

Technische Informatik 1 – Übung 10

Cache (Rechenübung)

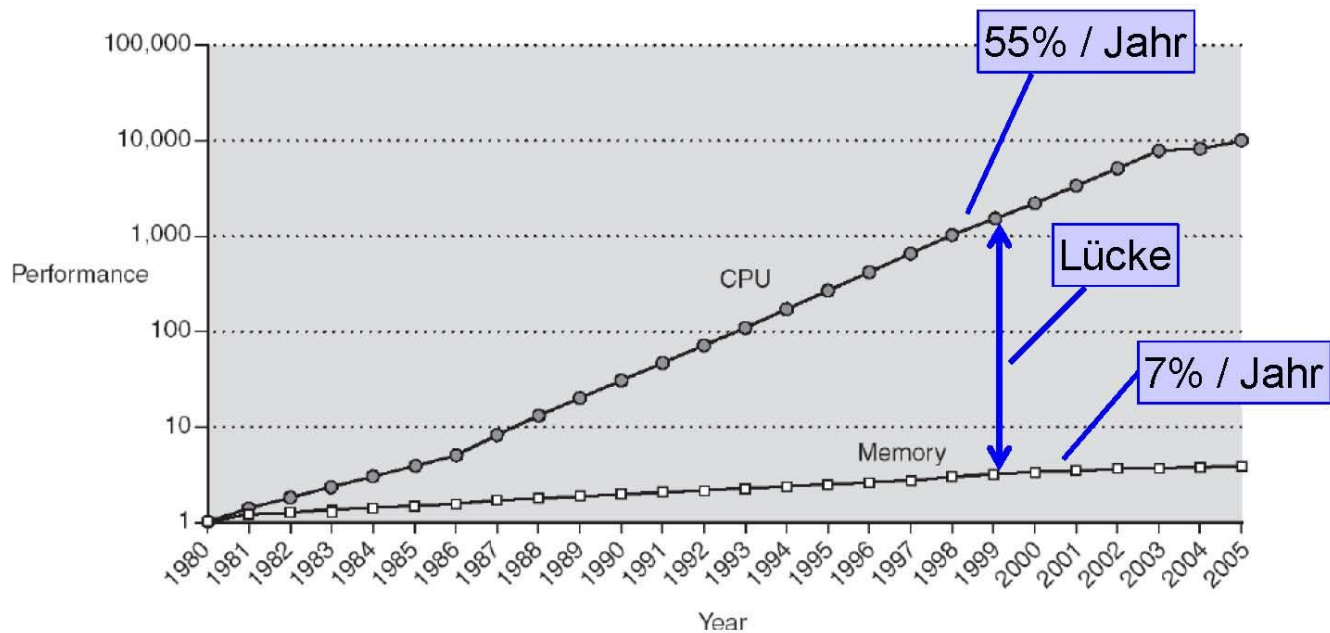
David Hasenfratz
15.-16. Dezember 2011



Motivation

Technologische Trends

► Zugriffszeit DRAM vs. Zykluszeit Prozessor

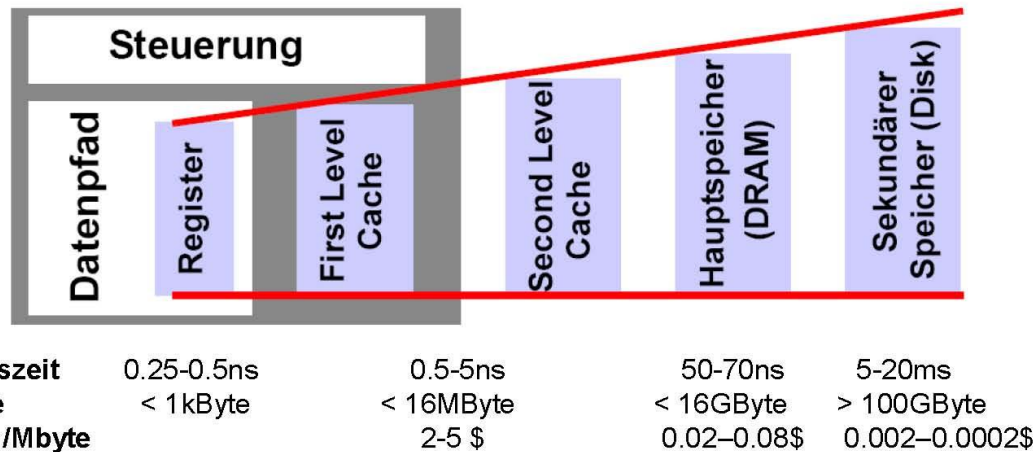


Verringerung der Zugriffszeit durch Caches

Übersicht

► Ziele:

- Dem Benutzer *möglichst viel* Speicherkapazität zur Verfügung stellen bei möglichst *kleiner Zugriffszeit*.
- Möglichst *kostengünstig*.



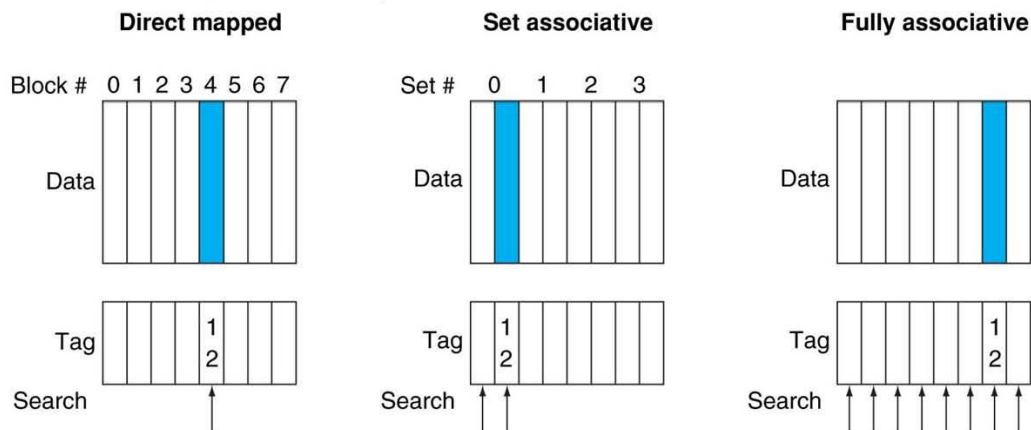
Ziele der Übung

- Verschiedene Cache-Konfigurationen
- Berechnung der Cache-Grösse
- Verwendung eines Caches
- Cache Misses

Wiederholung

Wo kann ein Block platziert werden?

- Platzierung eines Speicherblocks mit der Adresse 12:



Cacheblock, in dem der Speicherblock 12 liegt:

