

Arquitectura de Computadores  
**Práctica de Memorias Cache**  
Ejercicios

Curso 2011/2012

## 1. Información general

En esta convocatoria, la práctica consta de **2** ejercicios cuyos enunciados aparecen en el **cuestionario** adjunto. En todos los casos se suministra al alumno un programa escrito en el ensamblador del MC88110 y un conjunto de configuraciones de la memoria de cache de datos con los que se habrá de experimentar. Esta experimentación consistirá, fundamentalmente, en observar el comportamiento de la memoria **cache de datos** durante la ejecución del programa, dependiendo de su configuración: capacidad total, tamaño de los bloques, y políticas de ubicación y de escritura.

También se propondrá al alumno la modificación del programa original suministrado, con el fin de obtener una mejora en el rendimiento de la ejecución. En tal caso, se sugerirá el uso de algunas de las técnicas de mejora existentes (véase el apartado 3 de la documentación de la práctica) y se tratará de observar en qué medida se ha conseguido optimizar la versión original.

### 1.1. Valores de configuración comunes

En la mayoría de las configuraciones que se emplean en esta convocatoria hay un conjunto de parámetros que se mantiene fijo. Estos son:

- Memoria principal: 10 ciclos; entrelazado de 1 (no entrelazado).
- Cache de instrucciones: activada, 2 ciclos, 128 bloques, 16 bytes por bloque y ubicación directa.
- Cache de datos: activada, 2 ciclos.
- Registro de control: redondeo al más próximo, ejecución secuencial, no permite excepciones, *little endian*.

### 1.2. Cómo rellenar el cuestionario

Deberá entregar relleno el cuestionario que se suministra a continuación. Éste consta de una primera página en la que, además de la identificación de la práctica y de la convocatoria a que corresponde, deberán figurar los nombres de los integrantes del grupo.

Seguidamente aparecen los ejercicios, en cada uno de los cuales se plantea una serie de preguntas que es **obligatorio** contestar. Algunas de estas cuestiones exige, además, rellenar unas tablas. Para contestar las preguntas se recomienda editar la versión Word (**cuestionario.doc**) que se suministra en la distribución. Si no dispone de este procesador de textos, puede copiar y contestar las cuestiones con el editor que desee. En este caso se sugiere que complete directamente las tablas impresas, bien las incluidas en la documentación que se distribuye en el Servicio de Publicaciones, o bien imprimiendo las hojas correspondientes desde la versión pdf (**cuestionario.pdf**), que también se incluye en la distribución.

En cualquier caso, y tal y como se indica en el apartado de **Normas de Entrega** de la documentación, será necesario entregar un fichero conteniendo una versión ASCII (**cuestionario.txt**) de este cuestionario, en la que *no* es necesario incluir las tablas. *Este fichero debe estar presente en todas sus entregas*, aunque sólo se tendrá en cuenta el de la entrega definitiva, por lo que es válido que dicho fichero esté vacío salvo en el caso de esa última entrega.

## 1.4. Observaciones.

- En las configuraciones de cache asociativas o asociativas por conjuntos el alumno deberá tener en cuenta que, además de los tipos de fallos que aparecen explícitamente en las tablas (**forzosos** y de **capacidad**), pueden existir **otros** fallos debidos a la **política de reemplazo**, que es LRU.
- Los ficheros de código a entregar por el alumno **deberán incluir en su cabecera**, como comentarios, el programa modificado escrito en lenguaje de alto nivel. Asimismo, los comentarios que se incluyan en el código ensamblador deberán ajustarse a las modificaciones realizadas por el alumno.
- Al realizar las optimizaciones sobre los programas originales deberá conservarse el orden de los accesos a memoria según la precedencia de los operadores.

