **(Week 2 page 9)**
    - Increase the number of bits that are retrieved at one time
    - Reduce the frequency of memory access by incorporating increasingly complex and efficient cache structures between the processor and main memory
    - Change the DRAM interface to make it more efficient by including a cache or other buffering scheme on the DRAM chip
    - Increase the interconnect bandwidth between processors and memory by using higher speed buses and a hierarchy of buses to buffer and structure data flow
    -
3. Define the difference between MIPS and FLOPS Shadiq **(Week 2 page 25-26)**

    MIPS and FLOPS are both measures of processing power, but they have different things. which mean MIPS measures the number of instructions that a CPU can execute in a second, while FLOPS measures to the number of floating-point operations that an FPU can perform in a second.

    Example:
    A computer with **MIPS** 100 can execute 100 million instructions in one second.
    A computer with **FLOPS** 100 can perform 100 million floating-point operations in one second

4. Give explanation about benchmark concept Albert **(Week 2 page 27-28)**

    Benchmarks are used to summarize the performance of a system running examples of actual code.

    Benchmark Principles

    Desirable characteristics of a benchmark program:

    1. It is written in a high-level language, making it portable across different machines

    2. It is representative of a particular kind of programming domain or paradigm, such as systems programming, numerical programming, or commercial programming

    3. It can be measured easily

    4. It has wide distribution

    Here are some examples of how benchmarks can be used:

- A user is looking for a new laptop for gaming. They can use benchmarks to compare the performance of different laptops to see which one is the best for their needs.
- A company is considering upgrading its servers. They can use benchmarks to compare the performance of different servers to see which one would be the best fit for their needs.
- A software developer is developing a new game. They can use benchmarks to test the performance of their game on different hardware configurations.

Sub-CLO3: Top level view of computer - instruction execution

1. Describe the role of the registers (PC, MAR, MBR, IR, AC) during the instruction cycle (fetch and execution) M Prasetyo
   Answer:
   Fetch
   - Program Counter (PC): PC menyimpan alamat memori instruksi berikutnya yang akan dieksekusi. CPU membaca instruksi dari alamat ini kemudian dikirim ke MAR.
   - Memory Address Register(MAR): MAR adalah register yang berguna untuk menyimpan alamat memori yang akan diakses atau ditulis. Ketika CPU perlu membaca atau menulis data dari memori, alamat memori diambil dari MAR.
   - Memory Buffer Register(MBR) : MBR adalah register yang digunakan untuk menyimpan data yang telah diambil dari memori atau yang akan ditulis ke memori. MBR berfungsi sebagai perantara antara CPU dan memori untuk transfer data
   - Instruction Register (IR): Instruksi yang diambil dari memori dimasukkan ke dalam register IR. IR berisi opcode, yang merupakan kode operasi yang menentukan tindakan yang akan dilakukan, bisa juga berisi alamat memori(operand) yang diperlukan untuk mengeksekusi instruksi tersebut.
   - Accumulator(AC): AC adalah register untuk menyimpan hasil dari operasi aritmatika atau logika. Tempat ini disebut dengan penyimpanan sementara untuk hasil perhitungan dan nilai - nilai dalam AC untuk digunakan dalam operasi selanjutnya

   Execution
   - Program Counter (PC): Setelah mengambil alamat memori instruksi lalu dilanjutkan ke MAR kemudian PC akan melakukan increment 1 (+1) untuk mengambil memori instruksi selanjutnya
   - Memory Address Register(MAR): MAR mengambil data dari Program Counter kemudian membaca atau menulis data dari memori kemudian dikirim ke MBR
   - Memory Buffer Register(MBR) : MBR mengambil data dari MAR untuk menyimpan data yang telah diambil dari memori atau yang akan ditulis sebagai memori untuk transfer data
   - Instruction Register (IR): Instruksi yang diambil dari MBR yang berisi opcode yang telah dibaca atau ditulis kemudian diteruskan ke AC.
   - Accumulator(AC): AC mengambil data dari IR lalu disimpan sebagai tempat penyimpanan sementara dan nilai dalam AC akan digunakan dalam operasi selanjutnya.

2. List and briefly define the types of bus in computer structure David B **(Week 3 page 32-36)**
   Answer :
   - Data Bus : a Signal line that provides a path for moving / exchanging data between CPU and memory, and between CPU and I/O.
   - Control Bus : a signal line that used to control the access and the use of the data and address line.
   - Address Bus : a signal line that used to designate the source or destination of the data on the data bus.

