

Kurs 01609 Computersysteme II

Lösungsvorschläge zu Einsendearbeit 4

Aufgabe 1 (8 Punkte)

Moore'sches Gesetz

Das Moore'sche Gesetz besagt, dass alle 18 Monate die Anzahl der Transistoren verdoppelt wird.

a) Die SIA-Prognose sagt für das Jahr 2016 einen Hochleistungsprozessor-Chip mit 3,092 Milliarden Transistoren voraus. Rechnen Sie diese Prognose anhand des Moore'schen Gesetzes zurück auf das Jahr 2003. Ist die SIA-Prognose optimistisch oder pessimistisch? (3 P)

b) Nennen Sie Gründe für die Grenzen des Wachstums mit heutiger Technologie. (5 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 2 (6 Punkte)

a) Nennen Sie verschiedene Arten von Latenzen. Warum verringert oder beseitigt ein mehrfädiger Prozessor Latenzen? (3 P.)

b) Mehrfädigkeit kann auch mit einem "normalen" Prozessor durch den Einsatz eines geeigneten Betriebssystems erreicht werden. Welchen Vorteil bietet dagegen ein mehrfädiger Prozessor? (3 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 3 (16 Punkte)

Cycle-by-Cycle Interleaving Gegeben seien sechs Kontrollfäden (Threads) T_1, \dots, T_6 , die auf einem mehrfädigen Prozessor ablaufen, der das Cycle-by-Cycle Interleaving Verfahren verwendet. Die Threads haben folgende Eigenschaften:

- T_1 gibt zunächst einen Befehl aus, der einen Speicherzugriff erfordert. Danach werden zwei weitere Befehle ausgegeben, die jedoch keine Latenzen bewirken.
- T_2 gibt zunächst drei latenzfreie Befehle aus und dann einen Befehl der auf den Speicher zugreift
- T_3 besteht aus fünf latenzfreien Befehlen.
- T_4 gibt zunächst zwei latenzfreie Befehle aus. Der dritte Befehl bewirkt eine Pipeline-Latenz von zwei Taktzyklen. Darauf folgen wieder zwei latenzfreie Befehle.
- T_5 hat die gleichen Eigenschaften wie T_1 .
- T_6 besteht aus vier latenzfreien Befehlen.

Die Latenz eines Speicherzugriffs beträgt 10 Taktzyklen.

a) Skizzieren Sie das Laufzeitverhalten der Threads bei einfädiger Befehlsverarbeitung. (5 P.)

b) In welcher Reihenfolge werden die Threads auf dem mehrfädigen Prozessor abgearbeitet?

Der Prozessor soll mit T_1 beginnen und mit jedem Takt zyklisch zwischen den Threads umschalten (Round-Robin). Markieren Sie die Latenzzeiten! (5 P.)

c) Vergleichen Sie die Zahl der benötigten Taktzyklen bei ein- und mehrfädiger Verarbeitung! (3 P.)

d) Bestimmen Sie den Grad der Prozessoreffizienz für die beiden Fälle aus c)! Der Grad der Effizienz ist der Quotient aus Anzahl ausgeführter Befehle und Anzahl Takte der gesamten Ausführung (inklusive Latenz). (3 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 4 (20 Punkte)

Block-Interleaving

Gegeben seien vier Kontrollfäden (Threads), die auf einem mehrfädigen Prozessor ablaufen, der die Block-Interleaving-Technik verwendet. Die Threads haben folgende Eigenschaften:

- T_A gibt zunächst einen Befehl aus, der einen Speicherzugriff erfordert. Danach werden vier latenzfreie Befehle ausgeführt.
- T_B gibt zuerst einen Befehl aus, der zu einer Pipeline-Latenz von zwei Taktzyklen führt. Darauf folgt ein Befehl mit anschließender Speicher-Latenz. Daran schließen sich zwei latenzfreie Befehle an.
- T_C besteht aus zwei latenzfreien Befehlen, gefolgt von einem Befehl der eine Speicher-Latenz bewirkt. Dann folgen drei latenzfreie Befehle.
- T_D besteht aus drei latenzfreien Befehlen, gefolgt von einem Befehl der eine Speicher-Latenz bewirkt. Dann folgt ein latenzfreier Befehl.

