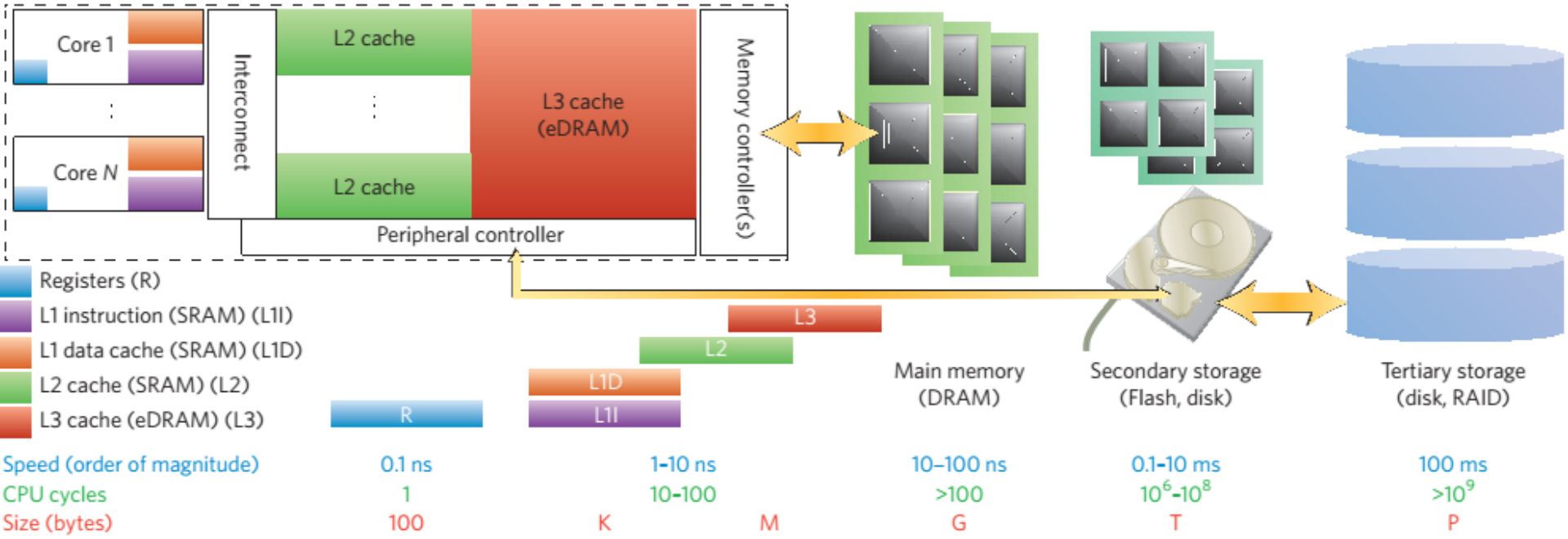
Key features for next-generation memories

Yuewei Yin

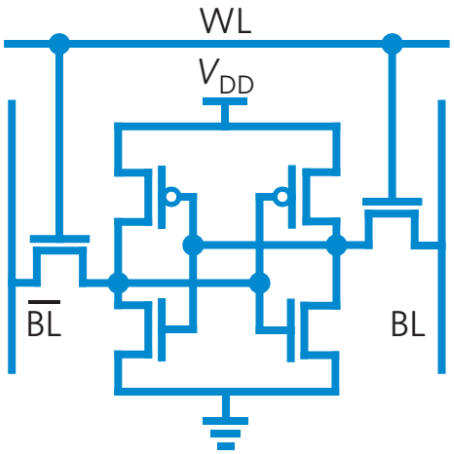
10-16-2015

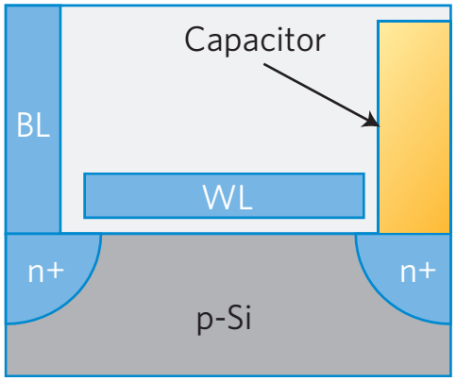
Two types of digital memories: memory and storage

