Budapest University of Technology and Economics
Department of Electron Devices

Technology of IT Devices

Lecture 8

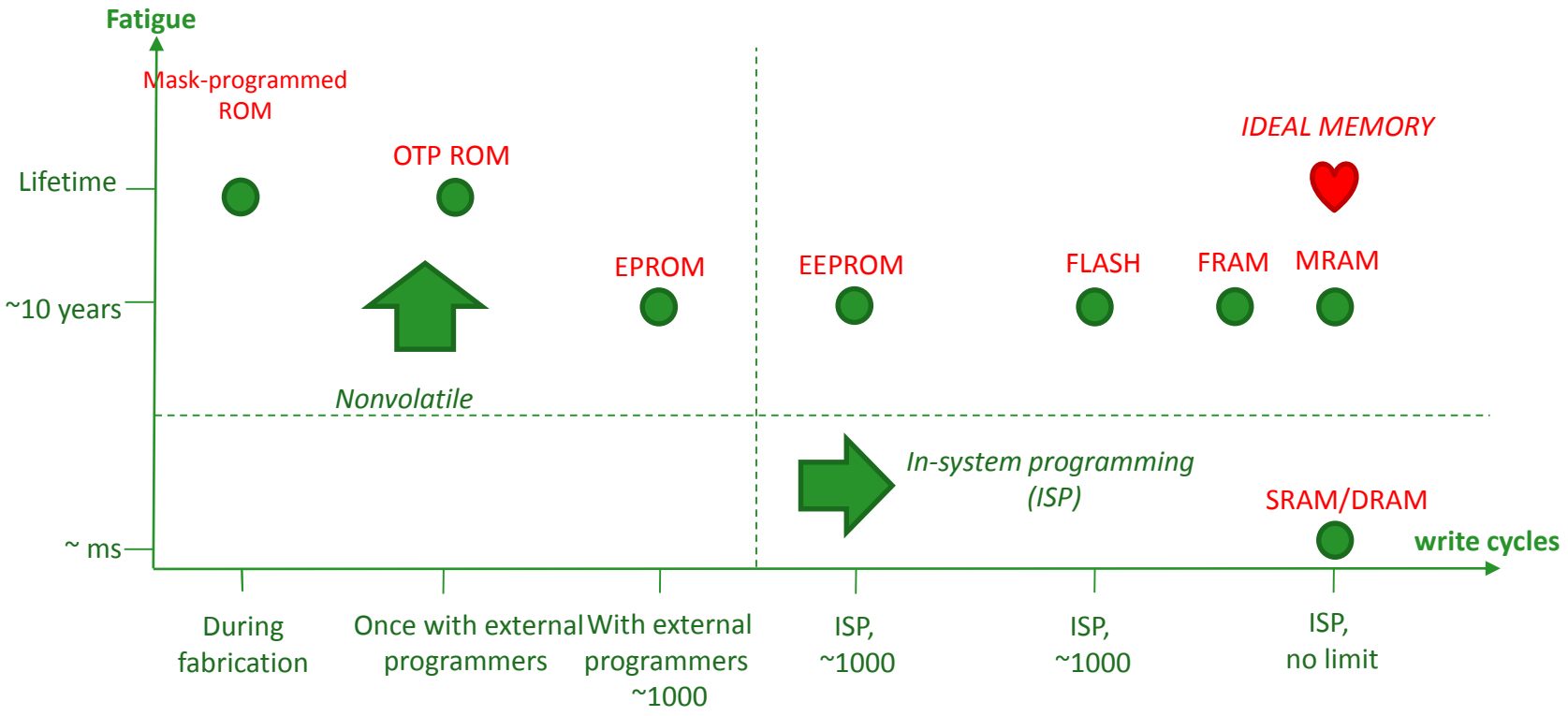
Memory II.

ROM memory

ROM memory

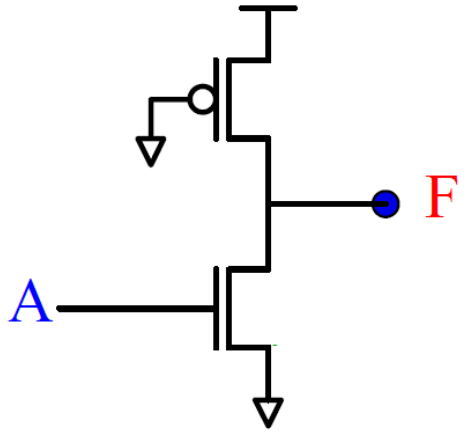
- Mask-programmed ROM
  - Logic gates with passive load
  - NOR or NAND type
- OTP ROM
  - fuse, antifuse
- EEPROM
- FLASH EEPROM

