



Karlsruher Institut für Technologie
Institut für Technische Informatik
Prof. Dr. Wolfgang Karl

Klausur Rechnerstrukturen
Wintersemester 2010/11
Aufgabenteil

Aufgabe 1: Parallelverarbeitung

11P

Quantitative Maßzahlen

3P

- a) Zur Bestimmung der Gesamtausführungszeit eines parallel ablaufenden Programms wird oft das Amdahlsche Gesetz herangezogen. Geben Sie dessen Formel an und erklären Sie die Bedeutung der unterschiedlichen Teile der Formel. 1P
- b) Eine Berechnung, die auf einem Einprozessorsystem $T(1) = 300$ Sekunden Rechenzeit benötigt, liefert auf einem Mehrprozessorsystem mit 10 Prozessoren nach $T(10) = 20$ Sekunden ein Ergebnis. Wie nennt man dieses Verhalten und welcher Abschätzung der Beschleunigung widerspricht es? 1P
- c) Geben Sie eine Formel für die Beziehung zwischen Auslastung und Effizienz an. 1P

Verbindungsstrukturen

3P

- d) Zeichnen Sie einen K -ären n -Kubus mit $K = 2$ und $n = 3$. 1P
- e) Das Netzwerk aus Aufgabe d) wird erweitert, so daß $n = 4$ ist. Bestimmen Sie die Anzahl von Knoten N . Wieviel Knoten müssen bei diesem Netzwerk für jede Erweiterung von n hinzugefügt werden ($K = 2$)? 1P
- f) Was versteht man unter einer vollständigen Verbindung? Was ist dabei der große Nachteil? 1P

Parallele Architekturen und Parallelverarbeitung:

3P

- g) Für was stehen die Abkürzungen NORMA, NUMA sowie UMA? 1P
- h) Zur Auswahl stehen 1P
- ein „Multiprozessor mit gemeinsamem Speicher“
 - ein „Multiprozessor mit verteiltem Speicher“
 - ein „Multiprozessor mit verteiltem gemeinsamen Speicher“

Ordnen Sie die Begriffe aus Aufgabe g) dem jeweils entsprechenden Multiprozessor-system zu.

- i) Geben Sie ein Programmiermodell für Multiprozessorsysteme mit verteiltem Speicher an und erklären Sie kurz die Kommunikationsarchitektur der Prozesse (Threads). 1P

Vektorverarbeitung:

2P

- j) In Vektorprozessoren können sowohl aus Multifunktions-Pipelines als auch aus spezialisierten Pipelines bestehen. Geben Sie für spezialisierte Pipelines mindestens einen Vor- und Nachteil an. 1P
- k) Welches Problem kann bei der Benutzung eines Vektorrechners durch einen „Vektor Stride“ vereinfacht werden? 1P

Aufgabe 2: Fragen des Rechnerentwurfs

10P

Low-Power-Entwurf

4P

- a) Geben Sie die Formel zur Ermittlung des Leistungsverbrauchs einer CMOS-Schaltung an. Erläutern Sie die Bestandteile kurz. 2,5P
- b) Zur Ermittlung der Schaltwahrscheinlichkeit einer Schaltung wird häufig ein statistisches Modell herangezogen. Geben Sie die allgemeine Formel zur Berechnung der Schaltwahrscheinlichkeit $\mathbb{P}_{\text{Schalt}}$ an und berechnen Sie diese für ein XOR-Gatter mit $\mathbb{P}_{\text{Eingang 1}}(1) = \frac{1}{2}$ und $\mathbb{P}_{\text{Eingang 2}}(1) = \frac{3}{4}$. 1,5P

Schaltungsentwurf

6P

- c) Eine Architecture in VHDL kann auf mindestens zwei verschiedene Arten implementiert sein. Nennen Sie zwei solche Prinzipien und beschreiben Sie diese. 2P
- d) In VHDL gibt es Signale und Variablen. Wo werden diese definiert und wie unterscheiden sie sich hinsichtlich ihres Verhaltens? 2P
- e) Gegeben ist folgende VHDL-Beschreibung eines Prozesses. 2P

```

1  ...
2  signal count : std_logic_vector(0 to 6) := "0000000";
3  ...
4
5  process (rst , clk) is
6  begin
7      if (rst='1') then
8          count <= (others => '0');
9          flag <= '0';
10     elsif (clk'event and clk='1') then
11         count <= count + 1;
12         if (count(0 to 6)=63) then
13             flag <= '1';
14         else
15             flag <= '0';
16         end if;
17     end if;
18 end process;
```

Was ist die Gesamtfunktionalität dieser Schaltungsbeschreibung? Begründen Sie.

