

Tentamen Computer Architectuur

15 januari 2007, 14.00 – 17.00h.

Dit tentamen bestaat uit 5 opgaven, waarmee in totaal 100 punten te behalen zijn. Iedere opgave valt uiteen in verschillende onderdelen. Het maximaal aantal te behalen punten per onderdeel staat vermeld tussen [..].

Het is niet toegestaan andere informatie te raadplegen dan eigen kennis en hetgeen eventueel tijdens het tentamen door de surveillanten verteld of toegelicht wordt.

Tenslotte: vergeet niet je naam te vermelden, beargumenteer ALLE antwoorden (losse getallen, ja/nee, en dergelijke worden NIET gehonoreerd!!!), en schrijf duidelijk. Succes!

Opgave 1 Performance

We beschouwen een load/store machine met een 1 GHz clock. Veronderstel de volgende karakterisatie van de operaties op deze machine

type	CPI	frequentie
ALU operaties	1	42%
Loads	2	22%
Stores	2	15%
Branches	3	21%

- a. Wat is de CPI van deze machine? [3]
- b. Veronderstel dat een optimizing compiler in staat is om 50% van de ALU operaties te verwijderen. Wat is dan de nieuwe CPI? [3]
- c. Wat is de speed-up als gevolg van deze compiler optimization? [2]
- d. Een andere veel gebruikte maat voor de efficiency van een machine is de MIPS (milions of instructions per second). Wat is de MIPS rating zonder en met optimization? (*Hint*: schrijf MIPS in termen van CPI en clock rate.) [4]
- e. Gegeven de speed-up en de MIPS ratings, wat is je conclusie wat betreft het gebruik van MIPS als maat voor efficiency? [3]

Opgave 2 Caches

- a. Leg uit wat temporal locality en spatial locality zijn en hoe deze door een cache uitgebuit worden. [7]

Veronderstel dat we een byte adresseerbare machine hebben, met 32 bits adressen, 4 bytes in een word en 4 words in een cache line. Veronderstel dat de cache 16 KByte groot is en direct mapped.

- b. Gegeven het hexadecimale adres 1234 1234. In welke cache line komt het byte op dit adres terecht? Leg precies uit. [4]
- c. Bespreek **gedetailleerd** wat er gebeurt als vervolgens de hexadecimale adressen 1234 1239 en 1234 5234 aangesproken worden. Leg precies uit hoe de hardware tot een conclusie ‘hit’ of ‘miss’ komt. [8]
- d. Er wordt onderscheid gemaakt tussen compulsory, capacity en conflict misses. Leg uit wat deze termen betekenen. [3]
- e. Beschrijf 3 manieren om miss rates te verminderen. Leg de voor- en nadelen uit. [3]

Opgave 3 Pipelining

De 5 stage pipeline voor de MIPS ISA bestaat uit: (1) Instruction fetch (IF); (2) Instruction decode/register fetch (ID); (3) Execution/effective address calculation (EX); (4) Memory access/branch completion (MEM); en (5) Write back (WB).

- a. Leg precies uit wat er in deze verschillende cycles gebeurt voor de verschillende soorten instructies die er voor de MIPS bestaan. [6]
- b. Bespreek de drie soorten data hazards die kunnen bestaan. [6]
- c. Leg uit wat bypassing of forwarding is en waarom dat gebruikt wordt. [4]
- d. Kunnen alle data hazards opgelost worden met behulp van forwarding? Zo nee, welke niet? [4]

Opgave 4 Branch prediction

- a. Leg uit wat een branch prediction buffer is en hoe die werkt. [5]
- b. Leg uit wat een branch target buffer is en hoe die werkt. [5]
- c. Leg uit wat een correlating of two-level predictor is en hoe die werkt. [5]

Opgave 5 Memory

Om de bandbreedte van het geheugen te vergroten zijn onder andere een ‘wide memory’ organisatie en een ‘interleaved memory’ organisatie mogelijk. Dit zijn verbeteringen ten opzichte van een ‘one-word-wide’ memory organisatie.

- a. Beschrijf de drie genoemde organisaties. [6]
- b. We bekijken nu waarom de eerste twee organisaties een verbetering vormen. Stel dat de performance van het geheugen als volgt is:
 - 2 clock cycles om een address te sturen
 - 24 clock cycles access time per woord
 - 2 clock cycles om een woord terug te sturen.

Als een cache line 4 woorden groot is, hoelang duurt het dan om een cache miss af te handelen in elk van de drie organisaties? Hoe groot is de bandbreedte naar het geheugen in elk van de drie organisaties? [6]

- c. Leg uit hoe virtual memory werkt en de rol van de Translation Look-aside Buffer (TLB) daarin. [5]
- d. Wat is het probleem als er zowel virtual memory als caches gebruikt worden? [4]
- e. De IBM 3033 heeft 16-way set associative cache ter grootte van 64KB en 4KB pages om dit probleem op te lossen. Waarom is dit een oplossing? [4]

#