Cache Coherency

MESI is a protocol use by Intel and others

- Each cache line can be in one of 4 states:
- M – modified (in one cache & was altered since read from memory)
- E – exclusive (in one cache & matches what's in memory)
- S – shared (in at least 2 caches and matches what's in memory)
- I – invalid (never filled or is outdated)

Lines change state based on accesses that are made

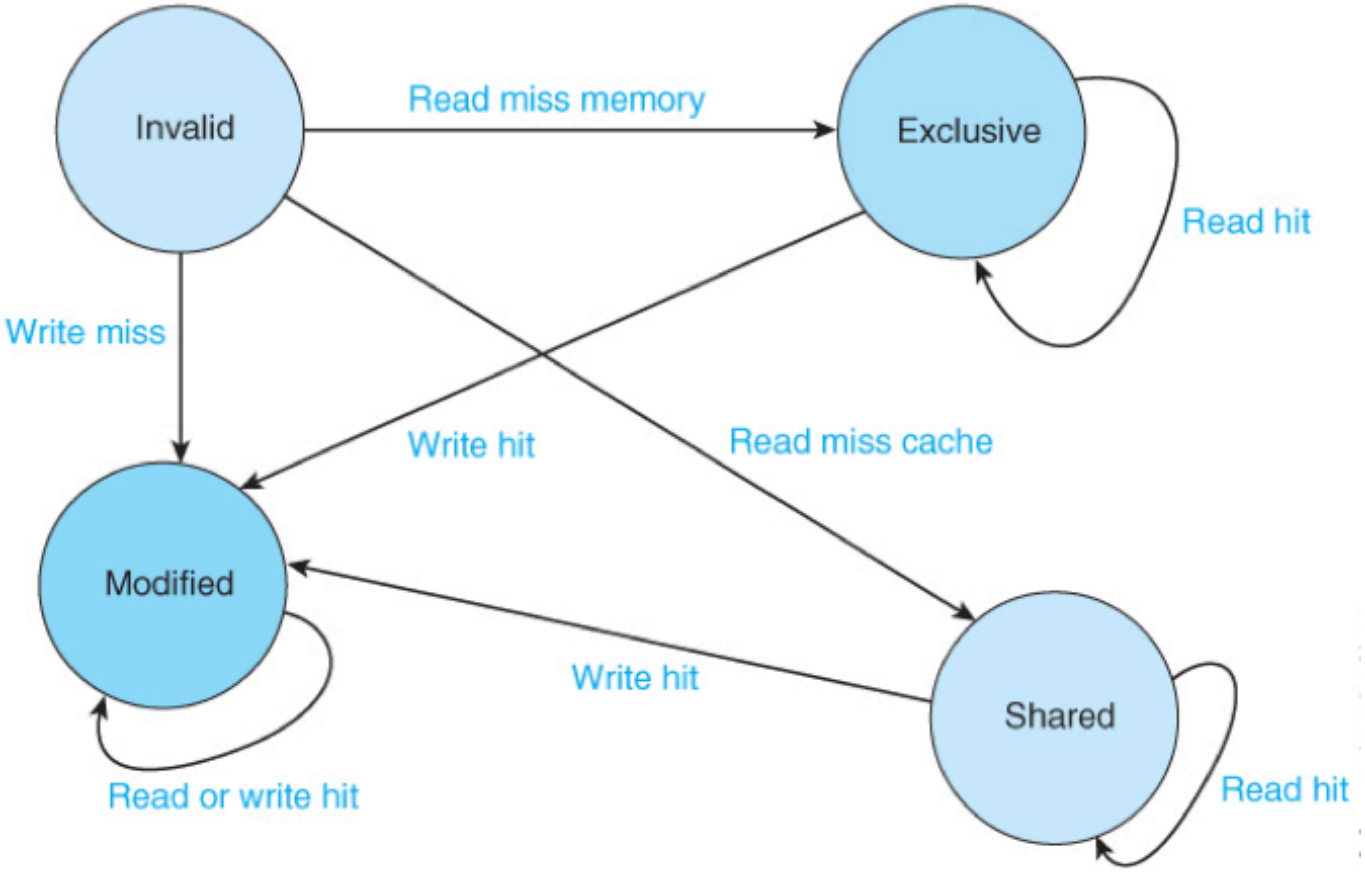
- From local processor or from other processors over the bus
- Cache controller listens (snoops) on the bus for accesses
- A *snoopy* protocol

Snoopy Protocol

- Writes are broadcast on the bus
  - All caches observe the write
  - Writes may invalidate copies of lines in other caches
- Does not scale beyond about 128 processors
  - The bus traffic due to snooping becomes too high
  - Exclusive and modified states reduces traffic
    - Writes to modified or exclusive lines need not be broadcast

