

PROBLEMES 2: Jerarquia de memòries

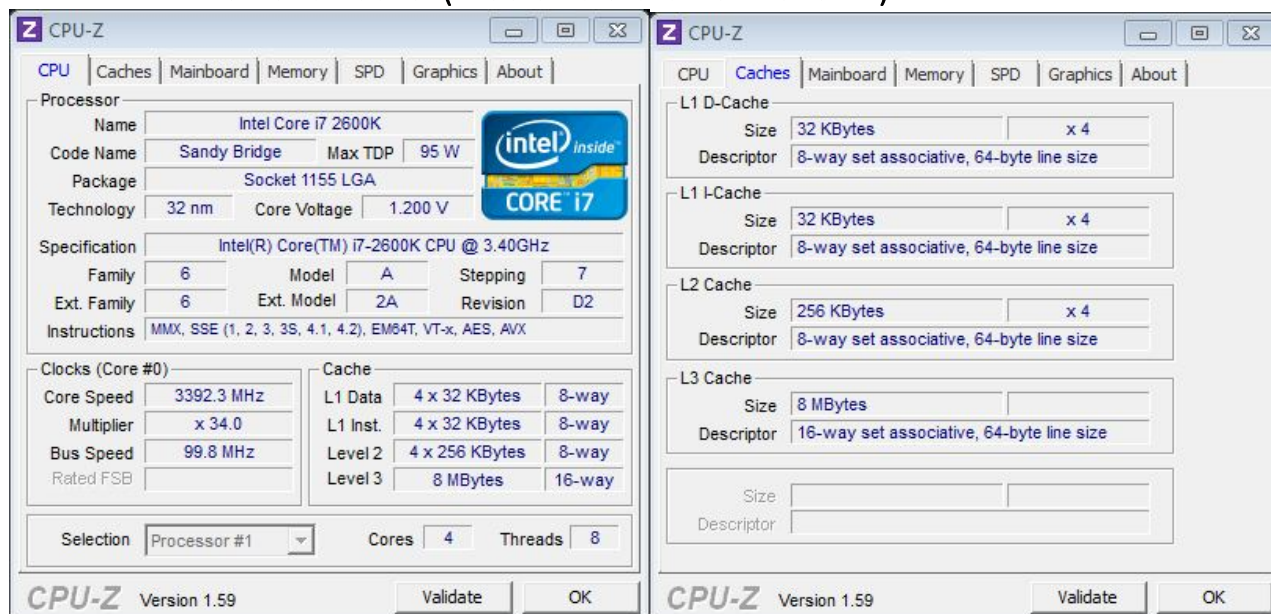
2.1. Caché (Problema d'examen 2020-21)

Considerem el següent “loop”, treballem amb un vector d'enters on N és una gran potència de 2. Cada enter té una mida de 4 bytes. El “loop” l'executem en una CPU de 32 bits. La memòria caché comença buida i la funció “process()” no introdueix cap pressió a la memòria caché (això vol dir que no afecta ni en miss ni en hit):

```
for (int j=0; j < 30; j++) {  
    for (int i=0; i < N; i += 1) {  
        process(arr[i]);  
    }  
}
```

- a) Suposem que tenim una memòria caché completament associativa de mida $4N$ B, mida de bloc de 4 B i que l'algoritme de reemplaçament és LRU (1 punt):
 - i. Quin és el nombre de hits ? (0.4 punts)
 - ii. Quin és el nombre de misses ? (0.4 punts)
 - iii. De quin tipus de localitat treu profit aquesta caché (tria una) ? (0.15 punts)
 - Quasi-balístic
 - Temporal
 - Cap
 - Espacial
 - iv. La resposta anterior canviaria si la caché fos del tipus associativa a 2 camins (mantenint la mateixa mida total) ? (0.05 punts)
- b) Suposem que tenim una memòria caché completament associativa de mida $2N$ B, mida de bloc de 4 B i que l'algoritme de reemplaçament és LRU (1 punt):
 - i. Quin és el nombre de hits ? (0.4 punts)
 - ii. Quin és el nombre de misses ? (0.4 punts)
 - iii. De quin tipus de localitat treu profit aquesta caché (tria una) ? (0.15 punts)
 - Quasi-balístic
 - Temporal
 - Cap
 - Espacial
 - iv. La resposta anterior canviaria si la caché fos del tipus associativa a 2 camins (mantenint la mateixa mida total) ? (0.05 punts)
- c) Suposem que tenim una memòria caché completament associativa de mida $2N$ B, mida de bloc de 8 B i que l'algoritme de reemplaçament és LRU (1 punt):
 - i. Quin és el nombre de hits ? (0.4 punts)
 - ii. Quin és el nombre de misses ? (0.4 punts)
 - iii. De quin tipus de localitat treu profit aquesta caché (tria una) ? (0.15 punts)
 - Quasi-balístic
 - Temporal
 - Cap
 - Espacial
 - iv. La resposta anterior canviaria si la caché fos del tipus associativa a 2 camins (mantenint la mateixa mida total) ? (0.05 punts)

