

# Cache Lab Q&A

we use 8-bit addresses

eg. 0 1 1 0 1 0 1 0  
7 6 5 4 3 2 1 0

spatial locality means programs will access addresses that differ mainly in their low-order bits

break the address up into 3 parts, starting at bit zero (the right side):

- line offset
- index
- tag

---

valid bit: does not come out of the address or the data.

It just records whether our cache entry holds cached data or not.

## How are threads executed?

1. Run all threads on a single processor by switching between them at some regular interval.

Useful for latency hiding, not for performance.

2. On a machine with multiple processors (or cores) run threads across all processors.

use threads to keep multiple cores busy

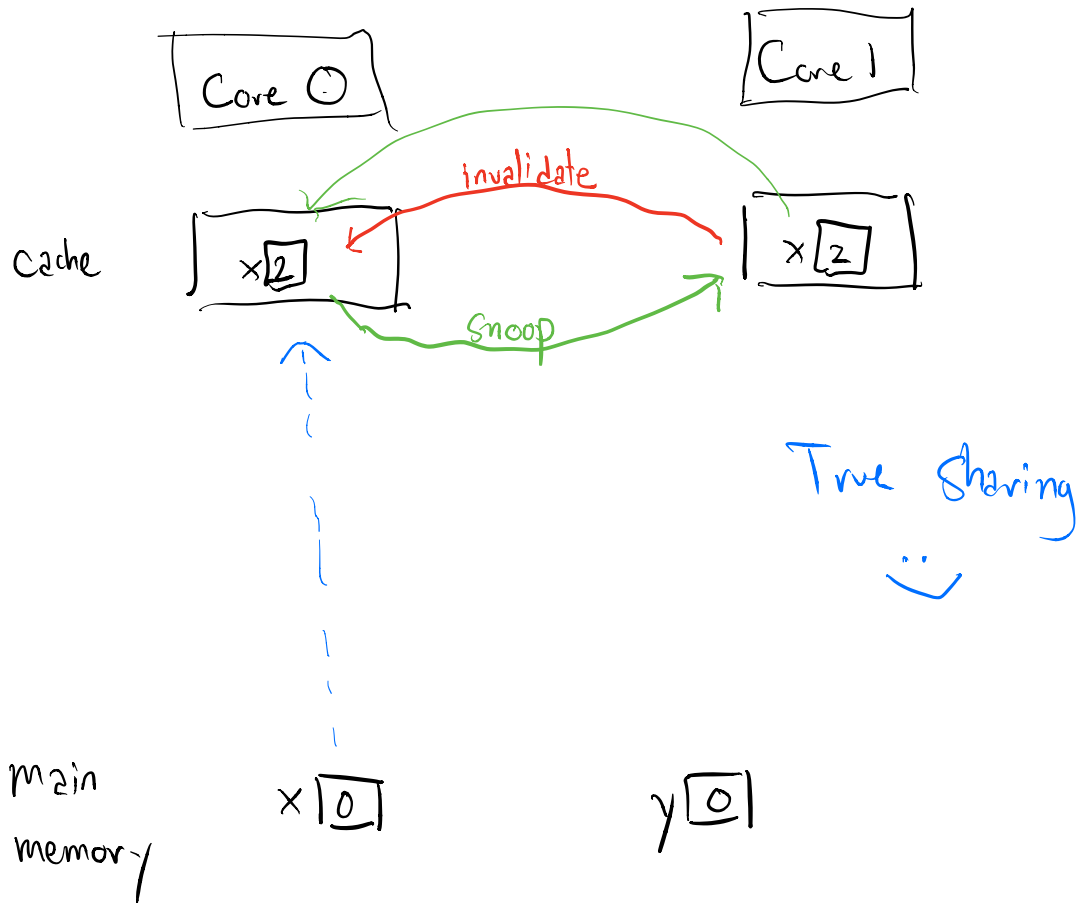
3. Use simultaneous multithreading to run more than one thread on a single core.