Aufgabe 3: Speicherhierarchie

10P

Cache-Leistung

3P

Bei dem Entwurf eines Systems stehen zwei Entwurfsalternativen zur Auswahl. In beiden Entwurfsalternativen kommt eine zwei-stufige Cache-Hierarchie zum Einsatz.

Entwurfsalternative A hat einen größeren L1-Cache mit einer Zugriffszeit von $t_{A-L1} = 2,5 ns$, Entwurfsalternative B besitzt einen kleineren L1-Cache mit einer Zugriffszeit von $t_{B-L1} = 2 ns$. Beide Entwurfsalternativen besitzen einen L2-Cache mit $t_{L2} = 10 ns$. Die Zugriffszeit des Hauptspeichers sei in beiden Entwurfsalternativen gleich und betrage $t_{Mem} = 100 ns$.

Um die Leistung der Cache-Hierarchie zu steigern, wird bei Entwurfsalternative B alle Hierarchieebenen parallel angefragt. Bei Entwurfsalternative A findet ein Zugriff auf die nächste Hierarchieebene erst statt, wenn die vorherige Ebene das angeforderte Datum nicht gespeichert hat.

- Geben Sie eine allgemeine Formel zur Berechnung t_a der mittleren Zugriffszeit in einer zwei-stufigen Cache-Hierarchie an. 1P
- Bei der Evaluation beider Entwurfsalternativen wurden folgende Hit-Raten gemessen: 2P
 - Alternative A: $r_{A-L1} = 80\%$, sowie $r_{A-L2} = 90\%$
 - Alternative B: $r_{B-L1} = 70\%$, sowie $r_{B-L2} = 90\%$

Für welche Entwurfsalternative würden Sie sich aus Gründen der Leistung entscheiden? Begründen Sie ihre Antwort.

Cache-Kohärenzprotokoll

3P

- Auf welchem Konzept basiert das MESI-Kohärenzprotokoll und was ist hierfür die technische Grundvoraussetzung? 1P
- Geben Sie an, welche Hardwareerweiterungen in den Caches, sowie im Cachecontroller vorhanden sein müssen, damit ein Prozessor in einem SMP-System eingesetzt werden kann, in dem das MESI-Kohärenzprotokoll zum Einsatz kommt. 2P

MESI-Kohärenzprotokoll

4P

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

- Vervollständigen Sie die auf dem Lösungsblatt angegebene Tabelle: Geben Sie jeweils Inhalt der Cache-Zeile und MESI-Zustand an. 4P

Aufgabe 4: Fehlertoleranz

10P

Ausfallverhalten

2P

Den Verlauf des Produktausfallverhaltens in Relation zum Produktalter lässt sich durch eine einfache Kurve illustrieren.

- a) Zeichnen Sie diese Kurve, benennen Sie die einzelnen Phasen dieser Kurve und geben Sie eine kurze Erläuterung zu diesen. 2P

RAID 0-System

4P

Ein RAID 0-System bestehe aus den folgenden Komponenten mit ihren jeweiligen MTTFs:

Komponente	Anzahl	MTTF
SATA-Festplatte FP	2	100 000 h
SATA-Kontroller K	1	100 000 h
Netzteil N	1	100 000 h
SATA-Kabel KB	2	200 000 h

- b) Zeichnen Sie das Zuverlässigkeitsblockdiagramm des RAID 0-Systems. Hinweis: Zum korrektem Betrieb sind alle Festplatten notwendig. 1P
- c) Stellen Sie die Systemfunktion SF des RAID 0-Systems auf. 1P
- d) Berechnen Sie die Ausfallrate und die MTTF des Gesamtsystems unter der Annahme, dass die Fehler der einzelnen Komponenten unabhängig voneinander sind. 2P

RAID 5-System

4P

Um die Ausfallsicherheit weiter zu erhöhen, wird dem System eine weitere SATA-Festplatte hinzugefügt und das System in ein RAID 5-System umgewandelt.