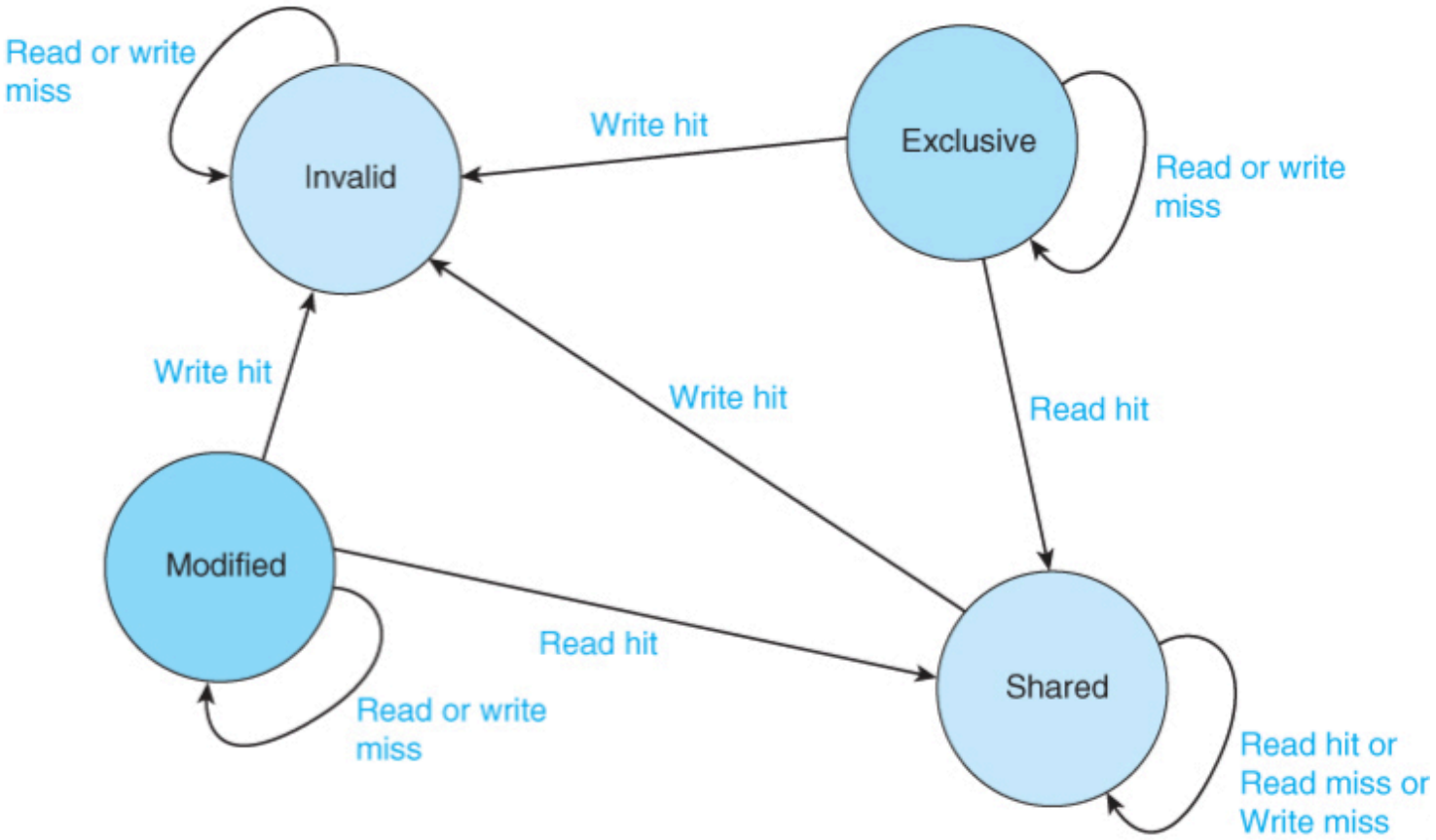
Snoopy Protocol

MESI state diagram as seen from the local CPU bus



Snoopy Protocol

MESI state diagram as seen from the system bus



MESI State Change

Bus Transaction	Action by local cache
Read hit	E➔ S, M➔ S or no change if S
Read miss	no change
Write hit	E ➔ I, S ➔ I, M ➔ I
Write miss	no change

Directory Based Protocol

Snoopy caches work well when connected to a single bus

Large shared-memory multiprocessors use interconnects

- These are networks such as rings or meshes
- Broadcasting cache operations to all processors would be inefficient

Cache directories can be used as an alternative to snooping

- Each memory module would have a directory
- This approach scales better than snooping on a shared bus

Directory Based Protocol

Directories identify which nodes contain a copy of a block

- The state of each cache line containing a copy is recorded

- Accesses to a module are intercepted by the directory

- The directory determines the action to take:
  - Reads are forwarded to the cache containing the copy
  - Writes are sent just to the nodes whose copies are affected
  - Misses cause the memory module to be accessed