

PROBLEMAS DE ESTRUCTURA DE COMPUTADORES

MEMORIA

Memoria Cache - Problemas básicos

1.- En un computador con caches separadas de datos e instrucciones, de 32KB cada una y bloques de 256 bytes, se quiere ejecutar el siguiente código C, que realiza la operación matricial $C = A+B$:

```
#define N 128
int A[N][N];
int B[N][N];
int C[N][N];

for (i=0; i < N; i++)
    for (j=0; j < N; j++)
        C[i][j] = A[i][j] + B[i][j];
```

Asumiendo que la matriz **A** se ubica en la dirección **0x0C000000**, que B y C se almacenan consecutivamente a continuación de A, y que en la traducción a ensamblador del código anterior la variable *i* se ubica en un registro y que las direcciones de comienzo de los arrays están también en registro, se pide:

- Calcular el número de fallos de cache en lectura (load) y en escritura (store) si la cache es de emplazamiento directo sin asignación en escritura.
- Repetir el cálculo para una cache asociativa por conjuntos, de 2 vías, con política de reemplazamiento LRU, en los siguientes casos:
 - sin asignación en escritura.
 - con asignación en escritura.

En cada caso indicar si el rendimiento mejoraría o empeoraría y en qué proporción.

- Un programador sugiere la siguiente modificación de código para una cache asociativa por conjuntos de 2 vías con asignación en escritura:

```
#define N 128
struct {
    int A;
    int B;
} AB[N][N];
int C[N][N];

for (i=0; i < N; i++)
    for (j=0; j < N; j++)
        C[i][j] = AB[i][j].A + AB[i][j].B;
```

Asumiendo que el array AB se ubica en la dirección 0x0C000000 y que el array C se ubica a continuación, calcule el número de fallos en lectura y escritura que se producirían, y compárelo con el resultado del apartado b para la misma configuración de cache.

- ¿Puede obtenerse un resultado similar al del apartado c con el código original y sin aumentar el tamaño de la cache? Razone la respuesta.

2.- Sea un computador con una memoria principal de 64K bytes, direccionable en bytes, al que se dota de una memoria cache de 2K bytes, con líneas de 256 bytes, y prebúsqueda bajo fallo (en caso de fallo se trae el bloque que lo provoca y el siguiente).

Sea la secuencia de acceso a memoria principal dada por las siguientes direcciones: 0x3345, 0x1244, 0x3374, 0x5BAC, 0x132A, 0x5C41 y 0x13AF.

Muestre la evolución de la caché de etiquetas indicando los fallos y las prebúsquedas que se producen en los siguientes casos:

- Una organización directa
- Una organización asociativa por conjuntos de 2 vías, con algoritmo de reemplazo LRU
- Una organización directa y un buffer de 512 bytes donde se almacena inicialmente el bloque prebuscado hasta que se le referencia. El buffer actúa con emplazamiento asociativo y política FIFO.

3.- Considere un computador con una memoria principal de 4MB, direccionable por bytes, al que se dota de una

memoria cache de 4KB, con líneas de 512B, y cache de víctima de 1024B. La memoria cache tiene dos vías mientras que la cache víctima es completamente asociativa. Para ambas el algoritmo de reemplazamiento es FIFO. Originalmente en las dos memorias se encuentran los bloques que aparecen en las siguientes tablas:

Conjunto	Etiqueta	FIFO
0	668	0
0	008	1
1	297	0
1	368	1
2	5FF	0
2	668	1
3	297	1
3	200	0

Etiqueta/ Conjunto de MP	FIFO
0A80	0
3FFF	1

Sea la secuencia de acceso a memoria principal dada por las siguientes direcciones: 0x334500, 0x14BF00, 0x150084, 0x004021, 0x0540AB y 0x0041F1.

Muestre la evolución de la caché de etiquetas indicando los fallos y las transferencias entre cache de víctima y cache.