Die Latenz eines Speicherzugriffs umfasst 10 Taktzyklen.

- a) Skizzieren Sie das Laufzeitverhalten der Threads bei einfädiger Befehlsverarbeitung. (5 P.)
- b) Der Prozessor verfüge über eine fünfstufige Pipeline mit den Stufen Instruction-Fetch (IF), Instruction-Decode (ID), Execute (EX), Memory Access (MEM) und Write-Back (WB). Die Kontextumschaltung erfolgt, sobald ein Speicherzugriff erforderlich ist. Wenn ein Thread beendet wird, erfolgt der Kontextwechsel ohne zusätzliche Taktzyklen. Wie viele Taktzyklen werden zur Kontextumschaltung benötigt, wenn der Rechner über Cache-Speicher verfügt? (4 P.)
- c) Skizzieren Sie den Ablauf der Threads auf einem mehrfädigen Prozessor unter der Annahme, dass stets ein Cache-Fehlzugriff vorliegt. Jeder Speicherzugriff führt demnach zu einer Speicher-Latenz von 10 Taktzyklen. (5 P.)
- d) Vergleichen Sie die Anzahl der benötigten Taktzyklen bei ein- und mehrfädiger Bearbeitung! (3 P.)
- e) Bestimmen Sie den Grad der Prozessoreffizienz für die beiden Fälle aus d). (3 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 5 (5 Punkte)

Amdahls Gesetz

$$S(p) = \frac{1}{f_s + \frac{1-f_s}{p}}$$

b) Berechnen Sie den Speedup, wenn 2, 4, 8 und 16 Prozessoren verwendet werden und man von einem seriellen Anteil von 60% ausgeht. (3 P.)

b) Bestimmen Sie die Effizienz $E(p)$ für die Ergebnisse aus A). Was kann man aus diesen Zahlen folgern? (2 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 6 (10 Punkte)

Gustafsons' Gesetz

Multiprozessorsysteme werden immer dann eingesetzt, wenn große Datenmengen zu verarbeiten sind. Je leistungsfähiger ein Parallelrechner ist, desto größer ist die Last, die er bewältigen kann. Im Folgenden gehen wir daher davon aus, dass die Last mit der verfügbaren Rechenleistung so skaliert, dass die Ausführungszeit konstant bleibt. Der Einfachheit halber nehmen wir an, dass die Ausführungszeit auf einem Multiprozessorsystem mit p Prozessoren 1 Sekunde (sec) betrüge, d.h. $t_M = t_s + t_p = 1 \text{ sec}$, wobei t_s der sequentielle Anteil und t_p der parallele Anteil der gesamten Ausführungszeit t_M ist. Entsprechend gilt $f_s + f_p = 1$, wobei f_s bzw. f_p die entsprechenden Anteile beschreiben.

- a) Bestimmen Sie damit die Ausführungszeit t_E auf einem Einprozessorsystem! (2 P.)
- b) Berechnen Sie mit dem Ergebnis aus a) den skalierten Speedup $S_{skaliert}$ als Funktion des sequentiellen Anteils f_s . (4 P.) Dabei wird die Funktion $S_{skaliert} = f(f_s)$ als Gustafsons Gesetz bezeichnet.
- c) Vergleichen Sie den Speedup nach Amdahls und Gustafsons Gesetz anhand eines Diagramms, das den Speedup als Funktion von f_s für ein Multiprozessorsystem mit 8 Prozessoren darstellt. (2 P.)
- d) Skizzieren Sie den Speedup nach Amdahl und Gustafson als Funktion der Prozessorzahl für ein Multiprozessorsystem mit 4 Prozessoren! Wählen Sie $f_s = 0$, $f_s = 0,8$ und $f_s = 1$ und diskutieren Sie die Kurven! (2 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 7 (15 Punkte)