- e) Geben Sie eine allgemeine Formel zur Berechnung der Funktionswahrscheinlichkeit eines n-aus-m Systems an. 1P
- f) Welcher Form der dynamischen Redundanz wird hier verwendet und was ermöglicht sie? 1P
- g) Zeichnen Sie das Zuverlässigkeitsblockdiagramm des RAID 5-Systems. Hinweis: Ein Raid 5-System ist auch dann noch funktionsfähig, wenn eine Festplatte ausgefallen ist. 2P

Aufgabe 5: Quantifizierung

10P

Analytische/empirische Leistungsbewertung:

3P

Ein Rechenzentrumsleiter bekommt ein neues Rechensystem A angeboten und möchte dieses unter Zuhilfenahme von einfachen analytischen Mitteln mit dem installierten Rechensystem B vergleichen.

- a) Das existierende Rechensystem B weist für einen Job eine durchschnittliche Wartezeit von 20 Minuten auf. Für System A gibt der Hersteller folgende Kennzahlen an: 22 Aufträge in der Warteschlange bei einem Durchsatz von $120 \frac{\text{Jobs}}{\text{Stunde}}$. Verwenden Sie das Gesetz von Little, um die Rechensysteme zu vergleichen. Würden Sie einen Wechsel der Systeme empfehlen? 1P
- b) Zur Beobachtung des Netzwerkverkehrs wurde das Betriebssystem um eine Komponente erweitert, welche diesen mitprotokolliert. Nennen Sie den allgemeinen Fachbegriff, den Sie in der Vorlesung für diese Methodik kennengelernt haben und die spezielle Art der Implementierung (jeweils $\frac{1}{2}$ P). 1P
- c) Es stellt sich heraus, dass der Netzwerkverkehr durch die zusätzliche Betriebssystemkomponente unter Last um 20% ausgebremst wird. Welche Möglichkeit zur Realisierung der gewünschten Funktionalität besteht außerdem, die eine geringere Beeinträchtigung des Netzwerkverkehrs verspricht? Nennen Sie eine Grundvoraussetzung für die von Ihnen vorgeschlagene Vorgehensweise. 1P

Leistungsbewertung der Gleitkommaarithmetik:

7P

Ein Programm wird auf einem Prozessor mit einer Taktfrequenz von 2 GHz ausgeführt und weist die folgenden Kenngrößen bei der Ausführung auf:

Befehl	Anzahl in 10^3	Zyklenzahl
Integer-Arithmetik	2.500	3
Fließkomma-Arithmetik	300	5
Speicherzugriff	100	5
Kontrollflußtransfer	250	10

Da eine der Anwendungen des Prozessors die Numerik ist, wird Ihnen vorgeschlagen, die Fließkommaeinheit durch eine spezielle Einheit zur Beschleunigung numerischer Verfahren zu ersetzen. Durch den Einsatz der Beschleunigereinheit beträgt die Anzahl der insgesamt benötigten Zyklen für Fließkommaoperationen 1.750.000, die Anzahl der Befehle bleibt gleich. Allerdings muss der Beschleuniger für seinen Einsatz vorbereitet werden, indem der zu bearbeitende Speicherbereich vorher kopiert wird. Hierfür ergeben sich zusätzlich 250.000 Zyklen für die benötigten Lade- und Speicheroperationen. Die restlichen Kenngrößen seien unverändert.

