Fonaments dels Computadors

Arquitectura del Processador (2/3)

Grau en Intel·ligència Artificial

Xavier Martorell

Xavi Verdú

Facultat d'Informàtica de Barcelona (FIB)

Universitat Politècnica de Catalunya (UPC)

Creative Commons License

Aquest document utilitza Creative Commons Attribution 3.0 Unported License

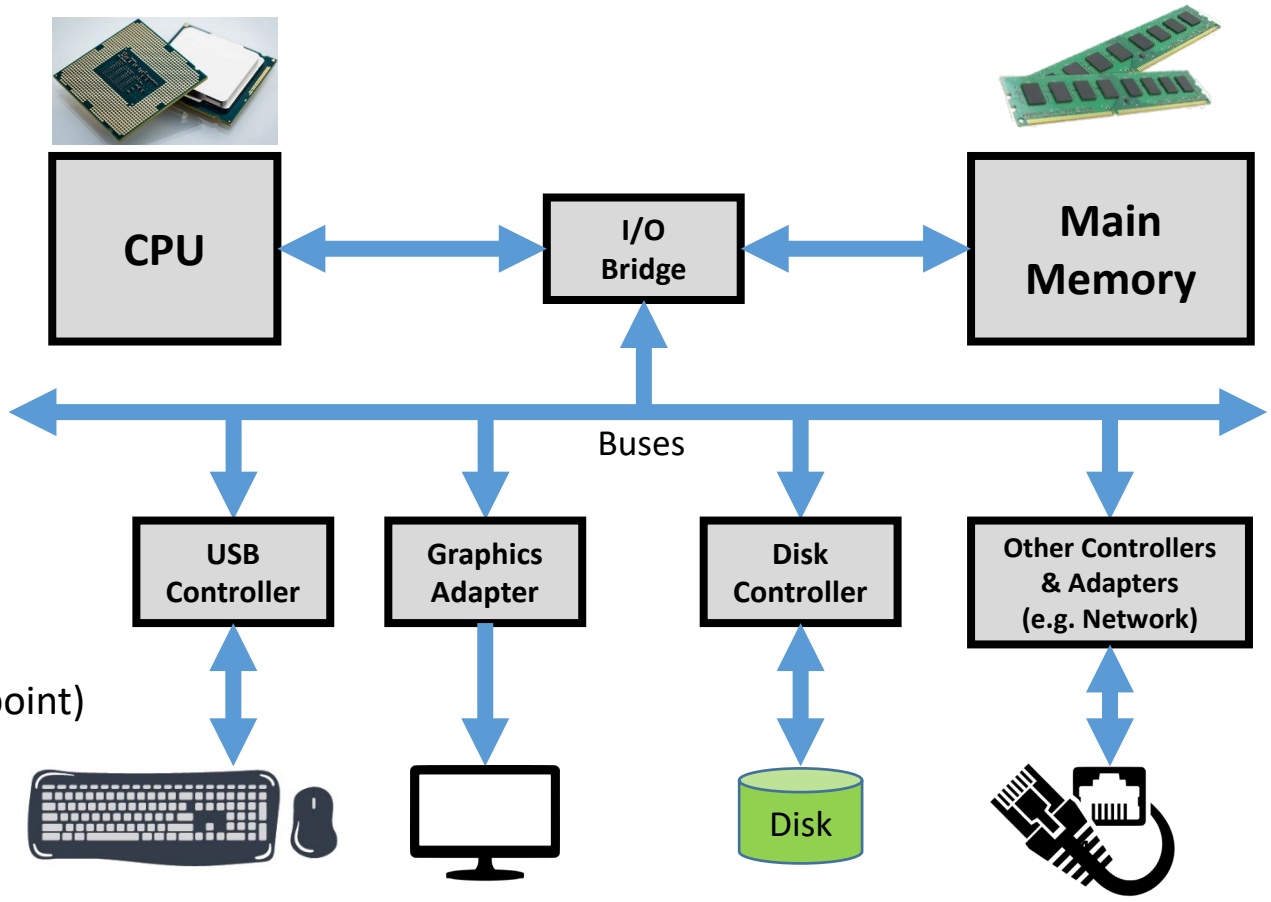
Els detalls d'aquesta licencia es poden trovar a https://creativecommons.org/licenses/by-nc-nd/4.0

Index

- Aquest tema té tres parts
  1) Conceptes Bàsics
  2) **Jerarquia de Memòria**
  3) Fluxe d'execució, rendiment, colls d'ampolla i consum d'energia

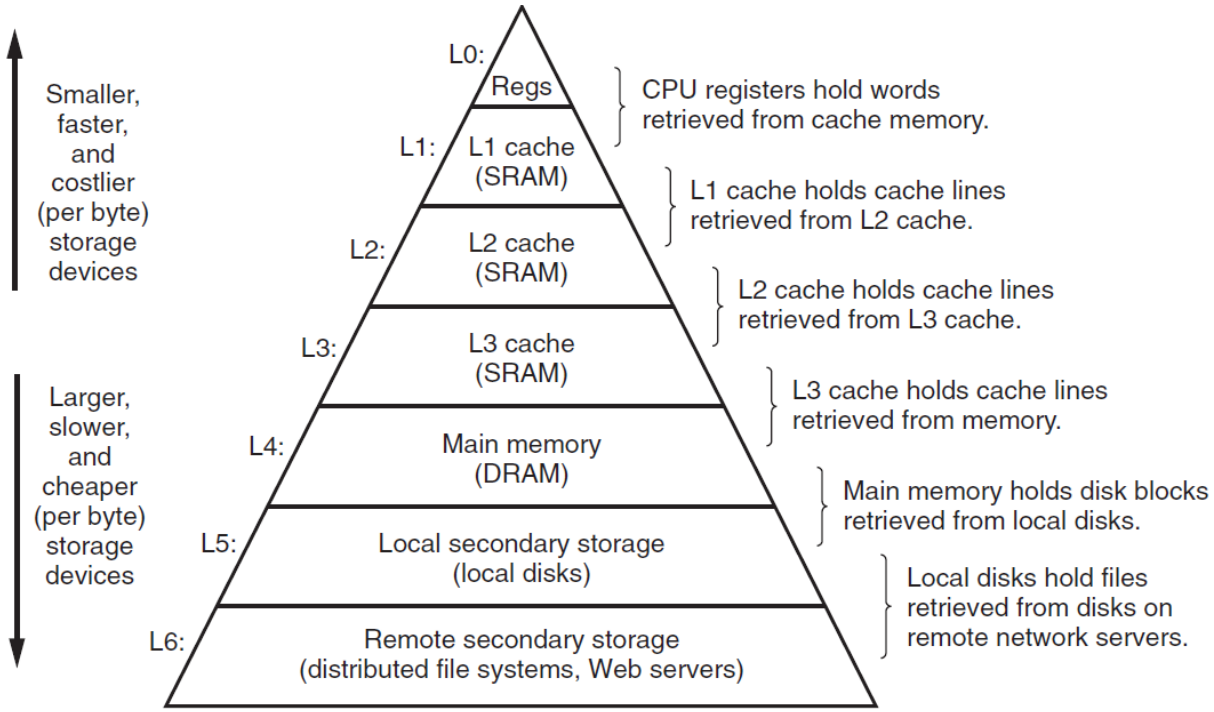
Bases...

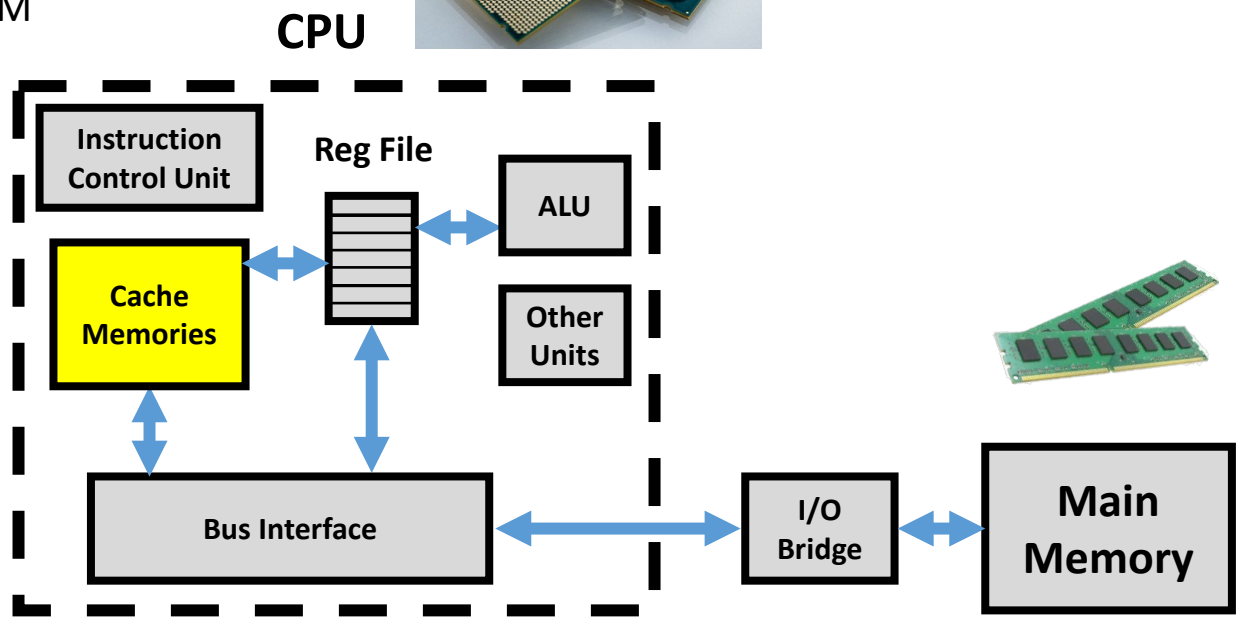
- La CPU només pot accedir a
  - Registres
  - Memòria
- No pot accedir directament a altres
  - Instrs i dades s'han de carregar en memòria per poder ser
    - Arquitectura Von Neumann
- Quan un programa es llença …
  - Es reserva memòria
  - Carrega l'executable des del disc a memòria
  - La CPU començ a executar la primera instr (i.e. entry point)



