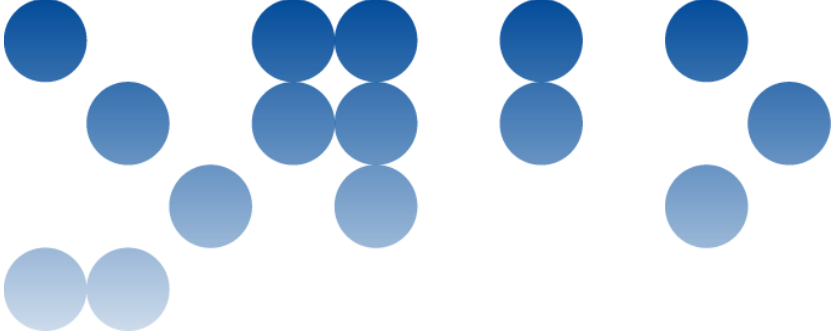

PEC2

Estructura de computadores

2019 s2

Estudios de informática, multimedia y comunicación

EIMT.UOC.EDU

UOC
Universitat Oberta de Catalunya

Presentación

La presente PEC2 contiene 2 preguntas y representa el 50% de la nota de la evaluación continua.

Como podréis ver, los ejercicios son muy parecidos a los que habéis hecho durante estos días, en los que además habéis podido dar las soluciones, comentarlas y plantear dudas en el foro. Esta PEC es **individual**, **evaluable** y por tanto no puede comentarse.

Competencias

Las competencias específicas que persigue la PEC2 son:

- [13] Capacidad para identificar los elementos de la estructura y los principios de funcionamiento de un ordenador.
- [14] Capacidad para analizar la arquitectura y organización de los sistemas y aplicaciones informáticos en red.
- [15] Conocer las tecnologías de comunicaciones actuales y emergentes y saberlas aplicar convenientemente para diseñar y desarrollar soluciones basadas en sistemas y tecnologías de la información.

Objetivos

Los objetivos de la siguiente PEC son:

- Conocer la organización del sistema de memoria de un computador.
- Conocer el funcionamiento de la memoria cache, así como los algoritmos de correspondencia y reemplazamiento.
- Conocer la organización del sistema de entrada/salida.
- Comprender las técnicas de entrada/salida (entrada/salida programada, Interrupciones y DMA).

Enunciado

Responder cada pregunta o apartado en el recuadro correspondiente.

Recursos

Podéis consultar los recursos disponibles en el aula, pero no hacer uso del foro.

Criterios de valoración

La **puntuación** y los **criterios de evaluación** los encontraréis en cada pregunta.

Formato y fecha de entrega

La PEC1 podéis entregarla en el apartado de **entrega de actividades** con el nombre **apellido1_apellido2_nombre_PEC2 (pdf / odt / doc / docx).**

La fecha **límite** de entrega es el **01/05/2020.**

Enunciado

Pregunta 1 (4 puntos)

Tenemos un sistema de memoria en el que todos los accesos se hacen a palabra (no nos importa cuál es el tamaño de una palabra). Supondremos que el espacio de direcciones de memoria se descompone en bloques de 8 palabras. Cada bloque comienza en una dirección múltiplo de 8. Así, el bloque 0 contiene las direcciones 0, 1, 2, 3, 4, 5, 6 y 7; el bloque 1, las direcciones 8, 9, 10, 11, 12, 13, 14 y 15, y el bloque N las direcciones 8*N, 8*N+1, 8*N+2, 8*N+3, 8*N+4, 8*N+5, 8*N+6 y 8*N+7.

Una fórmula para calcular el identificador numérico del bloque es la siguiente:

Bloque = dirección de memoria (dirección palabra) DIV 8 (tamaño del bloque en palabras)

Suponemos que el sistema también dispone de una memoria cache de 4 líneas (donde cada línea tiene el tamaño de un bloque, es decir, 8 palabras). Estas líneas se identifican como líneas 0, 1, 2 y 3. Cuando se hace referencia a una dirección de memoria principal, si esta dirección no se encuentra en la memoria cache, se trae todo el bloque correspondiente desde la memoria principal a una línea de la memoria cache (así si hacemos referencia a dirección 2 de memoria principal traeremos el bloque formado por las palabras 0, 1, 2, 3, 4, 5, 6 y 7).