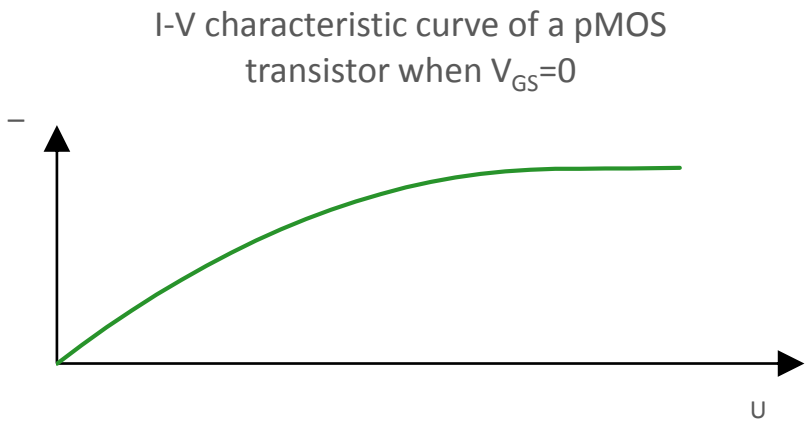
Semiconductor Memory Classification - Reminder



Pseudo nMOS logic gates

- In the case of CMOS logic gates two transistors switch in complementary phase
  - This requires 2N transistors (N is the number of inputs)
  - In fact, only one transistor and one resistor could be enough
  - However, in IC technology, the area occupied by resistors is very large
  - We can use a transistor in the linear region.
    - It acts like a non-linear resistor

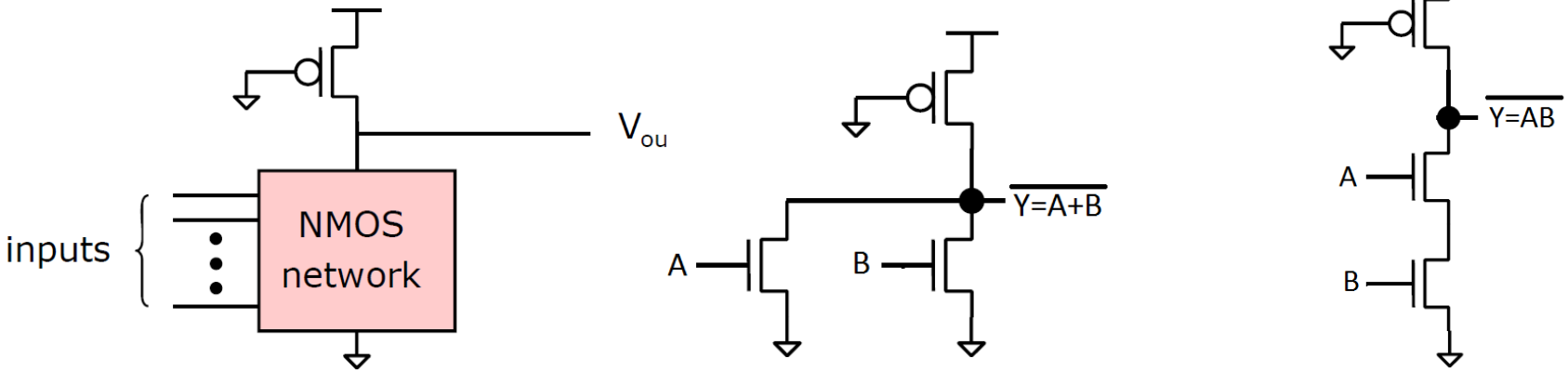




Pseudo nMOS logic gates

- The operation of the logic gate:
  - The switching transistor is called the driver, the upper transistor which acts like a resistor is called the load transistor.
  - If A=0, the nMOS transistor is closed (open circuit), the output is connected to VDD.
  - If A=1, the nMOS transistor conducts current (its resistance is low). The output voltage depends on the ratio of the transistor resistances.
  - $V_L = \frac{R_D}{R_D + R_L} V_{DD}$
  - Problems:
    - This inverter is not a rail-to-rail logic gate!
    - There is another name for this type of logic gate: ratio type logic gate.
    - The voltage level of the logic 0 depends on the ratio of the transistor sizes. If we would like to approximate the ground potential, $R_D << R_L$, so W/L driver >> W/L load