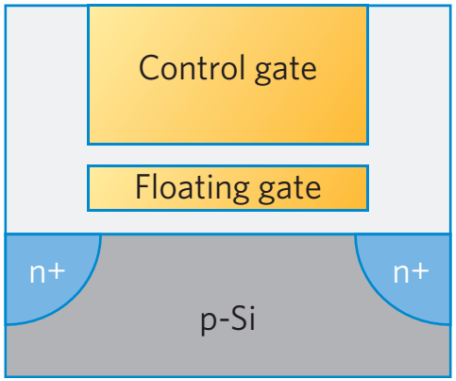


- Fast access used by computer processor: Random access memories: fast read/write (30 ns for DRAM, <1ns for SRAM): volatile to achieve high speed.
- long time storage with infrequently access: hard disk drive: direction of magnetization: high data density to order of terabyte slow read/write (~ms) with moving parts: non-volatile to save energy.

In used memories: SRAM DRAM and Flash







SRAM:
Volatile
Quick
Six transistor needed
High power consumption

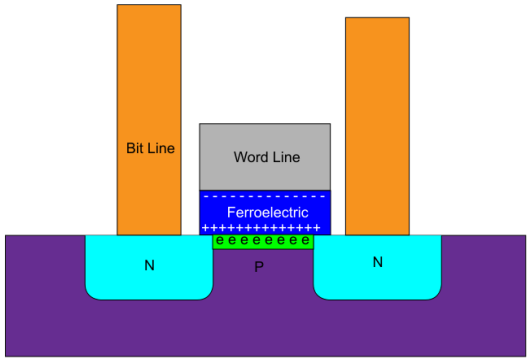
DRAM:
Volatile, periodical refresh
Cheaper, smaller, lower power consumption than SRAM.

Flash:
Non-Volatile, low endurance.
Slow write, large voltage needed.

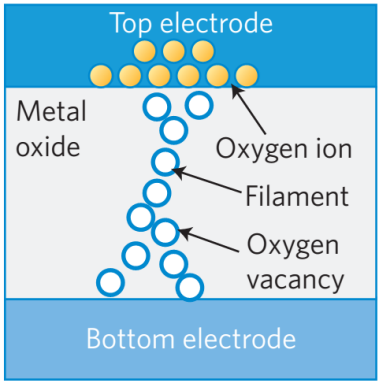
Universal: Fast, dense, low power, non-volatile