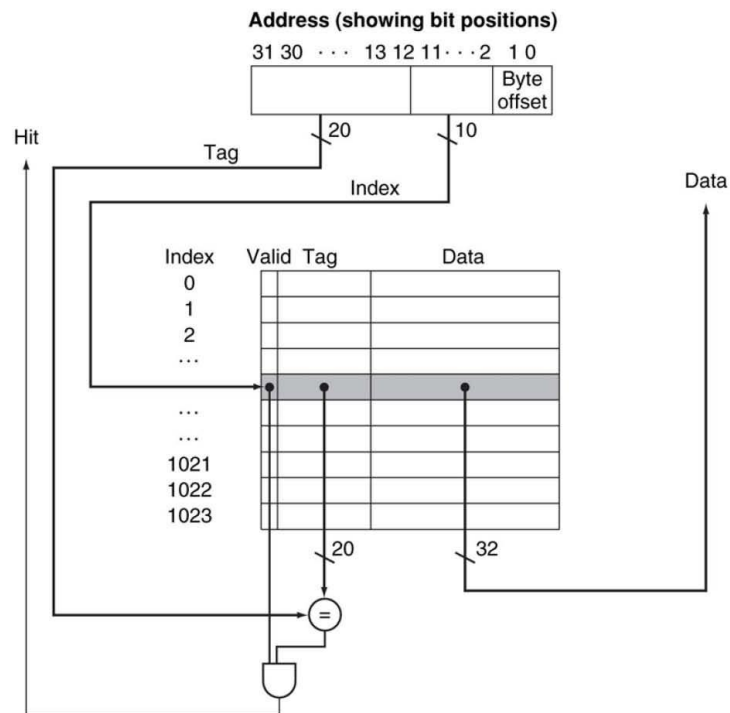
direkte Abbildung: $(12 \bmod 8) = 4$

teillassoziativ: $(12 \bmod 4) = 0$

vollassoziativ: beliebig

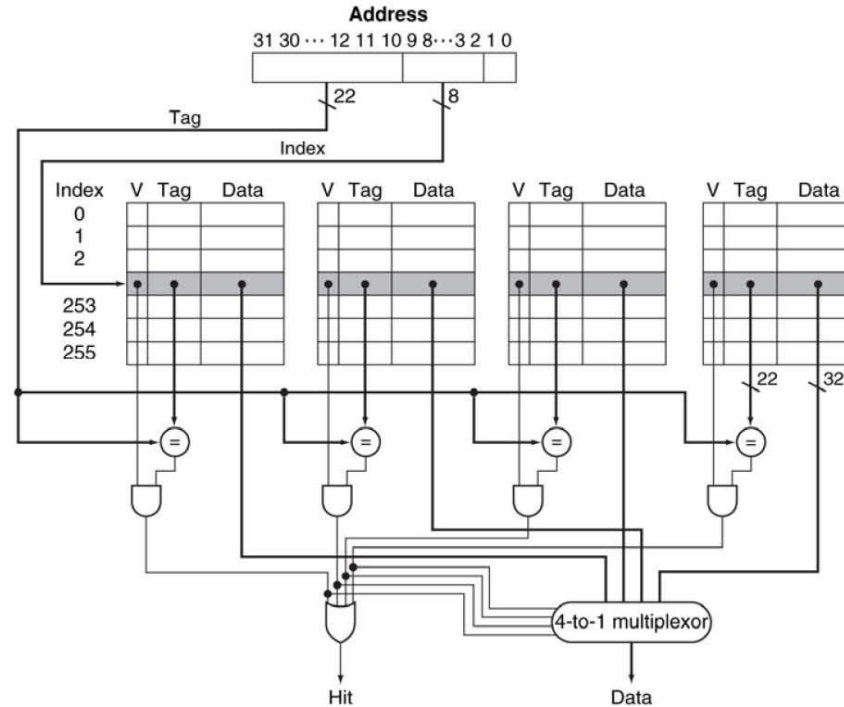
Wiederholung (Fort.)

Direkte Abbildung



Wiederholung (Fort.)

Diagramm eines teilassoziativen Caches



Aufgabe 1: Setup

Gegeben seien drei Rechner mit unterschiedlichen Cache-Kongurationen:

- Rechner A: direkte Abbildung (direct-mapped).
- Rechner B: teilassoziativ mit 2 Einträgen pro Index (2-way set associative).
- Rechner C: vollassoziativ (full associative).

Aufgabe 1 a)

- Jeder Cache hat eine Kapazität von 4 Blöcken à 2 Worte.
- Berechnen Sie für jeden Rechner die effektive Cachegrösse in Bit. Gehen Sie hierbei von einer 32-Bit-Byteadressierung aus. Von den Statusbits soll lediglich das Valid-Bit berücksichtigt werden, andere Statusbits (Dirty-Bit, LRU, etc.) sind in dieser Teilaufgabe nicht zu berücksichtigen.
- Tipp: Vorlesung Kapitel 11, Folie 16

Lösung 1 a)

Reduktion der Tag-Grösse:

- Im Cache sind nur Worte adressierbar: 2 Bit Byte-Offset
- Eine Cacheline speichert 2 Worte: 1 Bit Word-Offset
- Anzahl Bits für Index: 0-2 Bit, je nach Konfiguration

Ergebnisse:

- Rechner A: Tag-Grösse = $(32-3-2)$ Bit = 27 Bit
- Rechner B: Tag-Grösse = $(32-3-1)$ Bit = 28 Bit
- Rechner C: Tag-Grösse = $(32-3-0)$ Bit = 29 Bit

Lösung 1 a) (Fort.)