Jerarquia de Memòria

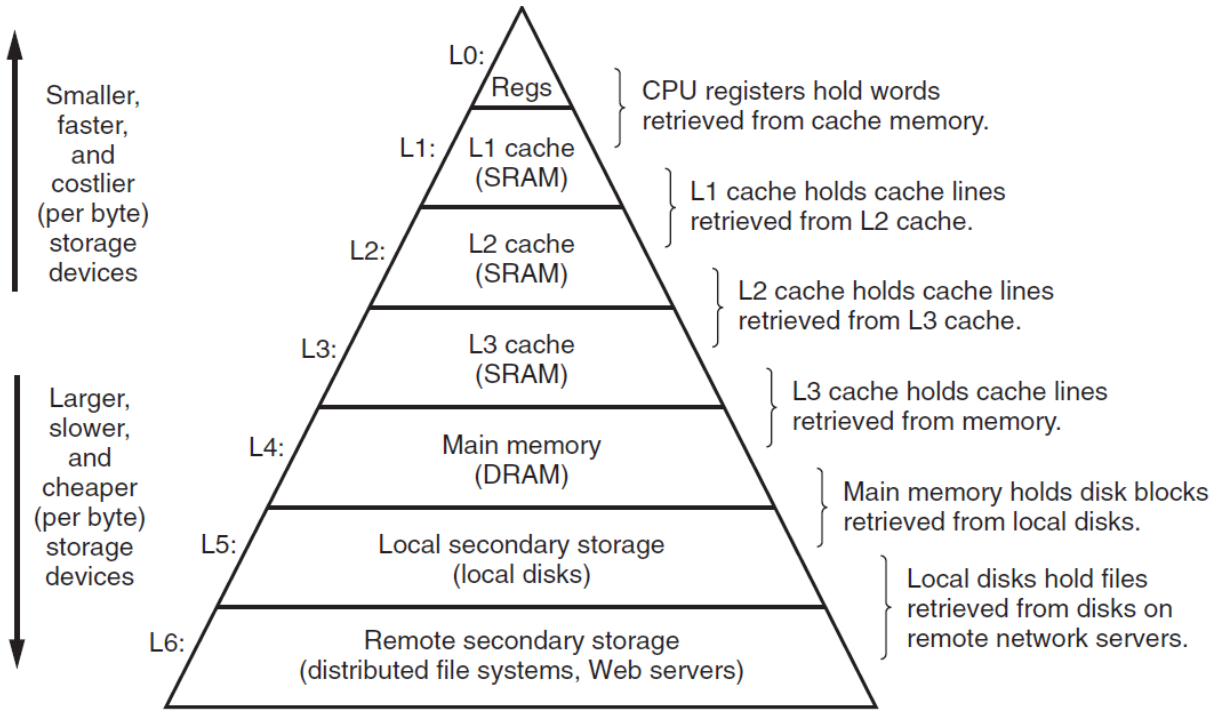
- **RAM**: L'unitat básica adreçable de storage és una supercel·la (1 Byte) composta per 8 cel·les (1 bit cada cel·la)
  - Static RAM (SRAM) per la **jerarquia de cache**
  - Variants de Dynamic (DRAM) per la **memòria principal**
    - E.g. SDRAM, DDR(2, 3, 4) SDRAM
  - SRAM és més ràpida (~10x) i més cara (~100x) que DRAM



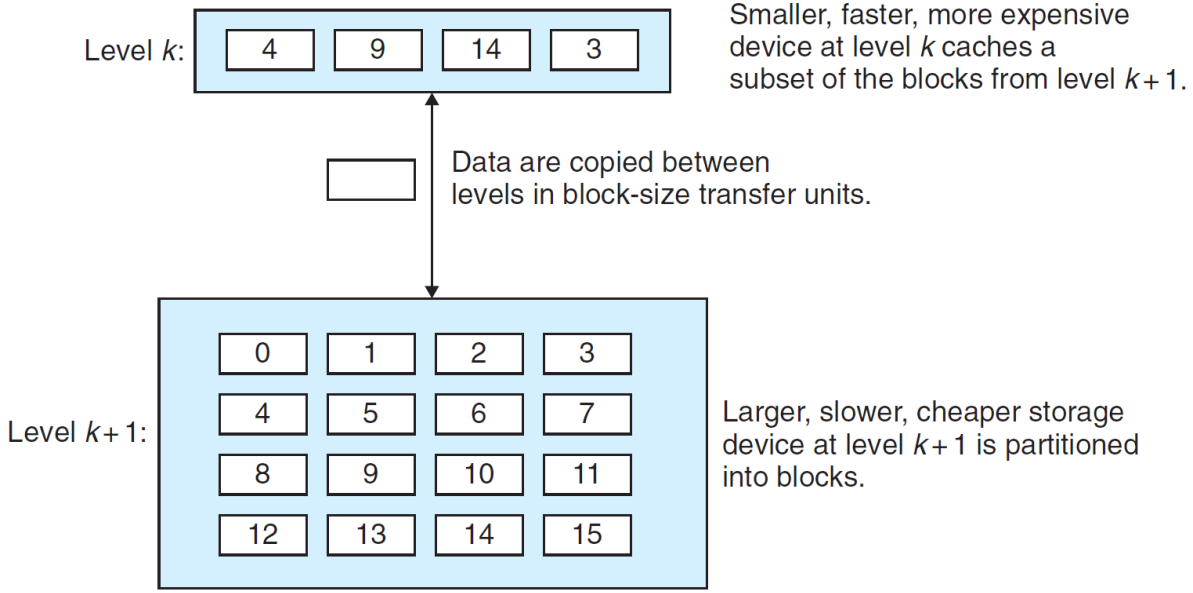


Jerarquia de Memòria

- **RAM**: L'unitat básica adreçable de storage és una supercel·la (1 Byte) composta per 8 cel·les (1 bit cada cel·la)
  - Static RAM (SRAM) per la **jerarquia de cache**
  - Variants de Dynamic (DRAM) per la **memòria principal**
    - E.g. SDRAM, DDR(2, 3, 4) SDRAM
  - SRAM és més ràpida (~10x) i més cara (~100x) que DRAM



- **Caching:** utilitza un dispositiu d'emmagatzamatge més petit, però més ràpid per portar elements guardats en memòries més grans i lentes
  - Normalment els processadors van des de L0 (registres) fins a L3
  - Explota la **Localitat…**



Localitat

- El principi de localitat és la tendencia a accedir elements que estan propers a altres en l'espai o temps
  - **Localitat Espacial**: des d'una adreça de memòria en concret, posteriorment s'accedeix a adreces en posicions properes
  - **Localitat Temporal**: des d'una adreça de memòria en concret, posteriorment s'accedeix a la mateixa adreça en un curt periode de temps

- Si hi ha una alta localitat espacial…portar no només els bytes sol·licitats, sinó també els de les posicions adjacents

- Si hi ha una alta localitat temporal…mantener els bytes accedits en nivells alts de la jerarquia de memoria (cache)

- Els diferents tipus de paral·lelisme també explota la localitat espacial
  - E.g. SIMD

Exemple de Localitat

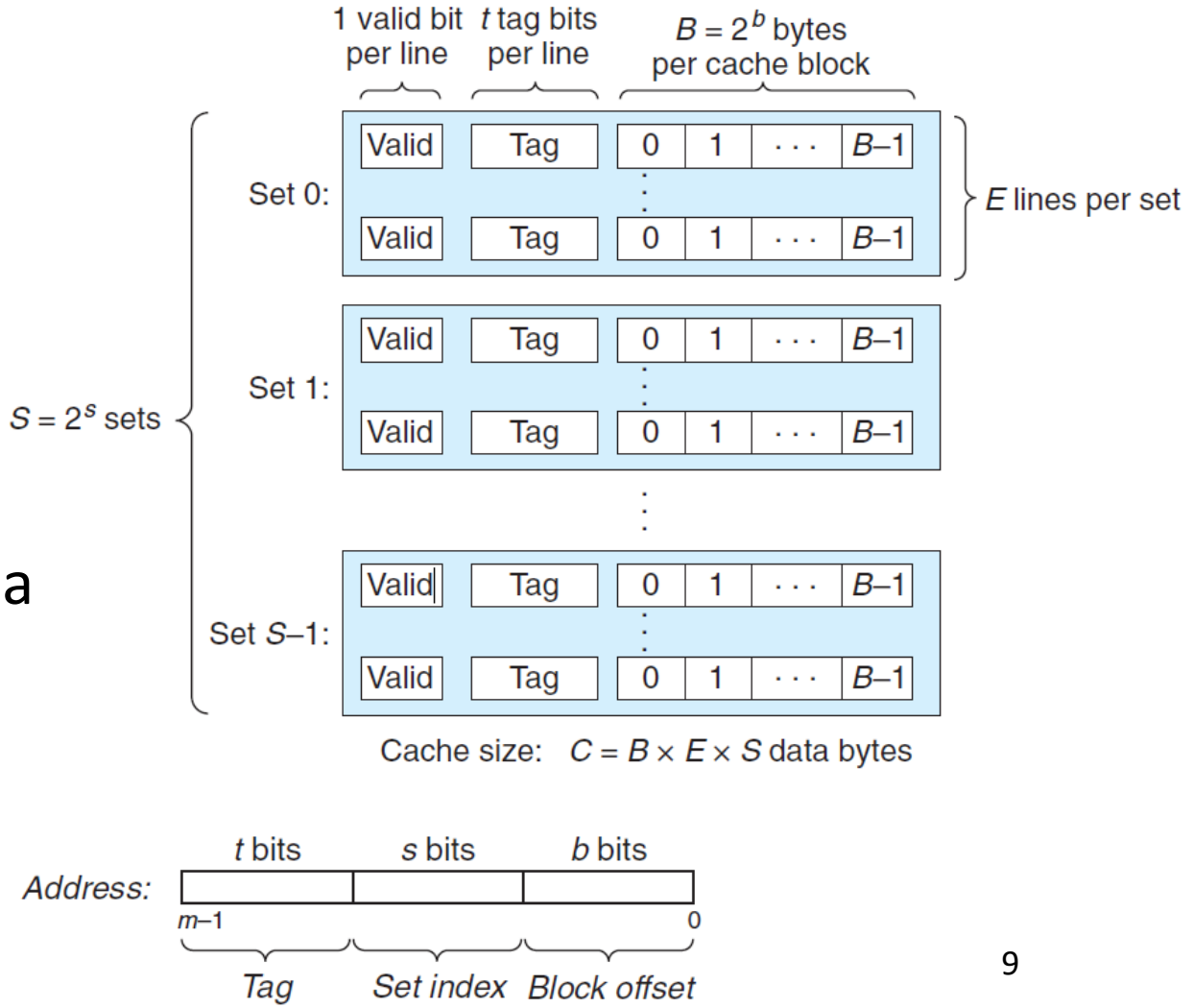
- Quins tipus de localitat identifiques en el següent fragment de codi...
  - ...en referències a instruccions?
  - …en referències a dades?

```
func (vec, n)
{
        tmp = 0;
        for (i=0; i<n; i++)
                tmp += vec[i];
        return tmp;
}
```

