Memory Devices and Revisiting Locality

HW1 due tonight;
Check-In 3 today!

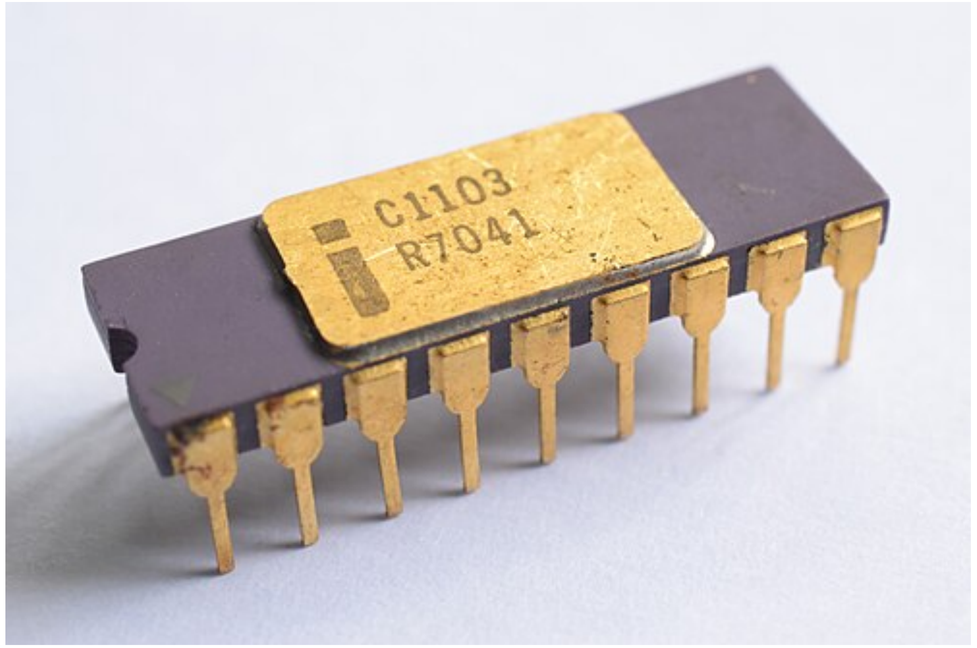


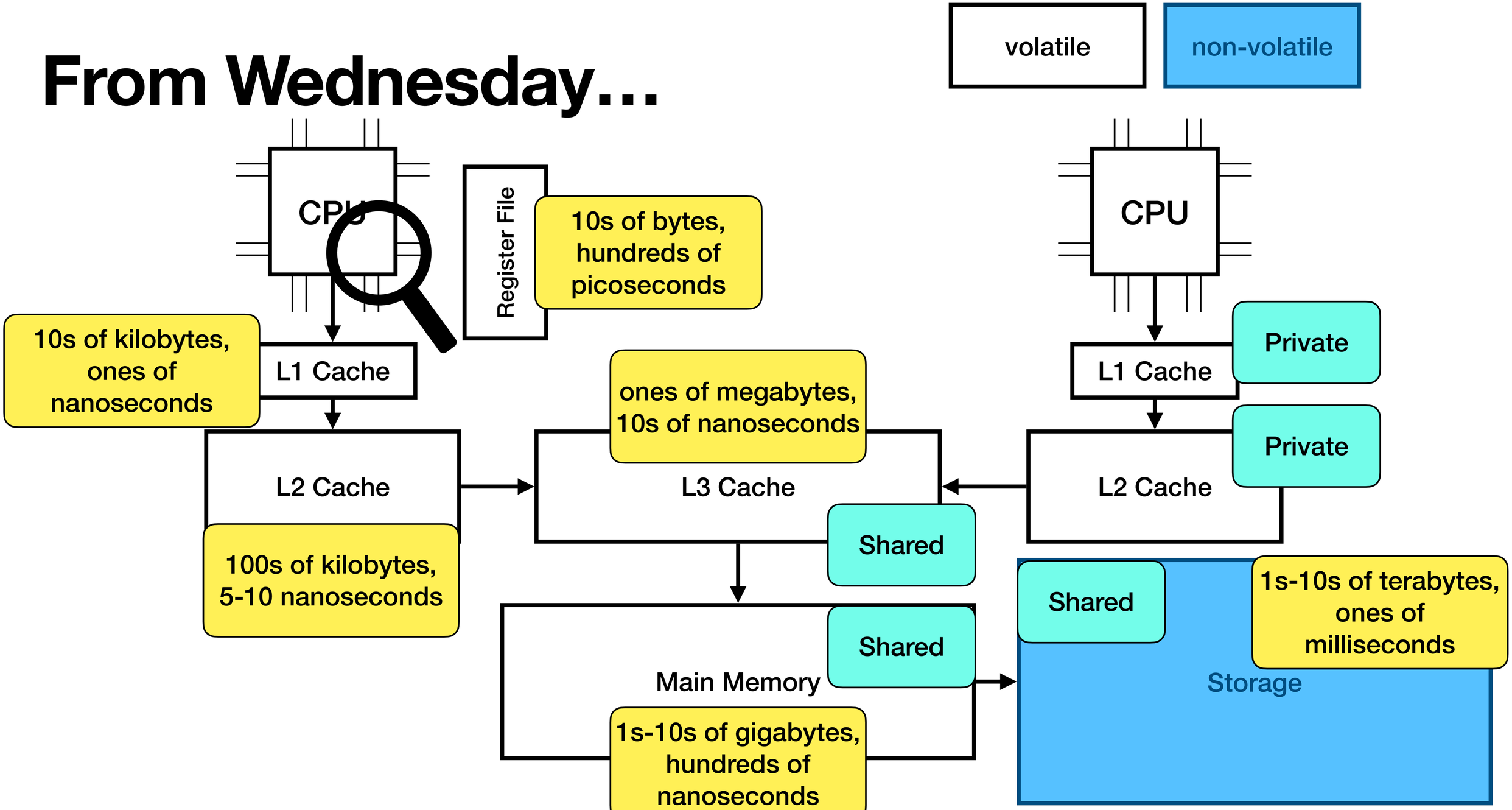
Image credit: https://en.wikipedia.org/wiki/Intel_1103