2.2. Tenim un ordinador amb les següents característiques de CPU i Memòria Caché: (Problema d'examen 2016-17)



A. Respecte al sistema de Caché, ompliu la següent taula: (0.6 punts)

Longitud Paraula del Processador (bits/Bytes)	64 bits
Bits totals de l'adreça	40 bits
CARACTERÍSTIQUES CACHÉ L1 DE DADES	
Nombre de Línies (blocs) C	
Nombre de Camins K	8
Nombre de Conjunts S	
Bits necessaris pel TAG	
Bits necessaris per definir el Conjunt (i)	
Bits necessaris per definir la Paraula (W)	
Bits necessaris per definir el Byte (B)	
CARACTERÍSTIQUES CACHÉ L2	
Nombre de Línies (blocs) C	
Nombre de Camins K	8
Nombre de Conjunts S	
Bits necessaris pel TAG	
Bits necessaris per definir el Conjunt (i)	
Bits necessaris per definir la Paraula (W)	
Bits necessaris per definir el Byte (B)	
CARACTERÍSTIQUES CACHÉ L3	
Nombre de Línies (blocs) C	
Nombre de Camins K	16
Nombre de Conjunts S	
Bits necessaris pel TAG	
Bits necessaris per definir el Conjunt (i)	
Bits necessaris per definir la Paraula (W)	
Bits necessaris per definir el Byte (B)	

- B. A la caché de nivell 2 del problema anterior, indica a quin conjunt pertany cada una de les dades amb les següents adreces de memòria principal:

Adreça a Memòria Principal (hexadecimal)	Conjunt
0000000000h	
0000FF0030h	
0000000FC0h	
FFFFFFFFFFFh	
FFFFFFF804Fh	
FFFFFFF8050h	
000FFF3080h	
0000FF3380h	
0000FF0F00h	

2.3. El nostre sistema (processador de 32 bits+ memòria cau + memòria principal) utilitza adreces de memòria de 32 bits i té una memòria principal de mida: 1 GiB. Té una memòria cau de 4 KiB organitzada en la manera associativa a k camins, amb 4 blocs per conjunt i 64 bytes per bloc.

- A. Calcula el nombre de bits en cada un dels camps d'etiqueta, “set” (conjunt) i paraula de l'adreça de la memòria.
- B. Supposeu que la memòria cau està inicialment buida. Supposem que el processador obté 1088 paraules cadascuna de les ubicacions de la paraula successiva a partir de la ubicació 0. Després repeteix aquesta seqüència nou vegades més. Si la memòria cau és 10 vegades més ràpida que la memòria principal, calculeu el factor de millora derivat de l'ús de la memòria cau. Assumeixi que l'algoritme LRU s'utilitza per a la substitució de blocs

2.4. Considerem un processador acompanyat d'un sistema de memòria “caché” i la memòria principal. Com que la memòria principal és molt lenta, suposarem que el Miss penalty del sistema equival al temps d'accés a la memòria principal:

- A. Model simple de “caché”, només tenim 1 nivell : Quan és el temps d'accés a la “caché” si el temps mig d'accés a dades, vist des del processador, són 5 ns; el hit rate és d'un 99% i sabent que el temps d'accés a la memòria caché és 1/400 el temps d'accés a la memòria principal ?
- B. Model avançat, sistema “caché” de dos nivells: Si al sistema anterior li afegim un segon nivell de “caché” amb un hit rate del 99%, quin temps d'accés necessitem que tingui la segona “caché” per a que el temps mig d'accés a dades, vist des del processador, siguin 2 ns ?