Organització de la cache

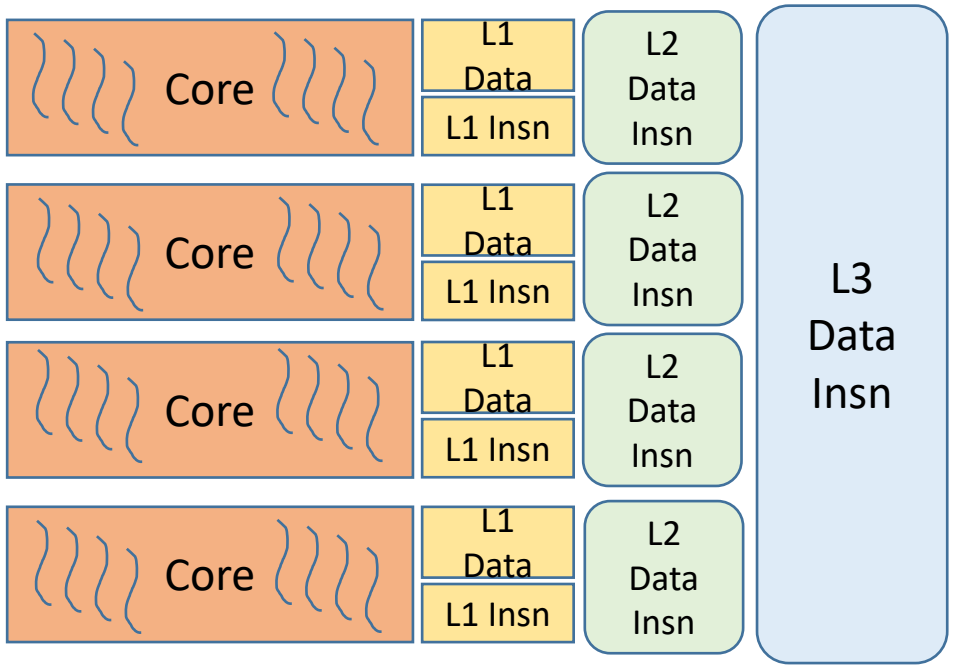
- Una cache consisteix en…
  - Un array de conjunts
  - Un conjunt (set) té una o més línies
  - Una línia té…
    - Un bit de validessa
    - Uns pocs bits per l'etiqueta (tag)
    - Varios bytes de dades

- Una adreça de memòria es decomposa
  - Una adreça de *m* bits
    - t bits pel tag
    - s bits pel set
    - b bits per seleccionar el Bloc (offset)



Jerarquia Cache

- **Instruction** cache, **data** cache i **unified** cache



Memòria Principal

Exemples

AMD

Private L1
Shared L2
Shared L3

Intel Core i7

Private L1
Private L2
Shared L3

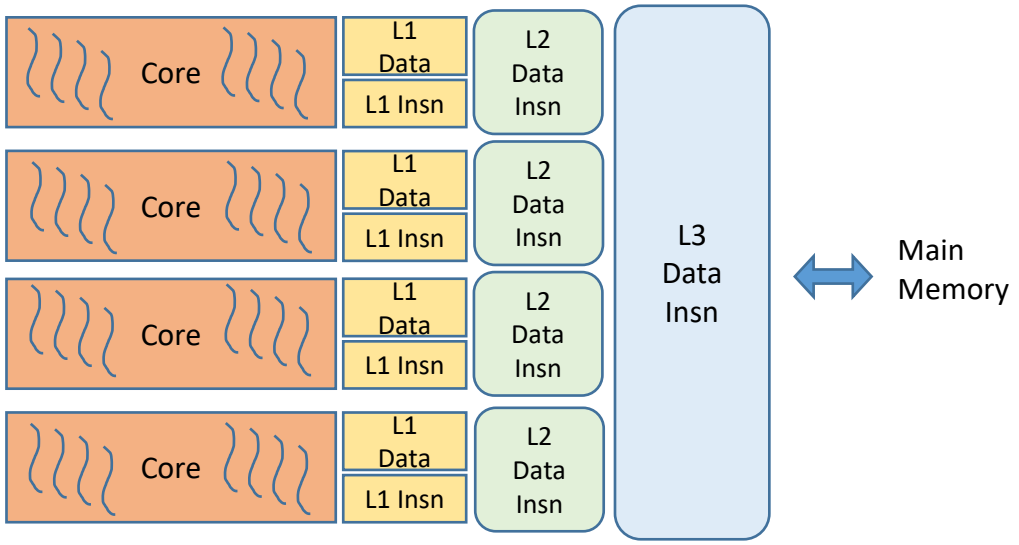
Cache type	Access time (cycles)	Cache size (C)	Assoc. (E)	Block size (B)	Sets (S)
L1 i-cache	4	32 KB	8	64 B	64
L1 d-cache	4	32 KB	8	64 B	64
L2 unified cache	10	256 KB	8	64 B	512
L3 unified cache	40–75	8 MB	16	64 B	8,192

Estratègies d'escriptura

- Què passa si hi ha un encert?
  - Write-through
    - Tant aviat com hi ha una escriptura en el nivell i-èssim de cache, s'envia la modificació al nivell i+1
    - Senzill d'implementar, però incrementa el tràfic en el bus
    - Es pot utilizar un bufer (write-buffer) per millorar el rendiment d'actualització de memòria
    - Les errades de lectura són menys costosos perquè no impliquen una escriptura
  - Write-back
    - Quan es fa una escriptura en el nivell i-èsim de cache, NO s'envia al nivell i+1. S'endarrereix tant com espugui
    - Menys transferències → per tant hi haurà més amplada de banda disponible per dispositius d'E/S
- Què passa si hi ha una errada?
  - Write-no-allocate (a.k.a. write around)
    - Les dades no es carreguen en la cache, sinó que s'escriuen directament al següent nivell de cache
    - Les dades només es carreguen quan hi han errades en lectures
  - Write-allocate (a.k.a. fetch on write)
    - Les dades es carreguen en la cache, seguides d'una operación d'escriptura que encert
    - Intenta explotar la localitat espacial d'escriptures, però incrementa el tràfic
- Normalment…
  - …nivells més baixos (e.g. L1) de cache són write-through; nivells més alts (e.g. L2) són write-back
  - …quan s'utilitza write-through també s'utilitza write-no-allocate; write-back també va amb write-allocate

Requisits: Consistència & Coherència

- Els accessos fan referència a posicions de memòria, no a posicions de la cache
  - Les memòries cache són gestionades de manera transparent pel hardware
  - **Coherència en Memòria**: qualsevol lectura des de qualsevol CPU a una posició en concret de memòria, retorna l'últim valor escrit en aquella posició
  - **Consistència en Memòria**: assegura que les escriptures a diferentes posicions de memòria es veuran en l'ordre correcte des de totes les CPUs
- Hi ha protocols (e.g. snoopy) per assegurar aquests requisits




```
void add_vectors( float * c, float * a, float * b )
{
  int i;
  for (i=0; i < N; i++) c[i] = a[i] + b[i];
}
```

Errada de Cache (Cache Miss)