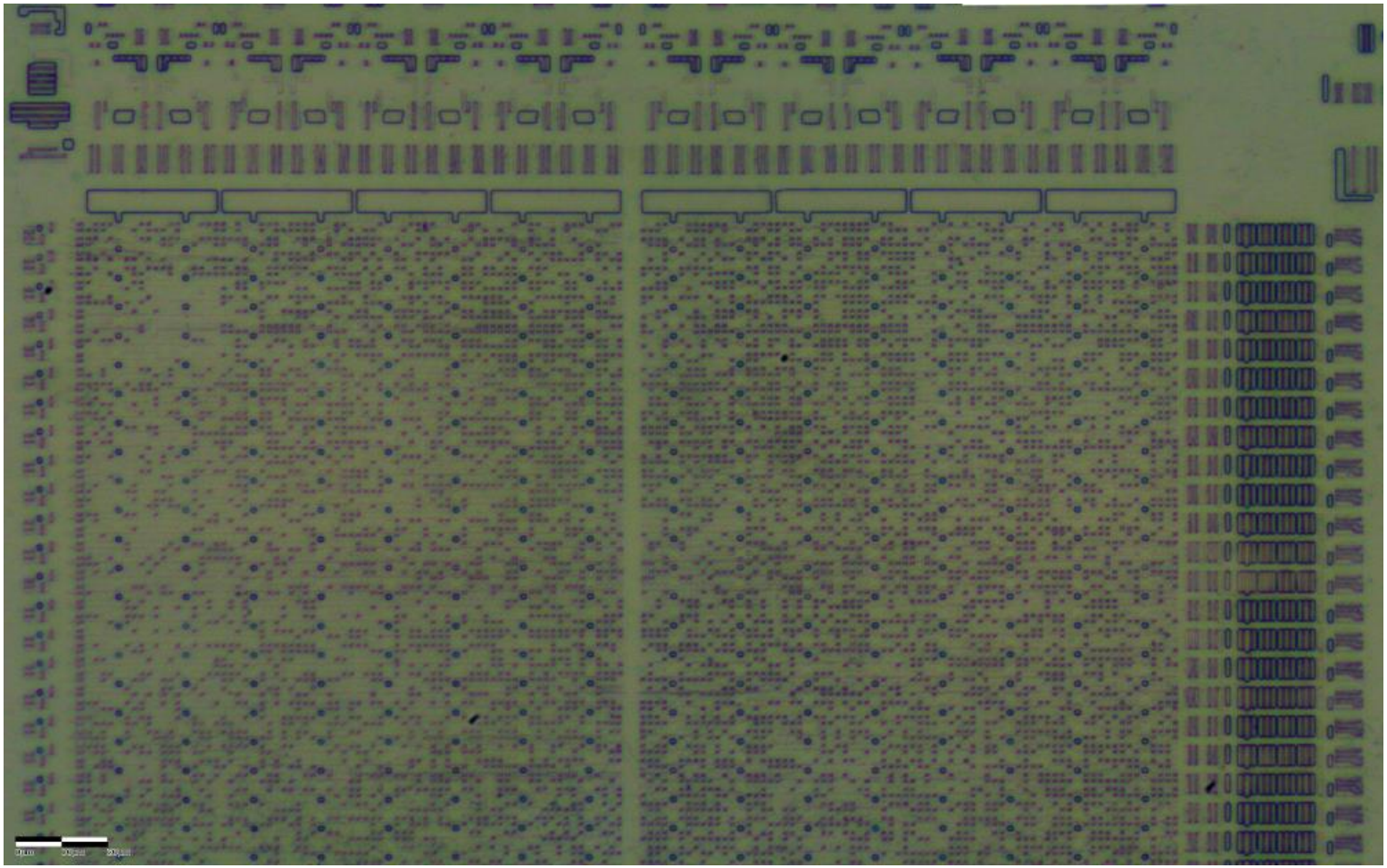
Pseudo nMOS logic gates



- Advantages
  - Simpler circuit
  - N+1 transistors are required (N is the number of inputs)
  - Lower input capacitance
- Drawbacks
  - Not rail-to-rail.
  - Transition time of 0->1 change is higher than 1->0
    - Caused by the higher resistance of the load transistor ($t_{pd} \sim RC$)
  - If the output is 0, it has static power consumption

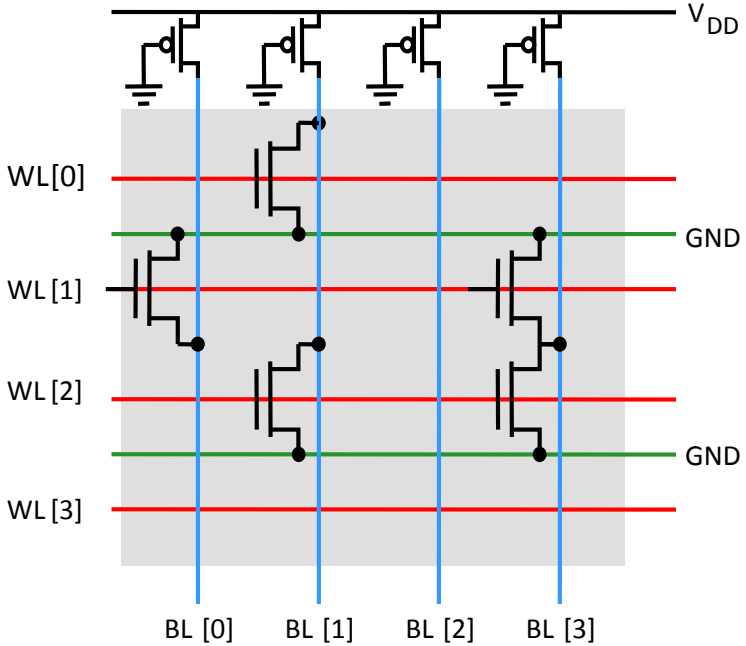
Mask-programmed ROM (MROM)

