



Karlsruher Institut für Technologie

Institut für Technische Informatik

Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

Prof. Dr. rer. nat. Wolfgang Karl

Klausur Rechnerstrukturen Wintersemester 2016/17 Aufgabenteil

29. März 2017

Aufgabe 1: Sprungvorhersage und Verbindungsstrukturen 10 P

Sprungvorhersage 4 P

- a) Was ist der Vorteil bei der Verwendung eines 2-Bit Prädiktors mit Hysteresezähler gegenüber einem 2-Bit Prädiktor mit Sättigungszähler? 1 P
- b) Gegeben sei ein globaler (1,2)-Korrelationsprädiktor mit einem Branch History Register für zwei bedingte Sprünge $S1$ und $S2$, deren Sprungverlauf mittels eines Profiling-Laufs erfasst wurde. Das Schieberegister sei mit *Not Taken*, die verwendeten 2-Bit-Prädiktoren mit Hysteresezähler jeweils mit *Weakly Taken* initialisiert. Füllen Sie die auf dem Lösungsblatt angegebene Tabelle aus und kennzeichnen Sie den jeweils ausgewählten Prädiktor deutlich. 3 P

Verbindungsstrukturen 6 P

- c) Zeichnen Sie das Grundmuster der Kreuzpermutation mit 8 Eingängen und 8 Ausgängen auf. Verwenden Sie hierfür als Hilfe die Vorgaben auf den Lösungsblättern. 1 P
- d) Welches mehrstufige Permutationsnetzwerk verwendet die Kreuzpermutation als Grundlage? 1 P
- e) Zu welcher Permutation ist die Kreuzpermutation bei $n = 2$ und $n = 3$ Bits identisch? 0,5 P
- f) Wo befindet sich die Steuerung des Verbindungsaufbaus bei dynamischen Verbindungsstrukturen? 0,5 P
- g) Für was steht die Abkürzung Phit und was wird damit angegeben? 1 P
- h) Aus welchen fünf Bestandteilen setzt sich die Latenz beim end-to-end packet latency model zusammen? 2 P

Aufgabe 2: Vektorrechner und Tomasulo

10 P

Vektorrechner

5 P

- a) Was versteht man unter einem Vektorrechner? 1 P
- b) Wie können IF-Anweisungen bei Vektorrechnern umgesetzt werden? 1 P
- c) Vektorbefehle werden oft durch ein spezielles Speichersystem mit Verschränkung und mehreren Speicherbänken unterstützt. Wie lange dauert ein Ladebefehl eines 64-elementigen Vektors bei 16 Speicherbänken und einer Latenz von 12 Zyklen bei einem Stride von 10? 2 P
- d) Was ist der Vorteil der Verkettung von Vektor-Pipelines? 1 P

Tomasulo

5 P

- e) Untenstehend finden Sie den Zustand der Reservierungstabelle und der Registerdatei eines Superskalarprozessors nach Abarbeitung des zweiten Taktes der in Listing 1 dargestellten Befehlsfolge. Geben Sie den Zustand der Reservierungstabelle, sowie der Registerdatei nach Ablauf von Takt 6, d.h. nach vier weiteren Takten, unter Berücksichtigung der in Listing 1 dargestellten Befehlsfolge wieder. Pro Takt kann ein Befehl in die Reservierungstabelle eingetragen werden. Eine Addition benötigt 2 Takte, eine Multiplikation 5 Takte und eine Division 9 Takte. 5 P

Takt	Befehlsfolge
1	mul R3, R2, R4
2	add R4, R1, R5
3	div R1, R4, R5
4	add R2, R3, R4
5	add R5, R3, R5
6	mul R3, R3, R4

Listing 1

(Format: Opcode Ziel, Quelle 1, Quelle 2)

Feld	R1	R2	R3	R4	R5
Value	(R1)	(R2)	–	–	(R5)
Valid	1	1	0	0	1
RS	–	–	Mul 1	Add/Sub 1	–

Registerdatei

Unit	Empty	InFU	Op	Dest	Src1	Vld1	RS1	Src2	Vld2	RS2
Add/Sub 1	0	0	add	R4	(R1)	1	–	(R5)	1	–
Add/Sub 2	1									
Mul 1	0	1	mul	R3	(R2)	1	–	(R4)	1	–
Div 1	1									

Reservierungstabelle

Aufgabe 3: Caches

10 P

Beim Entwurf von neuen Chip Multi-Core-Prozessorsystemen (CMP) ist die Organisation der Caches eine wichtige Fragestellung. Insbesondere die Abwägung zwischen privaten und geteilten Caches ist hierbei zu untersuchen. Untersuchen Sie für ein CMP, das vier Cores besitzt, welche der folgenden L1 Cache-Organisationen gewählt werden sollte.