3. List and briefly define two approaches to dealing with multiple interrupts. Indah **(Week 3 page 25-28)**
   Answer :
   - Disabled interrupts : setelah prosesor mulai menangani interupsi, prosesor akan mengabaikan interupsi yang lainnya. Jika interupsi lebih lanjut terjadi, biasanya interupsi tersebut akan tetap tertunda sehingga dapat ditangani setelah handler saat ini selesai. Dalam hal ini, memerlukan antrian interupsi. Pendekatan yang ini tidak memperhitungkan prioritas atau kebutuhan waktu yang kritis
   - Define interrupt priorities : yaitu interupsi dengan prioritas lebih tinggi dapat mengganggu pengendali interupsi dengan prioritas lebih rendah.

Sub-CLO 4 - cache memory

1. What is the general relationship among access time, memory cost, and capacity? Dimas Bagus **(Week 4 page 8)**
   Answer: For a practical system, the cost of memory must be reasonable in relationship to other components. Based on the memory hierarchy, the relationship goes like:
   - Faster access time, greater cost per bit;
   - Greater capacity, smaller cost per bit;
   - Greater capacity, slower access time;
2. A two-way set-associative cache has lines of 16 bytes and a total size of 8 kB. The 64-MB main memory is byte addressable. Show the format of main memory addresses. Kanza
3. Consider a direct mapped cache of size 16 KB with block size of 256 bytes. The size of the main memory is 128 KB. Find the number of bits in the tag. What are the tag and cache line address (in hex) for the main memory address (E201F)? Gilbert
4. with the same details as number 3 but the mapping technique change to fully-associative cache, find the tag address (in hex) for main memory address (CABBE) Gilbert

Sub-CLO 5 internal-external memory

1. What is the difference between DRAM and SRAM in terms of characteristics such as speed, size, density, cost, and power requirements? M Marshal
   Speed :
   - DRAM : DRAM memerlukan siklus penyegaran berkala untuk mempertahankan penyimpanan data. Karena karakteristik ini kecepatan DRAM lebih lambat dibanding SRAM

- SRAM : SRAM tidak memerlukan penyegaran data secara berkala seperti DRAM, ini membuat SRAM menjadi lebih cepat dan efisien dibanding DRAM. Biasanya SRAM digunakan sebagai cache memory karena karakteristiknya yang efisien dan cepat

Size :

- DRAM : Sel memory pada DRAM lebih kecil,karakter ini membuat DRAM lebih padat dalam kepadatan memory. Yang berarti sel ditempatkan ke fisik yang lebih kecil.
- SRAM : Sel memory pada SRAM lebih besar dibanding DRAM. Hal ini membuat sel SRAM secara fisik lebih besar dibanding DRAM.

Density :

- DRAM : mempunyai kepadatan lebih padat karena ukurannya yang lebih kecil per bit data dibanding SRAM.
- SRAM : mempunyai kepadatan yang lebih rendah dibanding DRAM karena ukuran yang lebih besar per bit data.

Cost :

- DRAM : DRAM lebih murah dibanding SRAM karena struktur memory yang lebih sederhana dibanding SRAM dan lebih efisien dalam penggunaan fisik seperti PC.
- SRAM : SRAM lebih mahal dibanding DRAM Karena SRAM mempunya struktur yang lebih kompleks dan mempunyai sel yang lebih besar dibanding DRAM.

Power Requirements :

- DRAM : DRAM memerlukan lebih banyak daya karena karakternya yang membutuhkan siklus penyegaran secara berkala.
- SRAM : dalam kebutuhan daya SRAM lebih efisien karena tidak memerlukan penyegaran secara berkala.

2. What are the differences among EPROM, EEPROM, and flash memory? Raden **(Week 5 page 11)**

EPROM	● Erasable programmable read-only memory ● Erasure process can be performed repeatedly ● More expensive than PROM but it has the advantage of the multiple update capability ● Erasure is performed by shining an intense ultraviolet light
EEPROM	● Electrically erasable programmable read-only memory ● Can be written into at any time without erasing prior contents

	● Combines the advantage of non-volatility with the flexibility of being updatable in place ● More expensive than EPROM
Flash Memory	● Intermediate between EPROM and EEPROM in both cost and functionality ● Uses an electrical erasing technology, does not provide byte-level erasure ● Microchip is organized so that a section of memory cells are erased in a single action or "flash"

3. Briefly define the seven RAID levels along with its advantages and disadvantages. Darryl **(Week 6 page 39-40)**

   Answer:

   RAID 0: splits data evenly across two or more disks without parity or redundancy.

   + No redundancy, data striped across all disk, round robin striping, increased speed
   - If one disk fails, all array is lost

   RAID 1: mirrors data across two or more disks.

   + Mirrored disks, data is striped accross disks, 2 copies of each stripe on separate disks, read from either, write to both, recovery is simple
   - Expensive, if one disk fails others can continue to operate

   RAID 2: uses a Hamming code to store error correction information across multiple disks.

