



Karlsruher Institut für Technologie

Institut für Technische Informatik

Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

Prof. Dr. rer. nat. Wolfgang Karl

Klausur Rechnerstrukturen

Sommersemester 2017

Aufgabenteil

14. August 2017

Aufgabe 1: Sprungvorhersage, Fehlertoleranz und Verbindungsstrukturen 10 P

Sprungvorhersage 5 P

- a) Warum ist die Sprungvorhersage bei der Pipeline-Ausführung wichtig? 1 P
- b) Was speichert der Branch Target Address Cache (BTAC)? 1 P
- c) 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 *Taken*, die verwendeten 2-Bit-Prädiktoren mit Sättigungszähler jeweils mit *Weakly Taken* initialisiert. Füllen Sie die auf dem Lösungsblatt angegebene Tabelle aus und kennzeichnen Sie den jeweils ausgewählten Prädiktor deutlich. 3 P

Fehlertoleranz 3 P

- d) Zeichnen Sie das Zuverlässigkeitsblockdiagramm für die Systemfunktion $S = P \wedge (M_1 \vee M_2 \vee M_3) \wedge (G_1 \vee G_2)$. 1 P
- e) Wie lautet die allgemeine Bezeichnung für Systeme wie die Subsysteme aus M_1, M_2 und M_3 und G_1 und G_2 der zuvor angegebenen Systemfunktion S ? 0,5 P
- f) Geben Sie alle möglichen Fehlerbereiche B bis auf die leere Menge der Systemfunktion $S = M_1 \vee M_2 \vee M_3$ an. 1,5 P

Verbindungsstrukturen 2 P

- g) Was bedeutet die Abkürzung Flit und auf welcher Ebene des QPI-Protokolls werden Flits verwendet? 1 P
- h) Wann sind die Übertragungsmodi Cut Through und Wormhole Routing der Paketvermittlung identisch? 1 P

Aufgabe 2: Vektorrechner, VLIW und Tomasulo **10 P****Vektorrechner** **2 P**

- a) Nennen Sie die vier Möglichkeiten der Parallelarbeit, die bei einem Vektorrechner möglich sind. *2 P*

VLIW **3 P**

- b) Es stehen ein VLIW-Prozessor zur Verfügung. Das System hat zwei universell einsetzbare Funktionseinheiten und eine VLIW-Breite von zwei Befehlen. Nehmen Sie vereinfachend an, dass alle Befehle innerhalb eines Taktzykluses abgearbeitet werden können. *3 P*

```
1  add r5, r3, r1 ; r5 = r3 + r1
2  ld r2, [r1]    ; load r2 from mem[r1]
3  ld r4, [r3]    ; load r4 from mem[r3]
4  mul r5, r5, r4 ; r5 = r5 · r4
5  add r1, r4, r2 ; r1 = r4 + r2
6  st [r6], r1    ; store r1 to mem[r6]
```

Tragen Sie die obigen Befehle möglichst kompakt in den Ablaufplan auf dem Lösungsblatt ein.

Tomasulo

5 P

- c) Untenstehend finden Sie den Zustand der Reservierungstabelle und der Registerdatei eines Superskalarprozessors nach Abarbeitung des 2. Taktes der in Listing 1 dargestellten Befehlsfolge. Geben Sie den Zustand der Reservierungstabelle, sowie der Registerdatei nach Ablauf von Takt 5, d.h. nach 3 weiteren Takten, unter Berücksichtigung der in Listing 1 dargestellten Befehlsfolge wieder. Pro Takt kann ein Befehl in die Reservierungstabelle eingetragen werden. Eine Addition/Subtraktion benötigt 2 Takte, eine Multiplikation 3 Takte und eine Division 4 Takte.

5 P

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

Listing 1

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

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

Registerdatei

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

Reservierungstabelle

Aufgabe 3: Caches

10 P

- a) Was versteht man unter Inklusionseigenschaft bei Cache-Speichern? *1 P*
- b) Erklären Sie das Gültigkeitsproblem im Kontext von Multiprozessorsystemen? *1 P*
- c) Benennen Sie zwei Speicherkonsistenzmodelle! *1 P*

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 FIFO-Strategie verdrängt. Als Cache-Kohärenzprotokoll komme das aus der Übung bekannte MOESI-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 MOESI-Zustand an. *2 P*

Betrachten Sie einen Multiprozessor mit gemeinsamen Speicher, der aus drei Prozessoreinheiten sowie Caches besteht. Jeder Cache bietet Platz für eine Cache-Zeile. Als Cache-Protokoll kommt MESI zum Einsatz. Gegeben sei ein Zugriffsmuster (siehe Lösungsteil) der Prozessoren auf vier Variablen (A,B,C,D), wobei die Variablen A,B und C dem selben Speicherblock angehören, also zusammen geladen werden und D einem anderen Speicherblock angehört. Beide Speicherblöcke werden auf die selbe Cache-Zeile abgebildet.