- The information is stored in it during fabrication
  - The terminology "mask" comes from integrated circuit fabrication, where regions of the chip are masked off during the process of photolithography.
  - The cost of a mask is high, so it is most suitable for mass production
    - Per bit, mask ROM is more compact than any other kind of semiconductor memory.
  - In SoC devices and microprocessors the microcode and the lookup tables are stored in this way
  - In the case of microcontrollers, the bootloader, the C runtime library, and the peripheral library are stored in a mask-programmed memory unit
    - In this way the user flash memory can be smaller

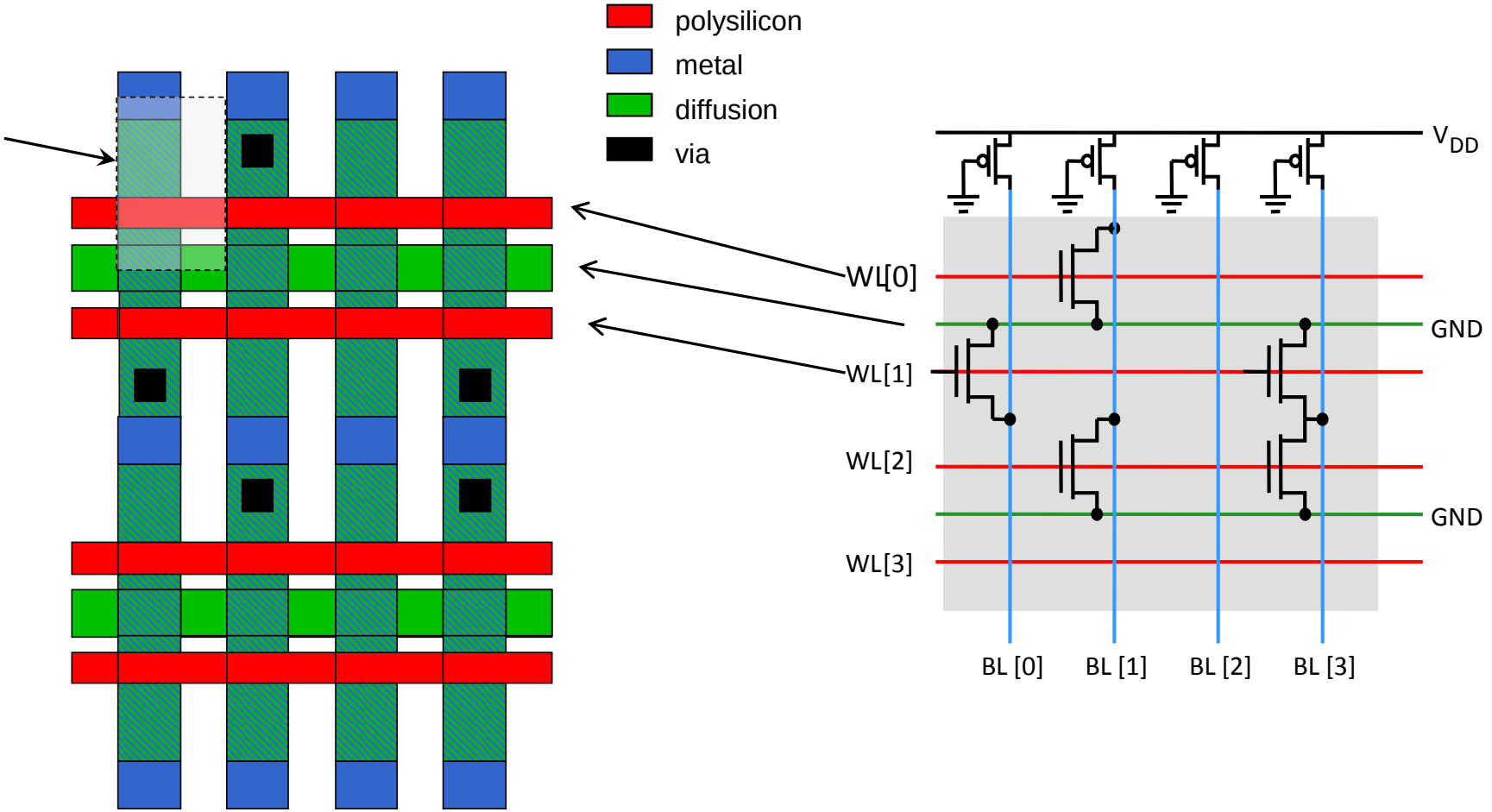


Source and history: http://zeptobars.ru/en/read/FTDI-FT232RL-real-vs-fake-supereal