Kommunikation und Speedup

Analysieren Sie den Einfluss der Kommunikationszeit in einem nachrichtengekoppelten Multiprozessor-system auf den erreichbaren Speedup. Das parallele Programm basiere auf einem Master-Slave Modell. Der Master und die Slaves laufen auf unterschiedlichen Knoten und es findet nur Kommunikation zwischen dem Masterprozess und den Slaveprozessen statt. Nehmen Sie an, dass die Knoten durch ein non-switched Ethernet miteinander verbunden sind. Die Zeitdauer für die Kommunikation zwischen Master und Slave betrage t_C . Es ist davon auszugehen, dass die Kommunikation zwischen Master und Slave in beiden Richtungen vor sich geht und somit $2t_C$ pro Slave für die Kommunikation anfallen. Die Rechendauer eines Slaves wird mit t_S bezeichnet. Der Master verteilt zu Beginn lediglich die Aufgaben auf die Slaves und sammelt am Ende der Berechnung deren Ergebnisse. Die Rechendauer t_M des Master ist vernachlässigbar, d.h. $t_M = 0$.

- a) Entwickeln Sie eine Gleichung für den Speedup $S_n(r)$ als Funktion des Quotienten $r = t_C/t_S$ (n steht für die Anzahl Slaves). (7 P.)
- b) Zeichnen Sie die Funktion für $S_n(r)$ mit $n=2,4,8,16$ und 32 für $0 \leq r \leq 0,2$. (3 P.)
- c) Bestimmen Sie den Wert für r_1 , bei dem die Leistung des Master-Slave Ansatzes dem eines nicht-parallelen Systems entspricht. (5 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.

Aufgabe 8 (20 Punkte)

Optimale Prozessoranzahl

In der vorangegangenen Aufgabe war die Anzahl der Knoten (Slaves) vorgegeben. Hier sollen Sie nun die optimale Anzahl von Knoten errechnen, wenn die Anzahl der auszuführenden Prozesse vorgegeben ist. Gehen Sie wieder von einem Master-Slave Modell aus, bei dem $p = 128$ Knoten (inklusive Master, der ebenfalls als Slave benutzt wird, ohne jedoch Kommunikationszeit zu verbrauchen) zur Verfügung stehen auf denen $n = 256$ Prozesse abgearbeitet werden sollen. Die Knoten sind mittels eines non-switching Ethernet verbunden. Die Übertragungsrate f_T beträgt 12 Mbyte pro Sekunde. Die durchschnittliche Latenz ist $t_L = 200 \mu s$ (inklusive Kollisionen). Ein Slave-Prozess erhält $d_{in} = 180$ Bytes Eingabedaten und produziert $d_{out} = 60$ Bytes Ausgabedaten. Die Berechnungsdauer eines Slaves beträgt $t_S = 3ms$.

- a) Entwickeln Sie Gleichungen für t_{seq} und t_{par} . (6 P.)
- b) Berechnen Sie $t_{parallel}$ für $p = 2, 4, 8, 16, 32, 64$ und 128 und fassen Sie die Ergebnisse tabellarisch zusammen. Welche Knotenanzahl ist optimal? (3 P.)
- c) Welcher Parameter begrenzt den Speedup? (3 P.)
- d) Leiten Sie eine Formel zur Berechnung der optimalen Knotenanzahl her. Die sich ergebende Funktion soll abhängig sein von: n , f_T , t_L , d_{in} , d_{out} und t_S . Vergleichen Sie das Ergebnis mit dem Aufgabenteil B) und zeichnen Sie den Verlauf von t_{par} abhängig von der Knotenzahl p als Funktionskurve oder Wertetabelle auf! (8 P.)

Zur Einsendung dieser Aufgabe erstellen Sie bitte eine PDF-Datei mit den von Ihnen bearbeiteten Lösungen. Laden Sie diese PDF-Datei in WebAssign hoch, so dass wir auf Ihre Lösungen zugreifen und korrigieren können.