- e) Klassifizieren Sie die auftretenden Cache Misses. Unterscheiden Sie dabei zwischen True- und False-Sharing Misses. *4 P*
- f) Welcher der Misses könnte ignoriert werden ohne die Korrektheit der Ausführung zu beeinflussen? *1 P*

Aufgabe 4: Hardware-Entwurf und VHDL**10 P****Hardware-Entwurf****4 P**

Hardware-Implementierungen für die Berechnung von Gleitkomma Quadratwurzeln können stark in ihrer Performanz variieren. Nehmen Sie an, dass 20% der Ausführungszeit eines Benchmarks mit der Berechnung von Quadratwurzeln verbraucht wird. Auf dem Markt ist eine Gleitkomma-Quadratwurzeleinheit vorhanden, die die Ausführungszeit der Gleitkomma-Quadratwurzeloperationen um den Faktor 10 beschleunigen kann. Demgegenüber steht die Aussage des Entwicklerteams, dass statt der ausschließlichen Beschleunigung der Gleitkomma-Quadratwurzelberechnung auch Gleitkommaoperationen im Allgemeinen um den Faktor 2 beschleunigt werden können. Gleitkommaoperationen verbrauchen 50% der Ausführungszeit des betrachteten Benchmarks. Vergleichen Sie diese beiden Alternativen und entscheiden Sie sich für eine der Varianten. Begründen Sie ihre Antwort.

- a) Mithilfe welchem aus der Übung und Vorlesung bekannten Gesetz können Sie obige Entscheidung treffen? Geben Sie die Formel mit an. 1,5 P
- b) Nutzen Sie obige Formel, um die beiden Alternativen zu vergleichen. Welche Alternative ist besser und warum? 2,5 P

Produktion**1 P**

- c) Die Formel für die Berechnung der erzielbaren Dies pro Wafern dpw hat die Form 1 P

$$dpw = A - B = \frac{\pi(d_{wafer} * \frac{1}{2})^2}{a_{die}} - \frac{\pi * (d_{wafer})}{\sqrt{2} * a_{die}}$$

Welche Bedeutung haben die Werte A und B?

VHDL**5 P**

Gegeben Sei folgende VHDL-Beschreibung:

```
library IEEE;
use IEEE_std_logic_1164.all;

entity counter is
port(
  clk, rst: in std_logic;
  count: out std_logic_vector(3 downto 0)
);
end entity;

architecture rtl of counter is
begin
  process(clk,rst)
  begin
    if rst='0' then
      count<=(others=>'1');
    elsif clk'event and clk='1' then
      count<=count+"0001";
    end if;
  end process;
end architecture;
```

- d) Ungeachtet eventueller Fehler im Quelltext: Welche Funktion wird in der Architecture beschrieben und was ist die genaue Aufgabe der in der Entity deklarierten Signale bzw. welche Aktionen werden wann hierdurch ausgelöst? (*Hinweis: "Auslösung der Abarbeitung des Prozesses" ist keine ausreichende Antwort.*) 3 P
- e) Die Beschreibung wird sich so nicht übersetzen lassen, da sie einen Zuweisungskonflikt enthält. Woraus resultiert dieser Konflikt (genaue Erklärung) und wie lässt sich der Fehler in der VHDL-Beschreibung **ohne Änderung der Verhaltensbeschreibung** (Architecture) beheben? 2 P

Aufgabe 5: Leistungsbewertung

10 P

- a) Sind die folgenden Aussagen wahr oder falsch? Begründen Sie. *4 P*
- (a.1) Die Verbesserung einer Komponente eines Rechnersystems schlägt sich direkt proportional auf die Leistung des Systems nieder.
 - (a.2) Die Verwendung einer Eigenschaft eines Systems (wie die Frequenz) als Leistungsmaß ist ausreichend.
 - (a.3) Der MIPS-Wert ist allgemein als Maßstab für den Vergleich von Computern verwendbar.
 - (a.4) Amdahls Gesetz besagt, dass die durch eine bestimmte Verbesserung mögliche Leistungssteigerung durch den von dieser Verbesserung nicht betroffenen Anteil einer Programmausführung beschränkt ist.
- b) Gegeben sei eine bedingte Verzweigung, die auf zwei verschiedene Arten implementiert werden kann: *6 P*
- CPU A: Mittels einer Vergleichsoperation wird zunächst explizit eine Bedingungsprüfung durchgeführt. Im Anschluss wird in einem weiteren Schritt eine Verzweigungsanweisung auf Basis des Prüfungsergebnisses ausgeführt.
 - CPU B: Die Vergleichsoperation wird nicht explizit implementiert, sondern ist impliziter Bestandteil der Verzweigungsanweisung.

Auf beiden CPUs dauern Verzweigungsanweisungen 2 Zyklen, alle übrigen Instruktionen nehmen 1 Zyklus in Anspruch. Der auf CPU A ausgeführte Code besteht zu 20% aus Verzweigungsoperationen und entsprechend der Implementierung zu weiteren 20% aus Vergleichsoperationen. Der Code für CPU B entspricht dem Code für CPU A mit entsprechender Implementierung der Verzweigungen. Die Clock Cycle Time von CPU A ist dabei um den Faktor 1,25 schneller als die von CPU B.