Private Cache-Organisation: Jeder Cache besitzt die Kapazität für 8 Blöcke und ist direct mapped. Bei einem Treffer entspricht die Zugriffsdauer einem Zyklus. Die Zeit für eine Cache-Kohärenz-Aktion beträgt 10 Zyklen. Bei einem Miss werden 100 Zyklen für den Erhalt des Cache-Blockes benötigt.

Geteilte Cache-Organisation: Die Kapazität dieses Caches ist genau so hoch, wie die Summe der Blöcke der privaten Cache-Organisation. Der Cache ist in B Bänke unterteilt, wobei B der Anzahl der Prozessoren entspricht. Die Abbildung von Speicherblöcken zu Bänken erfolgt im Round-Robin verfahren. Somit ist Blockadresse i auf Bank $(i \bmod B)$ abgebildet. Der Zugriff auf Bänke erfolgt über eine $B \times B$ cross-bar Switch. Ein Zugriff auf eine Bank dauert 2 Zyklen für den Fall, dass keine gegenseitige Beeinflussung vorliegt. Bei einem Miss werden 100 Zyklen für den Zugriffe auf den Hauptspeicher benötigt.

Die Blockgröße ist jeweils 16 Byte. Der Zugriff auf die nächste Hierarchieebene findet erst dann statt, wenn die vorherige Ebene das angeforderte Datum nicht gespeichert hat

- Ermitteln Sie die durchschnittliche Speicherzugriffszeit für den Fall, dass ein einzelner Prozessor zweimal sequentiell auf die Speicherblöcke 0 bis 15 zugreift. Welche der beiden Organisationen hat die kürze Zugriffszeit? Geben Sie den Faktor mit an! *3P*
- Ermitteln Sie die durchschnittliche Speicherzugriffszeit für den Fall, dass jeder Prozessor zweimal sequentiell auf die Speicherblöcke 0 bis 15 zugreift. Ignorieren Sie dabei den gegenseitigen Einfluss der Prozessoren. Welche der beiden Organisationen hat die kürze Zugriffszeit? Geben Sie den Faktor mit an! *2P*
- Nehmen Sie an, dass die Speicherblöcke 0 bis 7 bereits im jeweiligen Cache vorhanden sind. Nehmen Sie weiterhin an, dass Prozessor 1 alle 8 Blöcke modifiziert und Prozessor 2 diese nach und nach liest. Ignorieren Sie wiederum gegenseitige Beeinflussung der Prozessoren, aber beachten Sie den Aufwand des Cache-Kohärenz-Protokolls. Gehen Sie zudem davon aus, dass bei einem Kohärenz-Miss die Daten vom anderen Cache innerhalb der Cache-Kohärenz-Aktion übertragen werden. Welche der beiden Organisationen hat die kürze Zugriffszeit? Geben Sie den Faktor mit an! *2P*

Ein Dreiprozessorsystem sei speichergekoppelt. Die Caches haben je eine Größe von zwei Cache-Zeilen, welche je genau ein Speicherwort aufnehmen können. Die Füllung des Caches erfolgt von der niedrigsten Cache-Zeile aufwärts, sofern noch freie Zeilen zur Verfügung stehen, andernfalls wird gemäß LRU-Strategie verdrängt. Als Cache-Kohärenzprotokoll komme das aus der Vorlesung bekannte MESI-Protokoll zum Einsatz. Der Cache sei initial leer.

- d) Vervollständigen Sie die auf dem Lösungsblatt angegebene Tabelle: Geben Sie jeweils Inhalt der Cache-Zeile und MESI-Zustand an. *3P*

Aufgabe 4: Hardware-Entwurf und VHDL

8 P

Hardware-Entwurf

6 P

Betrachten Sie für einen Basisprozessor, der einen Load/Store Befehlssatz besitzt und bisher Fließkommaoperationen in Software durchführt, zwei verschiedene Erweiterungsmöglichkeiten. Zudem ist die Zykluszeit T_c bekannt. Die erste Erweiterung fügt eine Fließkommaeinheit hinzu, die Fließkommaoperationen um den Faktor 10 beschleunigt. Die zweite Erweiterung vergrößert den verfügbaren L1 Cache, um dadurch Lade und Speicheroperationen zu beschleunigen. Eine Schätzung legt nahe, dass bei gleichem Verbrauch von Fläche auf einem Chip verglichen mit der Fließkommaeinheit Lade und Speicherbefehle um den Faktor 2 beschleunigt werden. Nehmen Sie an, dass F_{fp} der Anteil der Ausführungszeit ist in der Fließkommaoperationen durchgeführt werden. F_{ls} gibt diesen Anteil für Lade und Speicherbefehle an. Während der Ausführung überlappen sich F_{fp} und F_{ls} zeitlich gesehen nicht.