- a) Stellen Sie das System mit Fließkommaeinheit dem System mit numerischem Beschleuniger gegenüber, indem Sie die jeweiligen MIPS-Werte und Ausführungszeiten berechnen. 2P

- b) Lässt sich für diese spezielle Anwendung ein Speedup gegenüber dem ursprünglichen System erzielen? Welches System schlagen Sie vor und warum? *1P*
- c) Es besteht die Möglichkeit die Fließkommaleistung anhand des MFLOPS-Werts zu berechnen, indem Sie entweder die Gesamtlaufzeit des Programms als Bezugsgröße wählen oder die Ausführungszeit der Fließkommaoperationen. Welche Aussagen können Sie bei der jeweiligen Vorgehensweise treffen und wie unterscheiden sich diese? *1.5P*
- d) Welche der unter Teilaufgabe a) vorgeschlagenen Fließkommaeinheiten hat eine höhere Effizienz bei der Berechnung? Begründen Sie Ihre Antwort. *0.5P*

Für die Livermore Loops (bekannt aus der Vorlesung) wurde die Bewertung der Fließkomma-rechenleistung erweitert: Fließkomma-befehle werden je nach Art mit einem Faktor gewichtet, um eine normalisierte Form zu erreichen.

- e) Erklären Sie beispielhaft an den Befehlen `add` und `div`, warum diese Gewichtung notwendig ist. *1P*
- f) Der Fließkommaanteil eines bestimmten Programms bestehe zu 60% aus `add/sub` und zu 40% aus `div` Befehlen. Die Gewichtung erfolgt mit Hilfe der untenstehenden Tabelle. Der bisherige MFLOPS-Wert des speziellen Programms betrage 300. Berechnen Sie den normalisierten MFLOPS-Wert $MFLOPS_{norm}$. *1P*

Befehl	Faktor
<code>add/sub</code>	2
<code>div</code>	4

Aufgabe 6: Prozessorarchitektur

9P

Pipelining und Leistungsbewertung

4P

- a) Zur Leistungsbewertung von Prozessorarchitekturen wird häufig der Speedup über die Ausführungszeit $T_{\text{pipe}} = n + k - 1$ bei k Pipelinestufen gegenüber $T_{\text{seq}} = n * k$ bei sequentieller Ausführung errechnet. Worin liegt die Formel zur Berechnung der sequentiellen Ausführungszeit begründet? 1P
- b) Eine Prozessorarchitektur verfüge über eine 7-stufige Pipeline, auf der ein Benchmark mit 100.000.000 Befehlen ausgeführt werde, wobei 4% der Befehle einen Datenkonflikt verursachen, der jeweils mit 2 zusätzlichen Zyklen behoben wird, und 4% der Befehle Sprungbefehle sind, die zu einer Verzögerung von jeweils 3 Takten führen. Berechnen Sie die effektive Ausführungsdauer T in Takten sowie Sekunden für eine Taktrate von 2,4 GHz, den erzielten Speedup gegenüber sequentieller Ausführung und die Effizienz (jeweils mit Formel). 2P

- c) 1P

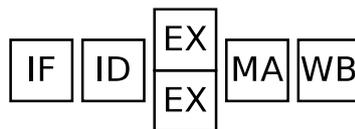


Abbildung 1: Pipeline mit verdoppelter Ausführungseinheit

Nehmen Sie an, eine weitere Architektur verfüge über zwei Ausführungseinheiten, so dass zwei unterschiedliche Daten mittels desselben Befehls gleichzeitig verarbeitet werden können. Die Taktfrequenz sei allerdings nur halb so hoch. Wo liegt der Nachteil bei dieser Architektur? Wie können Sie den Nachteil beheben, ohne die Taktfrequenz zu erhöhen?

Sprungvorhersage

3P

- d) Skizzieren Sie den Aufbau eines (2,2)-Korrelationsprädiktors mit globaler Indizierung. 1,5P
- e) Gegeben sei ein global indizierter (1,1)-Korrelationsprädiktors mit Tabelleneinträgen, die jeweils mit *Taken* (T) initialisiert seien; das globale Sprungverlaufsregister sei mit *Not Taken* (NT) belegt. 1,5P

Vervollständigen Sie die auf dem Lösungsblatt befindliche Tabelle. Pro korrekter Zeile erhalten Sie 0,5P.



