

EJERCICIO 8

a) El tiempo medio por acceso quedaría:

$$\begin{aligned} T_{MA} &= \frac{1}{1,3} \times (2 + 0,01 \times 100 + 0,02 \times 40) \\ &+ \frac{0,3}{1,3} \times (3 + 0,01 \times 100 + 0,1 \times (40 + 0,4 \times 40)) = 5,134 \text{ ciclos} \end{aligned}$$

Donde vemos que el tiempo de acierto es de 2 ciclos en el caso de la cache de instrucciones (etapas F_1 F_2) y 3 ciclos en la de datos (etapas M_1 M_2 M_3). La cache de datos puede tener bloques sucios, así que hay una probabilidad del 40 % de que tengamos que enviar un bloque a escribir. Ahora calculamos el CPI sin la jerarquía de memoria, quedando:

$$CPI_{sin} = 0,3 \times 1,4 + 0,5 \times 1,2 + 0,2 \times 1,2 = 0,42 + 0,6 + 0,24 = 1,26 \text{ ciclos}$$

y el CPI con la jerarquía de memoria quedaría:

$$CPI_a = 1,26 + 0,01 \times 100 \times 1,3 + 0,02 \times 40 + 0,3 \times 0,1 \times 56 = 5,04 \text{ ciclos}$$

donde $56 = 40 + 0,4 \times 40$ y el 1.3 que multiplica a la penalización por TLB es el número de accesos a memoria por instrucción, ya que tanto instrucciones como datos acceden al mismo TLB.

b) El tiempo de ciclo ha cambiado, así que el tiempo de penalización por fallo de cache también cambia debido a que la memoria sigue tardando lo mismo en reaccionar. Este tiempo pasa a valer $\frac{40}{1,1} = 36,36 \approx 37 \text{ ciclos}$. Podemos considerar que la penalización del TLB también varía o no, según queramos, ya que también entra parte del sistema operativo en dicha penalización. Si varía, el tiempo de penalización del TLB pasaría a valer $\frac{100}{1,1} = 90,9 \approx 91 \text{ ciclos}$. De todas formas, en lo que sigue vamos a considerar que no cambia. De esta forma el CPI queda:

$$\begin{aligned} CPI_b &= 1,26 + 1,3 \times 0,01 \times 91 + \\ &+ 0,01 \times (0,6 \times 1 + 0,4 \times 37) + \\ &+ 0,3 \times 0,05 \times (37 + 0,4 \times 37) \\ &= 3,374 \text{ ciclos} \end{aligned}$$

donde hay una probabilidad de 0.6 de que encontremos el bloque en el buffer de prelectura, y hemos considerado que nos penaliza un ciclo, aunque podríamos haber considerado otra cantidad. De la misma forma se han usado las nuevas tasas de fallos. Si ahora comparamos los CPIs tenemos:

$$\frac{CPI_a}{CPI_b} = \frac{5,04}{3,374} = 1,4937$$

El tiempo de ejecución queda como sigue:

$$T_{CPU_a} = NI \times 5,04 \times Ciclo_Relej$$

$$T_{CPU_b} = NI \times 3,374 \times 1,1 \times Ciclo_Relej = NI \times 3,7114 \times Ciclo_Relej$$

Y la mejora ahora es:

$$\frac{T_{CPU_a}}{T_{CPU_b}} = \frac{NI \times 5,04 \times Ciclo_Relej}{NI \times 3,7114 \times Ciclo_Relej} = \frac{5,04}{3,7114} = 1,358$$

- c) Con el nuevo pipeline tenemos un tiempo de acierto de 1 y 2 ciclos respectivamente para la cache de instrucciones y datos. Además, sólo se accede al TLB en caso de fallo de cache, así que ahora la expresión del tiempo medio por acceso a memoria queda:

- Para el caso b)

$$\begin{aligned} T_{MA_b} &= \frac{1}{1,3} \times (2 + 0,01 \times 91 + 0,01 \times (0,6 \times 1 + 0,4 \times 37)) + \\ &+ \frac{0,3}{1,3} \times (3 + 0,01 \times 91 + 0,05 \times (37 + 0,4 \times 37)) = 3,33 \text{ ciclos} \end{aligned}$$

- Para el caso actual

$$\begin{aligned} T_{MA_c} &= \frac{1}{1,3} \times [1 + 0,01 \times (0,01 \times 91 + (0,6 \times 1 + 0,4 \times 37))] + \\ &+ \frac{0,3}{1,3} \times [2 + 0,05 \times (0,01 \times 91 + (37 + 0,4 \times 37))] = \\ &= \frac{1}{1,3} \times [1 + 0,01 \times 16,31] + \frac{0,3}{1,3} \times [2 + 0,05 \times 52,71] = 1,96 \text{ ciclos} \end{aligned}$$

Donde podemos considerar que en la escritura se accede al TLB también o no. En este caso hemos considerado que no, ya que podemos guardar en la cache la dirección física una vez que se traduzca tras el fallo de cache. Cualquiera de las dos soluciones es válida.

