

Problema 6. Repaso Cache

L'empresa A.C.M.E. està dissenyant un nou processador del que sabem que genera **1.3 referències** a memòria per instrucció de les que **0.3 son de dades**. Els seus dissenyadors s'han adonat de que al mateix xip queda espai suficient per posar-hi un poc de cache. Després de fer alguns càlculs resulta que poden posar-hi una cache de 16k o dues (una de 8k i una de 4k), i han optat per una de les dues configuracions següents:

- Una sola cache **unificada de 16Kb**. En aquest cas (si la cache fos ideal) tenim un CPI_{ideal} de 1.5 cicles/instrucció degut a que quan accedim a la cache per buscar una dada no podem buscar una instrucció al mateix temps.
- Dues caches separades, una d'**instruccions de 4Kb** i una de **dades de 8Kb**. En aquest cas el CPI_{ideal} es de 1.2 cicles/instrucció ja que podem buscar dades i instruccions simultaniament.

En qualsevol de les dues configuracions:

- T_c (temps de cicle): 10ns
- T_{sa} (temps de servei en cas d'encert): 1cicle.
- T_{sf} (temps de servei en cas de fallada): 10 cicles.

La taxa de fallades, per diferents configuracions de cache, es mostra a la següent taula:

Mida	Instruccions	Dades	Unificada
4K	8.6%	8.7%	11.2%
8K	5.8%	6.8%	8.3%
16K	3.6%	5.3%	5.9%

Es demana:

- Quin serà el temps mig d'accés T_{ma} per cada configuració (en cicles)?
- Quin serà el temps d'execució T_{exec} de 1 instrucció real en cada cas?
- Per quina opció optarieus i perquè? (en una línia)
- Creus que es pot trobar alguna opció millor en base a les **dades de que disposem**? En cas afirmatiu, digueu quina i perquè.

$$a) T_{ma_{16}} = t_{sa} + m \cdot t_{sf} = h \cdot t_{sa} + m \cdot t_{sf} = (1 - 0.059) \cdot 1 + 0.059 \cdot 10 = 1.531 \text{ c}$$

$$T_{ma_{8+8}} = \frac{0.3 \cdot [(1 - 0.068) + 0.068 \cdot 10] + 1 \cdot [(1 - 0.086) + 0.086 \cdot 10]}{1.3} = 1.7366 \text{ c}$$

$$b) T_{exec} = N \cdot CPI \cdot T_c = T_c \cdot (CPI_{ideal} + CPI_{mem}) = T_c \cdot (CPI_{ideal} + mr(T_{ma} - t_{sa}))$$

$$T_{exec_{16}} = 10 (1.5 + 1.3 \cdot (T_{ma_{16}} - 1)) = 21.903 \text{ ms} \quad | \quad T_{exec_{8+8}} = 10 (1.5 + 1.3 \cdot (T_{ma_{8+8}} - 1)) = 21.576 \text{ ms}$$

$$c) T_{exec_{8+8}} < T_{exec_{16}} \Rightarrow T_{exec_{8+8}}$$

$$d) swap tornaros dades - inst.$$

- Menys referències a dades (no importa que fallem més)
 - Mes enerts a inst. (més utilitzat \rightarrow més millora)
- } Llei de Amotah.

Problema 9. Caches petites y simples

- a) Indiqueu quins accessos seran *hit* (amb una X) per cada una de les tres possibilitats per la següent seqüència de referències a bloc (en octal) on tots els accessos son lectures. En el cas de la directa + VC es considerarà *miss* si el bloc referenciat no es troba ni a MC ni a VC.

Bloc de memòria	73	55	43	45	73	45	13	43	73	55	45	73	15	43
Directa						X						X		
2-associativa					X	X				X	X	X		X
Directa + VC					X	X		X	X	X	X	X		

- b) Creus que hi hauria cap diferència si la VC fes servir un reemplaçament LRU? Perquè?