2.5. Suposem que una computadora té un processador amb dues caches L1, una per a instruccions i una per a dades, i una memòria cau L2. Sigui τ el temps d'accés per a les dues caches L1. Les penalitzacions són aproximadament 15τ per transferir un bloc de L2 a L1 i 100τ per transferir un bloc des de la memòria principal a L2. A l'efecte d'aquest problema, suposem que les taxes d'èxit són iguals per a les instruccions i les dades i que les taxes d'èxit de les memòries L1 i L2 són de 0,96 i 0,80, respectivament.

- A. Quina fracció dels accessos es perden tant a la memòria cau L1 com a la L2 i, per tant, requereixen accés a la memòria principal?
- B. Quin és el temps d'accés mitjà vist pel processador?
- C. Suposem que la memòria cau L2 té un “hit rate” ideal de 1. Per quin factor es reduirà el temps d'accés mitjà a la memòria tal com el processador veu?
- D. Tingueu en compte el següent canvi a la jerarquia de memòria. S'elimina la memòria cau L2 i s'incrementa la mida de les memòries L1 de manera que es redueix la meitat el “miss rate”. Quin és el temps d'accés mitjà de la memòria vist pel processador en aquest cas?

2.6. La memòria d'un ordinador és adreçable per bytes, i la longitud de la paraula és de 32 bits. Un programa consisteix en dos bucles nidats (nested-loops): un petit bucle interior i un bucle exterior molt més gran.

L'estructura general del programa es mostra a la figura 1. Les adreces de memòria decimal mostren la localització dels dos bucles i el principi i final del programa total.

Totes les ubicacions de memòria a les diverses seccions del programa, 8-52, 56-136, 140-240, etc., contenen instruccions per executar-les en seqüències. El programa s'executarà en una computadora que tingui una memòria cau d'instruccions organitzada amb mapejat directe amb els següents paràmetres:

- Mida de la memòria cau 1KiB
- Mida del bloc de 128 bytes

El “miss penalty” en el caché d'instruccions és 80τ , on τ és el temps d'accés de la memòria cau. Calculeu el temps total necessari per a obtenir instruccions durant l'execució del programa a la Figura 1. Supposeu que, després de portar un bloc de MP, triguem τ en portar la dada de la caché.

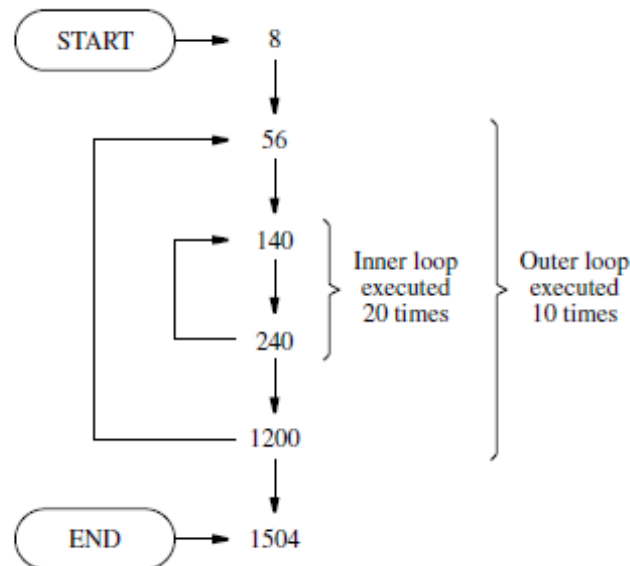
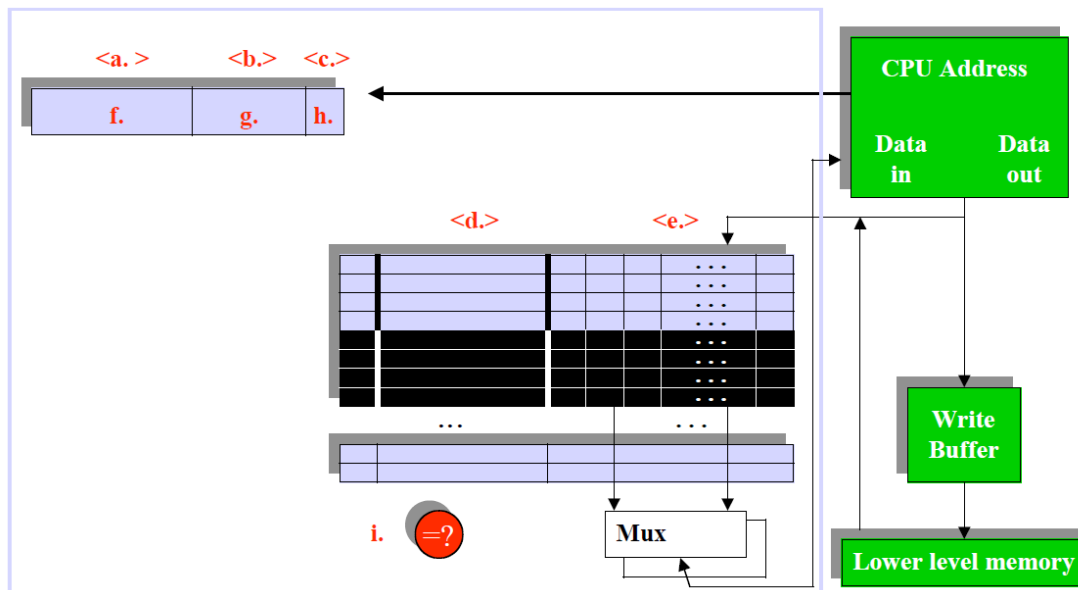


