



Profesores: Néstor González – Leo Medina
Fecha: 1 de agosto de 2018

Problema 1 (PEP 2, Verano 2018)

Se tiene un procesador con una memoria caché en el primer nivel L1 de 64KB, un segundo nivel L2 de 512KB y un tercer nivel L3 de 4MB. El procesador está en una placa que soporta 4GB de Memoria RAM. El diseñador de la CPU definió que los Bloques almacenarán hasta 8 Palabras. Las palabras son de 256 bits. Todas las caché son asociativas de 8 vías. Determine la estructura de la dirección para cada una de las tres caché indicadas.

Análisis de la estructura de la memoria RAM:

La memoria RAM tiene una capacidad 4GB (Direccionable por Byte) => se requieren 32 bits para las direcciones.

Cada Bloque es de $8 (2^3)$ palabras y cada palabra es de 256 bits = 32 (2^5) Bytes. Luego, cada Bloque es de $2^3 \times 2^5 = 2^8$ Bytes = 256 Bytes => se requieren 8 bits de “Byte Offset” en un Bloque (o bien 3 para Palabra y 5 para Byte)

Existen $2^{32} / 2^8$ Bloques = 2^{24} Bloques = 16 MBloques en la memoria RAM => se usan 24 bits para determinar un Bloque en la memoria RAM.

Diseño de cada nivel de Caché:

Todos los niveles han sido definidos de tipo asociativo de 8 vías. Esto significa que se deben usar 3 bits para determinar las Vías y que cada Conjunto alberga 8 Bloques (cada vía es un Bloque).

Nivel L1: Capacidad de la memoria Caché = 64 KBytes = $2^6 \times 2^{10}$ Bytes = 2^{16} Bytes. Cada Bloque es de 256 (2^8) Bytes, luego esta Caché puede contener $2^{16} / 2^8$ Bloques = $2^8 = 256$ Bloques (Líneas de Caché) agrupados de a 8 (vías) por Conjunto, entonces habrá $2^8 / 2^3 = 2^5 = 32$ Conjuntos. Luego, se requiere de 5 bits para los Conjuntos (Index). El resto de los bits constituye la etiqueta (tag), o sea, $32 - 5 - 8 = 19$

Tag	Index	Palabra	Byte Offset
19	5	3	5

Nivel L2: Capacidad de la memoria Caché = 512 KBytes = $2^9 \times 2^{10}$ Bytes = 2^{19} Bytes. Cada Bloque es de 256 (2^8) Bytes, luego esta Caché puede contener $2^{19} / 2^8$ Bloques = $2^{11} = 2$ KBloques (Líneas de Caché) agrupados de a 8 (vías) por Conjunto, entonces habrá $2^{11} / 2^3 = 2^8 = 256$ Conjuntos. Luego, se requiere de 8 bits para los Conjuntos (Index). El resto de los bits constituye la etiqueta (tag), o sea, $32 - 8 - 8 = 16$

Tag	Index	Palabra	Byte Offset
16	8	3	5

Nivel L3: Capacidad de la memoria Caché = 4 MBytes = $2^2 \times 2^{20}$ Bytes = 2^{22} Bytes. Cada Bloque es de 256 (2^8) Bytes, luego esta Caché puede contener $2^{22} / 2^8$ Bloques = $2^{14} = 2^4 \times 2^{10} = 16$ MBloques (Líneas de Caché) agrupados de a 8 (vías) por Conjunto, entonces habrá $2^{14} / 2^3 = 2^{11} = 2$ KConjuntos. Luego, se requiere de 11 bits para los Conjuntos (Index). El resto de los bits constituye la etiqueta (tag), o sea, $32 - 11 - 8 = 13$

Tag	Index	Palabra	Byte Offset
13	11	3	5



Problema 2 (PEP 2, Verano 2018)

Para los mismos datos de la pregunta 1, considere que el procesador tiene $CPI = 1,2$ cuando se mide sin considerar los accesos a la memoria. Considere también que se requieren 1,3 accesos a memoria por instrucción. La tasa de hit es de 90%, 60% y 40% para L1, L2 y L3 respectivamente. La espera (stall) es de 0, 3 y 6 ciclos para un hit en L1, L2 y L3 respectivamente. La espera es de 50 ciclos para la memoria RAM.

A) Calcule el nuevo CPI con sólo L1, luego con L1 y L2, luego con los tres niveles de caché.