Apartado 1.1 (2 puntos) Memoria Cache de Acceso Directo

Suponemos que el sistema utiliza una **política de asignación directa**, de manera que cada bloque de la memoria principal sólo se puede llevar a una línea determinada de la memoria cache. En este caso, el identificador del bloque determina la línea específica donde se puede guardar utilizando la siguiente fórmula (similar a la fórmula para determinar el bloque):

Línea = identificador de bloque MOD 4 (tamaño de la cache en líneas)

La ejecución de un programa genera la siguiente lista de lecturas a memoria:

28, 29, 30, 2, 3, 15, 16, 1, 44, 45, 46, 12, 33, 8, 3, 17, 39, 60, 4, 35

1.1.a) (1,2 puntos)

La siguiente tabla muestra el estado inicial de la cache, que contiene las primeras 32 palabras de la memoria (organizadas en 4 bloques). Completar la tabla para mostrar la evolución de la cache durante la ejecución del programa. Para cada acceso hay que llenar una columna indicando si se trata de un acierto o de un fallo.

Si es un acierto escribiremos E en la línea correspondiente delante de la direcciones del bloque, si es un fallo escribiremos F y en ese caso se indicará el nuevo bloque que se trae a la memoria cache a la línea que corresponda, expresado de la forma b:e(a_0 – a_7), donde b: número de bloque, e: etiqueta y (a_0 – a_7) son las direcciones del bloque, donde a_0 es la primera dirección i a_7 es la octava (última) dirección del bloque.

Línea	Estado Inicial	28	29	30	2	3
0	0:0 (0 - 7)					
1	1:0 (8 - 15)					
2	2:0 (16 - 23)					
3	3:0 (24 - 31)					

Línea	15	16	1	44	45	46
0						
1						
2						
3						

Línea	12	33	8	3	17	39
0						
1						
2						
3						

Línea	60	4	35			
0						
1						
2						
3						

1.1.b) (0,4 puntos)

¿Cuál es la tasa de fallos (T_f)?

1.1.c) (0,4 puntos)

Suponemos que el tiempo de acceso a la memoria cache, o tiempo de acceso en caso de acierto (t_a), es de 4 ns y el tiempo total de acceso en caso de fallo (t_f) es de 20 ns. Considerando la tasa de fallos obtenida en la pregunta anterior, ¿cuál es el tiempo medio de acceso a memoria (t_m)?

Apartado 1.2 Memoria cache de acceso completamente asociativo

Ahora suponemos que el mismo sistema utiliza una política de emplazamiento completamente asociativa, de manera que cualquier bloque de la memoria principal se puede llevar a cualquier bloque de la memoria cache.
Si encontramos que la cache ya está llena, se utiliza un algoritmo de reemplazamiento LRU, de manera que sacaremos de la memoria cache aquel bloque que hace más tiempo que no se referencia.

Consideremos la misma lista de lecturas a memoria:

28, 29, 30, 2, 3, 15, 16, 1, 44, 45, 46, 12, 33, 8, 3, 17, 39, 60, 4, 35

1.2.a) (1,2 puntos)

La siguiente tabla muestra el estado inicial de la cache, que contiene las primeras 32 palabras de la memoria (organizadas en 4 bloques).

Completar la tabla para mostrar la evolución de la cache durante la ejecución del programa. Para cada acceso hay que llenar una columna indicando si se trata de un acierto o de un fallo.

Si es un acierto escribiremos E en la línea correspondiente delante de las direcciones del bloque, si es un fallo escribiremos F y se indicará el nuevo bloque que se traerá a la memoria cache a la línea que le corresponda, expresado de la forma b (a_0 – a_7), donde b: número de bloque, y (a_0 - a_7) son las direcciones del bloque, donde a_0 es la primera dirección y a_7 es la octava (última) dirección del bloque.

