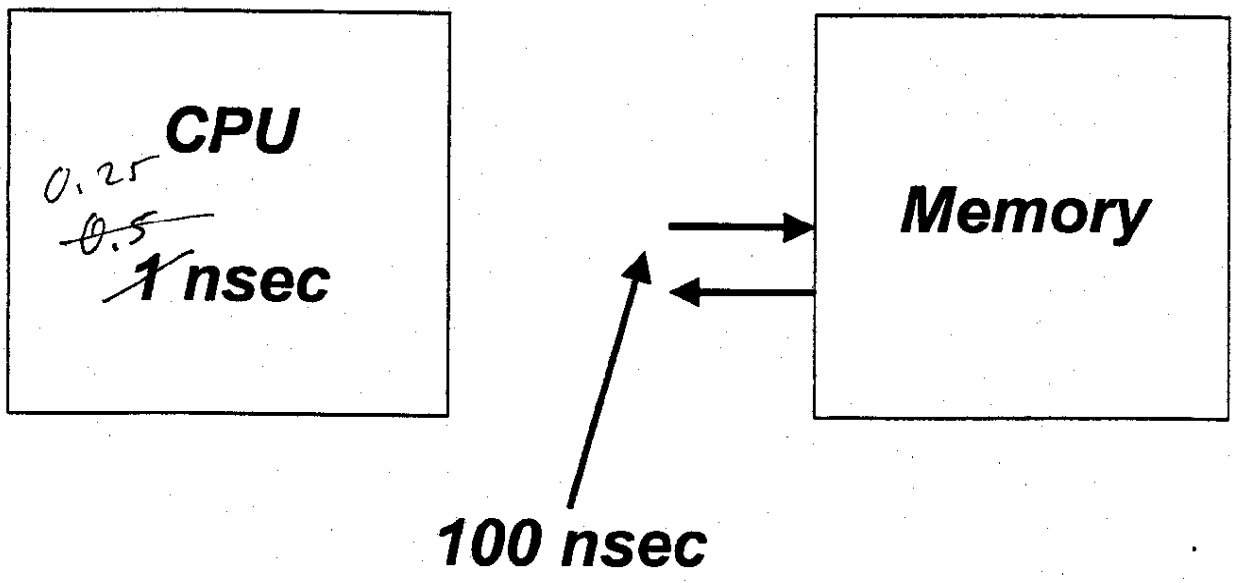
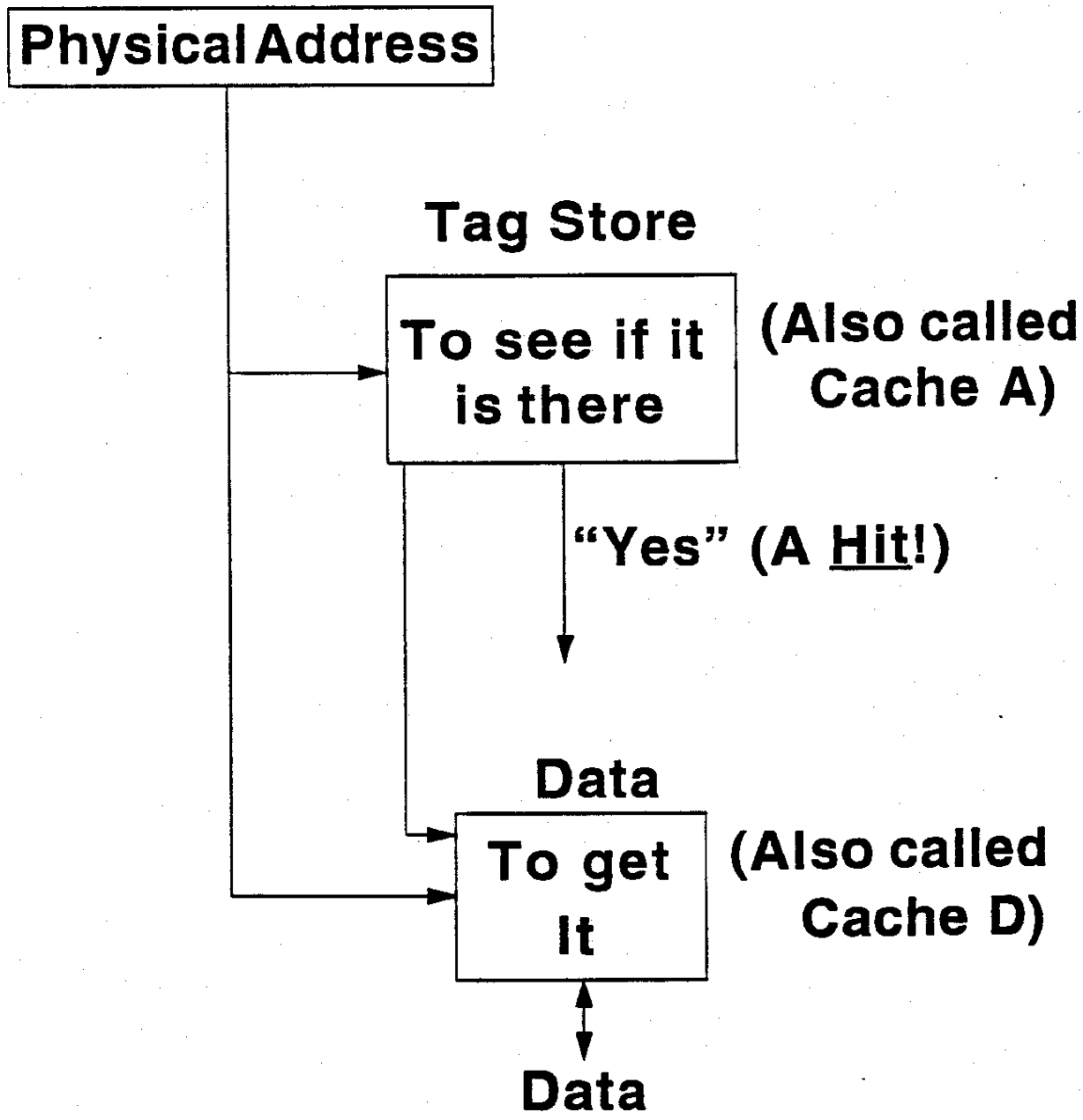
Cache/1