Welche CPU wird den Code schneller abarbeiten? Begründen Sie Ihre Einschätzung rechnerisch.

Ändert sich diese Einschätzung, wenn CPU A nur eine um das 1,1-fache höhere Clock Cycle Time aufweist?

Aufgabe 6: Parallelverarbeitung und Pipelining 10 P

Parallelverarbeitung 6 P

- a) OpenMP und MPI sind zwei voneinander verschiedene Konzepte um parallele Programmierung zu ermöglichen. Erläutern Sie kurz beide Ansätze und zeigen Sie die Unterschiede auf. Für welche Architekturen eignet sich welcher Ansatz besonders? *3 P*
- b) OpenMP und MPI stehen als Representanten für zwei bestimmte Modelle paralleler Programmierung. Nennen Sie zwei weitere Modelle der parallelen Programmierung, erläutern Sie diese kurz und nennen Sie eine Einsatzschwerpunkt. *3 P*

Pipelining 4 P

Gegeben sei ein C-Codesegment durch:

```
A = B + D;  
C = A - E;
```

Eine Implementierung dieses Segmentes im MIPS-Code finden Sie auf dem Lösungsblatt.

- c) Welche Pipelinekonflikte sind in der angegebenen Implementierung erkennbar, wenn von einer 5-stufigen Pipeline ausgegangen wird? Wie können diese ohne Veränderung der Pipeline effizient gelöst werden? Welche Konflikte kommen durch Ihre Lösung hinzu? *2 P*
- d) Definieren Sie den Begriff “forwarding” und erläutern Sie kurz, ob dieses Konzept zur Auflösung der neu entstehenden Konflikte beitragen kann. *2 P*



Karlsruher Institut für Technologie

Institut für Technische Informatik

Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

Prof. Dr. rer. nat. Wolfgang Karl

Klausur Rechnerstrukturen

Sommersemester 2017

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	/10	/10	/10
					Summe:	/60

Aufgabe 1: Sprungvorhersage, Fehlertoleranz und Verbindungsstrukturen 10 P

Sprungvorhersage 5 P

a) Antwort: 1 P

b) Antwort: 1 P

c) Tabelle: 3 P

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

Fehlertoleranz**3 P**

d) Zuverlässigkeitsblockdiagramm:

1 P

e) Antwort:

0,5 P

f) Fehlerbereiche:

1,5 P

Verbindungsstrukturen

2 P

g) Antwort:

1 P

h) Antwort:

1 P

Aufgabe 2: Vektorrechner, VLIW und Tomasulo **10 P****Vektorrechner** **2 P**a) *2 P*

•

•

•

•

Aufgabe 3: Caches**10 P**

a) Antwort:

1 P

b) Antwort:

1 P

c) Antwort:

1 P

-
-

d)

2 P

Proz.	Aktion	Proz. 1		Proz. 2		Proz. 3	
		Zeile 1	Zeile 2	Zeile 1	Zeile 2	Zeile 1	Zeile 2
-	init	-	-	-	-	-	-
2	wr 4						
1	rd 4						
3	wr 4						
3	rd 5						
3	wr 1						

e)

4P

Proz. 1	Proz. 2	Proz. 3	Art des Cache-Misses
1	Read A		
2	Read B		
3		Read C	
4	Write A		
5		Read D	
6	Read B		
7	Write B		
8		Read C	
9	Read B		

f) Antwort:

1P

Aufgabe 4: Hardware-Entwurf und VHDL**10 P****Hardware-Entwurf****4 P**

a)

1,5 P

b)

2,5 P

•

•

• Antwort:

Produktion**1 P**

c) A:

1 P

B:

VHDL**5 P**

d) Die Schaltungsbeschreibung modelliert:

*3 P*Funktion von `clk`:Funktion von `rst`:Funktion von `count`:

e) Antwort mit Begründung:

2 P

Aufgabe 5: Leistungsbewertung**10 P**

a)

4 P

(a.1) wahr / falsch
Begründung:

(a.2) wahr / falsch
Begründung:

(a.3) wahr / falsch
Begründung:

(a.4) wahr / falsch
Begründung:

b) Berechnung:

6P

Aufgabe 6: Parallelverarbeitung und Pipelining **10 P****Parallelverarbeitung** **6 P**a) Vergleich OpenMP und MPI: *3 P*b) Weitere parallele Programmiermodelle: *3 P*

Pipelining**4 P**

Der MIPS-Code für das gegebene 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	sub	$\$t5$, $\$t3$, $\$t4$
7	sw	$\$t5$, 16($\$t0$)

c) Konflikte, Lösungen und Folgekonflikte:

2 P

d) Definition forwarding und Beitrag zur Konfliktlösung:

2 P