No, porque a VC no es té en compte els accessos a MC, per tant, l'única vegada que una entrada es la més recent és en el moment d'enllaç. Es diu: FIFO = LRU (en VC)

Volem estudiar la implementació d'aquestes 3 caches com a cache de dades en un processador. Per un programa P que només fa lectures executat en aquest processador amb memòria ideal (on tots els accessos a memòria tarden 1 cicle) sabem que ha executat 10×10^9 instruccions en 12×10^9 cicles i s'han fet 3×10^9 accessos a memòria (dades).

- c) Calculeu el CPI amb memòria ideal (CPI_{ideal}).
d) Calculeu el ratio nr (accessos a memòria per instrucció).

$$c) CPI_{ideal} = \frac{12 \cdot 10^9}{10 \cdot 10^9} = 1.2 \text{ c/i}$$

$$d) nr = \frac{3 \cdot 10^9 \text{ refs}}{10 \cdot 10^9 \text{ i.}} = 0.3 \text{ accessos/i.}$$

Cache d'**emplaçament directe**: El temps de cicle (Tc) es de 10 ns/cicle, la taxa de fallades (m) es de 0.1 fallades/accés i el temps de penalització en cas de fallada (Tp) es de 10 cicles.

- e) Quants cicles tarda en executar-se el programa P?
f) Quin es el temps d'execució de P? (en segons)

$$e) \text{cycles} = m \cdot CPI = m (CPI_{ideal} + CPI_{mem}) = m \cdot (P_{ideal} + m \cdot mr \cdot m \cdot T_p) = 10 \cdot 10^9 \cdot (1.2 + 0.3 \cdot 0.1 \cdot 10) = 12 \cdot 10^9 + 3 \cdot 10^9 \cdot 0.1 \cdot 10 = 15 \cdot 10^9 \text{ cicles}$$

$$f) T_{exe} = T_c \cdot m \cdot CPI = T_c \cdot \#cycles = 10 \cdot 10^{-9} \cdot 15 \cdot 10^9 = 150 \text{ s}$$

Cache **2-associativa**: El temps de cicle (Tc) es de 12 ns/cicle, la taxa de fallades (m) es de 0.05 fallades/accés i el temps de penalització en cas de fallada (Tp) es de 9 cicles.

- g) Perquè creus que el temps de penalització en cas de fallada es de 9 cicles mentre que en el cas d'**emplaçament directe** era de 10 cicles si sabem que la memòria principal es la mateixa?
h) Quants cicles tarda en executar-se el programa P?
i) Quin es el temps d'execució de P? (en segons)

$$h) \# \text{cycles} = m \cdot CPI = m \cdot (CPI_{\text{ideal}} + CPI_{\text{mem}}) = 16 \cdot 10^9 (1.2 + 0.3 \cdot 0.05 \cdot 9) = 13'35 \cdot 10^9 \text{ cicles}$$

$$i) T_{\text{exe}} = \# \text{cycles} \cdot T_c = 13'35 \cdot 10^9 \cdot 12 \cdot 10^{-9} = 160'2 s$$

Cache **emplaçament directe + victim cache** amb accés simultani: A pesar que la victim cache es més ràpida que la d'emplaçament directe, la lògica i el multiplexor necessaris per controlar de quina cache obtindrem la dada fa que el temps d'accés sigui lleugerament més alt que en el cas de la cache directa. El temps de cicle (T_c) es de 11 ns/cicle, la taxa de fallades (m) global del conjunt MC+VC es de 0.06 fallades/accés i el temps de penalització en cas de fallada (T_{pf}) es de 10 cicles.