*Ejemplo1:*  $s[i] = s[i] + m[i][j]$

1<sup>er</sup> acceso: lectura de  $s[i]$ , 2<sup>o</sup> acceso: lectura de  $m[i][j]$  3<sup>er</sup> acceso: escritura en  $s[i]$

*Ejemplo2:*  $\text{if } (m[i][j] > \text{max}[i])$

1<sup>er</sup> acceso: lectura de  $m[i][j]$ , 2<sup>o</sup> acceso: lectura de  $\text{max}[i]$

*Ejemplo3:*  $m[i][j] = m[i][j] + r * n[i][j]$

1<sup>er</sup> acceso: lectura de  $n[i][j]$ , 2<sup>o</sup> acceso: lectura de  $m[i][j]$ , 3<sup>o</sup> acceso: escritura en  $m[i][j]$ , ya que el producto se realiza antes que la suma.

- Con el fin de unificar criterios todas las referencias a variables no escalares se traducirán en instrucciones ld o st, según sea lectura o escritura.
- **Nota General.** Las **optimizaciones** que se realicen en los diferentes ejercicios **deberán limitarse a las que se proponen**, ya que de realizarse optimizaciones adicionales no sería posible sacar conclusiones válidas sobre su influencia.

## 1.5. Material de la práctica.

Existe una dirección Web dedicada a esta práctica, desde la que se puede acceder tanto a información sobre la misma como al material que se distribuye: documentación, cuestionario, ficheros de código y ficheros de configuración. Su URL es:

[www.datsi.fi.upm.es/docencia/Arquitectura/caches](http://www.datsi.fi.upm.es/docencia/Arquitectura/caches)

También se pueden encontrar los ficheros de código y de configuración en la máquina **batman** en el directorio:

`/usr/local/datsi/practica_caches/`

Asimismo, está disponible en el **Servicio de Publicaciones** una copia impresa de la documentación y el cuestionario.



Arquitectura de Computadores  
**Práctica de Memorias Cache**  
Cuestionario

Curso 2010/11

**Identificador del grupo:** .....

**Apellidos, nombre, nº de matrícula:**

1. ....

2. ....

**Fecha:**.....

# Ejercicio 1

El programa **ej1.ens** suma los elementos de las columnas de una matriz de 16x32 elementos de 4 bytes, almacenada en memoria por filas.

## Cuestiones:

- 1) Ejecute el programa con los ficheros de configuración de la tabla 1 y conteste a las siguientes cuestiones:
  - a) ¿En qué bloques de memoria principal se encuentran la matriz y el vector utilizados en el programa?
  - b) En el caso de las caches directas o asociativas por conjuntos de la tabla 1 ¿en qué bloques o conjuntos de la cache de datos se ubicarán la matriz y el vector?
  - c) Rellene la tabla 1, clasificando los fallos según el criterio que se indica en el apartado 2 de la documentación de la práctica. **JUSTIFIQUE** los resultados obtenidos teniendo en cuenta tanto **la traza de ejecución** del programa como los bloques de memoria cache en los que se ubica la matriz (según la respuesta del apartado anterior). Asimismo, deberá justificar por qué, en algún caso, ocurre que con caches de distinto tamaño y misma política de ubicación se obtiene el mismo número de fallos pero distinto tiempo de ejecución.

		Fallos							Nº instrucciones		
		Nº total accesos		Forzosos		Capacidad		Otros		Totales	
		Total	%	Total	%	Total	%	Total	%	Hit ratio	Nº ciclos
16 B/bloque	copy128dir										
	copy128as										
	copy128ac2										
	copy64dir										
	copy64as										
	copy64ac2										

Tabla 1.- Resultados del programa ej1.ens

Escriba su respuesta a continuación:

- 2) ¿Cuántos fallos se producirían en la cache directa de 64 bloques si la política de escritura utilizada fuera *write through*? Para responder a esta cuestión genere un fichero de configuración **write64dir** a partir del fichero **copy64dir**, modificando únicamente la política de escritura. **Justifique, basándose en el tipo** de accesos a datos que se realizan en el programa, a qué se debe la diferencia en el número de fallos, y cuál de las dos políticas de escritura proporciona mejores prestaciones en este caso concreto. Tenga en cuenta que las prestaciones no solo se miden en función de la tasa de aciertos obtenida.