$$\frac{T_{MA_b}}{T_{MA_c}} = \frac{3,33}{1,96} = 1,6989 \approx 1,7$$

Para calcular ahora el CPI hay que recalcular el CPI sin la jerarquía de memoria:

$$CPI_{sin} = 0,3 \times 1,4 + 0,5 \times 1,2 + 0,2 \times 1,1 = 0,42 + 0,6 + 0,22 = 1,24 \text{ ciclos}$$

Con lo que el CPI con la configuración que nos dicen en este apartado sería:

$$\begin{aligned} CPI_c &= 1,24 + 0,01 \times [0,01 \times 91 + (0,6 \times 1 + 0,4 \times 37)] + \\ &+ 0,3 \times 0,05 \times [0,01 \times 91 + (37 + 0,4 \times 37)] = \\ &= 1,24 + 0,1631 + 0,79065 = 2,19375 \text{ ciclos} \end{aligned}$$

y por tanto el tiempo de ejecución es

$$T_{CPU_c} = NI \times 2,19375 \times 1,1 \times Ciclo_Relej = NI \times 2,413 \times Ciclo_Relej$$

y la ganancia es

$$\frac{T_{CPU_b}}{T_{CPU_c}} = \frac{NI \times 3,7114 \times Ciclo_Relej}{NI \times 2,413 \times Ciclo_Relej} = \frac{3,7114}{2,413} = 1,538$$

EJERCICIO 9

a) El CPI real para la configuración de cachés dada es:

$$CPI_{real} = 1.2 + \underbrace{0.02 \times 50}_{\text{deten. i-cache}} + \underbrace{0.20 \times 0.05 \times 50}_{\substack{\text{deten. loads} \\ \text{d-cache}}} + \underbrace{0.10 \times 50}_{\substack{\text{deten. stores} \\ \text{d-cache}}} = 1.2 + 1 + 0.5 + 5 = \underline{\underline{7.7 \text{ ciclos}}}$$

b) Con la caché de víctimas y teniendo en cuenta los bloques sucios de la caché de datos, el nuevo CPI real será:

$$CPI_{real} = 1.2 + \underbrace{0.02 \times 50}_{\text{i-cache}} + \underbrace{0.30 \times 0.05 \times [\underbrace{0.30 \times 2}_{\text{acierto C.V.}} + \underbrace{0.70 \times 50}_{\text{fallo C.V.}} + \underbrace{0.15 \times (0.10 \times 50 + 0.90 \times 2)}_{\text{sucios NO caben C.V.}}]}_{\text{d-cache}} = 1.2 + 1 + 0.549 = \underline{\underline{2.75 \text{ ciclos}}}$$

$CPI_a / CPI_b = 7.7 / 2.75 = 2.8 \Rightarrow$ la opción **b)** es **2.8** veces mejor que la opción **a)**

c) Si cambiamos la caché de víctimas por un *buffer* de pre-lectura en la caché de datos, el CPI será:

$$CPI_{real} = 1.2 + \underbrace{0.02 \times 50}_{\text{i-cache}} + \underbrace{0.30 \times 0.05 \times [\underbrace{0.30 \times 1}_{\text{acierto buffer}} + \underbrace{0.70 \times 50}_{\text{fallo buffer}} + \underbrace{0.15 \times 50}_{\text{bloques sucios}}]}_{\text{d-cache}} = 1.2 + 1 + 0.64 = \underline{\underline{2.84 \text{ ciclos}}}$$

$CPI_c / CPI_b = 2.84 / 2.75 = 1.032 \Rightarrow$ la opción **b)** es un **3.2%** mejor que la opción **c)**

d) Al TLB de datos se accede SIEMPRE que ejecutemos una instrucción de carga/almacenamiento puesto que a la d-cache se accede mediante dirección física y hay que realizar la traducción como paso previo. Sin embargo, al TLB de instrucciones sólo se accede cuando hay fallo de la i-cache ya que a ésta última se accede mediante dirección virtual.

$$CPI_{real} = 1.2 + \underbrace{0.02 \times 50}_{\text{i-cache}} + \underbrace{0.01 \times 100 \times 0.02}_{\text{deten. i-TLB}} + \underbrace{0.20 \times 0.05 \times 50}_{\substack{\text{deten. loads} \\ \text{d-cache}}} + \underbrace{0.10 \times 50}_{\substack{\text{deten. stores} \\ \text{d-cache}}} + \underbrace{0.30 \times 0.01 \times 100}_{\text{deten. d-TLB}} = \\ = 1.2 + 1 + 0.02 + 0.5 + 5 + 0.30 = \underline{\underline{8.02 \text{ ciclos}}}$$

$CPI_d / CPI_a = 8.02 / 7.7 = 1.041 \Rightarrow$ la opción **a)** es un **4.1%** mejor que la opción **d)**