- j) Quants cicles tarda en executar-se el programa P?
- k) Quin es el temps d'execució de P? (en segons)

$$j) \# \text{cycles} = m \cdot CPI = m \cdot (CPI_{\text{ideal}} + CPI_{\text{mem}}) = 16 \cdot 10^9 (1.2 + 0.3 \cdot 0.06 \cdot 10) = 13'8 \cdot 10^9 \text{ cicles}$$

$$k) T_{\text{exe}} = \# \text{cycles} \cdot T_c = 13'8 \cdot 10^9 \cdot 11 \cdot 10^{-9} = 151'8 s$$

Cache **emplaçament directe + victim cache** amb accés seqüencial: En aquesta segona implementació, els accessos que s'han de fer a la victim cache tenen una penalització addicional de un cicle, però el temps de cicle es el de la cache d'emplaçament directe. El temps de cicle (T_c) es de 10 ns/cicle, la taxa de fallades (m) global del conjunt MC+VC es de 0.06 fallades/accés, el temps de penalització en cas que fallem a MC però encertem a VC (T_{pvc}) es de 1 cicle i el temps de penalització en cas que fallem a totes dues (T_{pf}) es de 11 cicles.

- l) Perquè creus que el temps de penalització en cas de fallada es de 11 cicles mentre que en el cas d'emplaçament directe era de 10 cicles si sabem que la memòria principal es la mateixa?
- m) Calcular la probabilitat que un accés falli a MC però encerti a VC? (pista: es pot deduir a partir de la taxa de fallades global i la taxa de fallades que tenim quant només hi ha la cache d'emplaçament directe)
- n) Quants cicles tarda en executar-se el programa P?
- o) Quin es el temps d'execució de P? (en segons)

l) Perque s'ha d'accedir a la VC (+1 cicle)

$$m) 0.1 \cdot miss_{vc} = 0.06 \rightarrow miss_{vc} = 0.6 \rightarrow hit_{vc} = 1 - 0.6 = 0.4 \rightarrow \frac{miss_{mc} + hit_{vc}}{= 0.1 \cdot 0.4} = 0.04$$

$$n) \# \text{cycles} = m \cdot CPI = m \cdot (CPI_{\text{ideal}} + CPI_{\text{mem}}) = 16 \cdot 10^9 (1.2 + 0.3 (0.06 \cdot 11 + 0.04 \cdot 1)) = 14'1 \cdot 10^9 \text{ cicles}$$

$$o) T_{\text{exe}} = \# \text{cycles} \cdot T_c = 14'1 \cdot 10^9 \cdot 10 \cdot 10^{-9} = 141 s$$

$$12/a) \boxed{CPI_{ideal} = 5 \cdot 10^9 c / 2 \cdot 10^9 i = 2,5 \text{ c/i}}$$

$$b) 5 \cdot 10^9 / 50 \cdot 10^6 = \underline{\underline{100 \text{ ciclos entre fallo y fallo}}}$$

$$c) \boxed{CPI_B = (texec \cdot f) / n = 4 \cdot 2 \cdot 10^9 / 2 \cdot 10^9 = 4 \text{ c/i}}$$

$$d) \boxed{T_{pf} = (CPI_{tot} - CPI_{ideal}) \cdot 2 \cdot 10^9 / 50 \cdot 10^6 = 60 \text{ ciclos/fallo}}$$

$$e) \boxed{P_{fallo} \text{ en } 60 \text{ ciclos} = 1 - (99/100)^{60} = 0,453}$$

f) No, después del segundo fallo el procesador deja de ejecutar instrucciones

g) Fallo en la primera instrucción: 59 ciclos Y Fallo en la última instrucción: 0 ciclos

$$h) \boxed{M = \frac{59+0}{2} = 29,5 \text{ ciclos/fallo}}$$

$$i) \boxed{Ciclos_M = Ciclos_{ideal} + Ciclos_{fallo} = 5 \cdot 10^9 + 50 \cdot 10^6 \cdot 0,453 \cdot 29,5 = 5,67 \cdot 10^9 \text{ ciclos}}$$

$$j) \boxed{Speed Up = \frac{45}{5,67 \cdot 10^9} = 1,31 = 0,34\%}$$