4.- Considere un computador con direcciones de 32 bits, al que se dota de una memoria cache de 128 bytes, con bloques de 16 bytes y emplazamiento directo, sin asignación en escritura (No write-allocate). Suponga que la variable **a** está en el banco de registros y no en cache.

Se quiere ejecutar el siguiente código:

```
int nota[128]; // nota[0] está en la dirección 0x00000000
int media[128]; // media [0] está a continuación de nota[127]

for (i=0;i<128;i++) {
    if (i>7 && i<64) {
        nota[i] = media[i]/2;
    }
    else {
        a = nota[i]*media[i]
    }
}
```

a) Indique cuántos fallos de cache se producen.

b) Proponga dos optimizaciones de código (independientes) que mejoren el resultado, indicando cuántos fallos se producen en esos casos.

c) Si el tiempo de acceso a una palabra en cache es 1 ns, el tiempo de lectura o escritura de una palabra en memoria principal es 50 ns, el tiempo de transferencia de un bloque de MP a cache es 200 ns, ¿cuál es el tiempo necesario para acceder a todas las referencias del programa?

d) Si se añade al computador una memoria cache de segundo nivel de 1MB y asociatividad 4, con bloques de 16 bytes, con asignación en escritura y actualización diferida y si el tiempo de transferencia de un bloque desde memoria cache de segundo nivel a la de primer nivel son 15ns, ¿cuál es el tiempo necesario para acceder a todas las referencias del programa?



Memoria Cache - Problemas adicionales

1.- Sea un computador con memoria cache con las características siguientes:

Memoria física de 32 KB

Memoria cache de direcciones físicas de 512 bytes con bloques de 128 bytes y emplazamiento directo

- Indique el formato de la dirección para MP teniendo en cuenta las características de la MC
- Si en un momento dado el directorio de la cache tiene los contenidos (en hexadecimal) indicados en la tabla, exprese en hexadecimal el rango de direcciones físicas ubicadas en cada bloque de la memoria cache.
- Suponiendo que un programa realiza la siguiente cadena de referencias indique el número de aciertos que se producen: de 0x2080 a 0x209F, de 0x2880 a 0x289F y de 0x03F0 a 0x0410.

Etiqueta	Bloque
35	0
10	1
10	2
08	3

2.- Sea una memoria principal de 32M bytes direccionable por bytes a la que se dota de una memoria cache con las siguientes características:

- Tamaño de 2K bytes
- Líneas de 256 bytes,
- Prebúsqueda bajo fallo (en caso de fallo se trae el bloque que lo provoca y el siguiente)
- Algoritmo de reemplazamiento FIFO.

Sea la secuencia de acceso a memoria principal dada por las siguientes direcciones: 0x00023FA, 0x00014A2, 0x0003F02, 0x00040B1, 0x0005572 y 0x00023AA. Muestre la evolución del directorio cache indicando los fallos (F), las prebúsquedas (P) y los aciertos (A) que se producen suponiendo:

- emplazamiento directo
- emplazamiento asociativo por conjuntos de 2 vías

3.- Se dispone de un computador con una memoria direccionable por bytes y dotado de una cache de 256 bytes. Cada bloque es de 16 bytes.

a) Si la cache es de organización directa, indique para la siguiente secuencia de referencias a memoria si se trata de un acierto o un fallo: 0xA01, 0xB1F, 0x70A, 0x60F, 0xA70, 0xB11, 0xA7A, 0xA0B, 0x67A, 0xA7F, 0x071, 0x67F. Indique además qué fallos producen reemplazamiento.

b) Si la cache es asociativa por conjuntos con dos vías, con reemplazamiento LRU, indique para la misma secuencia de referencias del apartado anterior si se producen reemplazamientos.