Candidates for next generation memories

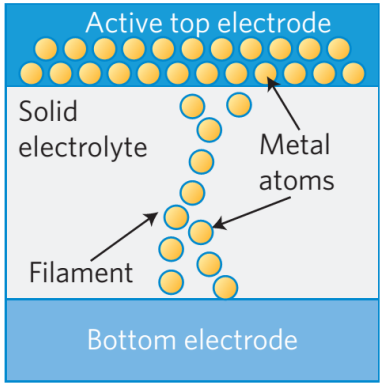
FeRAM



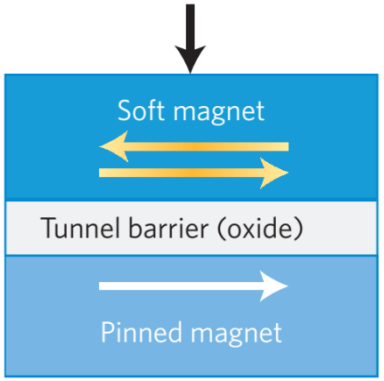
ReRAM

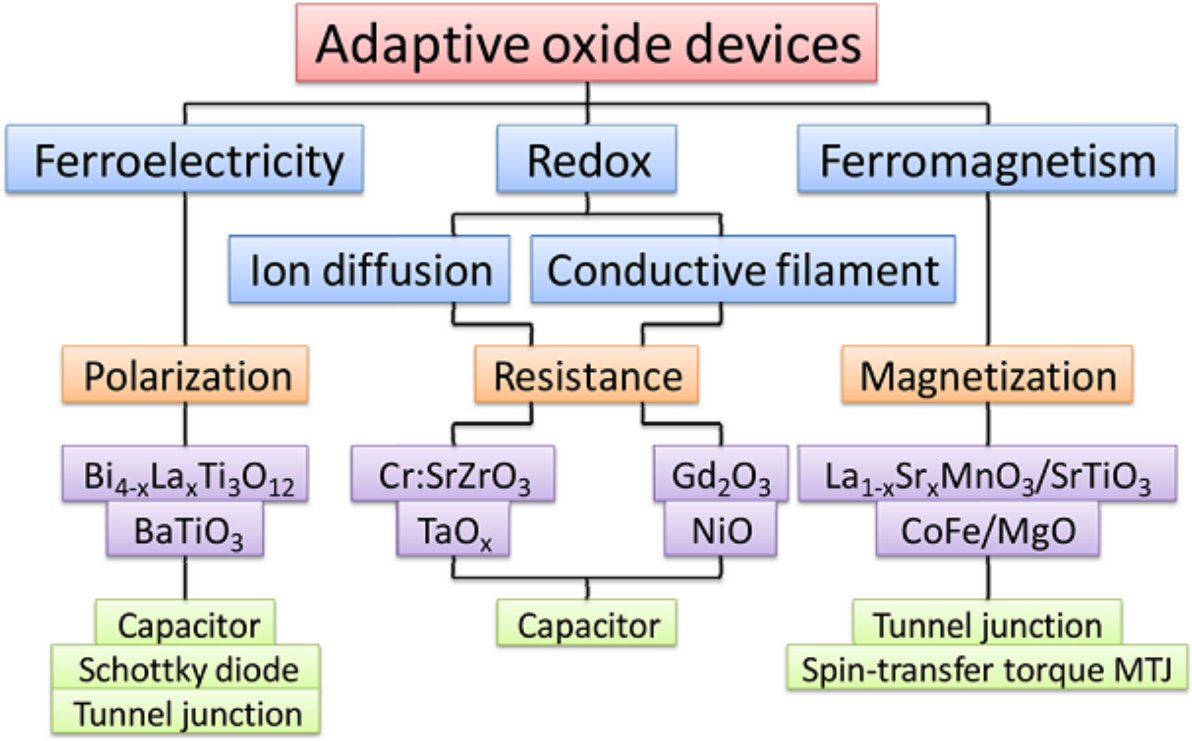


CBRAM

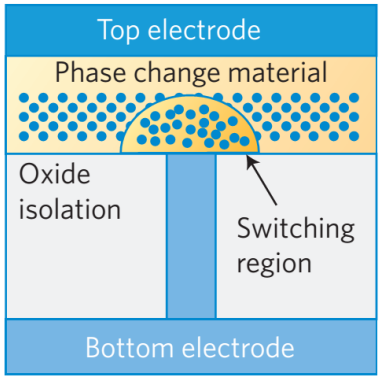


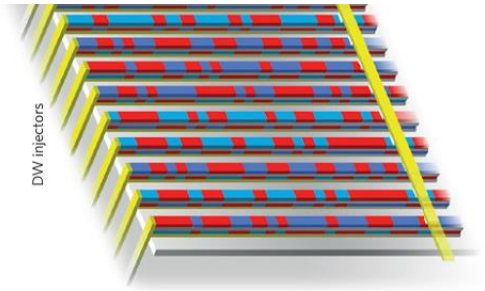
STT-MTJ



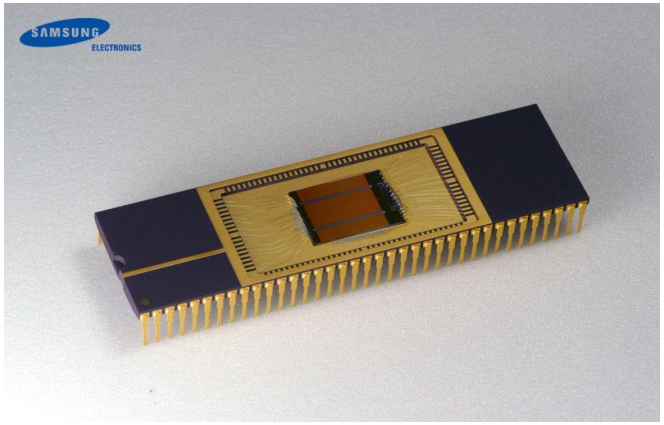


PCM