Image credit: https://en.wikipedia.org/wiki/3D_XPoint

Outline

- Overviewing memory system devices
- The (sad) story of Intel Optane
- Revisiting locality

From Wednesday…

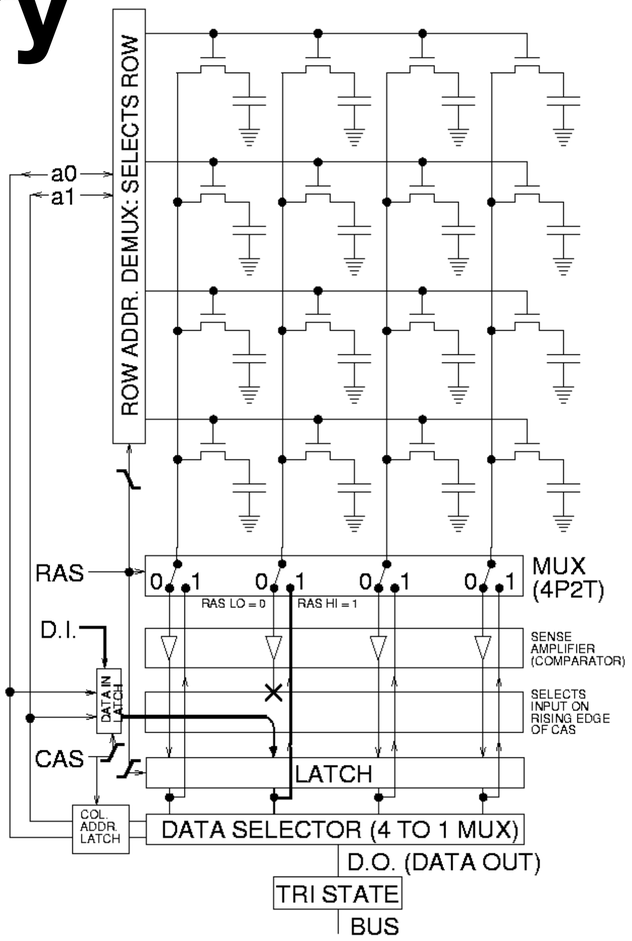


Memory Volatility

- A *volatile* storage device maintains its state during runtime, but data is lost after power loss
- A *non-volatile* storage device maintains its stored state at runtime *and* on a power loss
- In general, caches and main memory are *volatile* whereas long-term storage is *non-volatile*

DRAM, the Basis of Main Memory

- DRAM: *dynamic* random access memory
- Each bit is stored in a cell composed of a single *transistor* (which acts as a *capacitor*)
- This design means that the circuit is *unstable* at runtime, and data stored in a DRAM cell must be constantly *refreshed*
- Low transistor count per storage cell means that more bits of information can be stored in a small area ➡️ high data density!
- Low transistor count means that the cost per byte in DRAM is cheap!

