

Cachés

Ejercicio 5.6 (COD 5)

Ejemplo

- Ejercicio 5.6 de COD 5 (p. 487)
- Suponer lo siguiente:
- El acceso a la memoria principal es de 70 ns.
- Las instrucciones de memoria son el 36%.
- Hay dos CPUs, P1 y P2, que tienen un nivel de caché, L1, con los siguientes datos:

	L1 Size	L1 Miss Rate	L1 Hit Time
P1	2 KiB	8.0%	0.66 ns
P2	4 KiB	6.0%	0.90 ns

Pregunta 1

- Pregunta 1: Suponiendo que el tiempo de éxito (hit time) determina el tamaño del ciclo de P1 y P2, ¿cuáles son sus respectivas velocidades de reloj?
- $V1 = 1 / (0.66 \text{ ns / ciclo}) = 1.52 \text{ GHz}$.
- $V2 = 1 / (0.9 \text{ ns / ciclo}) = 1.1 \text{ GHz}$
- **Nota:** cuidado con las unidades.

Pregunta 2

- Pregunta 2: ¿Cuál es el AMAT de P1 y P2?
AMAT = tiempo de éxito + tasa de fallas x castigo por falla
- Se necesita calcular el castigo por falla.
Castigo por falla = Tiempo de acceso a memoria / Periodo
- Para P1:
- Castigo por falla = $(70 \text{ ns}) / (0.66 \text{ ns} / \text{ciclo}) = 106$ ciclos
- AMAT = $1 \text{ ciclo} + 0.08 \times 106 \text{ ciclos} = 9.48$ ciclos
- AMAT = $9.48 \text{ ciclos} \times 0.66 \text{ ns} / \text{ciclo} = 6.26 \text{ ns}$

Pregunta 2

- Para P2:
- Castigo por falla = $(70 \text{ ns}) / (0.9 \text{ ns} / \text{ciclo}) = 78$ ciclos
- AMAT = $1 \text{ ciclo} + 0.06 \times 78 \text{ ciclos} = 5.68 \text{ ciclos}$
- AMAT = $5.68 \text{ ciclos} \times 0.9 \text{ ns} / \text{ciclo} = 5.11 \text{ ns}$

Pregunta 3

- Pregunta 3: Suponiendo un CPI de 1.0 sin detenciones de memoria, ¿cuál es el CPI total de P1 y P2? ¿Qué CPU es más rápida y en cuánto?

CPI total = CPI base + ciclos de detención por memoria

Ciclos de detención por memoria = accesos a la memoria x tasa de fallas x castigo por falla

Pregunta 3

- Para P1 los ciclos de detención por memoria son:
- $1.36 \times I \times 0.08 \times 10^6 \text{ ciclos} = 11.53 \times I \text{ ciclos}$
- $\text{CPI total} = 1 + 11.53 = 12.53$

- Para P2 los ciclos de detención por memoria son:
- $1.36 \times I \times 0.06 \times 78 \text{ ciclos} = 6.36 \times I \text{ ciclos}$
- $\text{CPI total} = 1 + 6.36 = 7.36$

Pregunta 3

- Se calcula el tiempo de CPU usando:
- $T = I \times \text{CPI} \times P$
- Tiempo de CPU de P1:
- $T = I \times 12.53 \times 0.66 = 8.27 \times I$
- Tiempo de CPU de P2:
- $T = I \times 7.36 \times 0.9 = 6.62 \times I$
- $\text{Speedup} = (8.27 \times I) / (6.62 \times I) = 1.25$
- Conclusión: P2 es más rápida en 25%

Pregunta 4

- Pregunta 4: ¿Cuál es el AMAT de P1 si se le agrega un caché secundario, L2? ¿El AMAT es mejor o peor que el de P2?

L2 Size	L2 Miss Rate	L2 Hit Time
1 MiB	95%	5.62 ns

- La tasa de fallas de L2 que se indica es la *tasa de fallas local*.

Pregunta 4

- Se usa la fórmula:

AMAT = tiempo de éxito L1 + tasa de fallas L1 x (tiempo de éxito L2 + tasa de fallas L2 x castigo por falla)

- $AMAT = 0.66 \text{ ns} + 0.08 \times (5.62 \text{ ns} + 0.95 \times 70 \text{ ns})$
- $AMAT = 6.43 \text{ ns}$
- $AMAT = (6.43 \text{ ns}) / (0.66 \text{ ns} / \text{ciclo}) = 9.74 \text{ ciclos}$

Pregunta 4

- Conclusión: el nuevo AMAT de P1 (6.43 ns) es peor que el AMAT anterior de P1 (6.26 ns) y que el de P2 (5.11 ns).

Pregunta 5

- Pregunta 5: Suponiendo un CPI de 1.0 sin detenciones de memoria, ¿cuál es el CPI de P1 con la adición del caché L2?

CPI total = CPI base + detenciones primarias por instrucción +
detenciones secundarias por instrucción

- Castigo por falla a L2 = $5.62 / 0.66 = 8.5$ ciclos
- L2 tiene una tasa de fallas local de 0.95.
- La tasa de fallas global es $0.08 \times 0.95 = 0.076$.

Pregunta 5

- $CPI = 1 + 1.36 \times (0.08 \times 8.5 + 0.076 \times 106)$
- $CPI = 12.88$

Pregunta 6

- Pregunta 6: ¿Qué procesador es más rápido ahora que P1 tiene 2 niveles de caché? Si P1 es más rápido, ¿qué tasa de fallas debe tener P2 en su caché L1 para igualar el rendimiento de P1? Si P2 es más rápido, ¿qué tasa de fallas debe tener P1 en su caché L1 para igualar el rendimiento de P2?

Pregunta 6

- Tiempo de CPU de P1:
- $T = I \times 12.88 \times 0.66 = 8.50 \times I$
- Tiempo de CPU de P2 (el mismo de antes):
- $T = I \times 7.36 \times 0.9 = 6.62 \times I$
- Speedup = $(8.50 \times I) / (6.62 \times I) = 1.28$
- Conclusión: ahora P2 es más rápido 28% (antes era 25%).

Pregunta 6

- Para que P1 iguale el rendimiento de P2:
- $T P1 = I \times 12.88 \times 0.66 = 8.50 \times I$
- $T P2 = I \times 7.36 \times 0.9 = 6.62 \times I$
- El CPI de P1 (12.88) se obtuvo de:
- $CPI P1 = 1 + 1.36 \times (0.08 \times 8.5 + 0.076 \times 106) = 12.88$
- Se sustituye la expresión del CPI en la ecuación del tiempo para P1.

Pregunta 6

- $T P1 = I \times (1 + 1.36 \times (0.08 \times 8.5 + 0.076 \times 106)) \times 0.66 = 8.50 \times I$
- 0.08 es la talla de fallas del nivel L1.
- Si la tasa de fallas fuera 0:
- $T P1 = I \times (1 + 1.36 \times (0 \times 8.5 + 0.076 \times 106)) \times 0.66 = 7.89 \times I$
- No hay forma que el tiempo de P1 sea igual al de P2.