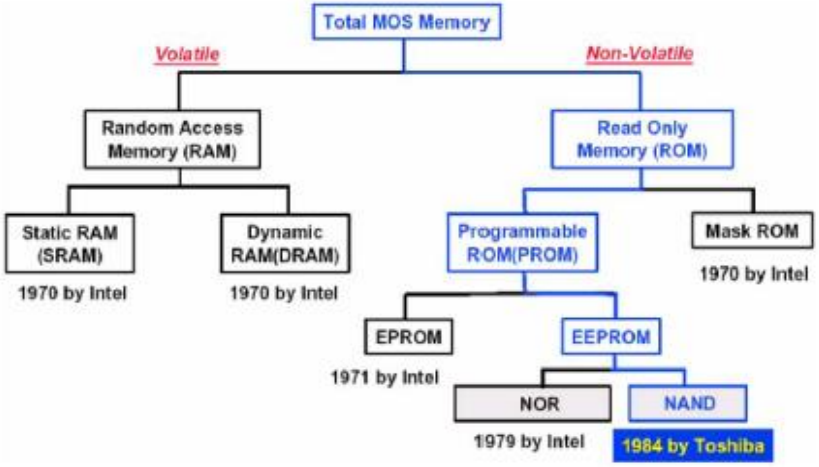
Disk Drive vs. Flash Memory



Read / Write

(+) Lost cost per bit

(-) Mechanical movement (SPM & VCM)
(-) High power consumption (10-15W)
(-) Heavy weight compared to flash



Read / Program / Erase

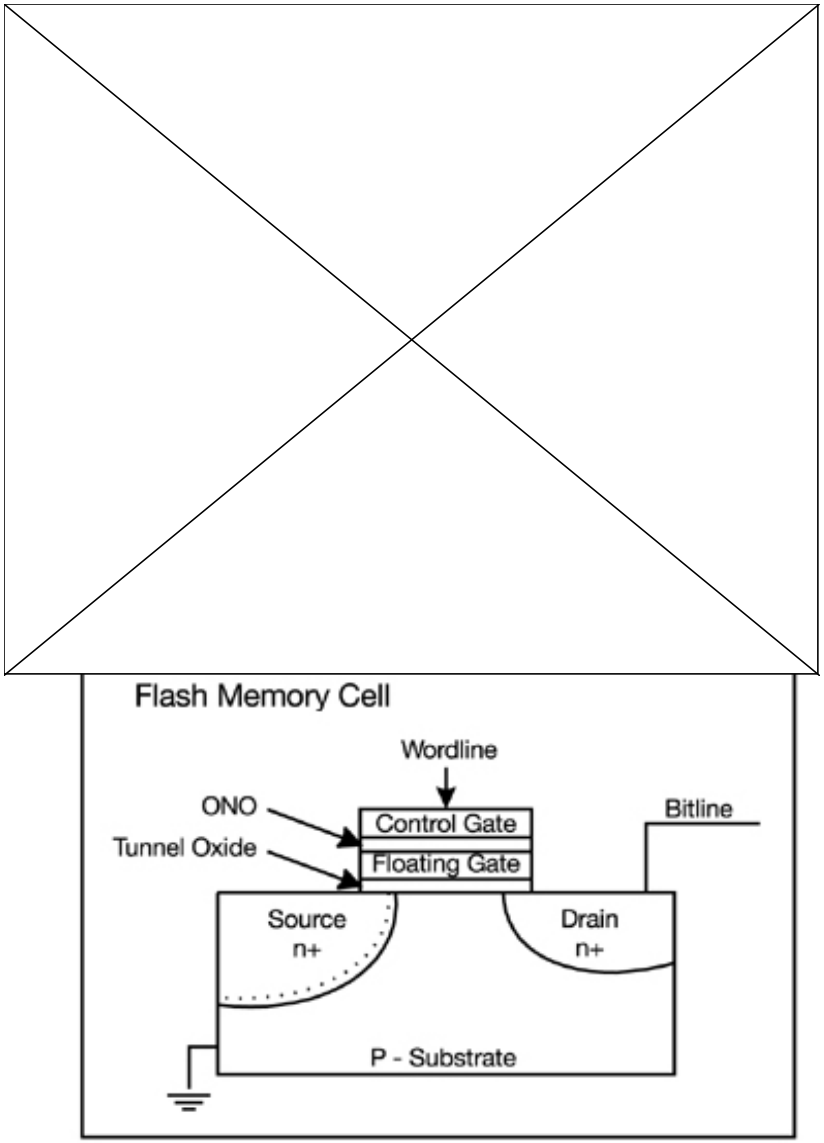
(+) Random Access
(+) Non-volatile
(+) Low Power Consumption (2W)