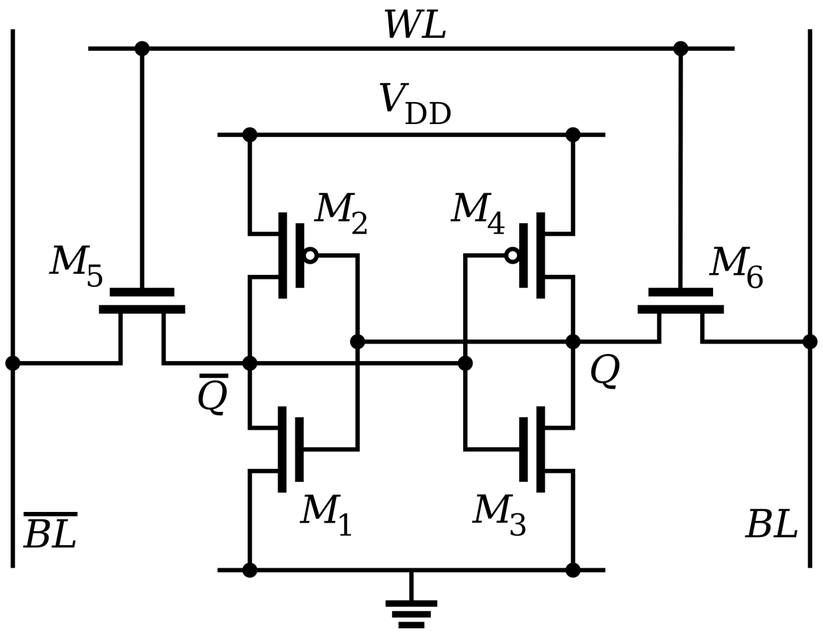


DRAM Diagram

Image credit: https://en.wikipedia.org/wiki/Dynamic_random-access_memory

SRAM, the Basis of Caches

- SRAM: *static* random-access memory
- Bit storage cells in SRAMs are composed of 6 transistors in a *stable* circuit. This stability means that access latency is very closely related to the cycle latency!
- Large number of transistors per cell means that the devices are dense!
- High transistor density ➡️ high cost per bit stored
- High transistor density ➡️ power-hungry component!



SRAM Transistor Diagram

Image credit: https://en.wikipedia.org/wiki/Static_random-access_memory

Storage Options

- Hard disk drives (HDDs) are based on *magnetic disks* that depend on a physical disk drive and magnet head ➡️ accessing data implies spinning the disk until the head points to the memory region of interest
- Solid state drives (SSDs) are based on *flash memory*, and generally support lower latency reads/writes to long-term storage (8-15 times faster than HDD)
  - Flash accesses are *sequential* to an entire page (coarse grained access)
  - Data must be *erased* before it is written in flash memory (i.e., “flashed!”)
  - Flash cells may only be written to ~10^8 times before the cell is worn out (limited *write endurance*)… bad if application has spatial locality!

Non-Volatile Memory: A Happy Medium?

- of the u … y
- hat if we … y for
- …rage is … n
- e betwee…
- faster th…
- power!
- duct that…

Announcement: Intel® Optane™ Persistent Memory 300 Series

Content Type: Product Information & Documentation | Article ID: 000093792 | Last Reviewed: 03/03/2025

We continue to rationalize our portfolio in support of our IDM 2.0 strategy. This includes evaluating divesting businesses that are either not sufficiently profitable or not core to our strategic objectives.

Effective January 31st, 2023, Intel intends to cancel the Intel® Optane™ Persistent Memory 300 Series (previously code-named *Crow Pass*) and not develop future Intel Optane products.

- As part of this plan, Intel does not intend to ramp production, take future orders, take last time buys, provide warranty coverage or technical support for the Intel® Optane™ PMem 300 Series.
- Intel does not intend to continue enablement of the Intel® Optane™ PMem 300 Series with 4th Gen Intel® Xeon® products.

NVMW 24 — Home

Session IV: Caches & Security
Chair: Steven Swanson

Rethinking Metadata Caches in Secure NVMs
Samuel Thomas (Brown University);Hammad Izhar (Brown University);Eliott Dinfotan (Boston University);Tali Moreshet (Boston University);Maurice Herlihy (Brown University);Iris Bahar (Colorado School of Mines);

Extended abstract Slides

NVMW 24 — Home F

Using a Fast Subtree for Efficient Secure NVMs
Samuel Thomas (Brown University);Kidus Workneh (University of Colorado, Boulder);Jac McCarty (Bryn Mawr College);Joseph Izraelevitz (University of Colorado, Boulder);Tamara Lehman (University of Colorado, Boulder);Iris Bahar (Colorado School of Mines);

Extended abstract Slides