Escriba su respuesta a continuación:

- 3) **Optimización 1: Intercambio de bucles.** Una posible optimización del programa original, **ej1.ens**, es recorrer la matriz en el mismo orden en el que está almacenada en memoria, es decir por filas. Realice dicha modificación, generando un nuevo programa, que deberá entregar en el fichero **ej1\_inter.ens**, debidamente comentado.

- a) Ejecute el nuevo programa, **ej1\_inter.ens**, empleando de nuevo la política de escritura **copy-back**, rellene la tabla 2 y **justifique debidamente** a qué se deben las diferencias que se observan (**fallos, ciclos, instrucciones, etc.**) con respecto al comportamiento del programa original.

		Fallos							Nº instrucciones		
		Nº total accesos	Forzosos		Capacidad		Otros		Totales		
			Total	%	Total	%	Total	%		Hit ratio	Nº ciclos
16 B/blq.	copy64dir										
	copy64as										
	copy64ac2										

Tabla 2.- Resultados de ej1\_inter.ens

**Escriba sus respuestas a continuación**

- b) Modifique el fichero de configuración **copy64as**, manteniendo la capacidad total de la cache pero aumentando el tamaño del bloque a 8, 16 y 32 palabras. Escriba los resultados obtenidos en la tabla siguiente y justifique a qué es debida la evolución observada en la tasa de aciertos. Haga lo mismo para el programa original y **justifique** por qué no se observa el mismo comportamiento en este último caso.

		ej1_inter.ens		ej.ens	
Tamaño bloque	Nº de bloques	Nº fallos	Hr	Nº fallos	Hr
<b>32B (8 palabras)</b>					
<b>64B (16 palabras)</b>					
<b>128B (32 palabras)</b>					
<b>256B (64 palabras)</b>					

**Escriba su respuesta a continuación:**

- 4) **Optimización 2: Unroll & jam.** Una posible optimización es reutilizar los datos mientras están en memoria caché, es decir, usar los datos presentes en la memoria caché, mientras queden operaciones pendientes con dichos datos, antes de que datos nuevos los desplacen. Para esta optimización se empleará de nuevo la política de escritura **copy-back**.

- a) Aplique al programa **ej1.ens** la técnica de **unroll&jam** y ejecute el programa resultante, que deberá entregar en el fichero **ej1\_unroll.ens** debidamente comentado, con los ficheros de configuración indicados en la tabla 3. Rellene dicha tabla y analice y compare los resultados con los obtenidos en la versión original. En particular analice y compare el número de ciclos, accesos, instrucciones ejecutadas, fallos y porcentajes de los distintos tipos de fallos. **JUSTIFIQUE** a qué se deben las diferencias.

		Fallos							Nº instrucciones		
		Nº total accesos	Forzosos		Capacidad		Otros		Totales		
			Total	%	Total	%	Total	%		Hit ratio	Nº ciclos
16 B/blq.	copy64dir										
	copy64as										
	copy64ac2										

Tabla 3.- Resultados de ej1\_unroll.ens

**Escriba sus respuestas a continuación**

- 5) Modifique el fichero de configuración **copy64as**, manteniendo el tamaño del bloque pero disminuyendo el tamaño de la memoria cache. Compruebe con tamaños de cache de 16, 8, 4 y 2 bloques la evolución de la tasa de aciertos del programa **ej1\_unroll.ens**. Escriba los resultados obtenidos en la tabla. Haga lo mismo para el programa **ej1\_inter.ens**.

	<b>ej1_unroll.ens</b>		<b>ej1_inter.ens</b>			
¿Cuán-ques se tan mínimo caso alcan-	<b>Nº de bloques</b>	<b>Nº fallos</b>	<b>Hr</b>	<b>Nº fallos</b>	<b>Hr</b>	tos blo-necesi-como en cada para zar la
tasa de aciertos máxima? ¿Cuántos se necesitan en el programa original? <b>Justifique debidamente su respuesta a continuación.</b>	<b>32</b>					
	<b>16</b>					
	<b>8</b>					
	<b>4</b>					

## Ejercicio 2

El programa **ej2.ens** consta de 3 bucles. En el primero se realiza una operación con tres matrices, mC, mA y mB, almacenando el resultado en mC, en el segundo se calcula la suma de las filas de la matriz mA, y en el tercero se calcula la media de las filas de la matriz mA. Las matrices utilizadas constan de 16x16 elementos de 4 bytes, y están almacenadas en memoria por filas.

