Försättsblad till skriftlig tentamen vid Linköpings universitet

Datum för tentamen	2017-01-12
Sal (1)	TER4(22)
Tid	8-12
Kurskod	TDTS08
Provkod	TEN1
Kursnamn/benämning **Provnamn/benämning**	Datorarkitektur Skriftlig tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	13
Jour/Kursansvarig Ange vem som besöker salen	Zebo Peng
Telefon under skrivtiden	013-282067
Besöker salen ca klockan	9:45
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Helene Meisinger, 013-281868,
Tillåtna hjälpmedel	Engelsk ordbok
Övrigt	
Antal exemplar i påsen	

TEKNISKA HÖGSKOLAN I LINKÖPING
Institutionen för datavetenskap (IDA)
Zebo Peng

Tentamen i kursen
TDTS08 Datorarkitektur

Examination of the course
TDTS08 Advanced Computer Architecture

2017-01-12, 8:00-12:00

Hjälpmedel:

Engelsk ordbok.

Poänggränser:

Maximal poäng är 40.
För godkänt krävs 21 poäng.

Supporting material:

English dictionary.

Points:

Maximum points: 40.
You need 21 points to pass the exam.

Jourhavande lärare (Teacher on duty)

Zebo Peng, tel. 013-28 2067

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

1. What is a Harvard architecture? What is the main feature of such an architecture? What are the advantage and disadvantage of such an architecture?

(2 p)

2. a) We have discussed three common replacement algorithms that are used in a virtual memory organization. Describe briefly each of these algorithms.

b) Why is the random replacement algorithm used sometimes as a cache replacement algorithm, but it should not be used for virtual memory replacement?

(3 p)

3. a) Why is instruction pipeline widely used to enhance performance of modern computers?

b) What are the three main types of hazards that can reduce the performance of an instruction pipeline? What are the impacts of each of these hazards?

c) In your opinion, which hazard causes the biggest problem for instruction pipeline? Why?

(4 p)

4. A computer has an instruction pipeline with four pipeline stages. As a designer, you are considering the possibility of increasing the number of pipeline stages of this computer in order to improve its performance. It turns out that one of the stages can't be divided into two or several shorter stages. Does it make sense to divide the other three stages so that the number of stages increased? Why?

(3 p)

5. a) Provide a taken/not-taken execution pattern consisting of exactly 4 branches where a bimodal predictor will perform better than a one-bit predictor. Assume that the one-bit predictor is initially set at 0 (not taken) while the bimodal predictor is initially set at 01 (weakly not taken). Your answer should be of the form {T, T, T, NT}, for example.

b) Explain why the bimodal predicator gives better prediction results than those of the one-bit predictor.

(3 p)

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

6. a) Describe the concept of overlapping register windows. How is a register window used?

b) Which problem is addressed by using the overlapping register windows?

c) Why can the overlapping register windows improve dramatically the performance of a computer system?

(3 p)

7. a) What is the most essential characteristics of a superscalar architecture?

b) Describe the following two policies for instruction execution, and explain which one gives better performance:

- 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?

(3 p)

8. a) Define the concept of loop unrolling. Why is loop unrolling very useful in the context of a VLIW processor?

b) Can the loop unrolling technique be used in the context of a superscalar architecture? Why?

c) If a loop is unrolled completely (i.e., there is no need for loop control any longer), what will happen? Discuss the negative side effects of unrolling a loop completely?

(4 p)

9. a) Identify all the different types of data dependencies in the following code. Indicate the type of dependency you have identified for each one, and give the reasons for your answers.

```
L1: move r3,r9        Note: r3 <- r9
    load r8,(r3)      Note: r8 <- memory location pointed by r3
    add r4,r3,4       Note: r4 <- r3 + 4
    load r9,(r4)      Note: r9 <- memory location pointed by r4
    ble r8,r9,L1      Note: branch to L1 if r8 <= r9
```

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

(3 p)

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

10. a) What is a NUMA computer system? What are the motivations for using such a system?
b) Draw a picture of a typical NUMA system. Use the picture to illustrate and discuss the important concepts and components of such a system.
c) What is the purpose of having the directories in a NUMA system?

(3 p)

11. a) There are two basic approaches to implement a snoopy protocol: write-invalidate and write-update. How do they work, respectively?
b) Both these approaches suffer from false sharing overheads. What does it mean by false sharing here? Give an example of a false sharing scenario.

(3 p)

12. a) What are the main features of a graphics processing unit (GPU)?
b) What does it mean by divergent execution in a GPU? Why is divergent execution bad for performance?
c) Discuss one technique that can be used to address the divergent execution problem.

(3 p)

13. We have discussed several techniques to reduce power consumption of processor cores. One of them is called pipeline gating.
a) What does it mean by pipeline gating?
b) Describe one way to implement the pipeline gating technique.

(3 p)