4.- Considere un computador con una memoria principal de 128MB, direccionable por bytes, al que se dota de una memoria cache de 1MB, con líneas de 256B, y cache de víctima de 1KB. La memoria cache tiene acceso directo, mientras que la cache víctima es completamente asociativa. Para ambas el algoritmo de reemplazamiento es FIFO. Sea la secuencia de acceso a memoria principal dada por las siguientes direcciones: 0x0334500, 0x014BF00, 0x4134501, 0x124BF00, 0x03345FE, 0x34345AC, 0x214BFAA y 0x014BF54.

Suponiendo la cache originalmente vacía muestre la evolución de la caché de etiquetas indicando los fallos y las transferencias entre cache de víctima y cache.

5.- Considere un computador con una memoria principal de 4MB, direccionable por bytes, al que se dota de una memoria cache de 2KB, con líneas de 512B. La memoria cache es de emplazamiento directo y con asignación de escritura. Se quiere ejecutar el siguiente código:

```
for(i=0;i<1024;i++) {  
    C[i]=A[i]-B[1023-i];  
}
```

}

Los datos son enteros y están almacenados desde la dirección 0x200000

- ¿Cuántos fallos de cache se producen?
- ¿Cuántos fallos se producen si aplicamos fusión de arrays? ¿Importa el orden de la fusión?
- ¿Cuántos fallos se producen si aplicamos alargamientos de arrays? ¿Importa la cantidad "alargada"?

6.- Considere un computador con direcciones de 32 bits, al que se dota de una memoria cache de 4KB, con líneas (bloques) de 16 bytes, de emplazamiento directo y con asignación de escritura.

Se quiere ejecutar el siguiente código:

```
int A[1024]; // A[0] están en la dirección 0x0C000000
int B[1024];
int C[1024];

for (i=0; i<10; i++) {
    for(j=0; j<1024; j++) {
        C[i] += A[j]*B[j];
    }
}
```

- ¿Cuántos fallos de cache se producen?
- ¿Qué asociatividad necesitamos (como mínimo) para evitar todos los fallos por conflicto con este código?
- Siguiendo con emplazamiento directo, ¿qué modificaciones de código podríamos hacer para evitar todos los fallos por conflicto?

7.- Considere el siguiente código que implementa una multiplicación de matrices:

```
int A[128][128];
int B[128][128];
int C[128][128];

for (i=0; i<128; i++){
    for (j=0; i<128; j++){
        C[i][j]=0;
        for (k=0; k<128; k++){
            C[i][j] += A[i][k] * B[k][j];
        }
    }
}
```

- ¿Qué matriz presenta peor localidad en este código? ¿Se puede mejorar mediante intercambio de bucles? Al mejorar la localidad de esa matriz, ¿empeora la de alguna otra?
- Busque información sobre la transformación llamada "loop tiling" (a veces "blocking" o bloqueo). Aplíquela a este código y discuta si disminuiría el número de fallo de cache (no es necesario cuantificar la mejora; sólo discutir si es una transformación conveniente).

8.- Se va a ejecutar el código que se muestra a continuación en una máquina con direcciones de 16 bits que dispone de una memoria cache de datos de acceso directo con 4 bloques de 16 bytes cada uno y política de escritura write-back con asignación en escritura. Teniendo en cuenta sólo la cache de datos (es decir, ignorando la lectura de instrucciones):

- Ignorando los accesos a la variable tmp (estará en registro), ¿cuántos fallos de cache se producirán? Justifique la respuesta, indicando además cuántos reemplazamientos se producen y en qué iteraciones.
- Recalcule el número de fallos si la cache fuese asociativa con 2 vías (2 bloques por conjunto) y la política de reemplazamiento fuese LRU.
- ¿Qué etiqueta (expresada en hexadecimal) tendrá el (marco de) bloque 2 de la cache tras la ejecución del código?

```
//Cada entero ocupa 4 bytes
int v[16]; // Dirección de comienzo 0x0000
int mask[4]; // Dirección de comienzo 0x0040

for(i=0; i<12; i++) {
    tmp = 0;
    for(j=0; j<4; j++){
        tmp = tmp + mask[j]*v[i];
    }
    v[i]=tmp;
}
```