Other names: hardware multithreading  
hyperthreading

use threads to keep all or most functional units in multiple-issue datapath busy.

# Cache Coherence

multiple cores with separate caches



Core 0 :

- ① read x
- ② write x
- ③ read x

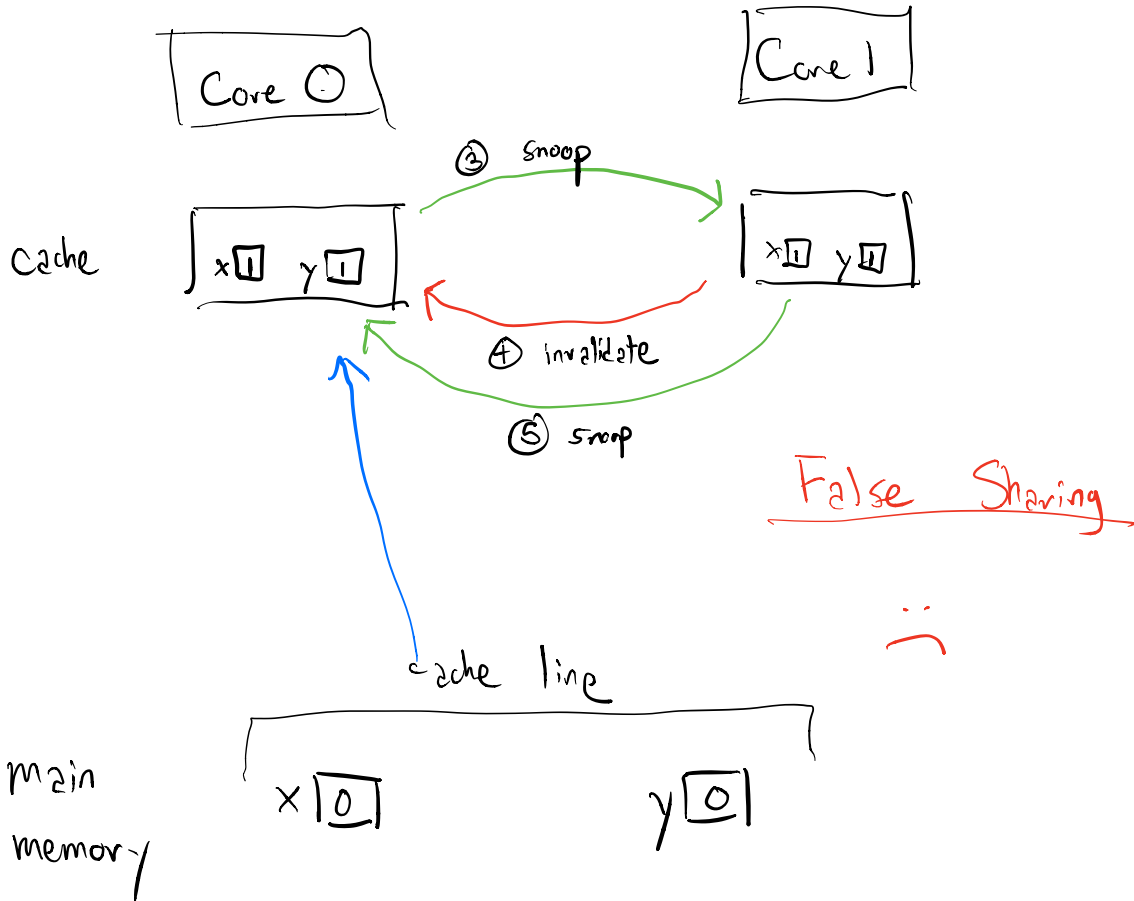
Core 1 :

- ③ read x
- ④ write x

# Cache Coherence

multiple cores with separate caches

MESI: modified exclusive shared invalid



Core 0:

- ① read x
- ② write x
- ③ read x

Core 1:

- ③ read y
- ④ write y