Karlsruher Institut für Technologie
Institut für Technische Informatik
Prof. Dr. Wolfgang Karl

Klausur Rechnerstrukturen

Wintersemester 2010/11

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 per Aushang (Matrikelnummer und Note) am schwarzen Brett
(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	/11	/10	/10	/10	/10	/9
Summe:						/60

Lösung 1: Parallelverarbeitung**11P****Quantitative Maßzahlen****3P**

a)

1P

b) •

1P

• Abschätzung:

c)

*1P***Verbindungsstrukturen****3P**

d)

*1P*e) $N =$ *1P*

Erweiterbarkeit:

f) •

1P

- Nachteil:

Parallele Architekturen und Parallelverarbeitung:**3P**

g) • NORMA:

1P

- NUMA:

- UMA:

h)

1P

Multiprozessor mit gemeinsamem Speicher	
Multiprozessor mit verteiltem Speicher	
Multiprozessor mit verteiltem gemeinsamen Speicher	

i) • Programmiermodell:

1P

- Kommunikationsarchitektur:

Vektorverarbeitung:**2P**

j) • Vorteil:

1P

- Nachteil:

k)

1P

Lösung 2: Fragen des Rechnerentwurfs*10P***Low-Power-Entwurf****4P**a) $P_{\text{Gesamt}} =$ *2,5P*

Erläuterung:

b) Formel:

1,5P

Berechnung:

Name:

Matrikelnummer:

6/16

e) Funktionalität der Schaltungsbeschreibung:

2P

Lösung 3: Speicherhierarchie**10P****Cache-Leistung****4P**

a) Formel:

1P

b) Antwort:

*2P***Cache-Kohärenzprotokoll***7P*

c) Antwort:

1P

d) Antwort:

2P

MESI-Kohärenzprotokoll**4P**

e) Tabelle:

4P

Prozessor	Aktion	Prozessor 1		Prozessor 2		Prozessor 3	
		Line 1	Line 2	Line 1	Line 2	Line 1	Line 2
	init	-	-	-	-	-	-
3	rd 3					3/E	
2	rd 2			2/E			
3	wr 4						4/M
1	rd 1	1/E					
1	wr 3						
2	rd 4						
3	wr 1						
3	wr 2						
2	rd 1						
1	rd 1						
3	rd 2						
2	rd 2						

Lösung 4: Fehlertoleranz**10P****Ausfallverhalten****2P**

a) Antwort:

*2P***RAID 0-System****4P**

b) Zuverlässigkeitsblockdiagramm:

1P

c) Systemfunktion:

1P

d) Ausfallrate und MTTF:

2P

Name:

Matrikelnummer:

10/16

RAID 5-System

4P

e) Formel:

1P

f) Antwort:

1P

g) Zuverlässigkeitsblockdiagramm:

2P

Lösung 5: Quantifizierung

10P

Analytische/empirische Leistungsbewertung:

a) Gesetz von Little:

1P

Rechnung & Empfehlung:

b) Fachbegriff:

1P

Art der Implementierung:

c) Antwort:

1P

Leistungsbewertung der Gleitkommaarithmetik**7P**

a) Nebenrechnung:

2P

$$t_{fpu} =$$

$$t_{num} =$$

$$MIPS_{fpu} =$$

$$MIPS_{num} =$$

b)

1P

$$Speedup_{num} =$$

Begründung:

Name:

Matrikelnummer:

13/16

c) Antwort:

1.5P

d) Antwort:

0.5P

e) Erklärung:

1P

f) Formel & Berechnung:

1P

Lösung 6: Prozessorarchitektur*9P***Pipelining und Leistungsbewertung****4P**

a) Antwort:

1P

b)

2P

- Ausführungsdauern:

- Speedup:

- Effizienz:

c) Antwort:

1P

Sprungvorhersage**3P**

d) Skizze:

1,5P

e) Sprungverlaufstabelle:

1,5P

Verlauf	Prädiktor	Vorhersage	Sprung	Akt. Prädiktor
			NT	
			T	
			NT	