Mask-programmed ROM - NOR type

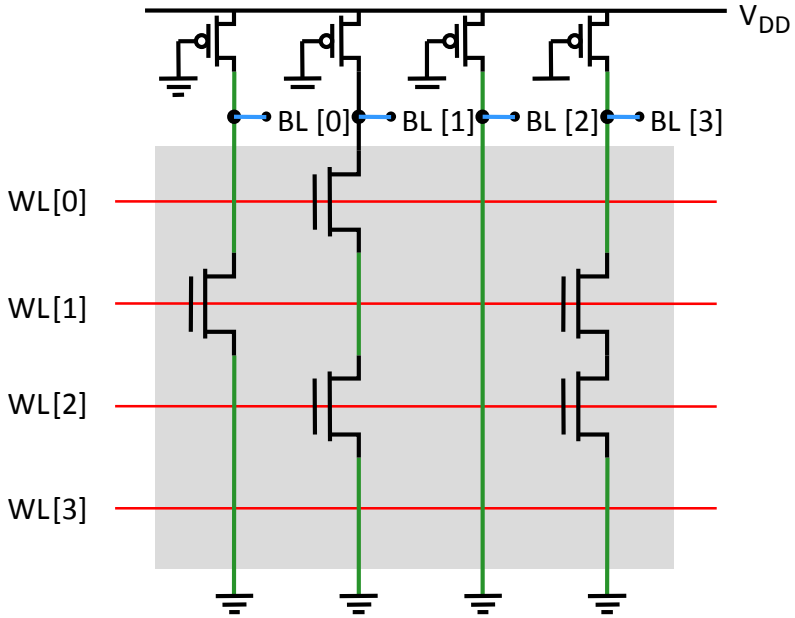


- The information is stored by the presence of a transistor in a cell:
  - If there is a transistor: it pulls the bit line to the ground when its line is selected (logic 1 on its gate),
  - If there is no transistor: the p-MOS (that is always open) at the top of the bit line pulls it up to logic 1.
  - The activated transistor connects the bitline to the ground
  - In the viewpoint of the bitline there is a pseudo NMOS NOR gate

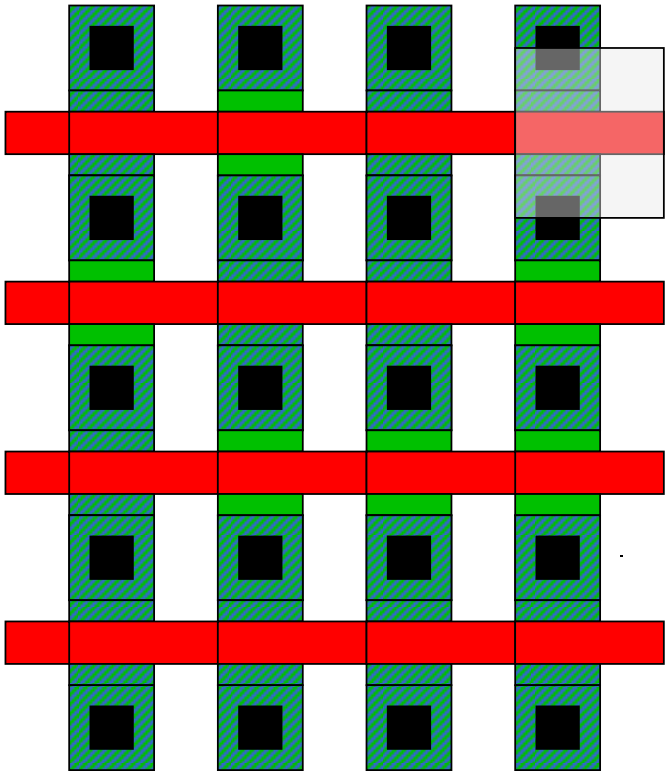


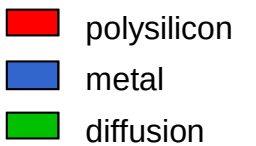
- Via programming.
  - The source of the n-type transistor is connected to the bit line with a via (black) if a transistor is needed in a cell.
  - The cell size is large due to the ground line

