

Reading 19

19.2 Cache basics: Part 2

- Upon System startup, a cache's entries are invalid.
- Thus, each cache entry has a valid bit that is initially set with 0, then set with 1 upon the first copy of an item into that entry (staying 1 from then on)
- If a program accesses a memory location, then the program will likely soon access adjacent locations too, known as spatial locality.
- Thus, if a memory item doesn't exist in cache, the system usually copies that item plus several adjacent items known as a block (or line) into the cache.
 - ↳ typical block sizes are 16, 32, or 64 bytes.
- Some address bits are used to offset into the block to get the proper item.
- Write through: a cache write approach
 - ↳ when a processor writes data to a cache, the system also writes that data to memory
- Write back: A faster cache write approach
 - ↳ when a processor writes data to the cache, the system marks the cache block as dirty, but doesn't write memory. Then, when a dirty block is replaced, the system writes the block to memory first.