Für jeden Block werden 1 Bit für das Valid-Feld, $2 \cdot 32$ Bit für die Daten (2 Worte) und eine gewisse Anzahl von Bits für das Tag-Feld benötigt.

Die Cachegrösse berechnet sich durch die Multiplikation von Cache-Zeilen mit effektiver Blockgrösse:

- Rechner A: $4 \cdot (1+64+27)$ Bit = 368 Bit
- Rechner B: $4 \cdot (1+64+28)$ Bit = 372 Bit
- Rechner C: $4 \cdot (1+64+29)$ Bit = 376 Bit

Aufgabe 1 b)

- Nach dem Einschalten des Rechners wird sequentiell auf die folgenden Wortadressen zugegriffen: 19, 10, 19, 14, 11, 15.
- Wie ist der Zustand der einzelnen Caches vor und nach jedem Speicherzugriff? Wann treten Fehlzugriffe (Misses) auf? Nennen Sie für jeden Cache die Zahl der insgesamt auftretenden Fehlzugriffe.
- Nehmen Sie an, dass als Blockersetzungsschema das LRU-Verfahren (Least Recently Used) benutzt wird, d.h. der am längsten unbenutzte Block wird ersetzt.

Hilfestellung 1 b)

Rechner A: ? Misses

Adr	Adresse binär				Hit/Miss	Ind 00	Ind 01	Ind 10	Ind 11
	tag	idx	wr-o	by-o					
19									
10									
19									
14									
11									
15									

Tag (tag), Index (idx), Word-offset (wr-o) und Byte-offset (by-o)

Lösung 1 b)

Rechner A: 5 Misses

Adr	Adresse binär				Hit/Miss	Ind 00	Ind 01	Ind 10	Ind 11
	tag	idx	wr-o	by-o					
19	10	01	1	00	Miss compul		18,19		
10	01	01	0	00	Miss confl		10, 11		
19	10	01	1	00	Miss confl		18, 19		
14	01	11	0	00	Miss compul		18, 19		14, 15
11	01	01	1	00	Miss confl		10, 11		14, 15
15	01	11	1	00	Hit		10, 11		14, 15

Tag (tag), Index (idx), Word-offset (wr-o) und Byte-offset (by-o)

Lösung 1 b) (Fort.)

Rechner B: 4 Misses

Adr	Adresse binär				Hit/Miss	Ind 0		Ind 1	
	tag	idx	wr-o	by-o		Blk 0	Blk 1	Blk 0	Blk 1
19	100	1	1	00	Miss compul			18, 19	
10	010	1	0	00	Miss compul			18, 19	10, 11
19	100	1	1	00	Hit			18, 19	10, 11
14	011	1	0	00	Miss confl			18, 19	14, 15
11	010	1	0	00	Miss confl			10, 11	14, 15
15	011	1	1	00	Hit			10, 11	14, 15

Tag (tag), Index (idx), Word-offset (wr-o) und Byte-offset (by-o)

Lösung 1 b) (Fort.)

Rechner C: 3 Misses

Adr	Adresse binär			Hit/Miss	Ind 0			
	tag	wr-o	by-o		Blk 00	Blk 01	Blk 10	Blk 11
19	1001	1	00	Miss compul	18, 19			
10	0101	0	00	Miss compul	18, 19	10, 11		
19	1001	1	00	Hit	18, 19	10, 11		
14	0111	0	00	Miss compul	18, 19	10, 11	14, 15	
11	0101	1	00	Hit	18, 19	10, 11	14, 15	
15	0111	1	00	Hit	18, 19	10, 11	14, 15	

Tag (tag), Word-offset (wr-o) und Byte-offset (by-o)

Aufgabe 1 c)

- Betrachten Sie die folgenden Fehlzugriffsraten:
 - Rechner A: Bei Instruktionen 4%, bei Daten 8%.
 - Rechner B: Bei Instruktionen 2%, bei Daten 5%.
 - Rechner C: Bei Instruktionen 2%, bei Daten 4%.
- Bei den drei Rechnern seien $\frac{3}{5}$ aller Instruktionen Speicherinstruktionen, die Miss-Strafe betrage 7 Zyklen bei Rechner A und 11 Zyklen bei den Rechnern B und C. Welcher Rechner wendet am meisten Zyklen für die Behandlung von Cache-Misses auf?
- Tipp: Vorlesung Kapitel 11, Folie 29