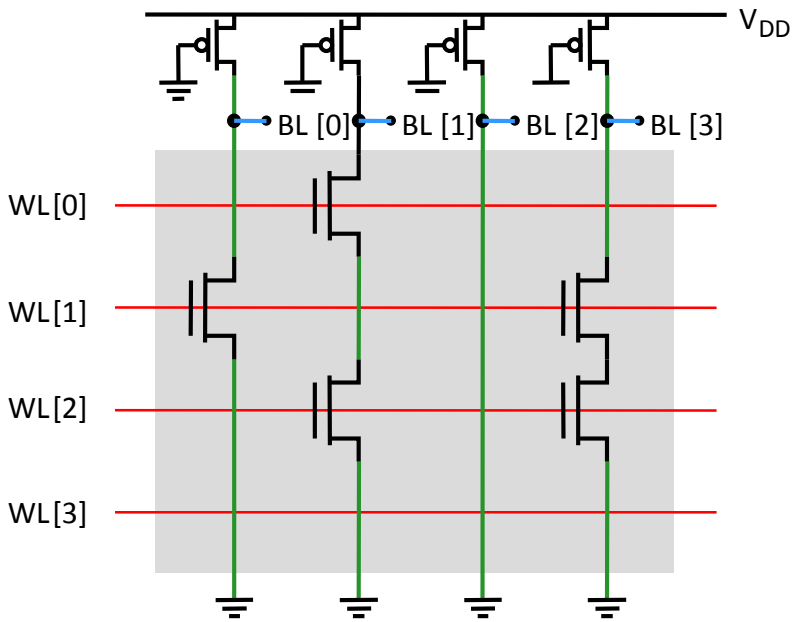
NAND ROM



- The transistors are connected in series, so we get a pseudo NMOS NAND gate
  - During the readout every word line is logic 1, except for the one selected.
  - If there is no transistor in a cell, the potential is pulled down to logic zero.
  - If a transistor is present, it stays closed, so the bit line is pulled up to logic 1.







- The layout
  - No need for a ground in every cell, so the cells are smaller than in NOR-type ROMs.
  - Due to the transistors connected in series, this type of ROM is slower.
  - Task: there is an error in the layout. Please find it.

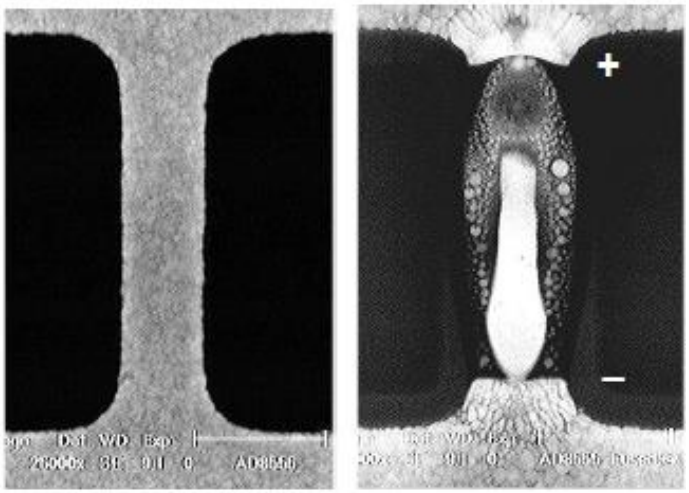
One-time programmable ROM

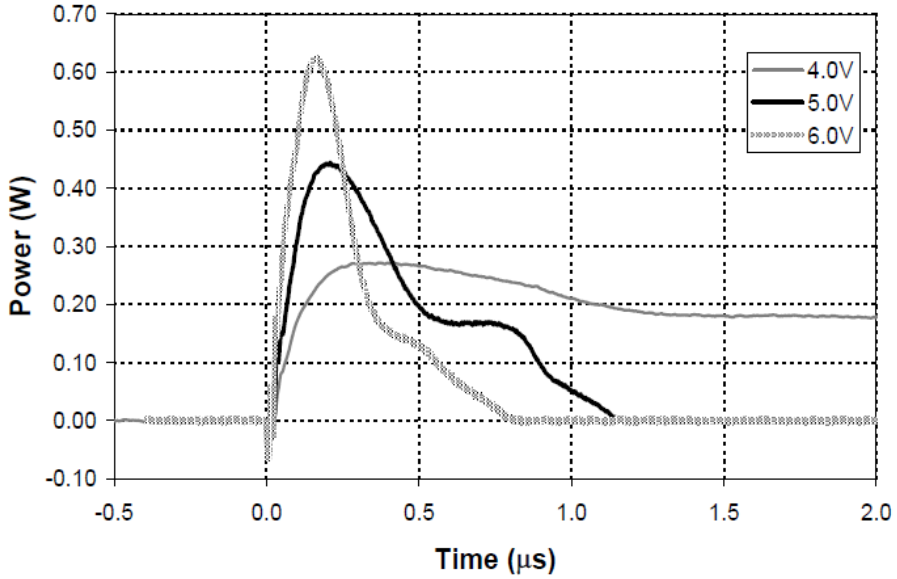
- Firmware
- On-chip configuration, even during operation
  - Calibration constants
  - Encryption keys
  - chip serial number
- The data in them is permanent and cannot be changed.
- It is a form of digital memory where the setting of each bit is locked by a fuse or antifuse.
- Fuse: short circuit, if it is burnt out (by a high energy impulse) it becomes an open circuit
  - Thinned high resistivity polysilicon layer
  - Antifuse: It conducts if it is burnt out, otherwise it is an open circuit. The lower the resistance, the lower the propagation delay