Sin memoria caché: $CPI = 1,2 + 1,3 \times 50 = 66,2$

-Con solo L1: Si hay un Hit, que tiene tasa de 90%, entonces no hay demora porque hay 0 esperas en Hit en L1, pero si no hay Hit, o sea, si hay Miss, que tiene tasa de 10%, se debe leer desde la RAM con 50 ciclos de espera (esto es, se lee desde la RAM con 10% de probabilidad), luego, una instrucción aporta, por lectura $1,3 \times 0,1 \times 50 = 6,5$ ciclos \Rightarrow **CPI con solo L1 = $1,2 + 6,5 = 7,7$**

-Con L1 y L2: Ahora, un Miss de L1 lleva a leer desde L2 con tasa de Hit 60% y 3 ciclos de espera si se lee desde L2. Si no se lee desde L2, con tasa de Miss de 40%, se debe leer desde la RAM (es decir, se lee desde la RAM con un 10% por Miss en L1 y 40% por Miss en L2). Si hay Hit en L1 no hay espera, si hay Hit en L2 hay 3 ciclos de espera. Luego, una instrucción aporta, por lectura $1,3 \times (0,1 \times 0,6 \times 3 + 0,1 \times 0,4 \times 50) = 1,3 \times (0,18 + 2) = 2,83$ entonces, **CPI con L1 y L2 = $1,2 + 2,83 = 4,034$**

-Con L1, L2 y L3: Ahora, un Miss de L1 lleva a leer desde L2 con tasa de Hit 60% y 3 ciclos de espera si se lee desde L2. Si no se lee desde L2, con tasa de Miss de 40%, se debe leer desde L3 con tasa de Hit de 40% y 6 ciclos de espera. Si no se lee de L3, con tasa de Miss de 60%, se debe leer desde la RAM (es decir, se lee desde la RAM con un 10% por Miss en L1, 40% por Miss en L2 y 60% de Miss en L3). Si hay Hit en L1 no hay espera, si hay Hit en L2 hay 3 ciclos de espera, si hay un Hit en L3 hay 6 ciclos de espera. Luego, una instrucción aporta, por lectura: $1,3 \times (0,1 \times 0,6 \times 3 + 0,1 \times 0,4 \times 0,4 \times 6 + 0,1 \times 0,4 \times 0,6 \times 50) = 1,3 \times (0,18 + 0,096 + 1,2) = 1,92$. Entonces, **CPI con los tres niveles de Caché = $1,2 + 1,92 = 3,12$**

B) Calcule la aceleración lograda al ir agregando niveles de Caché.

Al agregar L1: aceleración = $66,2 / 7,7 = 8,6$ veces

Al agregar L2: aceleración = $7,7 / 4,034 = 1,9$ veces

Al agregar L3: $= 4,034 / 3,12 = 1,3$ veces

aceleración total con los tres niveles respecto de no tener Caché = $66,2 / 3,12 = 21,2$ veces.

Se aprecia que se logra una aceleración cada vez que se agrega un nivel de Caché, pero la aceleración lograda cada vez, va disminuyendo.

C) ¿Habrá alguna condición que cumplan L1 y/o L2 que permita asegurar que el nivel L3 no sea necesario? Explique.

Se requeriría que la tasa de Hit de L1 y/o de L2 fuese de 100%. Si la tasa de Hit de L1 fuese 100%, no se requeriría ningún otro nivel de caché.



Problema 3 (PEP 2, Verano 2018)

Considere un procesador para el que se define una memoria virtual de hasta 64TB y un espacio de direccionamiento de memoria física con direcciones de 36 bits. Las páginas son de 4KB.

A) ¿Cuál es la máxima cantidad de memoria física que se puede disponer con este procesador?

Con 36 bits se pueden direccionar 2^{36} bytes = $2^6 \times 2^{30} = 2^6$ GBytes = 64 GBytes

B) ¿Cuántos bits se requieren para el espacio de direcciones virtuales?

64 TB = $2^6 \times 2^{40}$ Bytes => se requieren 46 bits

C) Determine la estructura de una dirección virtual y de una dirección física. Es decir, determine cuántos bits se usan para cada campo dentro de la memoria virtual y de la memoria física.

Las páginas son de 4KB = 2^{12} bytes, luego, en la memoria física habrán $2^{36} / 2^{12} = 2^{24} = 2^4 \times 2^{20} = 16$ MPáginas. Para determinar una página (page offset) se requiere 12 bits

MV => 46 bits separados en 34 bits para páginas virtuales y 12 bits para el offset de página.

MF => 36 bits separados en 24 bits para marcos de página y 12 bits para el offset de página

D) ¿Cuántas entradas tendrá la Tabla de Páginas?

Como hay 34 bits para la cantidad de páginas virtuales, la tabla debe tener 2^{34} entradas, esto es, $2^4 \times 2^{30} = 16$ G entradas.



Problema 4

Para una caché de mapeo directo con direcciones de memoria de 32 bits, los siguientes bits de la dirección son usados para acceder la caché:

Tag	Index	Offset
31-10	9-5	4-0

A) ¿Cuál es el tamaño del bloque (en palabras)?