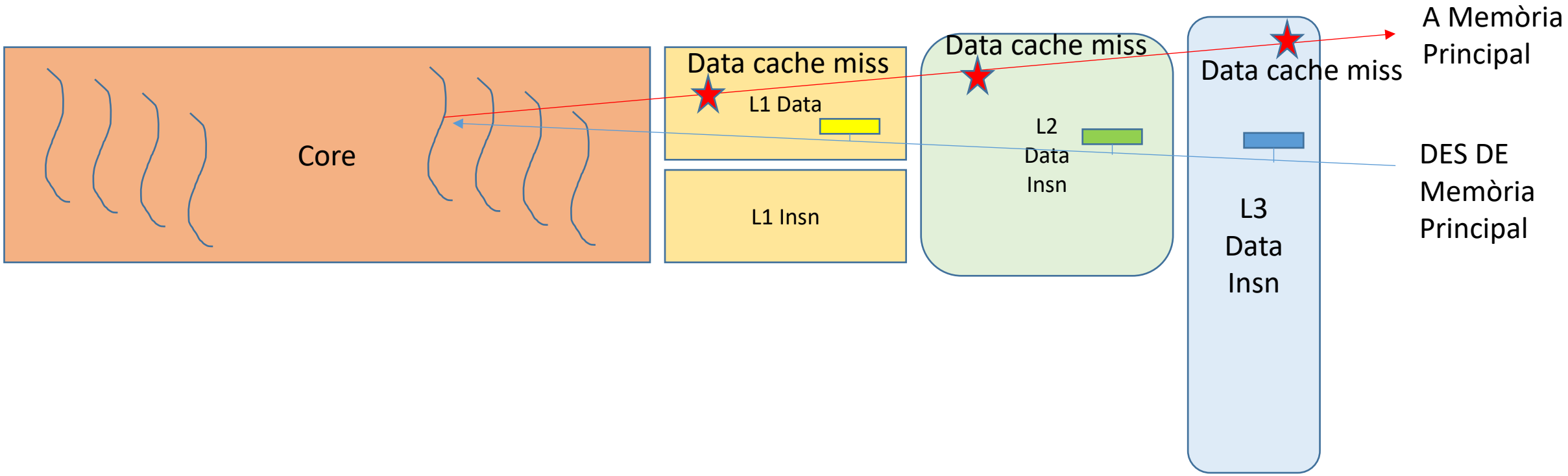
- Penalització per errada: temps addicional degut a una errada en el nivell-k de cache

- Tipus d'errades de Cache
  - **Cold or compulsory (obligatoria/inicial) miss**…
    - Passa quan la cache està buida
  - **Conflict (conflicte) miss**…
    - Succeix quan varis elements coincideixen en el mateix bloc
      - E.g. (i mod 4)…0, 4, 8, 12, 16 … tots aquests elements coincideixen en el bloc 0
  - **Capacity (capacitat) miss**…
    - Succeix quan el **conjunt de treball** (i.e. blocs actius de cache) és més gran que el nombre de blocs de cache

- Impacte directe de...
  - **Política d'emplaçament**: determina on va el bloc que es carrega
  - **Política de reemplaçament**: determina quin bloc es fa fora per carregar un de nou

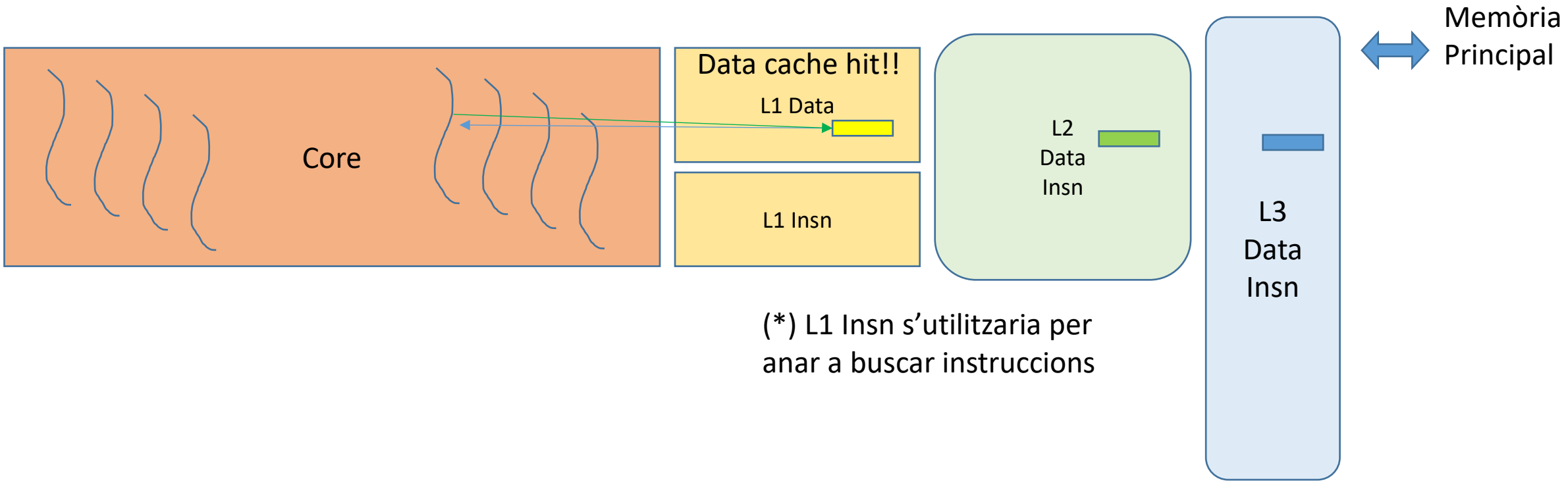
Accés Detallat a Memòria (MISS: errada)

- Instrucció LOAD: carregar dades que NO estan en les caches
  - Similar a anar a buscar instruccions que NO estan en les caches



Accés Detallat a Memòria (HIT: encert)

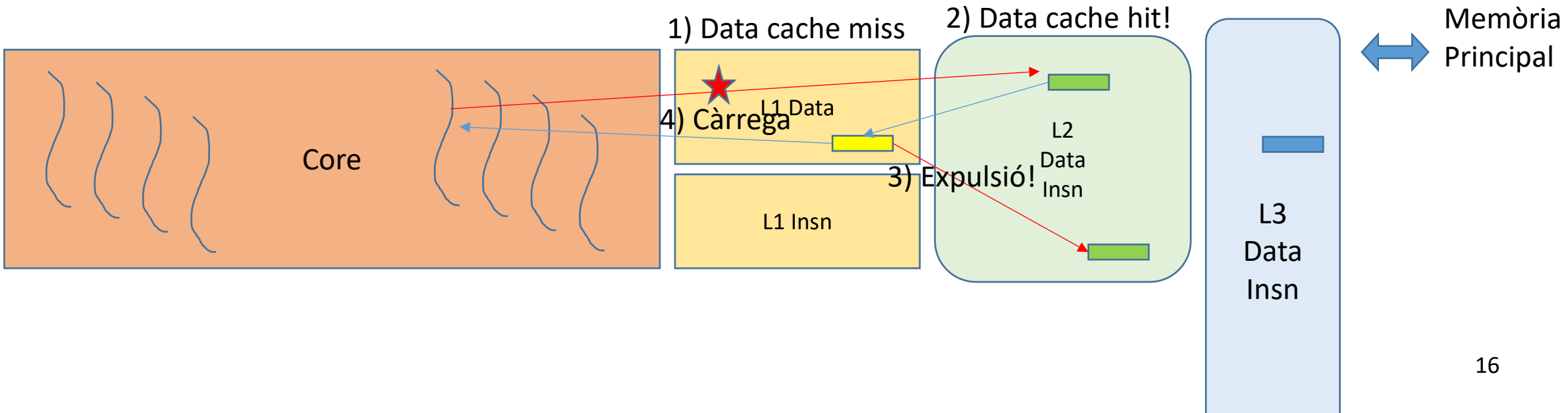
- Instrucció LOAD: carregar dades que estan en la L1 Dades
  - Similar a anar a buscar instruccions que estan en la L1 Insn (*)



(*) L1 Insn s'utilitzaria per anar a buscar instruccions

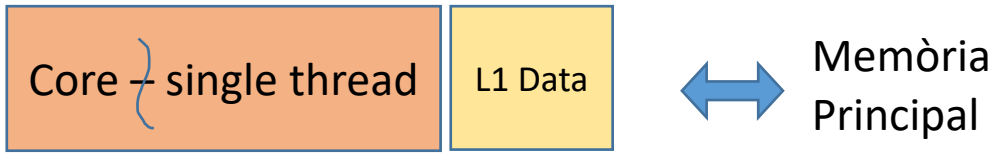
Accés Detallat a Memòria (Eviction: expulsió)

- La gestió de la cache és una característica Hw complexa
  - Què passa quan la cache està plena...
    i s'han d portar més dades/instrs?
    - Cache eviction... Els valors que fa més temps han sigut accedits poden ser expulsats al següent nivell de cache

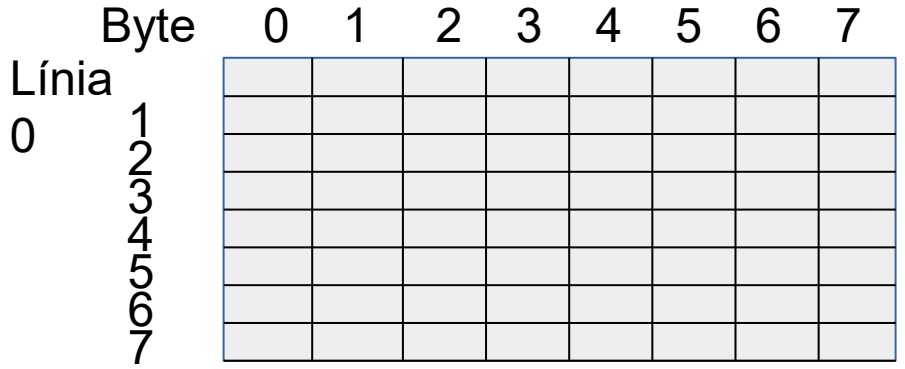


Alguns exemples

- Tenim un ordinador senzill amb la següent estructura interna:



  - Level 1 cache
    - 8 línies de cache
    - 8 bytes per línia de cache
    - 64 bytes en total



Exemple 1

- Executem el següent programa:

```
double V[16];       // using addresses to 0x100 to 0x17f
int i;              // using a CPU register

for (i=0; i < 16; i++)
    V[i] = 0.3;
```

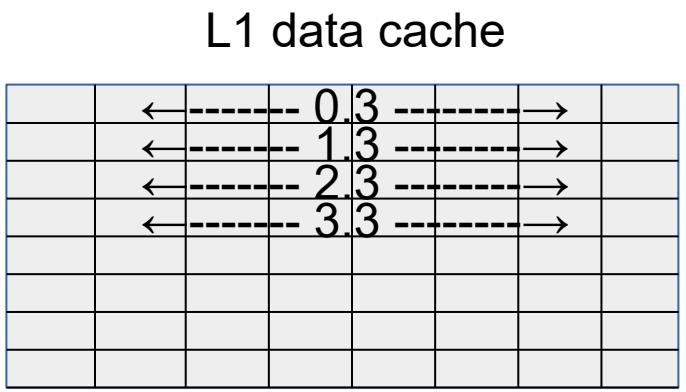
- Determina els ***hits*** i ***misses*** i ***errades de capacitat*** que succeeixen en la cache de dades L1
- Assumim que no hi ha interacció amb una possible cache d'instruccions L1

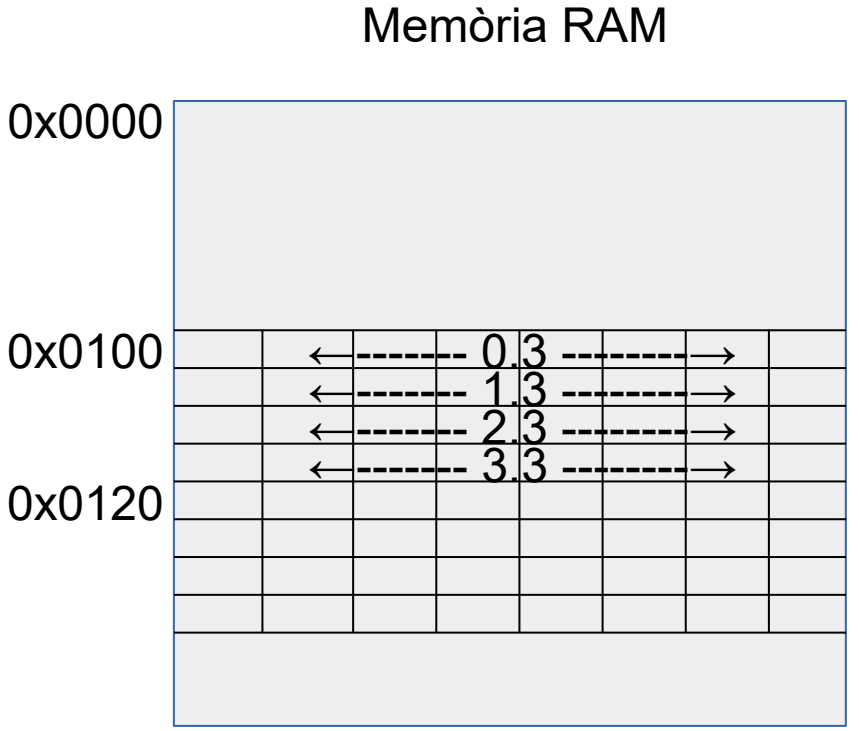
Exemple 1

- Executem el següent programa:

```
double V[16];      // using addresses 0x100 to 0x17f
int i;             // using a CPU register

for (i=0; i < 16; i++)
    V[i] = 0.3 + (double) i;
```

i = 0	–	–
V[0] = 0.3	0x0100	miss
i++	–	–
V[1] = 1.3	0x0108	miss
i++	–	–
V[2] = 2.3	0x0110	miss
i++	–	–
V[3] = 3.3	0x0118	miss
i++	–	–



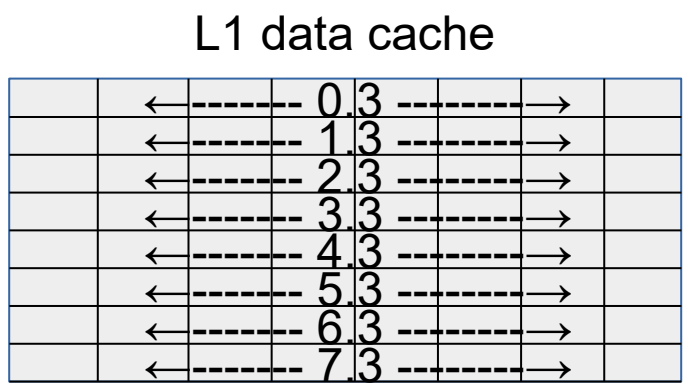


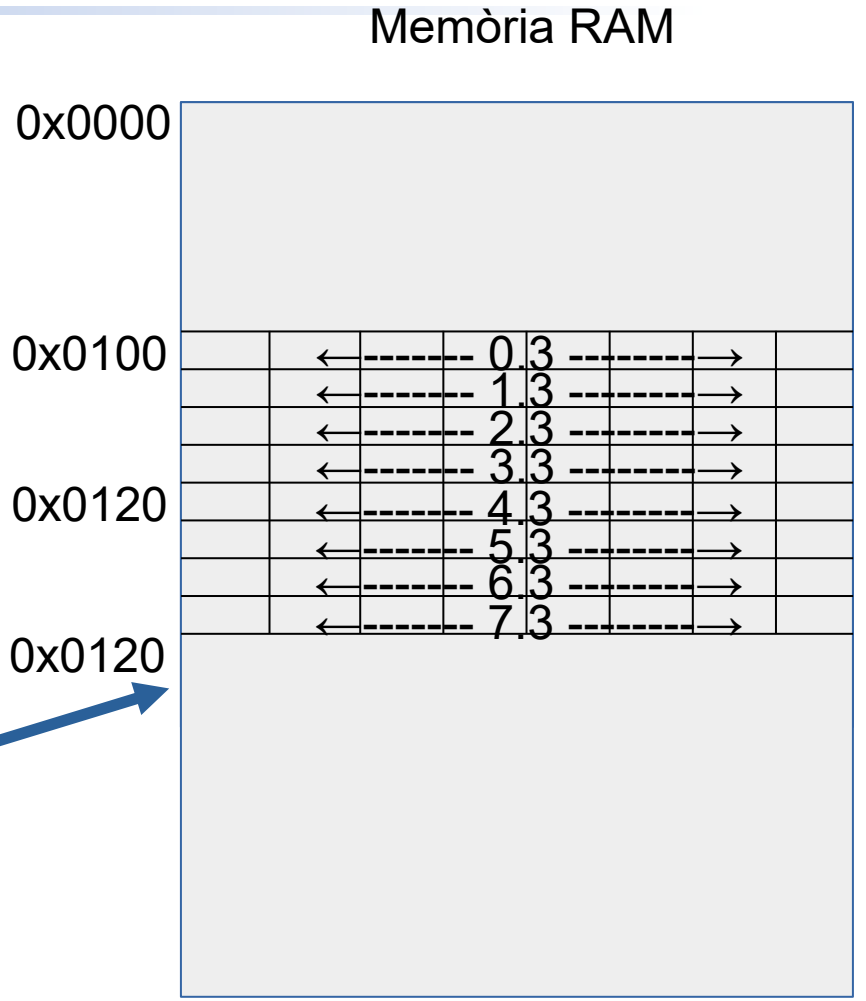
Exemple 1

- Executem el següent programa:

```
double V[16];      // using addresses 0x100 to 0x17f
int i;             // using a CPU register

for (i=0; i < 16; i++)
    V[i] = 0.3 + (double) i;
```

V[4] = 4.3	0x0120	miss
i++	–	–
V[5] = 5.3	0x0128	miss
i++	–	–
V[6] = 6.3	0x0130	miss
i++	–	–
V[7] = 7.3	0x0138	miss
i++	–	–
V[8] = 8.3	0x0140	???



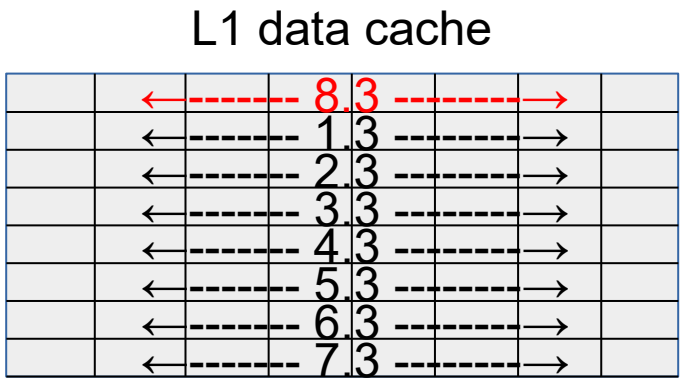


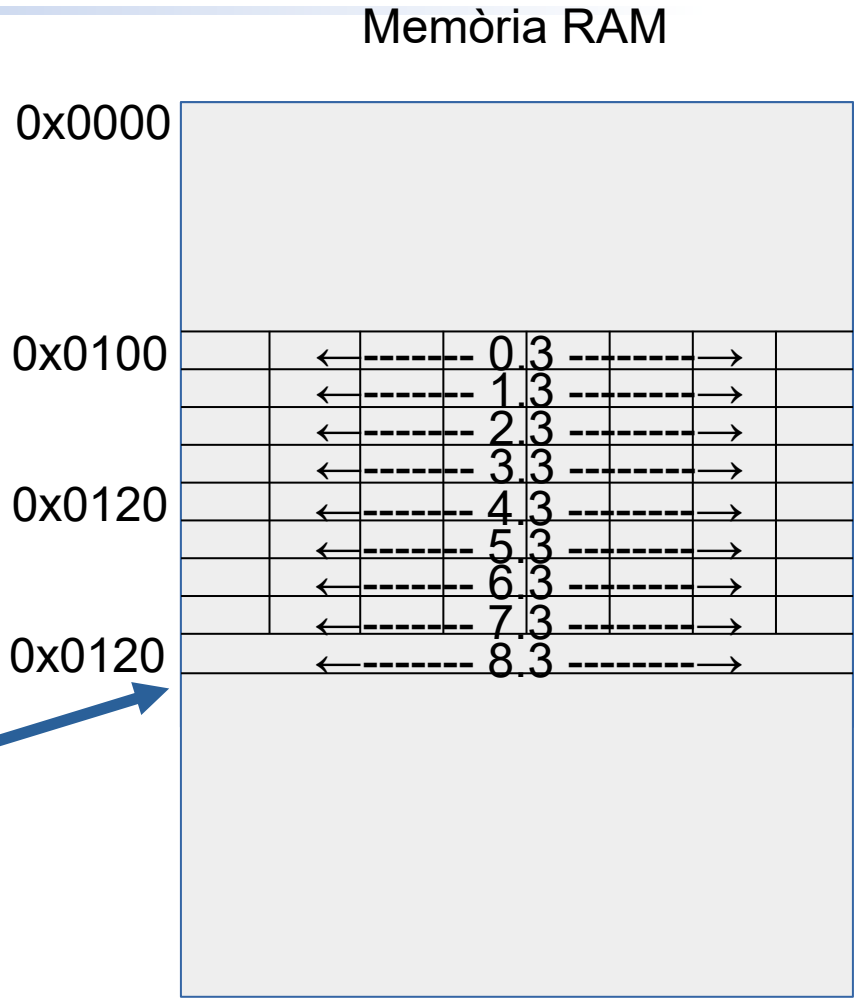
Exemple 1

- Executem el següent programa:

```
double V[16];       // using addresses 0x100 to 0x17f
int i;              // using a CPU register

for (i=0; i < 16; i++)
    V[i] = 0.3 + (double) i;
```

V[4] = 4.3	0x0120	miss
i++	–	–
V[5] = 5.3	0x0128	miss
i++	–	–
V[6] = 6.3	0x0130	miss
i++	–	–
V[7] = 7.3	0x0138	miss
i++	–	–
V[8] = 8.3	0x0140	miss & errada de capacitat



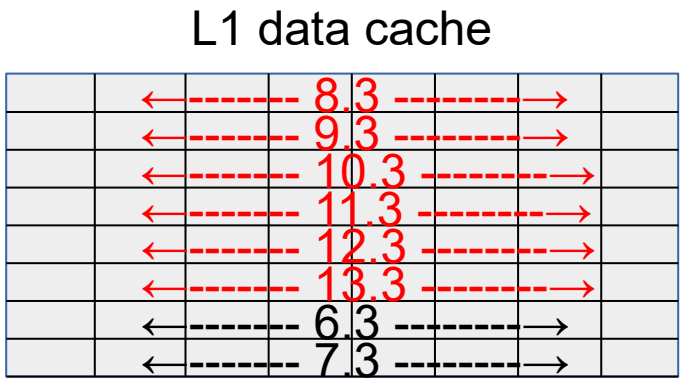


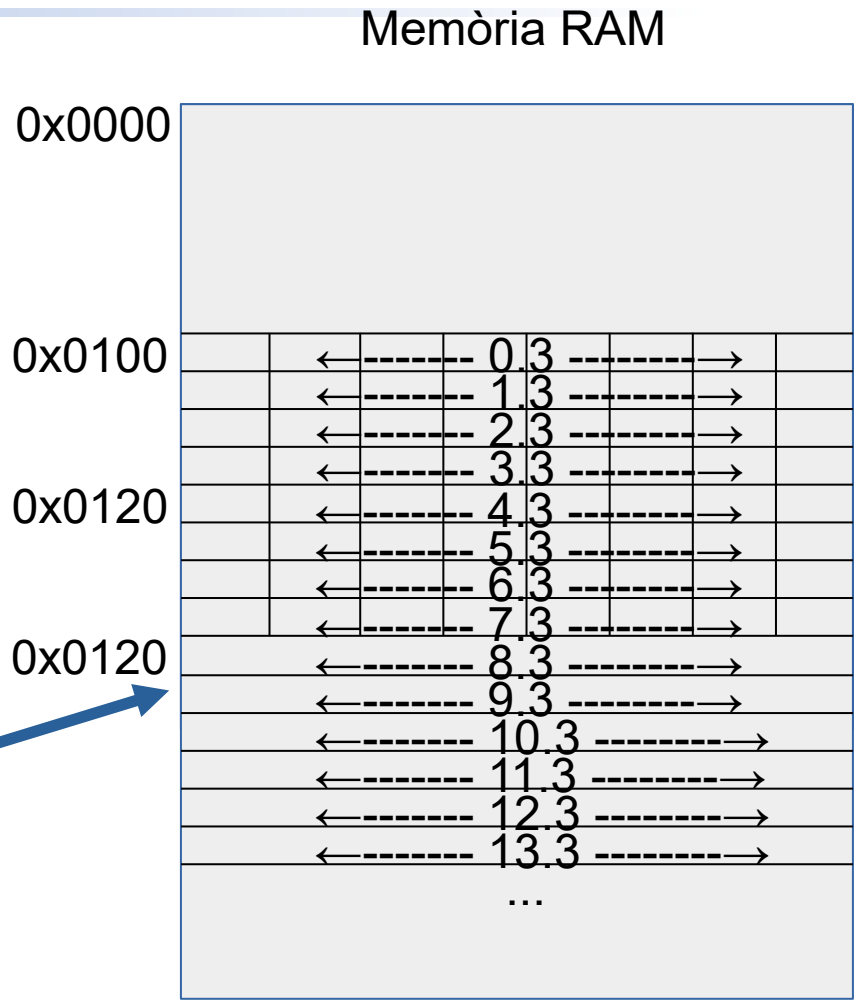
Exemple 1

- Executem el següent programa:

```
double V[16];       // using addresses 0x100 to 0x17f
int i;              // using a CPU register

for (i=0; i < 16; i++)
    V[i] = 0.3 + (double) i;
```

V[9] = 9.3	0x0148	miss & e.c.
i++	–	–
V[10] = 10.3	0x0150	miss & e.c.
i++	–	–
V[11] = 11.3	0x0158	miss & e.c.
i++	–	–
V[12] = 12.3	0x0160	miss & e.c.
i++	–	–
V[13] = 13.3	0x0168	miss & errada de capacitat
…		



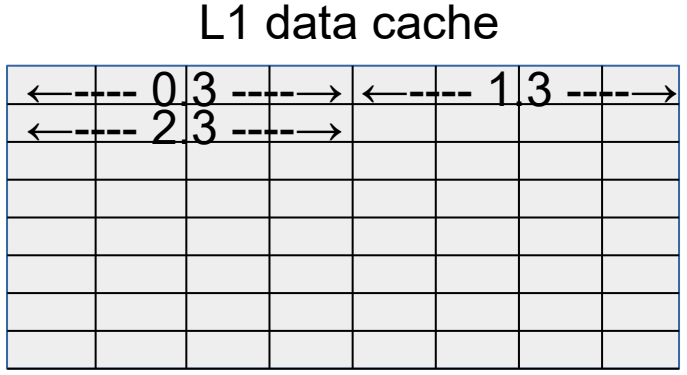


Exemple 2

- Ara canviem el programa per utilitzar «float» de precisió senzilla:

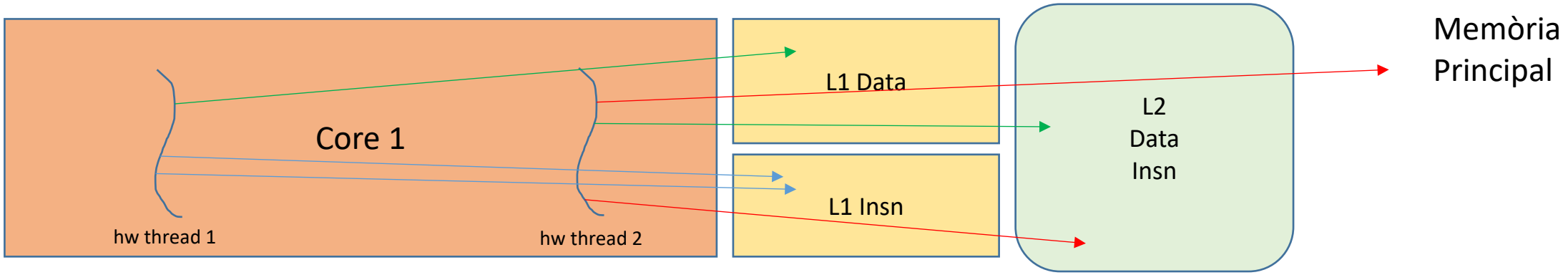
```
float V[32];          // using addresses 0x100 to 0x17f
int i;                // using a CPU register

for (i=0; i < 32; i++)
    V[i] = 0.3 + (float) i;
```

- Quin és el comportament del programa respecte a l'ús de la cache (hits, misses, errades de capacitat)?
- ...i respecte la memòria RAM?



Exemple 3

Assumim que l'arquitectura d'aquest processador executa un únic procés amb varis software threads



SW1 corrent en HW1
SW2 corrent en HW2

Exemple 3

- Shared L1, line cache size 128 (0x080) bytes
- Shared L2, line cache size 256 (0x100) bytes

Ordre d'execució	Core	HW Thread	Instr/dada	@ Memòria	L1 Insn	L1 Data	L2 Insn/Data
1	1	1	Instrucció	0x001000			
2	1	1	Dada	0x002000			
3	1	1	Instrucció	0x001008			
4	1	1	Dada	0x002080			
5	1	1	Instrucció	0x001010			
6	1	1	Dada	0x002100			
7	1	1	Instrucció	0x001018			
8	1	1	Instrucció	0x001020			
9	1	1	Data	0x002180			

Exemple 3

- Shared L1, line cache size 128 (0x080) bytes
- Shared L2, line cache size 256 (0x100) bytes

Ordre d'execució	Core	HW Thread	Instr/dada	@ Memòria	L1 Insn	L1 Data	L2 Insn/Data
10	1	2	Instrucció	0x001000			
11	1	2	Dada	0x002000			
12	1	2	Instrucció	0x001008			
13	1	2	Dada	0x002080			
14	1	2	Instrucció	0x001010			
15	1	2	Dada	0x002100			
16	1	2	Instrucció	0x001018			
17	1	2	Instrucció	0x001210			
18	1	2	Data	0x002280			

Exemple 3

- Shared L1, line cache size 128 (0x080) bytes
- Shared L2, line cache size 256 (0x100) bytes

Ordre d'execució	Core	HW Thread	Instr/dada	@ Memòria	L1 Insn	L1 Data	L2 Insn/Data
1	1	1	Instrucció	0x001000	Miss 0x1000	--	Miss 0x1000
2	1	1	Dada	0x002000	--	Miss 0x2000	Miss 0x2000
3	1	1	Instrucció	0x001008	Hit 0x1000	--	--
4	1	1	Dada	0x002080	--	Miss 0x2080	Hit 0x2000
5	1	1	Instrucció	0x001010	Hit 0x1000	--	--
6	1	1	Dada	0x002100	--	Miss 0x2100	Miss 0x2100
7	1	1	Instrucció	0x001018	Hit 0x1000	--	--
8	1	1	Instrucció	0x001020	Hit 0x1000	--	--
9	1	1	Data	0x002180	--	Miss 0x2180	Hit 0x2100

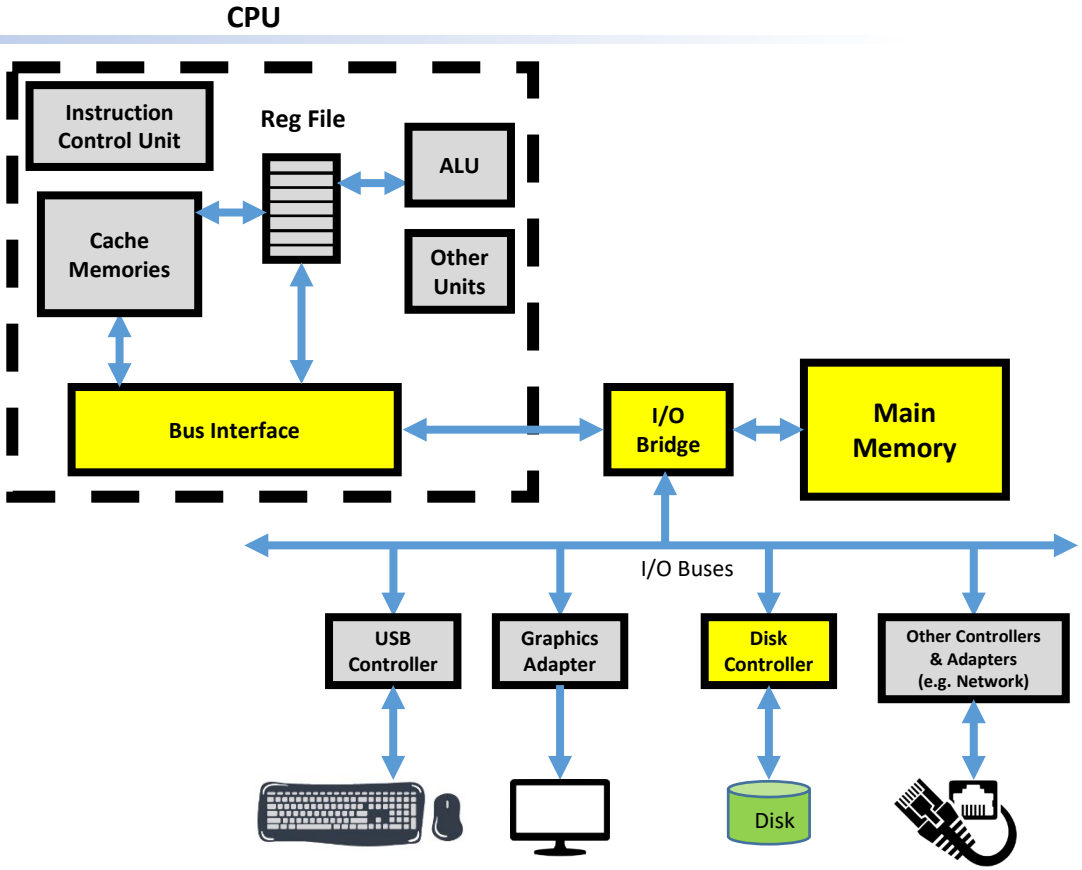
Exemple 3

- Shared L1, line cache size 128 (0x080) bytes
- Shared L2, line cache size 256 (0x100) bytes

Ordre d'execució	Core	HW Thread	Instr/dada	@ Memòria	L1 Insn	L1 Data	L2 Insn/Data
10	1	2	Instrucció	0x001000	Hit 0x1000	--	--
11	1	2	Dada	0x002000	--	Hit 0x2000	--
12	1	2	Instrucció	0x001008	Hit 0x1000	--	--
13	1	2	Dada	0x002080	--	Hit 0x2000	--
14	1	2	Instrucció	0x001010	Hit 0x1000	--	--
15	1	2	Dada	0x002100	--	Hit 0x2100	--
16	1	2	Instrucció	0x001018	Hit 0x1000	--	--
17	1	2	Instrucció	0x001210	Miss 0x1200	--	Miss 0x1200
18	1	2	Data	0x002280	--	Miss 0x2280	Miss 0x2200

Accés al Bus i E/S

- Un bus és una agrupació de cables
  - Porta adreces, dades i senyals de control

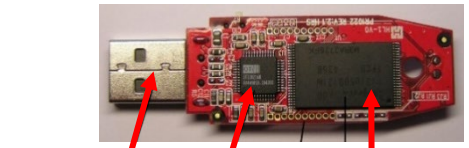


- La CPU pot comunicar-se amb dispositius d'E/S
  - Tècnica d'E/S: mapeig de memòria
    - Un dispositiu està "mapejat" a una adreça concreta de memoria (a.k.a. port d'E/S)

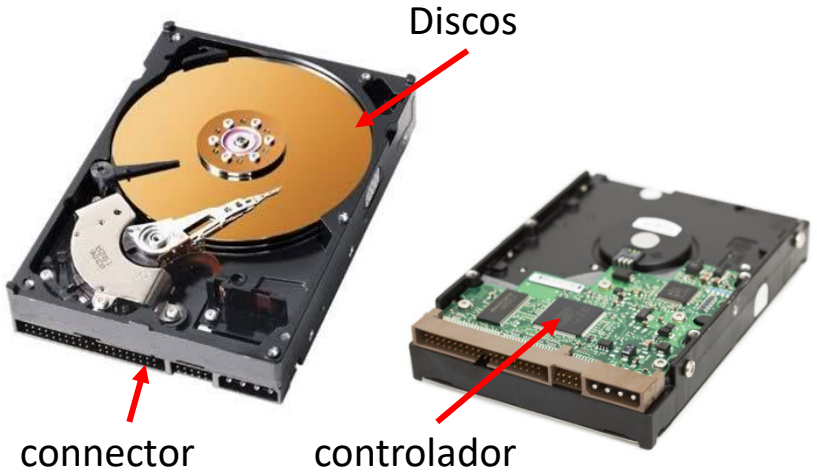
- DMA: Direct Memory Access
  - Els perifèrics poden llegir/escriure directament a/des d'adreces de memòria

Discos

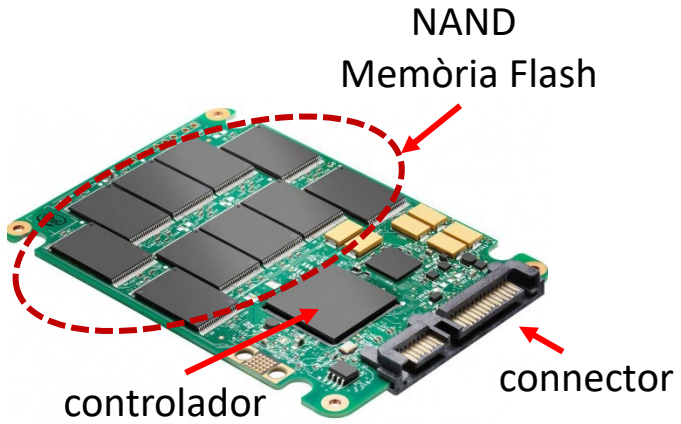
- Memòria no volàtil
  - Emmagatzema valors (dades/instrs) inclús quan no hi ha energia
    - Hi han altres memòries no volàtils a més a més del disc (e.g. ROM)
- Components similars vs diferents
  - Impacte en el rendiment i la capacitat



USB Drive



Hard Disk Drive



Solid State Drive (SSD)

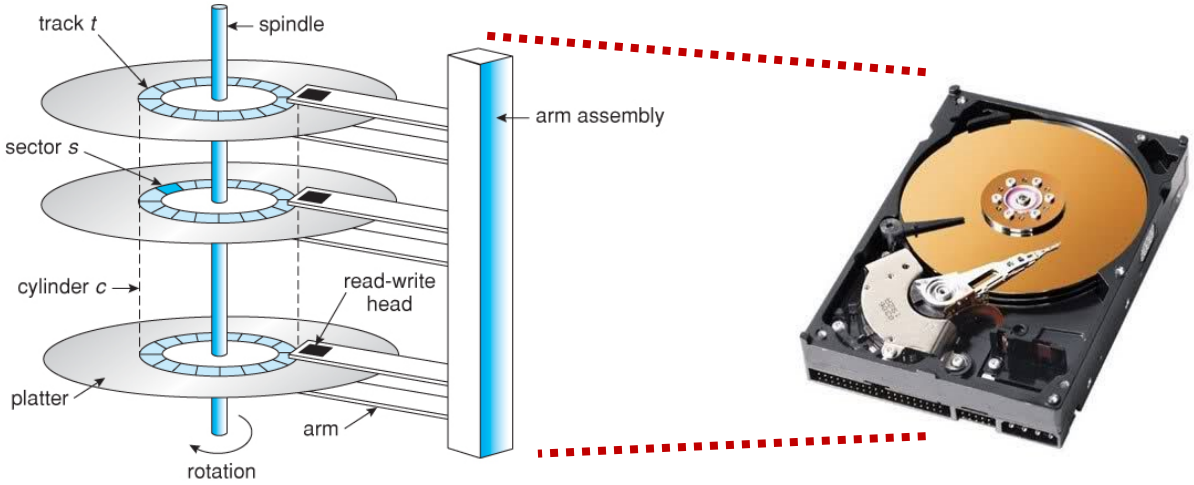
Com es guarden físicament els valors en els discos?

- Qualsevol dispositiu d'emmagatzematge ha d'organitzar la memòria
  - E.g.: DVD, hard-disk, pen-drive, etc.
- **Capacitat**: Nombre màxim de Bytes que pot guardar en un disc

$$\text{Capacity} = \frac{\#\text{ bytes}}{\text{sector}} \times \frac{\text{average }\#\text{ sectors}}{\text{track}} \times \frac{\#\text{ tracks}}{\text{surface}} \times \frac{\#\text{ surfaces}}{\text{platter}} \times \frac{\#\text{ platters}}{\text{disk}}$$

- **Sector**: L'unitat més petita d'emmagatzematge que pot ser llegida/escrita
  - **Definit per hardware**
  - Mida fixa (normalment 512 Bytes)

Alguns paràmetres impacten en el rendiment

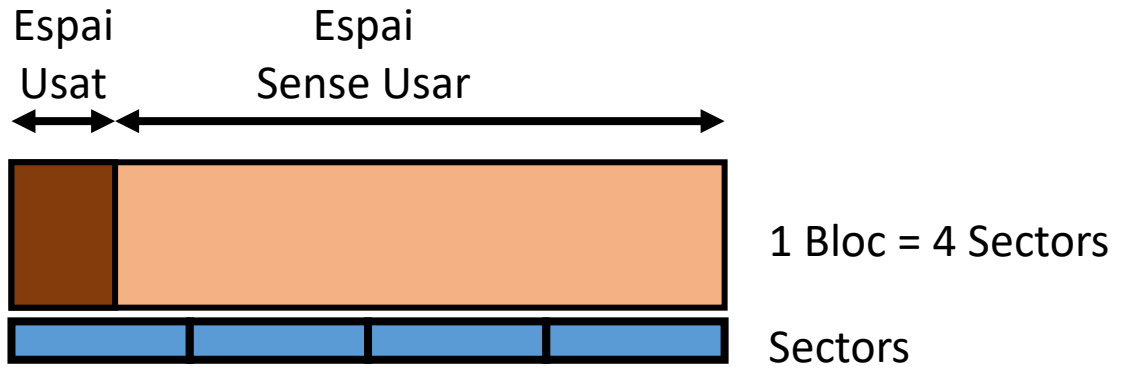


Velocitat: Revolucions Per Min.
e.g. 5400 RPM-15000 RPM
Max Amplada de Banda: Bytes Per Sec.
e.g. pocs MB/s – cents MB/s
(Mbps **no és el mateix que** MBps)

Com s'organitza un dispositiu de storage?

- **Bloc**: Un grup de sectors (l'unitat més petita per assignar memòria)
  - **Definit pel SO (quan es dona format al dispositiu)**

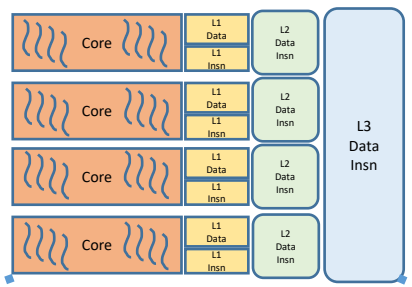
- Però, quin és la millor mida de bloc????
  - Si és probable usar fitxers grans…
    - Blocs grans
  - Si és probable usar fitxers petits…
    - Blocs petits

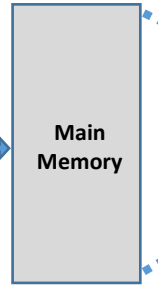


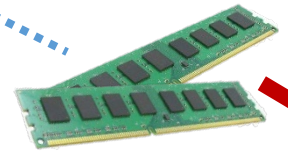
- Quin és l'impacte d'una mida incorrecte de bloc???
  - Massa gran: **fragmentació** (i.e. malgastar espai)
  - Massa petit: pèrdua de rendiment per fer masses accesos al dispositiu

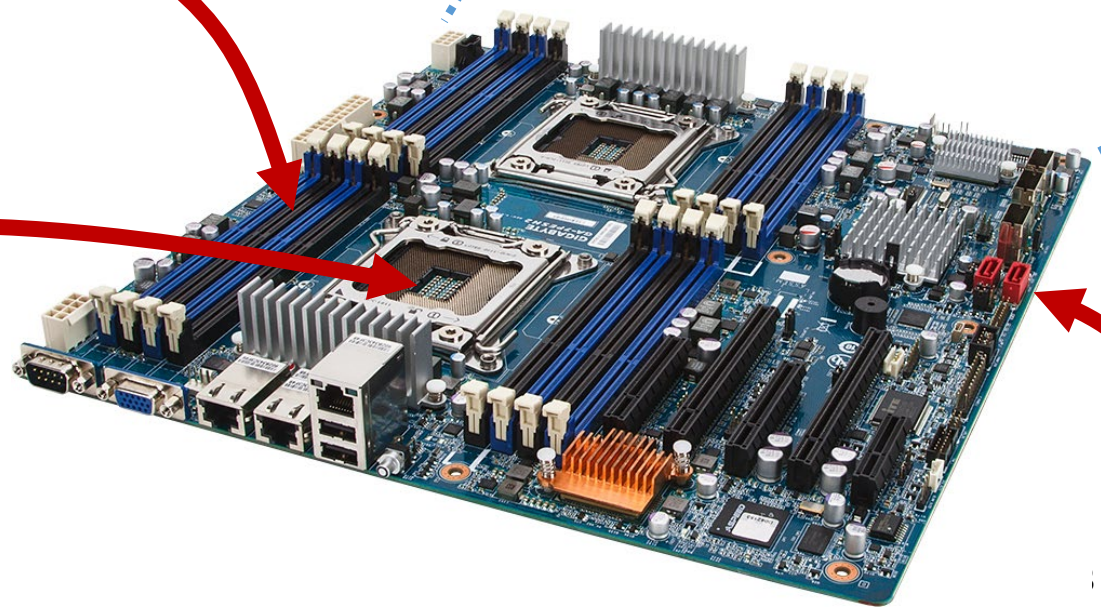
- Blocs grans tenen l'advantatge d'una de les formes de **Localitat** (localitat espacial)
  - Parlarem sobre localitat en transparències posteriors…

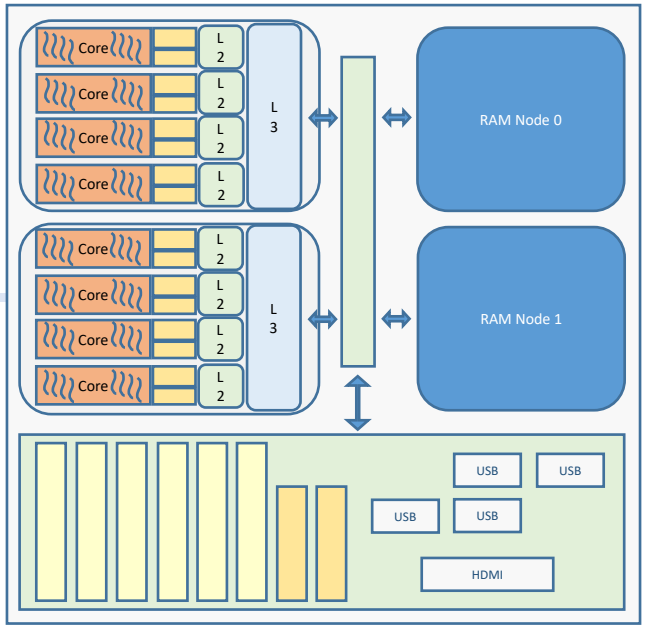
En Resum...


Core
L1 Data
L1 Insn
L2 Data Insn
Core
L1 Data
L1 Insn
L2 Data Insn
Core
L1 Data
L1 Insn
L2 Data Insn
Core
L1 Data
L1 Insn
L2 Data Insn
L3 Data Insn


Main Memory






Core
L 2
L 3
RAM Node 0
RAM Node 1
USB
HDMI

Bibliografia

- Computer Systems – A Programmer's perspective (3rd Edition)
  - Randal E. Bryant, David R. O'Hallaron, Person Education Limited, 2016
  - https://discovery.upc.edu/permalink/34CSUC_UPC/11q3oqt/alma991004062589706711
    - Several Chapters

- Computer Organization and Design (5th Edition)
  - D. Patterson, J. Hennessy, and P. Alexander
  - http://cataleg.upc.edu/record=b1431482~S1*cat
    - Several Chapters