Note: You can give the answers in English or Swedish.

1.   a) There are several write policies that are used to keep the cache contents and the contents of the main memory consistent. Describe briefly each of these policies and discuss the advantages and disadvantages of each of them, respectively.

         b) Describe the additional problems we have when applying these policies in a multiprocessor system.

                (3p)

2.   a) What does it by mean by a memory of sequential access type?

         b) Give an example of a sequential access memory.

         c) Can a memory of sequential access type be used as the main memory of a computer system? Why?

                (3p)

3.   What are the most important issues to be considered when designing an instruction set for a computer architecture? Explain in which way each of these issues has an impact on the performance of the computer.

                (3p)

4.   a) What is a data hazard in a pipelined unit? Illustrate this problem by an example and show how penalties are produced (consider a 6-stage pipeline as an example).

         b) How can this penalty be reduced with the forwarding (bypassing) technique? Draw figures to illustrate the pipelined executions without and with forwarding.

                (3p)

5.   a) What does it mean by dynamic branch prediction? What is the basic principle that is used to implment dynamic branch prediction?

         b) What are the advantages and disadvantages of dynamic branch prediction, as compared to static branch prediction?

                (3p)

6.   a) What are the most essential characteristics of a superscalar architecture?

   b) Explain the following two policies for instruction execution:

   - in-order issue with out-of-order completion, and

   - out-of-order issue with out-of-order completion.

   c) Why is the window of execution an important mechanism for a superscalar architecture?

                                                                    (3p)

7.   a) Identify all the true data dependencies, output dependencies and anti-dependencies on the following code. Provide the reasons for your answers.

```
L2: move r3,r9        Note: r3 <- r9
    load r8,(r3)       Note: r8 <- memory location pointed by r3
    add r3,r3,4        Note: r3 <- r3 + 4
    load r9,(r3)       Note: r9 <- memory location pointed by r3
    ble r8,r9,L3       Note: branch to L3 if r8 less than/equal r9
```

   b) Which of the identified dependencies can be eliminated? How?

                                                                    (3p)

8.   a) A VLIW architecture is said to support explicit parallel instruction execution. Define the concept of explicit parallelism. What are the advantages of exploiting explicit parallelism?

   b) What is the main problem of a VLIW computer? How is this problem addressed by the IA-64 architecture?

                                                                    (3p)

9.   a) There are two basic approaches to implement a snoopy protocol: write-invalidate and write-update. How do they work, respectively?

   b) Describe the situation when the write-invalidate approach works better, and the situation when the write-update works better, respectively.

   c) Both these approaches suffer from false sharing overheads. What does it mean by false sharing here?

                                                                    (4p)

Note: You can give the answers in English or Swedish.

10. a) Describe the concept and main features of a computer cluster.

    b) What are the main advantages of using a computer cluster?

    c) Discuss the features of the two typical configurations of a computer cluster, respectively.

    (3p)

11. a) What is a thread? What does it mean by thread-level parallelism?

    b) Describe the different multithreading approaches and discuss how they are applied in the context of superscalar architectures. What are the advantages and disadvantages of these different approaches, respectively?

    (3p)

12. a) What are the main features of a graphics processing unit (GPU)?

    b) Which of the GPU features have contributed to its high performance? How?

    c) Discuss the concept of divergent execution in a GPU processor. What is the main impact of such divergent execution?

    (3p)

13. Describe all low-power techniques and principles that can be used for architecture design (Note: we are not interested in low-power techniques that are used at circuit, logic, micro-architecture, and software levels). Explain why the techniques and principles you have described can help to reduce the power comsumption of a computer.

    (3p)