## Cuestiones:

- 1) Ejecute dicho programa con los ficheros de configuración que aparecen en la tabla 4. Rellene dicha tabla y **justifique los resultados** obtenidos teniendo en cuenta **la traza de ejecución** del programa así como **los bloques de memoria cache** en los que se ubican las distintas estructuras de datos utilizadas.

Para clasificar los fallos deberá calcular previamente el número de fallos que se obtendrían con una cache asociativa y **política de reemplazo óptima**. En este caso, basta con que determine el número mínimo de bloques necesario para que se produzcan únicamente fallos forzosos en una cache asociativa con política de reemplazo óptima.

		Fallos						Nº Instrucciones		
		Forzosos		Capacidad		Otros		Totales	Hit ratio	Nº ciclos
		Total	%	Total	%	Total	%			
16B/16q	copy128dir									
	copy128ac2									
	copy128ac4									
	copy128as									

Tabla 4.- Resultados del programa ej2.ens

Escriba sus respuestas a continuación:

- 2) **Optimización 2: Fusión de bucles.** Modifique el programa **ej2.ens**, aplicando la técnica de fusión de bucles. Genere un nuevo programa, ejecútelo con los ficheros de configuración de la tabla 5 y rellene dicha tabla. Compare los resultados con los obtenidos para la versión original. ¿Es necesario en este caso utilizar como referencia, para la clasificación de los fallos, la política de reemplazo óptima? ¿Por qué?

La modificación realizada se deberá entregar en el fichero **ej2\_fus.ens** debidamente comentada.

		Fallos						Nº Instrucciones		
		Forzosos		Capacidad		Otros		Totales	Hit ratio	Nº ciclos
		Total	%	Total	%	Total	%			
16B/16q	copy128ac4									
	copy128as									
	copy32ac4									
	copy32as									
	copy16as									
	copy8as									

Tabla 5.- Resultados de ej2\_fus.ens

Escriba su respuesta a continuación:

- 3) **Optimización 3: Reemplazo escalar.** Modifique la versión del programa obtenida en el apartado anterior (**ej2\_fus.ens**), aplicando la técnica de reemplazo escalar. Genere un nuevo programa, ejecútelo y rellene la tabla 6. Compare los resultados con los obtenidos en la ejecución de los programas anteriores **ej2.ens** y **ej2\_fus.ens**: número total de ciclos en la ejecución, número total de accesos, número de instrucciones ejecutadas y número de fallos. ¿Qué tipo de fallos aparecen con **copy8as**?

La modificación realizada se deberá entregar, debidamente comentada, en el fichero **ej2\_rees.ens**.

		Fallos						Nº Instrucciones			
		Nº total accesos		<i>Forzosos</i>		<i>Capacidad</i>		<i>Otros</i>		<i>Totales</i>	
		Total	%	Total	%	Total	%			<i>Hit ratio</i>	Nº ciclos
16B/16q	<b>copy32ac4</b>										
	<b>copy32as</b>										
	<b>copy16as</b>										
	<b>copy8as</b>										

Tabla 6.- Resultados de **ej2\_rees.ens**

Escriba su respuesta a continuación:

## Comentarios

Este apartado también es **obligatorio**. Incluya aquí todos los comentarios personales sobre la práctica y su realización:

1. Valoración general de la práctica y en qué medida la considera útil cómo complemento de la parte teórica correspondiente.
2. Esfuerzo invertido en su realización en horas de trabajo, indicando el esfuerzo de cada uno de los integrantes del grupo.
3. Aspectos que considera mejorables en la práctica, tanto en los medios utilizados como en los ejercicios propuestos
4. Cualquier comentario o sugerencia que crea pertinente para mejorar las futuras ediciones de la práctica.