Lösung 1 c)

Bei den drei Rechnern seien $\frac{3}{5}$ aller Instruktionen Speicherinstruktionen:

$$\begin{aligned}
 \text{Miss-Zykle } n &= \text{Instruktionen-Miss-Zyklen} + \text{Daten-Miss-Zyklen} \\
 &= \# \text{ Instruktionen} \times \text{Instruktionen-Miss-Rate} \times \text{Miss-Straf-Zyklen} \\
 &\quad + \frac{3}{5} \times \# \text{ Instruktionen} \times \text{Daten-Miss-Rate} \times \text{Miss-Straf-Zyklen} \\
 &= \# \text{ Instruktionen} \times \text{Miss-Straf-Zyklen} \times \left(\text{Instruktionen-Miss-Rate} + \frac{3}{5} \times \text{Daten-Miss-Rate} \right)
 \end{aligned}$$

$$\frac{\text{Miss-Zykle } n}{\# \text{ Instruktionen}} = \text{Miss-Straf-Zyklen} \times \left(\text{Instruktionen-Miss-Rate} + \frac{3}{5} \times \text{Daten-Miss-Rate} \right)$$

Lösung 1 c) (Fort.)

- Es gilt:

$$\frac{\text{Miss-Zykle } n}{\# \text{Instrukti onen}} = \text{Miss-Straf -Zyklen} \times \left(\text{Instruktio ns-Miss-Ra te} + \frac{3}{5} \times \text{Daten-Miss -Rate} \right)$$

- Rechner A: Miss Instruktionen 4%, Daten 8%

$$\frac{\text{Miss-Zykle } n}{\# \text{Instrukti onen}} = 7 \times \left(0,04 + \frac{3}{5} \times 0,08 \right) = 0,616$$

- Rechner B: Miss Instruktionen 2%, Daten 5%

$$\frac{\text{Miss-Zykle } n}{\# \text{Instrukti onen}} = 11 \times \left(0,02 + \frac{3}{5} \times 0,05 \right) = 0,55$$

- Rechner C: Miss Instruktionen 2%, Daten 4%

$$\frac{\text{Miss-Zykle } n}{\# \text{Instrukti onen}} = 11 \times \left(0,02 + \frac{3}{5} \times 0,04 \right) = 0,484$$

Aufgabe 1 d)

- Wir möchten Rechner C um den Faktor 4 beschleunigen unter der Annahme, dass die Speicherzugriffszeiten unverändert bleiben. Um welchen Faktor muss dann die Taktrate erhöht werden? Der CPI bei perfektem Cache sei 3,388, bezüglich der Miss-Raten gelten die Angaben aus Teilaufgabe c).
- Tipp: Vorlesung Kapitel 11, Folie 30

Lösung 1 d)

Sei k der gesuchte Faktor, $\text{CPUZeit}_{\text{normal}}$ die normale Rechenzeit von Rechner C und $\text{CPUZeit}_{\text{schnell}}$ die Rechenzeit des beschleunigten Rechners. Da sich die Speicherzugriffszeiten nicht ändern, hat der beschleunigte Rechner k -mal so viele Miss-Zyklen wie Rechner C.

$$\begin{aligned}
 4 &= \frac{\text{CPUZeit}_{\text{normal}}}{\text{CPUZeit}_{\text{schnell}}} = \frac{\text{CPI}_{\text{normal}} \times \text{Instruktionen} \times \frac{1}{\text{Taktrate}}}{\text{CPI}_{\text{schnell}} \times \text{Instruktionen} \times \frac{1}{k \times \text{Taktrate}}} \\
 &= \frac{\text{CPI}_{\text{perfekt}} + 0,484}{(\text{CPI}_{\text{perfekt}} + k \times 0,484) \times \frac{1}{k}} \\
 &= \frac{k \times 3,872}{3,388 + k \times 0,484} \\
 \Rightarrow k &= 7
 \end{aligned}$$