- a) Geben Sie zunächst das Gesetz von Amdahl an und beschreiben Sie die einzelnen Komponenten. Bestimmen Sie unter zur Hilfenahme von Amdahls Gesetz das Verhältnis zwischen F_{fp} und F_{ls} , bei dem die Erweiterung mittels einer Fließkommaeinheit sinnvoller ist als das Vergrößern des Caches. 3 P
- b) 3 P
- i) Nehmen Sie nun an, dass Sie anstelle der zeitlichen Anteile F_{fp} und F_{ls} die Anteile an Instruktionen der jeweiligen Operationen haben. IC_{fp} gibt die Anzahl Fließkommaoperationen an und IC_{ls} die Anzahl an Lade und Speicherbefehle. Zudem haben Sie die jeweilige durchschnittliche Anzahl an Zyklen, CPI_{fp} und CPI_{ls} , für die jeweilige Operation gegeben. Können Sie anhand dieser Werte immer noch bestimmen, welche Erweiterung besser ist? Wenn ja, wie?
- ii) Können Sie darüber hinaus auch mittels Amdahls Gesetz den maximalen Speedup berechnen? Begründen Sie ihre Antwort.

VHDL

2 P

- c) Geben Sie die 4 Schritte an, die zu Realisierung eines fertigen Chips mittels einer Hardwarebeschreibungssprache durchgeführt werden. 2 P

Aufgabe 5: Low-Power-Entwurf und Leistungsbewertung

10 P

Low-Power-Entwurf

2 P

- a) Geben Sie die allgemeine Formel zur Berechnung von Schaltwahrscheinlichkeiten sowie die Formel zur Berechnung der Leistungsaufnahme P_{total} in CMOS-Schaltungen an. Welcher Bestandteil der Leistungsaufnahme war früher vernachlässigbar, spielt heute jedoch eine zentrale Rolle? Begründen Sie diesen Bedeutungszuwachs. 2 P

Leistungsbewertung

8 P

- b) Nehmen Sie an, Sie haben zwei Implementierungen derselben Befehlssatzarchitektur vorliegen. Computer A verfügt über eine Taktzykluszeit von 250ps und einen CPI-Wert von 2,0 für ein Programm und Computer B verfügt über eine Taktzykluszeit von 500ps und einen CPI-Wert von 1,2 für dasselbe Programm. Welcher Computer ist bei diesem Programm schneller und um wie viel? 2 P
- c) Sie haben die Wahl zwischen vier verschiedenen Arbeitsplatzrechnern: PC A mit einem Intel i5-Prozessor, PC B mit einem i7-Prozessor, PC C mit einem AMD-Prozessor (unter der Annahme, dass alle den selben Compiler verwenden) und einem Apple Computer. Welche der folgenden Aussagen sind wahr? Begründen Sie. 3 P
- (c.1) Der schnellste Rechner ist der mit der höchsten Taktfrequenz.
- (c.2) Da AMD zum Ausführen von Befehlen andere Techniken als Intel verwendet, weisen diese Prozessoren möglicherweise andere CPI-Werte auf. Dennoch lässt sich anhand der Taktfrequenz feststellen, welcher der beiden Intel-PCs schneller ist.
- (c.3) Nur anhand der Ergebnisse von Benchmarks für Aufgaben ähnlich denen ihrer realen Last lässt sich die wahrscheinliche Leistung genau vorhersagen.
- d) Gegeben sei ein Computer mit den 3 Befehlsklassen A, B und C. Sie messen den Code für ein Programm, das von zwei verschiedenen Compilern erzeugt wurde. Sie erhalten die folgenden Daten: 3 P

Code	A	B	C
Compiler 1	5 Mrd. Befehle	1 Mrd. Befehle	1 Mrd. Befehle
Compiler 2	10 Mrd. Befehle	1 Mrd. Befehle	1 Mrd. Befehle
CPI-Wert	1,0	2,0	3,0

Zudem sei die Taktfrequenz von 4GHz gegeben. Welche Codesequenz erzielt einen höheren MIPS-Wert? Welche wird entsprechend der Ausführungszeit schneller ausgeführt?

Aufgabe 6: Parallelverarbeitung und Pipelining 12 P

Parallelverarbeitung 4 P

- a) Benennen Sie vier Ebenen der Parallelität und erläutern Sie diese kurz. 4 P

Pipelining 8 P

- b) Gegeben sei ein C-Codesegment durch: 6 P

```
A = B + E;
C = B + F;
```

Der MIPS-Code für dieses Segment lautet unter der Voraussetzung, dass sich alle Variablen bereits im Speicher befinden und als Offset von Register `$t0` adressierbar sind, wie folgt:

```
1      lw   $t1, 0($t0)
2      lw   $t2, 4($t0)
3      add  $t3, $t1, $t2
4      sw   $t3, 12($t0)
5      lw   $t4, 8($t0)
6      add  $t5, $t1, $t4
7      sw   $t5, 16($t0)
```

Welche Pipelinekonflikte treten in diesem Codesegment auf, wenn von einer 5-stufigen Pipeline ausgegangen wird? Welche nicht? Ordnen Sie die Befehle so um, dass die Pipeline möglichst wenig angehalten werden muss. Nennen Sie weitere Lösungsansätze neben dem Umsortieren von Codesegmenten, die sich eignen, um Pipelinekonflikte zu lösen, ohne die Pipelinearchitektur anzupassen. Erläutern Sie diese Lösungsansätze kurz.