Cache Memory

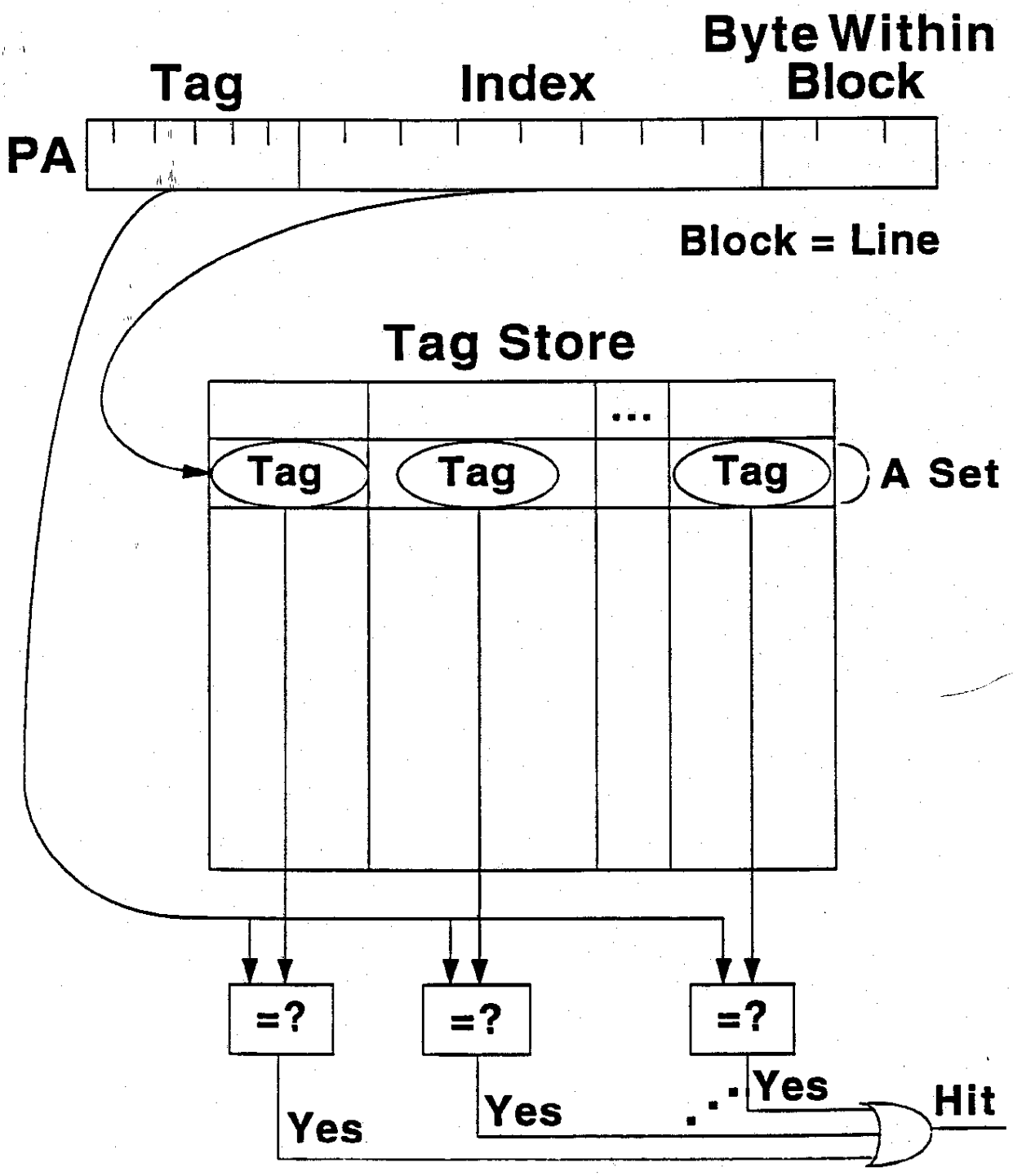
WHY?

