1)

a)

1. Central processing unit (CPU)
2. Main memory
3. I/O
4. System interconnection

b)

1. Control unit
2. Arithmetic and logic unit (ALU)
3. Registers
4. CPU interconnection

2)

a) Pipelining: With piplinia, a processor can simultaneously work on multiple instructions. The processor overlaps operations by moving data or instructions into a conceptual pipe with all stages of the pipe processing simultaneously.

Branch prediction: The processor look ahed in the instruction code fetched from memory and predicts which branches or groups of instructions are likely to be processed next. If the processor guesses right most of the time, it can prefetch the correct instructions and buffer them so that the processor is kept busy.

Data flow analysis: The processor analyzes which instructions are dependent on each other's result or data to create an optimized schedule of instructions. In fact, instructions are scheduled to be executed when ready, independent of the original program order. This prevents unnecessary delay.

Speculative execution: Using branch prediction and data flow analysis some processors speculatively execute instructions ahead of their actual apperance in the program execution holding the results in the temporary locations. This enables the processor to keep its execution engines as busy as possible by executing instructions that are likely to be needed.

2)

b)

Average CPI = $0,65 + 2 \times 0,15 + 4 \times 0,15 + 8 \times 0,05 = 1,95$

MIPS $\cong$ 410

c)

$$\frac{1}{(1-0,4) + \frac{0,4}{4}} \cong 1,4$$

3)

1. Sequential interrupt processing
2. Nested interrupt processing

4)

1. Centralized arbitration
2. Distrubuted arbitration

5)

Sequential acces: Memory is organized into units of data called records. Access must be made in a specific linear sequence. Stored addressing information is used to seperate records and assist in the retrieval process. A shared read-write mechanism is used and this must be moved from its current location to the desired location passing and rejeting each intermediate record. Thus the time to access an arbitar record is highly variable.

e.g. Tape units

Direct Acces: As with sequential access, direct acces involves a shared read-write mechanism. However, individual blocks or records have a unique address based on physical location.

e.g. Disk units

Random access: Each addressable location in memory has a unique, physically wired-in addressing mechanism.
e.g. Main memory and some cache systems are random access

Associative: This is a random access type of memory that enables one to make a comparison of desired bit locations within a word for a specified match and to do this for all words simultaneously.
e.g. Cache memories

6)

a) During the course of execution of a program memory references by the processor for both instructions and data tend to cluster. Programs typically contain a number of iterative loops and subroutines. Once a loop or subroutine is entered there are repeated references to a small set of instructions

b) The average access time of a computer system can be improved considerably by use of a cache. If the hit ratio is high enough so that most of the time the CPU access the cache instead of main memory the average access time is closer to the access time of the fast cache memory.

c)
- A dynamic memory cell is simler and smaller than a static memory cell. Thus, a DRAM is more dense and less expensive than a corresponding SRAM.

- DRAM requires the supporting the refresh circuitry. For larger memories the fixed cost of the refresh circuitry is more than compensated for by the smaller varible cost of DRAM cells. Thus, DRAMs tend to be favored for large memory requirements.

- SRAMs are somewhat faster than DRAMs. Because of these relative characteristics, SRAM is used for cache memory and DRAM is used for main memory.

6)

d) $(0,9)(0,01) + (0,1)(0,01\,\mu s + 0,1\,\mu s) = 0,02\,\mu s$

e) $2^k - 1 \geq M + k$

$32 + 1 + k \leq 2^k \Rightarrow k = 6$

6

7)

a)

216 line