One-time programmable ROM

- This type of memory is frequently used in:
  - Microcontrollers,
  - Video game consoles,
  - Mobile phones,
  - Radio-frequency identification (RFID) tags,
  - Implantable medical devices,
  - High-definition multimedia interfaces (HDMI)
  - And in many other consumer and automotive electronics products

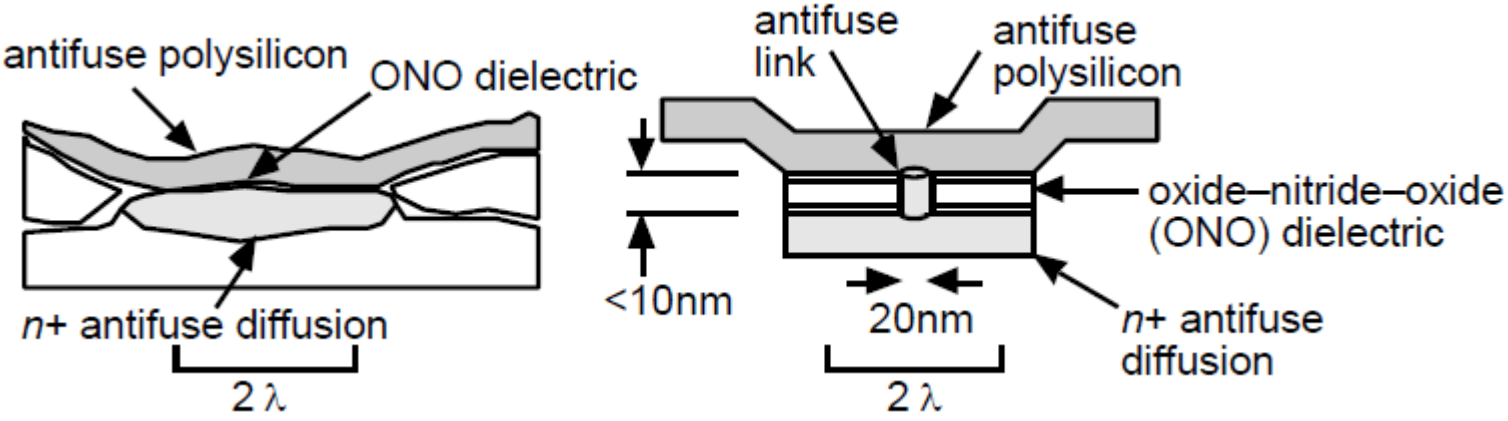
Fuse





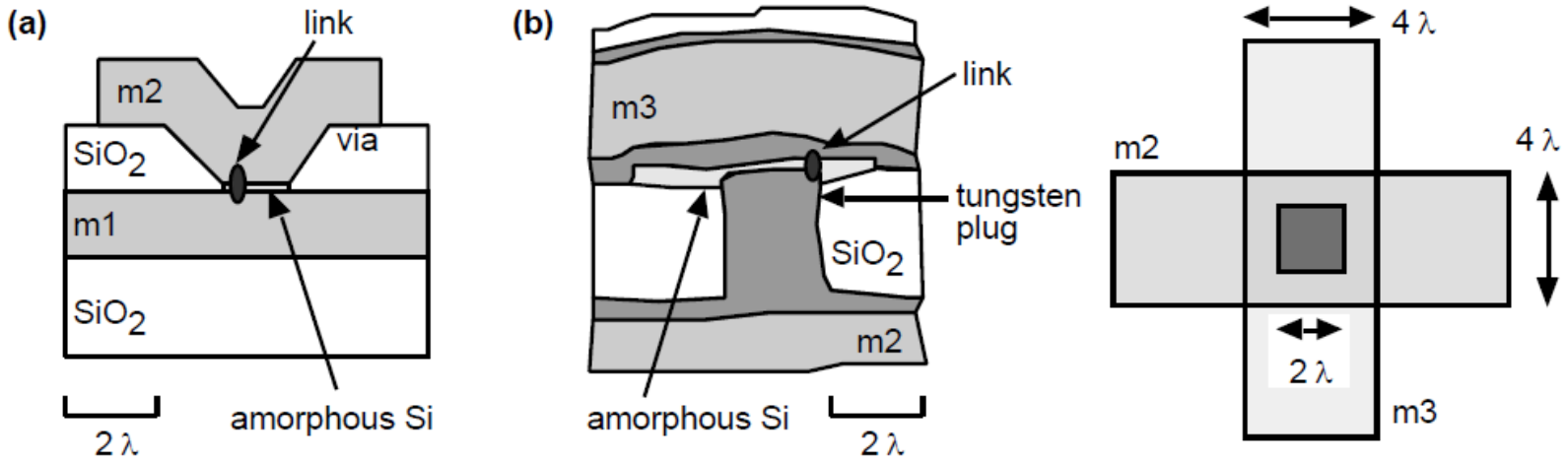
- Polysilicon or silicide, e.g. NiSi (nickel silicon alloy)
- The thin part of it has the highest resistance
  - If too high a current flows,
  - the element rises to a higher temperature and melts,
  - and it becomes an open-circuit

