Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2014-04-24
Sal	TER 2
Tid	14-18
Kurskod	TDTS08
Provkod	TEN1
Kursnamn/benämning **Provnamn/benämning**	Datorarkitektur Skriftlig tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	13
Antal sidor på tentamen (inkl. försättsbladet)	5
Jour/Kursansvarig Ange vem som besöker salen	Ke Jiang
Telefon under skrivtiden	013-282691
Besöker salen ca kl.	16:00
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Åsa Kärrman, tel 285760, asa.karrman@liu.se
Tillåtna hjälpmedel	Engelsk ordbok
Övrigt	
Vilken typ av papper ska användas, rutigt eller linjerat	rutigt
Antal exemplar i påsen	

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

Tentamen i kursen

TDTS08 Datorarkitektur

(Examination on TDTS 08 Advanced Computer Architecture)

2014-04-24, kl. 14-18

Hjälpmedel:

Engelsk ordbok.

Supporting material:

English dictionary.

Poänggränser:

Maximal poäng är 40.
För godkänt krävs 21 poäng.

Points:

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

Jourhavande lärare (Teacher on duty):

Ke Jiang, tel. 013-28 2691

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

1. a) What are the differences between direct-mapping and associative mapping for a cache?
   b) For a set-associative mapping cache, a main memory address is viewed as consisting of three fields. Define these three fields, and explain how they are used.
   c) What is the main advantage of set-associative mapping?

   (3p)

2. a) Locality of reference is an important feature of a program in the context of memory hierarchies. Explain what locality of reference means and why it is important.
   b) In the IBM 360 Model 65 computers, memory addresses are assigned to two separate main memory units in such a way that all even-numbered words are in one unit and all odd-numbered words in the other. What might be the purpose of this particular arrangement? Explain also why this memory organization is useful.

   (3p)

3. a) What is branch prediction? How does it work? Illustrate your answer by showing how a certain instruction sequence, of your choice, passes a pipelined unit, in the case of a correct prediction and of an incorrect prediction, respectively.
   b) Why is it important to have a branch prediction mechanism in a computer?
   c) How does dynamic branch prediction work? What is the basic principle that is used to implement dynamic branch prediction?

   (3p)

4. a) What is the most essential characteristics of a superscalar architecture?
   b) What are the differences between the superscalar and super-pipelining approaches? Compare these two approaches to each other, and discuss their advantages and disadvantages, respectively.
   c) Why is data dependence a bigger issue in a superscalar architecture than in an ordinary computer?

   (3p)

5. a) Even though both superscalar architecture and VLIW architecture exploit instruction level parallelism, a VLIW architecture can usually achieve a larger degree of parallelism than a superscalar architecture. Why?
   b) What are the other advantages of a VLIW architecture, as compared with a superscalar architecture?

   (3p)

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

6. Someone says that since both CISC computers and VLIW computers try to perform many computations with a single instruction, they are very similar from an architectural point of view. Do you agree with this statement? Why? Provide all the arguments you have to support your answer. (3p)

7. We have discussed two features that are characteristic to parallel computations: f (the ratio of computations that have to be executed sequentially) and f_c (the fractional communication overhead).

   a) Discuss how these two features influence the performance of a parallel architecture in terms of speedup and efficiency.

   b) How should these features be considered when determining the number of processors and the size of processes for a parallel machine?

   (3p)

8. 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?

   (3p)

9. a) Why does a vector processor have very high performance for vector computations, even though it doesn't contain several CPUs working in parallel?

   b) What is the role of the vector length register in a vector processor? Why is it important to have it?

   c) What is the role of the mask register in a vector unit? Give an example to illustrate the use of the mask register.

   (4p)

10. 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.

    (3p)

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

11. a) What does it mean by parametric computing in the context of parallel processing? Give an example to illustrate the features of such an approach.

    b) Why is the parametric computing approach appropriate to be used in a cluster computer?

    (3p)

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

    b) Which of these features have contributed to its high performance? Why?

    c) Which of these features have contributed to its power efficiency? Why?

    (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)