Línea	Estado Inicial	28	29	30	2	3
0	0 (0 - 7)					
1	1 (8 - 15)					
2	2 (16 - 23)					
3	3 (24 - 31)					

Línea	15	16	1	44	45	46
0						
1						
2						
3						

Línea	12	33	8	3	17	39
0						
1						
2						
3						

Línea	60	4	35			
0						
1						
2						
3						

1.2.b) (0,4 puntos)

¿Cuál es la tasa de fallos (T_f)?

1.2.b) (0,4 puntos)

Suponemos que el tiempo de acceso a la memoria cache, o tiempo de acceso en caso de acierto (t_a), es de 4 ns y el tiempo total de acceso en caso de fallo (t_f) es de 20 ns. Considerando la tasa de fallos obtenida en la pregunta anterior, ¿cuál es el tiempo medio de acceso a memoria (t_m)?

Criterios de valoración. Para los apartados 1.1.a y 1.2.a cada error en los fallos o aciertos de la memoria cache o en la colocación de un bloque en la cache resta 0,6. Los apartados restantes se puntuarán con los 0,4 puntos cada uno de ellos si la solución es correcta y coherente con vuestra respuesta a los apartados a) correspondientes.

Pregunta 2 (5 puntos)

Se desea analizar el rendimiento de la comunicación de datos entre la memoria de un procesador y un port USB con un dispositivo conectado, que tienen las siguientes características:

- Velocidad de transferencia del dispositivo de E/S ($\mathbf{v}_{transf}$)= 8MB/s
- Tiempo de latencia promedio del dispositivo ($t_{latencia}$) = 0
- Direcciones de los **registros de datos y de estado** del controlador de E/S: 0400h y 0404h
- El bit del **registro de estado** que indica que el controlador del port de E/S está disponible es el bit 3, o el cuarto bit menos significativo (cuando vale 1 indica que está disponible)
- Procesador con una frecuencia de reloj de 1 GHz, y el procesador puede ejecutar 1 instrucción cada dos ciclos de reloj (t_{instr}= 2*t_{ciclo})
- Tiempo de programación y finalización de la transferencia de 1000 ns (t_{prog} + t_{final})
- Transferencia de **escritura** desde memoria al port de E/S
- En cada escritura de un dato se transfieren 4 Bytes
- Transferencia de N_{datos}= 800.000 datos, es decir, 800.000 × 4 Bytes = 3.200.000 Bytes
- Dirección inicial de memoria donde residen los datos: 80000000h

Apartado 2.1 (2 puntos) E/S programada

El siguiente código realizado con el repertorio CISCA realiza la transferencia descrita antes mediante la técnica de E/S programada.

```
1.              MOV   R3, VALOR1
2.              MOV   R2, VALOR2
3. Bucle:       IN    R0, [VALOR3]      ; leer 4 bytes
4.              AND   R0, VALOR4
5.              JE    Bucle
6.              MOV   R0,[R2]           ; leer 4 bytes
7.              ADD   R2, VALOR5
8.              OUT   [VALOR6], R0      ; escribir 4 bytes
9.              SUB   R3, 1
10.             JNE   Bucle
```

2.1.a) (0,5 puntos)

Substituir por los valores apropiados:

VALOR1= **VALOR2**=

VALOR3= **VALOR4**=

VALOR5= **VALOR6**=

2.1.b) (0,5 punto)

¿Cuánto tiempo dura la transferencia?
¿Qué porcentaje de este tiempo dedica la CPU a la transferencia?

2.1.c) (1 punto)

Si quisiéramos utilizar el mismo procesador y el mismo programa, pero con un dispositivo de E/S más rápido, ¿Cuál es la máxima tasa o velocidad de transferencia del nuevo dispositivo que se podría suportar sin que el dispositivo se tuviera que esperar?

Apartado 2.2 (2 puntos) E/S por Interrupciones

Suponer que el siguiente código CISCA es una rutina de servicio a las interrupciones (RSI) para transferir a través del dispositivo de E/S anterior, el mismo número de datos que antes con E/S programada, pero ahora mediante la técnica de E/S por interrupciones.