(-) Erase before Write
(-) Erasing operation in the unit of block (not page)
(-) Maximum # of erase operations per cell
(-) High cost per bit

How Flash Memory Works

One typical kind:

- The floating gates only link to the row, or **wordline**, is through the control gate.
- Tunneling can be used to alter the placement of electrons in the floating gate.
  - An electrical charge, usually 10-13 V, is applied to the floating gate. The charge comes from the column, or bitline, enters the floating gate and drains to ground
  - This charge causes the floating gate transistor to act like an electron gun. The excited electrons are pushed through and trapped on other side of the thin oxide layer, giving it a negative charge. These negatively charged electrons act as a barrier between the control gate and the floating gate.

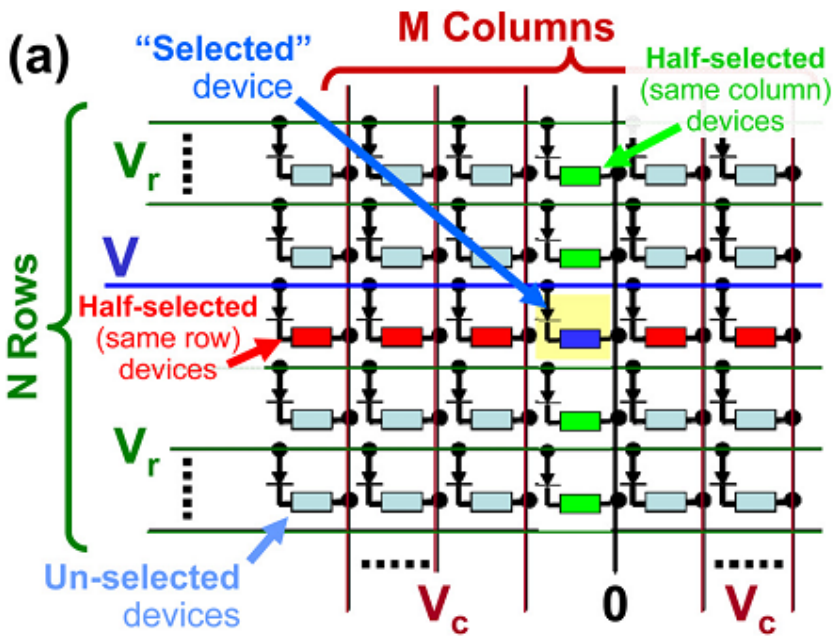


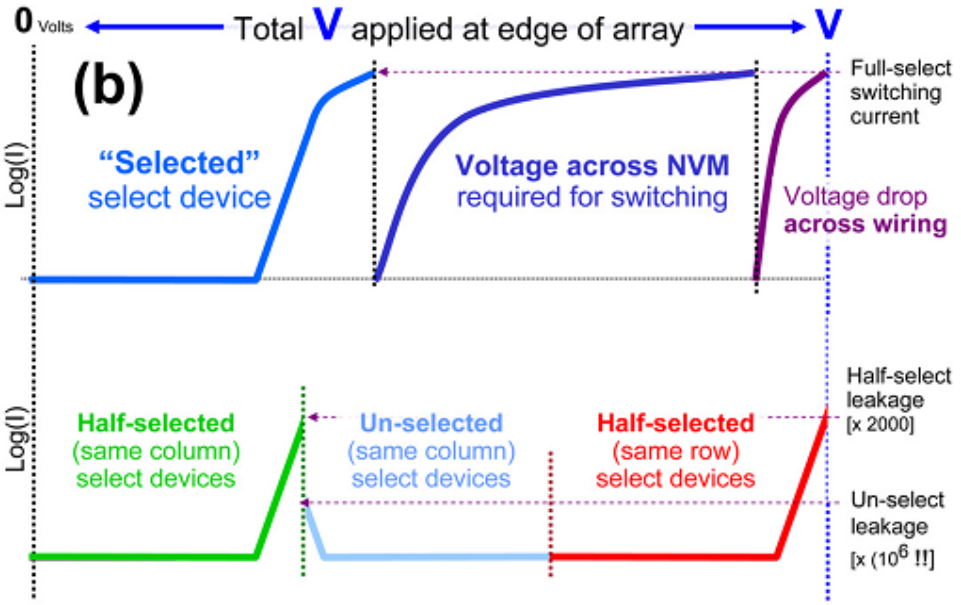
Key features for most studied RAMs

	SRAM	eDRAM	DRAM	eFlash (NOR)	Flash (NAND)	FeRAM	PCM	STT-MRAM	RRAM
Endurance (cycles)	Unlimited	Unlimited	Unlimited	**10^5**	**10^5**	10^{14}	**10^9**	Unlimited	**10^9**
Read/write access time (ns)	<1	1–2	30	10/**10^3**	100/**10^6**	30	10/100	2–30	1–100
Density	**Low (six transistors)**	Medium	Medium	Medium	High (multiple bits per cell)	**Low (limited scalability)**	High (multiple bits per cell)	Medium	High (multiple bits per cell)
Write power	Medium	Medium	Medium	**High**	**High**	Medium	Medium	Medium	Medium
Standby power	**High**	Medium	Medium	Low	Low	Low	Low	Low	Low