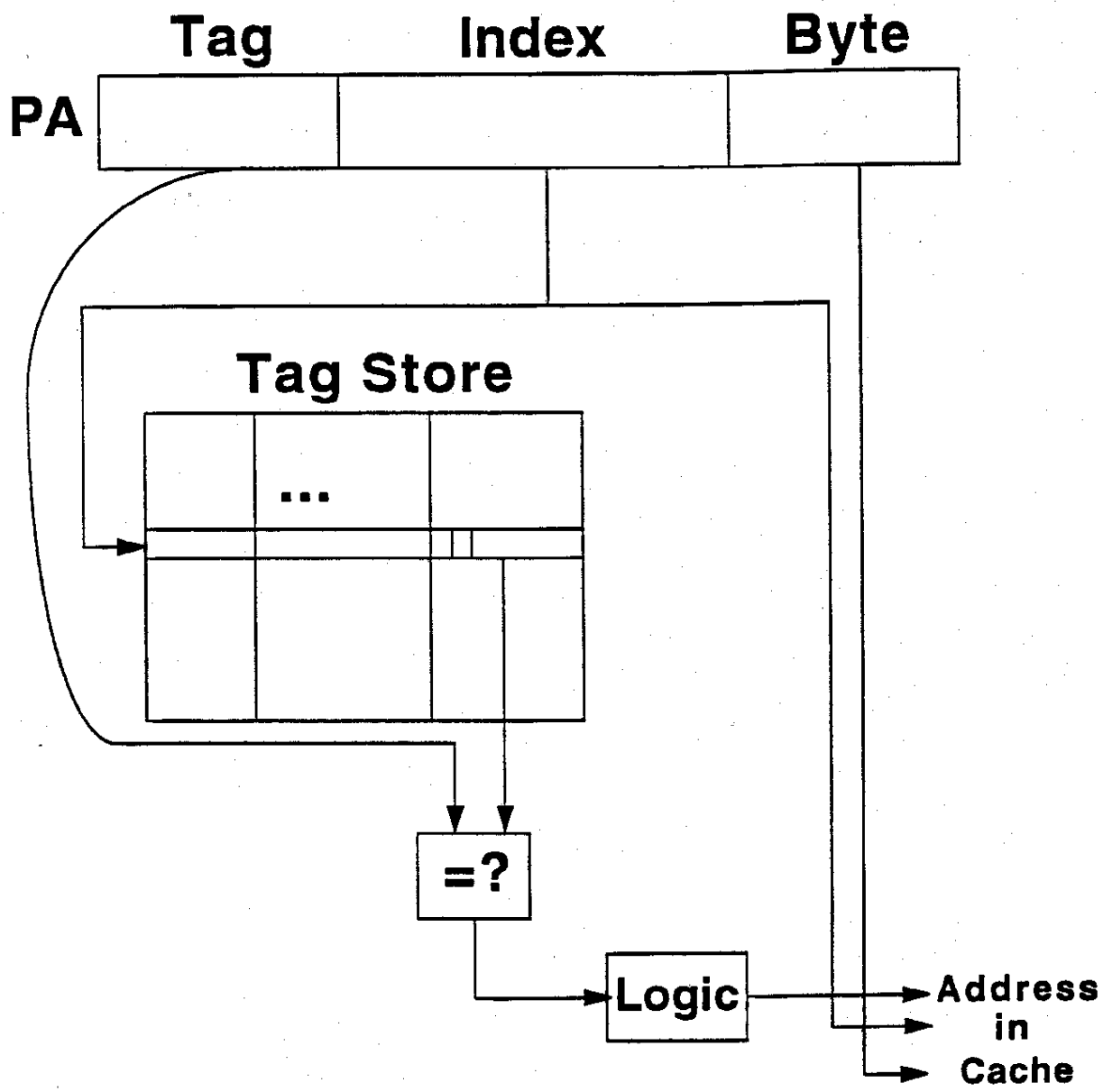


The Abstraction

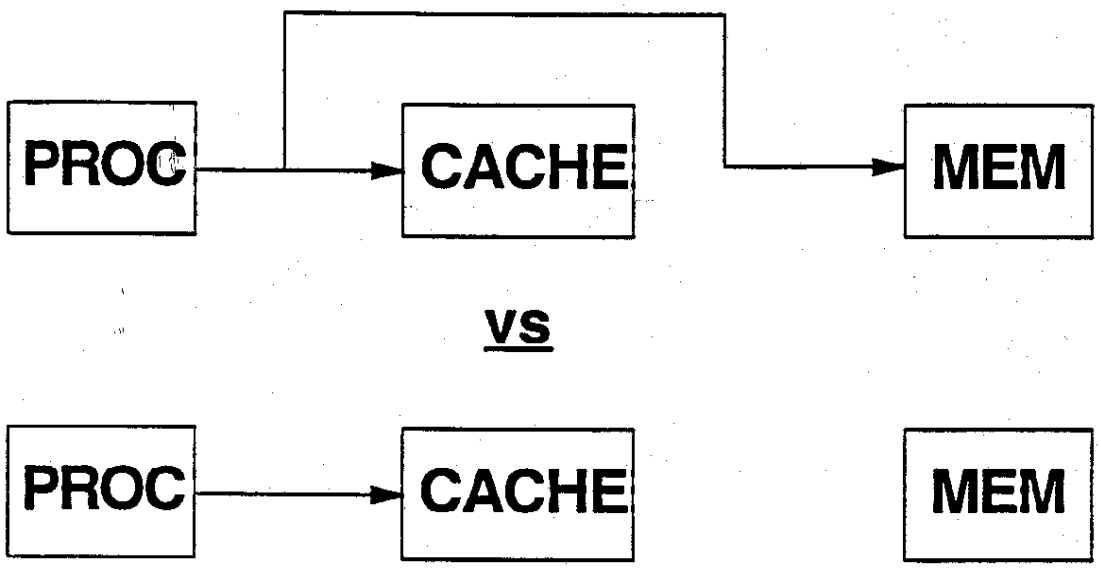


Note: Hit Ratio = $\frac{\text{Hits}}{\text{Hits + Misses}}$


Byte Within Block
Tag
Index
PA
Block = Line
Tag Store
...
Tag
Tag
Tag
A Set
=?
=?
=?
Yes
Yes
Yes
Hit


Tag
Index
Byte
PA
Tag Store
...
=?
Logic
Address
in
Cache

Write Through/Write Back



Issues

* Simplicity of Design
* Bus Traffic
* Application Environment (Stack Frame)
* Allocate on Write Miss
  - Sector Cache

Characteristics

* **Set Associative (Set Size)**
  - **Fully Associative**
  - **Direct Mapped**

* **Write Back, Write Through**

* **Replacement Algorithm**
  - **LRU**
  - **FIFO**
  - **Random**

* **Instructions/Data**

* **Superviser/User**

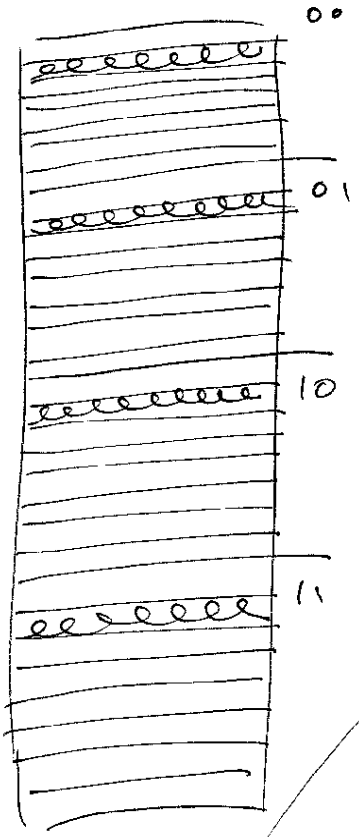
* **Virtual/Physical**

CACHE

1971
4004
2300 Transistors

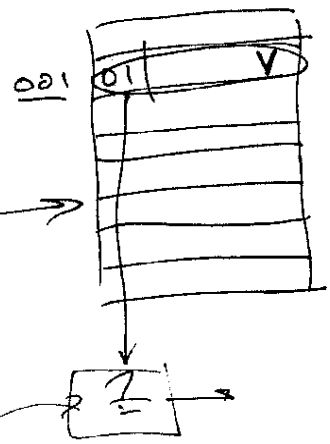
198x
Motorola 68020
256 B

PM
256 B

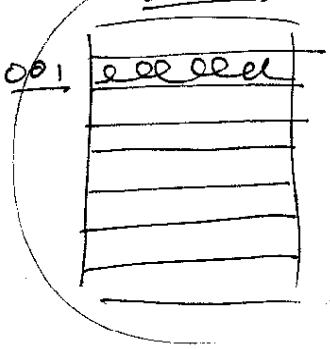


CACHE_A

TAG STORE

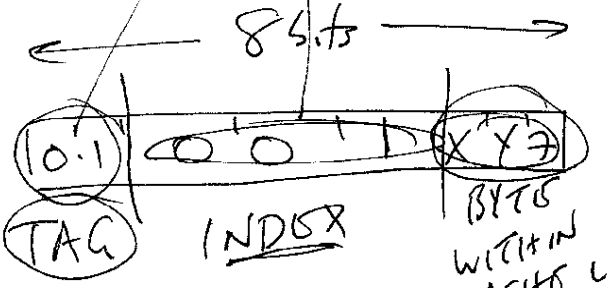


64 B CACHE_D



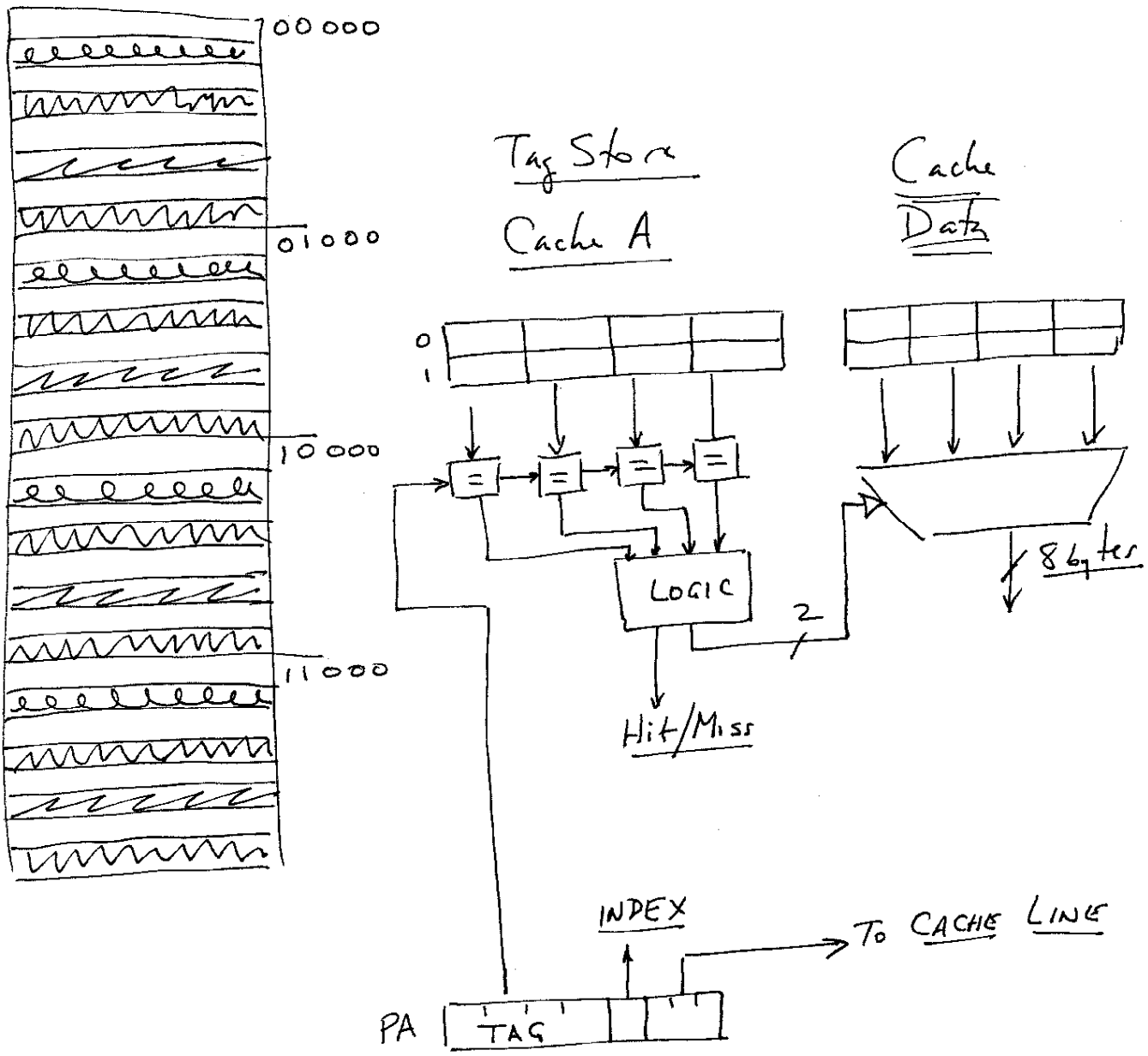
TEMPORAL LOCALITY

CACHE LINE
" BLOCK



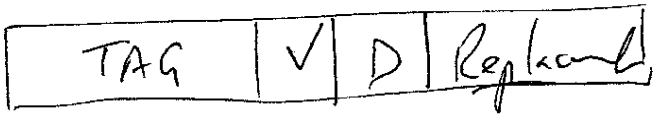
SPATIAL LOCALITY

MOV A, R1
=
MOV B, R2



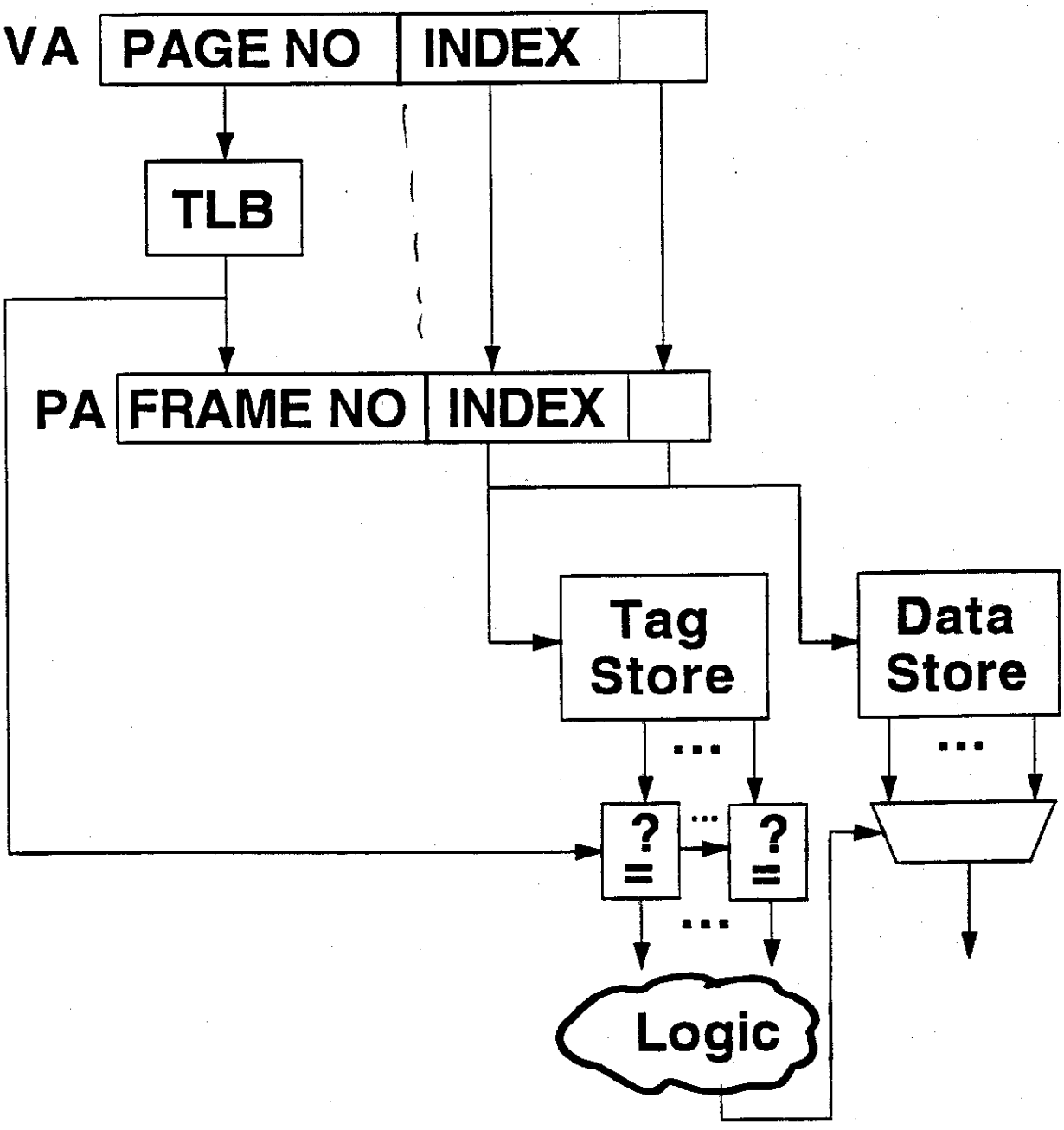
TAG STORE ENTRY:

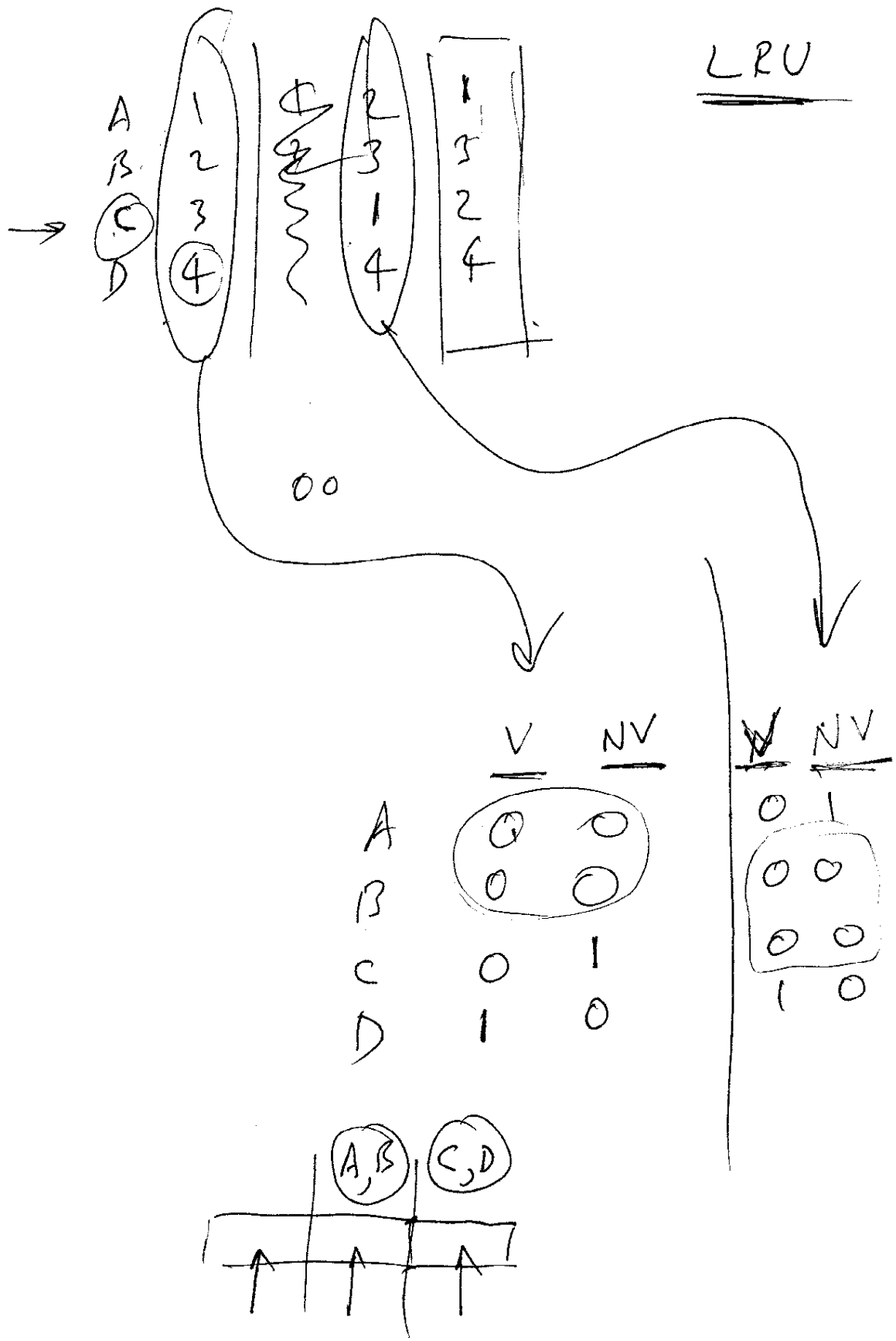
LRU



Cache 1

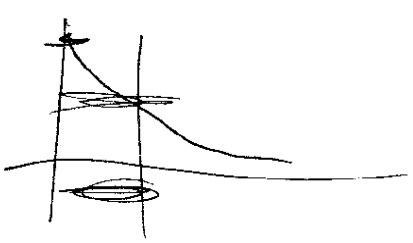
Virtual/Physical




LRU
A
B
C
D
1
2
3
4
2
3
1
4
1
3
2
4
00
V
NV
V
NV
A
B
C
D
0
0
0
1
1
0
0
1
0
0
0
0
0
1
1
0
A,B
C,D

Interrupts / Exceptions

Interrupts		Exceptions
External	Causes	Internal
depends	Prior	Same
When Convenient	When	Now
System	Context	process
Almost Always	Maskable	Almost never



1. Stop what you're doing
2. Consistent State
3. Vector to handler