   + Disks are synchronized, very small stripes, error correction, multiple parity disks store Hamming Codes
   - Lots of redundancy, expensive

   RAID 3: stripes data across multiple disks at the byte level and stores parity information on a dedicated disk.

   + Similiar to RAID 2, Stripes size are small, only one redundant disk, simple parity bit, data on failed drive can be reconstructed, very high transfer rates
   - Poor performance for random access and small transfers

RAID 4: stripes data across multiple disks at the block level and stores puarity information on a dedicated disk.

+ Each disk operate independently, good for high I/O rates, large stripes, bit by bit parity, parity stored on parity disk
- Poor performance for write operations due to the bottleneck of the parity disk

RAID 5: stripes data and parity information across multiple disks.

+ Similiar to RAID 4, parity striped across all disks, round robin allocation for parity stripe, avoids RAID 4 bottleneck at parity disk, used in network servers
- Has high overhead for parity calculation and disk rebuilds

RAID 6: stripes data and double parity information across multiple disks.

+ Two parity calculations, stored in separate blocks on different disks, user requirement of N disks need N + 2, high data availability
- Lower performance and higher overhead than RAID 5

Sub-CLO 6 input output **(Week 7 page 9)**

1. List and briefly define three techniques for performing I/O Frederico

Answer =
1 . Programmed I/0
Data are exchanged between the processor and the I/O module.
Processor executes a program that gives it direct control of the I/O operation.
When the processor issues a command it must wait until the I/O operation is complete.
If the processor is faster than the I/O module this is wasteful of processor time.

2 . Interrupt-driven I/O
Processor issues an I/O command, continues to execute other instructions, and is interrupted by the I/O module when the latter has completed its work.

3 . Direct Memory Access (DMA)
The I/O module and main memory exchange data directly without processor involvement.

2. What is the difference between memory-mapped I/O and isolated I/O? Aditya **(Week 7 page 14)**

Perbedaan Memory-Mapped I/O dan Isolated I/O

Memory-Mapped I/O dan Isolated I/O adalah dua cara berbeda bagi CPU untuk berkomunikasi dengan perangkat input-output (perangkat I/O).

Memory-Mapped I/O memperlakukan perangkat I/O sebagai lokasi memori. Ini berarti CPU dapat mengakses perangkat I/O menggunakan instruksi dan pengalamatan yang sama yang digunakan untuk mengakses memori. Memory-Mapped I/O lebih sederhana untuk diimplementasikan dan lebih cepat

daripada Isolated I/O, tetapi dapat membatasi jumlah ruang alamat yang tersedia untuk memori dan perangkat I/O.

Isolated I/O menggunakan ruang alamat dan instruksi terpisah untuk memori dan perangkat I/O. Ini berarti CPU harus menggunakan instruksi I/O khusus untuk mengakses perangkat I/O. Isolated I/O lebih kompleks untuk diimplementasikan dan lebih lambat daripada Memory-Mapped I/O, tetapi menyediakan ruang alamat I/O yang lebih besar dan fleksibilitas yang lebih besar.

3. When a device interrupt occurs, how does the processor determine which device issued the interrupt? Dylan **(Week 7 page 20)**
   Answer: There are many ways for a processor to determine the devices. One of those ways is for the processor to do a poll function that individually checks the connected devices for any services that it needs until it finds the ones that caused the interrupt.

   Another way is to Multiple Interrupt Lines that assigns each devices with a unique interrupt line index that makes it easier to sort and identify.

   A Third way is through a daisy chain method. This links all the devices into a single sequence where each output starts an input on the next device so by processing the interrupt it then passes the signal through the devices which responds accordingly with an output until it finds itself to the intended devices that needs attention.

   The Fourth way is through Bus Arbritation which makes the device need to take control of the Bus first before their Interrupt could be processed. By defining some protocol towards requesting control for the Bus, it creates a priority list on which interrupt needs attention first in an organized manner.

4.Describe what Direct Memory Access (DMA) is, and what problem it solves. Jonathan
Answer:

DMA Controller is a hardware device that allows I/O devices to directly access memory with less participation of the processor. A DMA controller needs the same old circuits of an interface to communicate with the CPU and Input/Output devices. One of the problems the DMA solves is that the The I/O module and main memory exchange data directly without the processor involvement. The reason why DMA was built do exchange data directly without the involvement of the processor because both form of of I/O suffer from two inherent drawbacks:

1) The I/O transfer rate is limited by the speed with which the processor can test and service a device

2) The processor is tied up in managing an I/O transfer; a number of instructions must be executed for each I/O transfer

Meanwhile DMA has a technique that can not only The I/O module and main memory exchange data directly without the processor involvement, but also a large volumes of data to be moved with a much better efficiency

Link Buku:

https://drive.google.com/file/d/1XgNwB955ETl1iAxYIMp6V2ApYtfaSKZZ/view?usp=drive_link