Other	**Volatile**	**Volatile. Refresh power and time needed**	**Volatile. Refresh power and time needed**	**High voltage required**	**High voltage required**	**Destructive readout**	**Operating T<125°C**	**Low read signal**	**Complex mechanism**

Write voltages should be from a few 100mV (compatible with CMOS) to few V (compete Flash. The time is desired to be <100ns to compete with DRAM and to outperform Flash, or even <10 ns to approach SRAM. Read voltages need to be significantly smaller than write voltages but cannot be less than one tenth of write. Read currents are typically at least 10 times smaller than write and in ON-state, current should not be less than approximately 1 mA. The read time must be in the order of write or preferably shorter.
An ON/OFF ratio of only 1.2 to 1.3 can be utilized as shown in MRAM, >10 are required to allow for small and highly efficient sense amplifiers.
A data retention time of >10 years is required for universal NVM kept at 85°C and a constant stream of Vrd pulses. U/(kT) > 40.

Further requirements: access





Write: non-perturbing and low power. Deliver high currents to selected cells and limit leakage through half-/non- selected cells. Current to write NVM can be delivered by access device.

Read: Accurate sense. Avoid sneak paths through half-/un-/half- series.

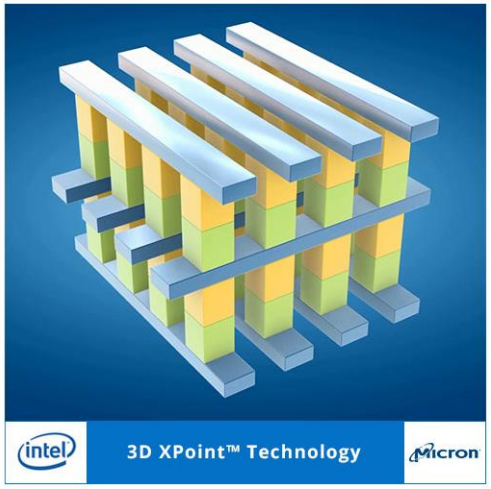
Specifications for access device depends on and also determined the specifications of NVM: voltage compatibility.

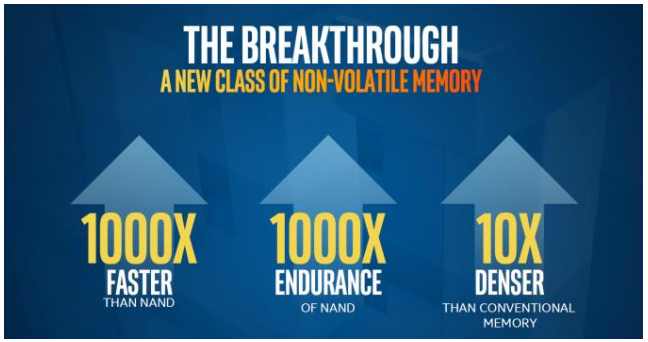
High On-state current density: several MA/cm^2; ON/OFF ratio depends on array size: 10^6 for 1000×1000 array.