8

B) ¿Cuántas entradas tiene la caché?

32

C) ¿Cuál es la razón entre el total de bits requeridos para esta implementación de caché y los bits usados para almacenar datos?

$1 + (23 / (8 * 32)) = 1,09$. Notar que se requieren 23 bits (22 tag + 1 valid) por cada línea de caché adicionales a los $8 * 32$ bits de datos.

Después de encender el procesador, se realizan las referencias a la caché de las siguientes direcciones de memoria expresadas en *byte*: 0, 4, 16, 132, 232, 160, 1024, 30, 140, 3100, 180, 2180.

D) ¿Cuántos bloques son reemplazados?

3. Notar que los *bloques* referenciados son: 0, 0, 0, 4, 7, 5, **32**, **0**, 4, 96, 5, **68**; y estos a su vez son mapeados en la línea de caché $\text{dir_bloque} \bmod 32$: 0, 0, 0, 4, 7, 5, 0, 0, 4, 3, 5, 4. Los reemplazos están marcados en rojo, y corresponden a los número de bloque “nuevos” que llegan a ocupar una línea ya ocupada.

E) ¿Cuál es la tasa de hit?

0,25 (3 / 12)

F) Muestre el estado final de la caché, con cada entrada **válida** representada como un registro de <index, tag, dato>

<0000₂, 00000000000000000000₂, mem[0-31]>
<00100₂, 00000000000000001000100₂, mem[2176-2207]>
<00111₂, 00000000000000000000111₂, mem[224-255]>
<00101₂, 00000000000000000000101₂, mem[160-191]>
<00011₂, 00000000000000001100000₂, mem[3072-3103]>

Nota: El sub-índice “2” denota expresión en binario.



Problema 5

En general, el tiempo de acceso a la caché es proporcional a su capacidad. Asuma que el tiempo de acceso a memoria principal es de 70 ns y que los accesos a memoria son el 36% de todas las instrucciones. La tabla siguiente muestra los tamaños, tasas de desacierto (miss) y tiempo de acierto (hit) para dos procesadores, P1 y P2.

	L1 Size	L1 Miss Rate	L1 Hit Time
P1	2 KiB	8%	0,66 ns
P2	4 KiB	6 %	0,90 ns

- A) Asumiendo que el tiempo de acierto (hit) de L1 determina los tiempos de ciclo de reloj de P1 y P2, ¿cuáles son las respectivas tasas de reloj?

P1: 1.52 GHz, P2: 1.11 GHz

- B) ¿Cuál es el AMAT para P1 y P2?

P1: 6,26 ns (0,66 ns + 8% * 70 ns) o 9,48 ciclos. P2: 5,1 ns o 5,66 ciclos.

- C) Asumiendo un CPI base de 1 sin esperas de memoria, ¿cuál es el CPI total para P1 y P2? ¿Cuál es mas rápido?

P1: CPI=4,05. P2: CPI 2,68

En las siguientes partes, consideraremos la adición de un caché L2 a P1 para compensar por su supuesta capacidad limitada en L1. A continuación se muestra la tasa de desacierto *local* y el tiempo de acierto de L2 de distintos tamaños.

L2 Size	L2 Miss Rate	L2 Hit Time
512 KiB	98%	4,28 ns
1 MiB	95%	5,62 ns
2 MiB	90%	6,42 ns
4 MiB	80%	7,02 ns

- D) ¿Cuál es el AMAT de P1 con la adición de L2? ¿Es peor o mejor con la caché L2?

AMAT 512 KiB = $0,66 + 8\% * (4,28 + 98\% * 70) = 6,49 \text{ ns} \rightarrow \text{peor.}$

AMAT 1 MiB = $0,66 + 8\% * (5,62 + 95\% * 70) = 6,42 \text{ ns} \rightarrow \text{peor.}$

AMAT 2 MiB = $0,66 + 8\% * (6,42 + 90\% * 70) = 6,21 \text{ ns} \rightarrow \text{mejor.}$

AMAT 4 MiB = $0,66 + 8\% * (7,02 + 80\% * 70) = 5,70 \text{ ns} \rightarrow \text{mejor.}$

- E) ¿Cuál es el CPI total de P1 después la adición de L2 en el mejor de los casos?

CPI = 3,75

- F) ¿Cuál procesador es mas rápido ahora que P1 tiene L2? Si P1 es mas rápido, ¿qué tasa de desacierto requeriría P2 en su caché L1 para igualar el desempeño de P1? Si P2 es mas rápido, ¿qué tasa de desacierto requeriría P1 en su caché L1 para igualar el desempeño de P2?

Sigue siendo más rápido P2. P1 requiere una tasa de desacierto de L1 de:

$5,1 = ,66 + MR * (7,02 + 80\% * 70) \rightarrow MR = 7\%$