Figura 1.

2.7. Contesteu les següents preguntes si el nostre sistema té les següents propietats:

- A. Dissenyeu un sistema de memòria amb una memòria cau amb les següents propietats:
- Les paraules (per cada dada) són de 32 bits cadascuna
 - Un bloc de memòria cau contindrà 2048 bits de dades
 - La memòria cau té mapejat directe
 - L'adreça subministrada des de la CPU és de 32 bits
 - Hi ha 2048 blocs a la memòria cau
 - Les adreces indiquen la paraula que s'agafa de MP

A continuació es mostra l'estructura general d'una caché. Hi ha 8 camps (etiquetats com a, b, c, d, e, f, g i h). A l'espai inferior, heu d'indicar el nom o el nombre correcte de bits per a una part determinada d'aquesta configuració de la memòria cau.



- 1) Quines 3 coses s'han de comparar? per determinar si l'entrada de la memòria cau es pot utilitzar o no?
 - 2) Quina és la mida total de la memòria cau?
- B. Ara, considerem què passa si fem la nostra memòria cau del tipus associada a 2 camins. Les dades donades a l'apartat A que encara són vàlides són:
- Les paraules (per cada dada) són de 32 bits cadascuna
 - Un bloc de memòria cau contindrà 2048 bits de dades
 - L'adreça subministrada des de la CPU és de 32 bits
 - Hi ha 2048 blocs a la memòria cau
 - Les adreces indiquen la paraula que s'agafa de MP
- a) Nombre de bits en W
 - b) Nombre de bits per conjunt
 - c) Nombre de bits en el TAG
- C. Ara, considerem què passa si fem la nostra memòria cau del tipus associada a 4 camins. Les dades donades a l'apartat A que encara són vàlides són:
- Les paraules (per cada dada) són de 32 bits cadascuna
 - Un bloc de memòria cau contindrà 2048 bits de dades
 - L'adreça subministrada des de la CPU és de 32 bits
 - Hi ha 2048 blocs a la memòria cau
 - Les adreces indiquen la paraula que s'agafa de MP
- a) Nombre de bits en W
 - b) Nombre de bits per conjunt
 - c) Nombre de bits en el TAG
- D. Ara, considerem què passa si les dades a la MP són "byte adreçables" (cada direcció indica un byte). La memòria cau continua sent associativa a 4 camins. Les dades donades a l'apartat A que encara són vàlides són:
- Les paraules (per cada dada) són de 32 bits cadascuna
 - Un bloc de memòria cau contindrà 2048 bits de dades

- L'adreça subministrada des de la CPU és de 32 bits
- Hi ha 2048 blocs a la memòria cau

- Nombre de bits en W
- Nombre de bits per conjunt
- Nombre de bits en el TAG

E. Ara, considerem què passa si la mida de la nostra adreça física canvia de 32 bits a 64 bits. Les adreces permeten accedir a un byte de la MP. La memòria cau continua sent associativa a 4 camins. Les dades donades a l'apartat A que encara són vàlides són:

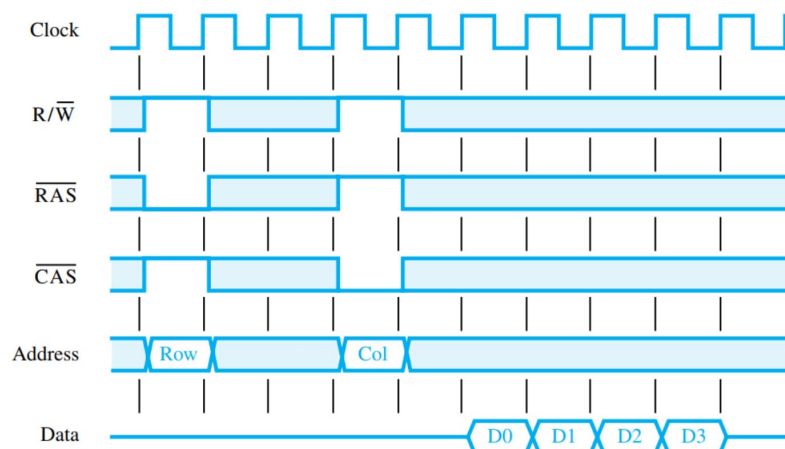
- Un bloc de memòria cau contindrà 2048 bits de dades
- L'adreça subministrada des de la CPU és de 32 bits
- Hi ha 2048 blocs a la memòria cau

- Nombre de bits en W
- Nombre de bits per conjunt
- Nombre de bits en el TAG

2.8. Considereu la cel·la de memòria dinàmica (DRAM). Suposem que $C = 30$ femtofarads (10^{-15} F) i que el corrent de fugues a través del transistor és de 0,25 picoampers (10^{-12} A). El voltatge a través del condensador quan està carregat completament és de 1,5 V. La cel·la s'ha de "refrescar" abans que aquesta tensió baixi de 0.9 V. Estimeu la freqüència mínima d'actualització.

2.9. Considereu una memòria principal construïda amb xips SDRAM. Les dades es transfereixen en ràfegues tal com es mostra a la figura, llevat que la longitud de ràfega és de 8. Suposem que 32 bits de dades es transfereixen en paral·lel. Si s'utilitza un rellotge de 400 MHz, quant de temps triga a transferir:

- 32 bytes de dades
- 64 bytes de dades
- Quina és la latència en cada cas?



- 2.10. Descriu una estructura similar a la de la figura 4 per a una memòria $8M \times 32$ utilitzant xips de memòria de $512K \times 8$.
- 2.11. Descriu una estructura similar a la de la figura 4 per a una memòria de $16M \times 32$ amb $1M \times 4$ xips de memòria.
- 2.12. Critiqueu la següent afirmació: "L'ús d'un xip del processador més ràpid produeix un augment corresponent del rendiment d'una computadora, fins i tot si la velocitat de la memòria principal segueix sent la mateixa".

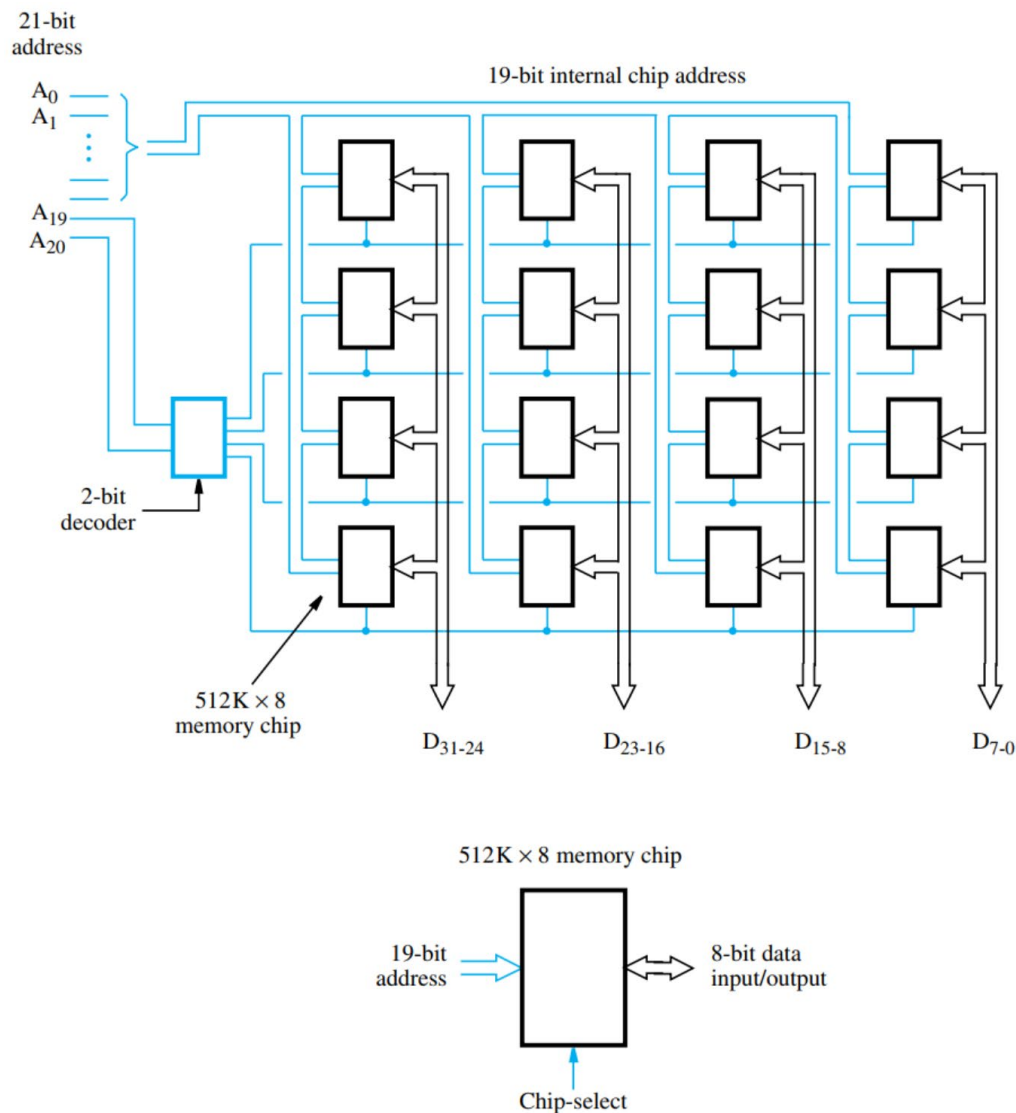
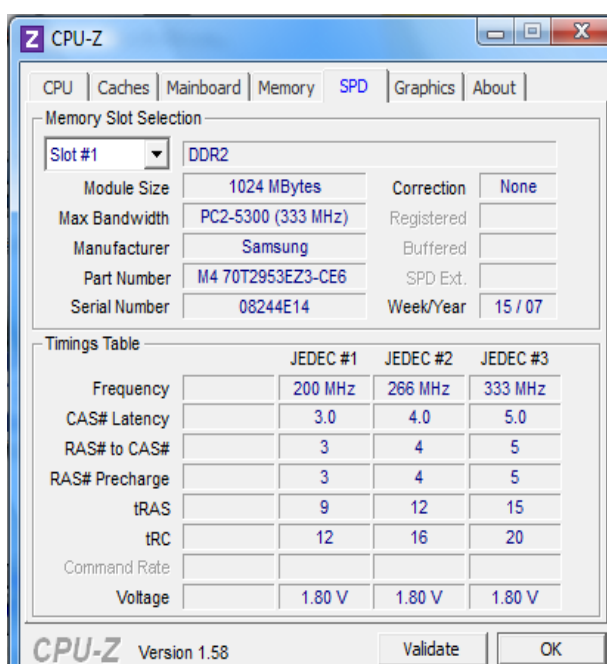
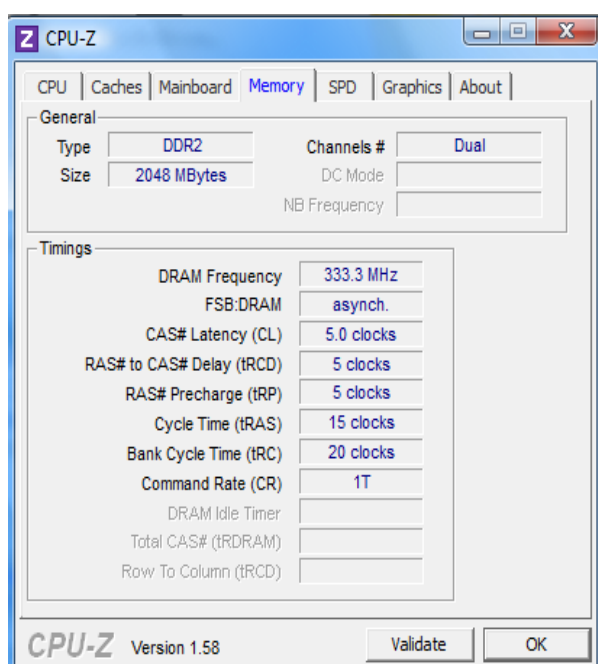


Figura 4

2.13. Tenim un ordinador amb les següents característiques de Memòria Principal, veure imatge: (Problema d'examen)

A. Ompliu la següent taula:

Quin tipus de RAM fa servir?	
Quantitat total de memòria principal?	
Número de mòduls de memòria?	
Freqüència real del BUS?	
Ample banda teòric de cada mòdul de M.P.?	
Ample de banda teòric del sistema M.P.?	
Aquest processador fa servir un sistema de comunicació "Single Channel", "Dual Channel" o "Triple Channel"?	
Quants cicles de bus es necessiten per fer la transferència de 8 Bytes?	
Quants cicles de bus es necessiten per fer la transferència de 16 Bytes contigus?	
Quants cicles de bus es necessiten per fer la transferència de 64 Bytes contigus?	
Podem afegir un altre mòdul de memòria que treballi a 3 V ?	



- B. Quina diferència hi ha entre un sistema Single Channel i un Dual, Triple o Quad Channel ? Dibuixa un esquema
- C. Enumera el problemes que podem trobar en memòries tipus DRAM. Hi ha alguna solució per minimitzar-los ?

2.14. Tenim un ordinador amb les següents característiques de Memòria Principal: (Problema d'examen)

General Memory Information:

- Type: DDR3
- Size: 8072 MBytes
- Channel #: Dual

Timings:

- DRAM Frequency: 798.7 MHz
- FSB:DRAM: 1:6
- CAS# Latency (CL): 11.0 clocks
- RAS# to CAS# Delay (tRCD): 11 clocks
- RAS# Precharge (tRP): 11 clocks
- Cycle Time (tRAS): 28 clocks
- Bank Cycle Time (tRC):
- Command Rate (CR): 1T
- DRAM Idle Timer:
- Total CAS# (tRDRAM):
- Row To Column (tRCD):

Memory Slot Selection: Slot #1

Module Information:

- Module Size: 4096 MBytes
- Max Bandwidth: PC3-12800 (800 MHz)
- Manufacturer: Hyundai Electronics
- Part Number: HMT351S6CFR8C-PB
- Serial Number: 1A13AAF9

JEDEC Timings Table:

	JEDEC #4	JEDEC #5	JEDEC #6	JEDEC #7
Frequency	609 MHz	685 MHz	761 MHz	838 MHz
CAS# Latency	8.0	9.0	10.0	11.0
RAS# to CAS#	8	9	10	11
RAS# Precharge	8	9	10	11
tRAS	22	24	27	30
tRC	30	33	37	41
Command Rate				
Voltage	1.50 V	1.50 V	1.50 V	1.50 V

Tenint en compte que al Slot #2 hi ha un mòdul exactament igual al Slot #1. Ompliu la taula:

Quin tipus de RAM fa servir?	
Quantitat total de memòria principal?	
Número de mòduls de memòria?	
Freqüència real del BUS?	
Ample banda teòric de cada mòdul de M.P.?	
Ample de banda teòric del sistema M.P.?	
Aquest processador fa servir un sistema de comunicació "Single Channel", "Dual Channel" o "Triple Channel"?	
Quants cicles de bus es necessiten per fer la transferència de 8 Bytes?	
Quants cicles de bus es necessiten per fer la transferència de 16 Bytes contigus?	
Quants cicles de bus es necessiten per fer la transferència de 64 Bytes contigus?	
Podem afegir un altre mòdul de memòria que treballi a 3 V ?	

2.15. Considereu una llarga sèrie d'accessos a un disc amb un temps mitjà de cerca de 6 ms i un retard de rotació mitjà de 3 ms. La mida mitjana d'un bloc al què s'accedeix és de 8K bytes. La velocitat de transferència de dades del disc és de 34 Mbytes / segon.

- A. Suposant que els blocs de dades estan ubicats a l'atzar al disc, calculeu el percentatge mitjà del temps total ocupat per operacions de cerca i retards de rotació.
- B. Repetiu la part (a) per la situació en què es disposen els accessos al disc, de manera que en el 90 per cent dels casos, el següent accés serà a un bloc de dades del mateix cilindre.