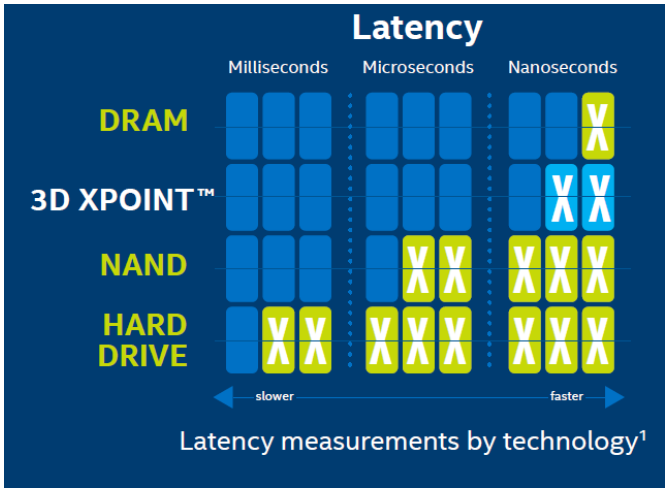
Bidirectional operation.

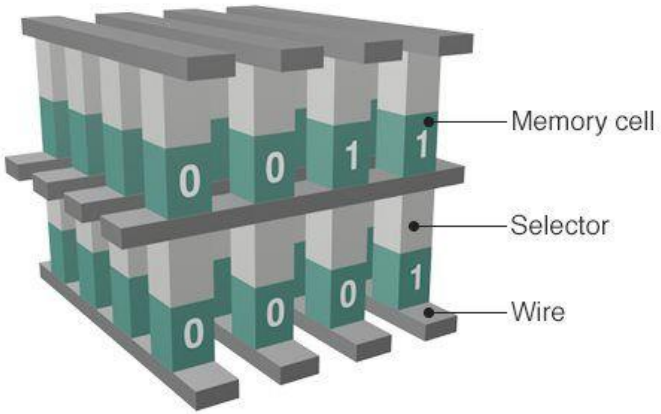
400°C process temperature.

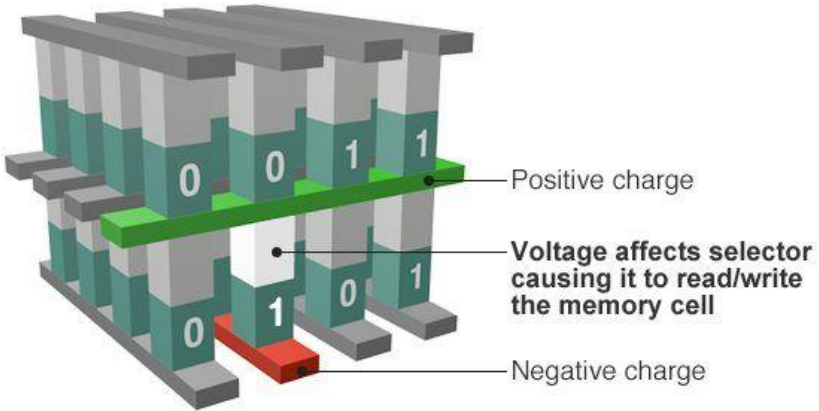
3D Xpoint memory











3D XPoint is a non-volatile memory technology announced by Intel and Micron in July 2015. Though details of the materials and physics of operation were not disclosed, storage density is claimed to be similar to flash memory, durability better and operating speed faster than flash memory though slower than dynamic RAM.

Bit storage is based on a change of bulk resistance, in conjunction with a stackable cross-gridded data access array.

3D XPoint technology uses its new material to switch the resistance state, so it doesn't rely on less reliable and more expensive elements, such as Memristor's titanium dioxide and platinum films or PCM's silver filaments, which wear out over time.

To be continued: