Trabajo Práctico Número 7

Arquitectura de Computadoras 27/05/2026

Instrucciones

- Los ejercicios propuestos en el presente trabajo práctico pueden ser resueltos en forma individual o grupal.
- El planteo de la solución debe realizarse basándose en lo aprendido en las clases teóricas.
- Puede utilizar las clases de consulta para consultar ejercicios de enunciado y verificar la validez de la solución obtenida.
- Posteriormente, se tomará una evaluación con ejercicios similares a los de este práctico, la cual deberá ser resuelta en forma individual. La fecha de evaluación se encuentra en el sitio web de la materia.

Conceptos involucrados

- Memorias caché como parte de la jerarquía de memorias.
- Políticas de escritura: write-back y write-through.
- Métricas de performance de cachés: tasa de fallos, tiempo medio de acceso a memoria.
- Impacto del uso de cachés en el CPI de un programa, y en el tiempo de ejecución del mismo.
- Coherencia de caché. Protocolo MSI. El problema de la falsa coherencia.

Problemas propuestos

1) Responda las siguientes preguntas teóricas, justificando la respuesta en todos los casos.
   a) La tasa de accesos a memoria por instrucción, ¿puede ser menor que 1? ¿Y puede ser mayor que 2?
   b) Si en un futuro la memoria principal de un sistema es fabricada con tecnología SRAM, teniendo la misma velocidad que un caché, pero con un costo muy bajo, ¿usted dejaría de usar memorias caché?
   c) Los cachés que trabajan con acierto bajo fallo tienen por objeto reducir la penalidad de fallos.
   d) En un sistema con varios niveles de caché, ¿en los cachés de primer nivel es más importante el tiempo de acierto que la tasa de fallos?
   e) En un sistema con varios niveles de caché, ¿en los cachés de segundo nivel es más importante el tiempo de acierto que la tasa de fallos?

2) *[Recomendado]* Dada la mezcla de instrucciones para gcc vista en prácticos anteriores, se intentará calcular la influencia del uso de una memoria caché en el CPI del programa.

Tipo de Instrucción	% gcc	$\overline{CPI}$
Aritméticas	48 %	1,0
Transferencias de datos	33 %	1,4
Saltos condicionales	17 %	1,7
Saltos incondicionales	2 %	1,2

   a) Calcule el CPI promedio de gcc corriendo en una máquina con un caché, donde la tasa de fallos es del 2 % para las instrucciones y del 4 % para datos, y la penalización por fallos es de 20 ciclos en ambos casos.
   b) Calcule cuánto más rápida es la máquina sin fallos que se venía viendo en los prácticos con respecto a la máquina del apartado anterior.
   c) Suponga que ahora Ud. realiza mejoras a la arquitectura interna del procesador de manera tal que reduce los CPI de cada instrucción a la mitad. Repita los apartados anteriores y saque conclusiones.

3) Dados dos sistemas de caché, el primero con dos caché de 16 KiB cada uno, uno para datos y el otro para instrucciones, y el segundo con un caché de 32 KiB unificado (para datos e instrucciones). Suponga que en todos los caché, el tiempo de acceso cuando hay un acierto

es de 1 T, y que un fallo cuesta 50 T. Suponga además que en el caché unificado, las instrucciones de transferencia de datos requieren un ciclo extra. Además, suponga que el 75 % de los accesos a memoria es para buscar instrucciones. Finalmente, la tasa de fallos del caché de instrucciones de 16 KiB es de 0.64 %, la del caché de datos de 16 KiB es de 6.47 %, y la del caché unificado de 32 KiB es de 1.99 %.

a) Explique por qué se realiza la suposición de que las instrucciones de transferencia de datos requieren un ciclo extra en el caché unificado.
b) Si tuviese que elegir uno de los dos sistemas basándose en la tasa de fallos promedio, ¿cuál elegiría?
c) Si tuviese que elegir uno de los dos sistemas basándose en el tiempo medio de acceso, entonces ¿cuál elegiría?
d) ¿Cuál de las dos alternativas anteriores considera que es una mejor métrica?

4) *[Recomendado]* Considere un caché de mapeo directo, con 16 palabras por bloque, las cuales se cargan de una memoria SDRAM que requiere un ciclo para enviar la dirección, un ciclo para enviar el tamaño del bloque, cinco ciclos de latencia y un ciclo por cada palabra a transferir. Calcule el tiempo de penalización por fallos para las siguientes situaciones, tomando siempre como referencia los porcentajes de gcc:
   a) Fallo de Lectura en un caché de instrucciones.
   b) Fallo de Lectura en un caché de instrucciones con Early Restart (considerando que la mitad de los saltos condicionales son tomados).
   c) Fallo de Lectura en un caché de datos con política de escritura write-back, suponiendo que el 20 % de las veces, el bloque a reemplazar ha sido modificado.
   d) Fallo de escritura en un caché de datos con política write-through sin buffer.
   e) Fallo de escritura en un caché de datos con política write-through con buffer que nunca desborda.

5) *[Recomendado]* Se conoce de un sistema de computación, donde cada instrucción demora en promedio 2 ciclos y hay en promedio 1,5 referencias a memoria por instrucción. Un programa de 100.000 instrucciones se ha ejecutado, y con un caché unificado de 32 KiB de tamaño ha resultado una tasa de aciertos de 95 %, con un tiempo de hit de 2 ns y una penalización en caso de fallo de 18 ns. Se plantea una mejora con un nuevo caché de 64 KiB de memoria, que resultaría en una tasa de aciertos del 97 %, con una penalización en caso de falla igual a la anterior pero un tiempo de hit de 3 ns.
   a) Explique por qué el caché más grande tiene menor tasa de fallos.
   b) Explique por qué el caché más chico tiene menor tiempo de acceso.
   c) Calcule el tiempo de ciclo del procesador con cada uno de los cachés.
   d) Calcule el AMAT para ambos casos. ¿Cuál caché es mejor?
   e) Calcule el CPI para ambos casos. ¿Cuál caché es mejor?
   f) Calcule el tiempo de ejecución del programa para ambos casos. ¿Cuál caché es mejor? ¿Por cuánto?

6) *[Recomendado]* Suponga un procesador con las siguientes características: frecuencia de reloj de 2 GHz, CPI ideal de 2 ciclos, 1,3 referencias a memoria por instrucción, tiempo medio de acceso a memoria principal de 125 ns y una tasa de fallos del caché L1 de 5 %.
   a) Calcule el CPI de este procesador.
   b) Se decide agregar al procesador un caché L2, con un tiempo de acceso de 10 ciclos y una tasa de fallos del 50 %. Calcule nuevamente el CPI.
   c) Ahora se decide agregar un caché L3, con un tiempo de acceso de 30 ciclos y una tasa de fallos del 30 %. Calcule nuevamente el CPI.

7) *[Recomendado]* Una jerarquía de memoria de un procesador está compuesta por dos cachés L1 separados (para instrucciones y datos) y por un caché L2 unificado. El caché L1 de instrucciones es de 16 KiB, posee una tasa de fallos del 3 % y un tiempo de acierto de 1 ns. El caché L1 de datos es idéntico, pero con una tasa de fallos del 8 %. El caché L2 es de 64 KiB, posee un tiempo medio de acceso de memoria de 12 ns. La frecuencia de trabajo del procesador es de 2 GHz.
   Luego de aplicar sobre el caché L2 varias técnicas vistas en la asignatura, se pudo reducir a la mitad su tiempo medio de acceso a memoria, sin aumentar su tamaño y sin que sea necesario agregar un caché L3.
   a) Sabiendo que el 75 % de los accesos a memoria se corresponden con accesos a instrucciones y el resto son accesos a datos, calcule el tiempo medio de acceso a memoria del

sistema original.

b) Mencione dos técnicas que pueden haber sido utilizadas para reducir el AMAT del caché L2.

c) Calcule el tiempo medio de acceso a memoria del sistema con el caché L2 mejorado.

d) A un integrante del equipo se le ocurre que, dado el caché L2 mejorado, sería posible reducir el tamaño de los cachés L1. Por ejemplo, para cachés de 4 KiB se obtiene una tasa de fallos del 8 % para instrucciones y del 19 % para datos, mientras que el tiempo de acierto se reduce a la mitad. Calcule el tiempo medio de acceso a memoria para el sistema con esta modificación propuesta.

e) ¿Tiene sentido la modificación propuesta en el apartado anterior? Justifique su respuesta describiendo dos ventajas y/o desventajas de la propuesta, si las hubiere.

8) Considere un microprocesador cuyo sistema de memoria está compuesto por dos cachés L1, uno para instrucciones y uno para datos, conectados directamente a la memoria principal. Ambos cachés L1 son asociativos por conjuntos de 2 vías, de 8 KiB de tamaño total con bloques de 64 Bytes, con un tiempo de acierto de 1 ciclo y una tasa de fallos del 1,3 % para instrucciones y de 4,1 % para datos. El caché de datos tiene una política de escritura write through con un buffer que nunca se llena. El tiempo de acierto de la memoria principal es de 80 ciclos.
Sobre este procesador se usa gcc como programa para evaluar la performance, el cual posee la siguiente mezcla de instrucciones:

Tipo de Instrucción	% gcc	$\overline{CPI}$
Aritméticas	48 %	1,0
Transferencias de datos	33 %	1,4
Saltos condicionales	17 %	1,7
Saltos incondicionales	2 %	1,2

Se pide que:

a) Calcule el CPI real del microprocesador al considerar el sistema de memoria.

b) Se desea duplicar la performance del programa de prueba en este microprocesador, y para ello se decide incorporar un caché L2, el cual tiene un tiempo de acierto de 8 ciclos, bloques del mismo tamaño que los L1 y la misma política de escritura. ¿Cuál debería ser la tasa de fallos de este caché para cumplir con el objetivo de performance?

9) Se tienen cuatro procesadores A, B, C y D con direccionamiento a palabras de 32 bits, que ejecutan tareas simultáneamente en un esquema de memoria compartida. Todos los procesadores tienen cachés del tipo write-back con bloques de cuatro palabras de 32 bits y se encuentran inicialmente vacíos. El protocolo usado para mantener la coherencia de los cachés es el visto en clase. Describa en cada paso cómo se va modificando el estado de los bloques en los cachés de los procesadores para la siguiente secuencia de actividades:

Paso	**Procesador**	**Actividad**	**Dirección**
1	A	Lectura	102
2	B	Lectura	101
3	A	Escritura	105
4	C	Lectura	103
5	B	Lectura	107
6	A	Escritura	100
7	B	Escritura	106
8	C	Lectura	104

10) Un programa realiza un procesamiento sobre un cierto conjunto de datos. Cada conjunto de datos está compuesto por 8 palabras, y se tienen 2048 conjuntos distintos, independientes y ubicados contiguamente en memoria. El procesamiento consiste en sumar las primeras 7 palabras de cada conjunto, y guardar el resultado en la octava palabra. Este programa se ejecutará en un procesador que posee un CPI ideal igual a 1, y un caché de datos con bloques de 32 palabras, en el que cada fallo tiene una penalización de 16 ciclos. En este procesador, y sin considerar los fallos, el procesamiento de cada conjunto de datos demora 41 ciclos.

a) Calcule cuánto demora el procesamiento de todos los conjuntos de datos considerando los

fallos del caché de datos.

b) Suponga que el procesador permite multithreading, y que se reescribe el programa para que utilice dos hilos, de manera que cada hilo se encargará del procesamiento de un conjunto de datos alternadamente (el primer hilo procesa el primer conjunto, el segundo hilo procesa el segundo conjunto, el primer hilo procesa el tercer conjunto, etc.). Calcule cuánto demora el procesamiento de todos los conjuntos de datos en este caso, considerando los fallos del caché de datos.

c) Suponga ahora que el procesador tiene dos núcleos, cada uno con su caché de datos, de tipo snooping. Cada núcleo se encarga de procesar un conjunto de datos, de manera alternada como en el apartado anterior. Calcule cuánto demora el procesamiento de todos los conjuntos de datos en este caso, recordando que el procesamiento del conjunto incluye una escritura, y suponiendo que el primer núcleo arranca un ciclo antes que el segundo.

d) Proponga dos alternativas para mejorar el tiempo de procesamiento calculado en el apartado anterior.

11) En las tablas se muestra el estado en que se encuentran dos cachés de mapeo directo asociados a dos núcleos de un mismo procesador, los que comparten una memoria única centralizada. La coherencia de los cachés se mantiene mediante el protocol MSI de “snooping”.

- **Núcleo N0:**

Bloque	Etiqueta	Palabra 1	Palabra 2	Palabra 3	Palabra 4	Estado
0	1000	10	20	30	40	M
1	4000	500	600	700	800	S
...	...	...	...	...	...	...
N	3000	2	4	6	8	S

- **Núcleo N1:**

Bloque	Etiqueta	Palabra 1	Palabra 2	Palabra 3	Palabra 4	Estado
0	1000	10	10	10	10	I
1	8000	500	600	700	800	S
...	...	...	...	...	...	...
N	3000	2	4	6	8	S

Se pide que responda las siguientes preguntas, justificando brevemente su respuesta:

a) Si N0 escribe en bloque 0, ¿cómo queda el estado de este bloque?

b) Si N1 escribe en bloque 1, ¿se debe invalidar el bloque 1 en P0?

c) Si N1 lee un nuevo bloque de memoria, J, ¿en qué estado debe quedar en su caché?

12) Considere dos microprocesadores, P1 y P2, ambos con un CPI base igual a 1.5 (sin considerar fallos de memoria) a los cuales se conecta una memoria principal cuyo tiempo medio de acceso es de 100 ns. El procesador P1 posee un caché L1 de 2 KiB con un tiempo de acierto de 0.66 ns y una tasa de fallos del 8 %, mientras que el procesador P2 posee un caché L1 de 4 KiB con un tiempo de acierto de 0.9 ns y una tasa de fallos del 6 %.

a) Asumiendo que el tiempo de ciclo de los procesadores está determinado por el tiempo de acierto de sus cachés L1, indique cuáles serán sus respectivas frecuencias de reloj.

b) Calcule el tiempo medio de acceso a memoria (en ciclos) para ambos procesadores.

c) Considerando que el 36 % de las instrucciones accede a memoria de datos, calcule el CPI de ambos procesadores. ¿Cuál de los dos procesadores es más rápido? ¿Cuánto más rápido?

d) Para mejorar la performance del procesador más lento, se desea agregarle un caché de segundo nivel L2, y se proponen dos opciones: un caché de mapeo directo con un tiempo de acierto de 12 ciclos y una tasa de fallos de 3.5 % y otro caché asociativo por conjuntos de 8 vías con un tiempo de acierto de 28 ciclos y una tasa de fallos de 1.5 %. Calcule el CPI para ambos casos, e indique con cuál de las opciones se quedaría, y cuánto mejora la performance gracias a esta adición.

e) Suponga que ahora el tiempo medio de acceso de la memoria principal se duplica, y es de 200 ns, ¿sigue siendo conveniente el caché elegido en el apartado anterior? Justifique cuantitativamente su respuesta.

f) Como ahora la performance empeoró a causa de la memoria principal, se propone agregar un tercer nivel de caché, con un tiempo de acierto de 50 ciclos y una tasa de fallos del 13 %.

Calcule cuánto será el nuevo CPI con este caché, y cuánto será la mejora de performance que se obtiene con respecto al apartado anterior.

Ejercicio a Desarrollar en Clase Práctica

Problema 1 del segundo parcial de 2022

13) Se tienen dos procesadores A y B que ejecutan tareas simultáneamente en un esquema de memoria compartida. Ambos procesadores tienen cachés del tipo write-back con bloques de una palabra de 32 bits. El protocolo usado para mantener la coherencia de los cachés es el visto en clase, y supondremos por comodidad que los datos ya están en el caché (pero que no fueron modificados). Para la siguiente secuencia de actividades:

Paso	Procesador	Actividad	Dirección de Memoria
1	A	Escritura	100
2	B	Escritura	104
3	A	Lectura	100
4	B	Lectura	104
5	A	Lectura	104
6	B	Lectura	100

Se pide que:

a) Describa qué acciones tienen lugar en cada paso, para ambos procesadores.
b) Calcule la cantidad de transacciones por el bus que se realizan.
c) Calcule la cantidad de transacciones por el bus que se realizan, pero ahora suponiendo que los bloques de caché son de 4 palabras.

Soluciones propuestas

1) ...

2)
   a) CPI = 1,919
   b) A = 1,53
   c) CPI = 1,291; A = 2,06

3)
   a) ...
   b) La tasa de fallos del sistema con cachés separados es 2.10 %, y la del sistema con caché unificado es 1.99 %, por lo que se elegiría esta última.
   c) El AMAT de los cachés separados es de 2,05 T, mientras que el del caché unificado es de 2,25 T, por lo que elegiría los cachés separados.
   d) ...

4)
   a) Penalización = 23 ciclos
   b) Penalización = 8,7875 ciclos
   c) Penalización = 27,6 ciclos
   d) Penalización = 8 ciclos
   e) Penalización = 1 ciclo

5)
   a) ...
   b) ...
   c) Caché de 32 KiB = 2 ns; Caché de 64 KiB = 3 ns.
   d) Caché 32 KiB = 2,9 ns; Caché 64 KiB = 3,54 ns. Es mejor el Caché de 32 KiB.
   e) Caché 32 KiB = 2,675; Caché 64 KiB = 2,27. Es mejor el caché de 64 KiB.
   f) Caché 32 KiB = 535 us ; Caché 64 KiB = 681 us. El caché de 32 KiB es 27

6)
   a) CPI = 18,25
   b) CPI = 10,77
   c) CPI = 6,06

7)
   a) AMAT = 3.02 T
   b) ...
   c) AMAT = 2.51 T
   d) AMAT = 2.29 T
   e) ...

8) a) $CPI = 3.375$
   b) $m_{L2} = 10.4\,\%$

9)
   Paso 1: El procesador A tiene un fallo de lectura en 102 y deja al bloque (100 a 103) en S. Los demás permanecen idénticos.
   Paso 2: El procesador B tiene un fallo de lectura en 101 y deja al bloque (100 a 103) en el estado S. Los demás permanecen idénticos.
   Paso 3: El procesador A tiene un fallo de escritura en 105 y deja al bloque (104 a 107) en el estado M.
   Paso 4: El procesador C tiene un fallo de lectura en 103 y deja al bloque (100 a 103) en el estado S. Los demás permanecen idénticos.
   Paso 5: El procesador B tiene un fallo de lectura en 107. Esto produce que el procesador A realice un write-back a memoria, y el procesador B deja al bloque (104 a 107) en el estado S para ambos procesadores. Los demás permanecen idénticos.
   Paso 6: El procesador A realiza una escritura en 100 y deja al bloque (100 a 103) en el estado M. Esto invalida el bloque en el procesador C.
   Paso 7: El procesador B realiza una escritura en 106 y deja al bloque (104 a 107) en el estado

M. Esto deja en estado I el bloque para el resto de los procesadores.

Paso 8: El procesador C tiene un fallo de lectura en 104. Esto produce que el procesador B realice un write-back a memoria, para que finalmente dejen al bloque (104 a 107) en el estado S.

10)

a) 92.160 ciclos.
b) 83.984 ciclos.
c) 66.560 ciclos.
d) Con una alternativa se puede reducir a 50.176 ciclos, y con una mejor alternativa a 42.000 ciclos.

11)

a) El bloque 0 en N0 queda en estado M.
b) No se debe invalidar el bloque 1 de N0.
c) Si N1 lee un nuevo bloque debe quedar en estado S.

12)

a) f(P1) = 1,51 GHz y f(P2) = 1,11 GHz.
b) AMAT(P1) = 13,16 y AMAT(P2) = 7,72.
c) CPI(P1) = 18,04 y CPI(P2) = 10,64. P2 es un 24.3 % más rápido que P1.
d) CPI(P1MD) = 3,38 y CPI(P1AC) = 4,79. Elegimos el L2 de mapeo directo, que produce una mejora igual a 5.33.
e) Sigue siendo conveniente el L2 de mapeo directo.
f) CPI = 3.15 y produce una mejora del 25,4 %.