Suponer:

- El tiempo para atender la interrupción (t_{rec_int}), o tiempo adicional desde que la CPU detecta la interrupción hasta que comienza a ejecutarse la primera instrucción de la RSI es de 10 ciclos de reloj.
- Se utiliza una variable global que se representa con la etiqueta **Dir,** y que al principio del programa contiene la dirección inicial de memoria donde residen los datos a transferir.

```
1.    CLI
2.    PUSH  R0
3.    PUSH  R1
4.    IN    R0, [VALOR1] ; leer 4 bytes
5.    AND   R0, VALOR2
6.    JE    Error          ; salta a un código de tratamiento de error no
                           ; descrito, se ha producido la petición por parte
                           ; del dispositivo, pero el dato no está disponible.
7.    MOV   R1, [VALOR3]
8.    MOV   R0, [R1]
9.    OUT   [VALOR4], R0 ; escribir 4 bytes
10.   ADD   R1, VALOR5
11.   MOV   [VALOR3], R1
12.   POP   R1
13.   POP   R0
14.   STI
15.   RETI
```

2.2.a) (0,5 puntos)

Sustituir por los valores adecuados:

VALOR1= **VALOR2**=

VALOR3= **VALOR4**=

VALOR5=

2.2.b) (1 punto)

¿Cuál es el tiempo total que dedica la CPU a la tarea de Entrada/Salida, t_{cpu}? ¿Cuál es el porcentaje que representa el tiempo de transferencia del bloque t_{transf_bloque} con respecto al tiempo de transferencia del bloque por parte del periférico t_{bloque}?

2.2.c) (0,5 puntos)

Si quisiéramos reducir la frecuencia de reloj del procesador para reducir su consumo energético, hasta qué frecuencia lo podríamos hacer sin reducir la velocidad de transferencia con el dispositivo de 'E/S?

Apartado 2.3 (1 punto) E/S por DMA

Supondremos que el controlador de E/S puede funcionar en modo DMA (Acceso Directo a Memoria). La suma del **tiempo de cesión** del bus y del **tiempo de recuperación** del bus es de 40 ns ($t_{cesión}$ + t_{recup}= 40 ns). El **tiempo de la transferencia** por el bus es de 1 ns (t_{mem} = 1 ns).

2.3.a) (0,5 puntos)

Consideremos que en la transferencia por DMA, los datos se envían entre el controlador de DMA y la memoria, en modo ráfaga, i dispone de un buffer de medida m_{buffer} = 1600 bytes. Calcular el tiempo total de ocupación del bus por parte del controlador de DMA para llevar a cabo la transferencia que venimos analizando.

2.3.b) (0,5 puntos)

La CPU no puede hacer ninguna tarea durante todo el tiempo en que el bus está ocupado por parte del controlador de DMA. ¿Cuál es el porcentaje de tiempo que tiene disponible la CPU para ejecutar código efectivo de otros programas durante la transferencia?

Criterios de valoración. En los apartados 2.1.a y 2.2.a cada valor erróneo resta 0,25. El resto de apartados están bien o están mal. No hay gradación.

Pregunta 3 (1 punto)

Preguntas teóricas

3.a) (0,25 puntos)

¿Por qué el sistema de memoria se organiza en una estructura jerárquica? ¿Cuáles son los niveles de esta estructura?

3.b) (0,25 puntos)

En la memoria caché, ¿Cuándo debe utilizarse un algoritmo de reemplazo? ¿Cuáles son los principales algoritmos de reemplazo?

3.c) (0,25 puntos)

¿Cuáles son las tres partes básicas de un módulo de E/S? ¿Qué tipos de registros incluye un módulo de E/S?

3.d) (0,25 puntos)

En un sistema de E/S gestionada por DMA ¿Qué diferencia hay entre el funcionamiento normal y el modo ráfaga, qué ventajas tiene un modo con respecto al otro?