- c) Welche zusätzlichen Eigenschaften müsste die obige Pipeline besitzen, damit ein Anhalten der Pipeline für das aufgeführte Beispiel komplett umgangen werden kann? Erklären sie diese kurz. 2 P



Karlsruher Institut für Technologie

Institut für Technische Informatik

Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

Prof. Dr. rer. nat. Wolfgang Karl

Klausur Rechnerstrukturen

Wintersemester 2016/17

Lösungsteil

Name: _____

Vorname: _____

Matrikelnummer: _____

Tragen Sie bitte auf jedem Blatt Ihren Namen und Ihre Matrikelnummer ein. Bitte tragen Sie alle Lösungen und Rechenwege an den vorgesehen Stellen ein und geben Sie keine zusätzlichen Blätter ab, ohne dies dem Aufsichtspersonal mitzuteilen.

Hinweis: Bei Rechenaufgaben ist die Angabe des Rechenwegs zwingend erforderlich. Ergebnisse ohne Rechenweg werden **nicht** gewertet.

Zum Bestehen der Klausur sind mindestens 20 Punkte erforderlich.

() Ich wünsche **keine** Notenveröffentlichung über einen anonymisierten Code auf der Webseite des Lehrstuhls.

(Bei Ankreuzen kann die Note erst in der Klausureinsicht erfragt werden.)

Erreichte Punkte (wird vom Institut ausgefüllt):

Aufgabe	1	2	3	4	5	6
Punkte	/10	/10	/10	/8	/10	/12
					Summe:	/60

Aufgabe 1: Sprungvorhersage und Verbindungsstrukturen

10 P

Sprungvorhersage

4 P

a) Antwort:

1 P

b) Tabelle:

3 P

	Globaler Prädiktor	S1		Globaler Prädiktor	S2	
		Vhs.	Sprung		Vhs.	Sprung
1	(WT , WT)		T	(,)		NT
2	(,)		NT	(,)		T
3	(,)		T	(,)		NT
4	(,)		T	(,)		T
5	(,)		T	(,)		NT
6	(,)		NT	(,)		T

Verbindungsstrukturen**6 P**

c) Kreuzpermutation:

1 P

0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	0
1	1	1	1	1	1

d) Antwort:

1 P

e) Antwort:

0,5 P

f) Antwort:

0,5 P

g) Antwort:

1 P

h) Antwort:

2 P

Aufgabe 2: Vektorrechner und Tomasulo **10 P****Vektorrechner** **5 P**a) *1 P*b) *1 P*c) *2 P*d) *1 P*

Aufgabe 3: Caches

10 P

a) Antwort:

3P

b) Antwort:

2P

c) Antwort:

2P

d)

3P

Proz.	Aktion	Proz. 1		Proz. 2		Proz. 3	
		Zeile 1	Zeile 2	Zeile 1	Zeile 2	Zeile 1	Zeile 2
-	init	-	-	-	-	-	-
2	rd 4						
1	rd 4						
3	wr 4						
2	rd 4						
1	rd 1						
1	wr 2						
2	wr 2						
2	rd 3						

Aufgabe 4: Hardware-Entwurf und VHDL

8 P

Hardware-Entwurf

6 P

a)

3 P

b)

3 P

i)

ii)

VHDL

2P

c)

2P

- 1.
- 2.
- 3.
- 4.

Aufgabe 5: Low-Power-Entwurf und Leistungsbewertung

10 P

Low-Power-Entwurf

2 P

a)

2 P

- Schaltwahrscheinlichkeit:

- Leistungsaufnahme:

- Bestandteil:

- Begründung:

Leistungsbewertung

8 P

b) Berechnung:

2 P

c) Antwort:

3P

(c.1) Wahr / Falsch
Begründung:

(c.2) Wahr / Falsch
Begründung:

(c.3) Wahr / Falsch
Begründung:

d) Berechnung:

3P

Aufgabe 6: Parallelverarbeitung und Pipelining **12 P****Parallelverarbeitung** **4 P**a) Ebenen der Parallelität: *4 P***Pipelining** **8 P**b) Konflikte: *6 P*

Umsortiertes Codesegment:

Weitere Lösungen:

c) Weitere Eigenschaften:

2P