PLICE



- The antifuse technology used in CMOS PROMs.
- An antifuse is an element that starts to conduct when it's been melted, otherwise it is an insulator
- PLICE – programmable low impedance circuit element
  - Thin Si_3N_4 – SiO_2 insulator.
  - an impulse breaks down the dielectric, it melts and a connection is made between the poli-Si and the diffusion.

Metal-to-metal antifuse



- A ViaLink is two electrodes separated by a dielectric
- ViaLink uses undoped amorphous silicon as the ViaLink dielectric
- When a sufficiently high voltage is applied across the amorphous silicon it is turned into a polycrystalline silicon-metal alloy with a low resistance, which is conductive.
- ~80 Ω per contact

EPROM

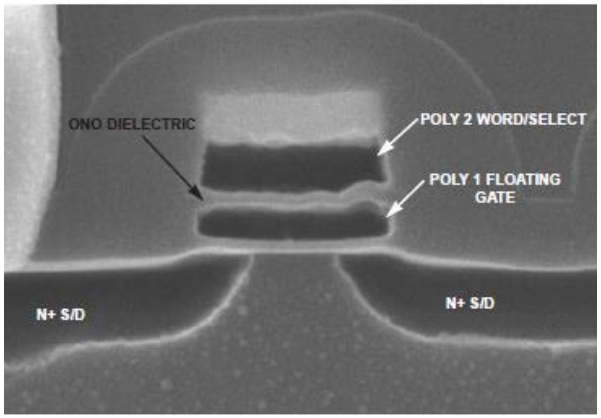
- Erasable programmable read-only memory
- Retains its data when its power supply is switched off
- It is an array of floating-gate transistors individually programmed by an electronic device that supplies higher voltages than those normally used in digital circuits

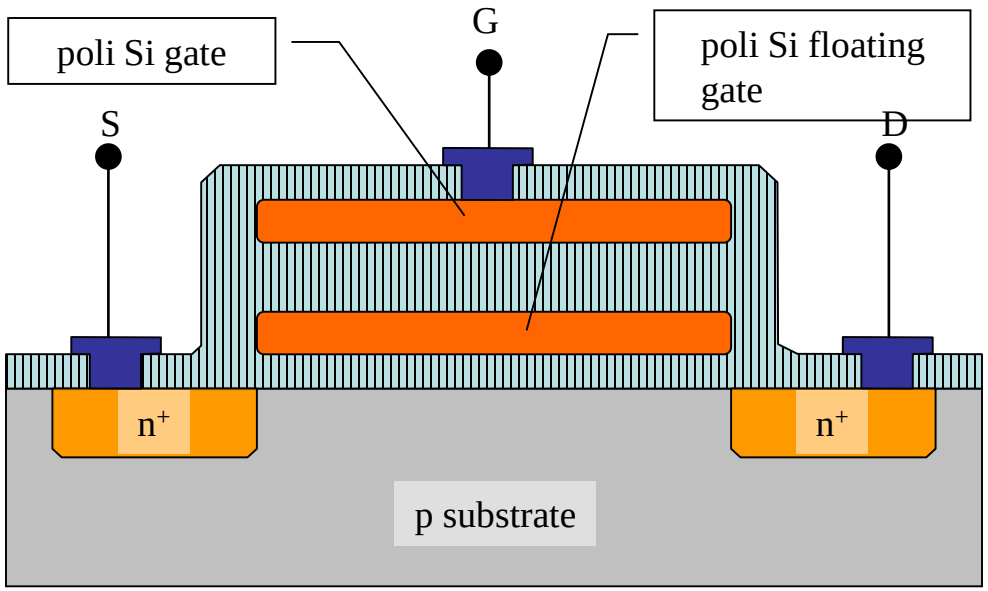
The principle of data storage

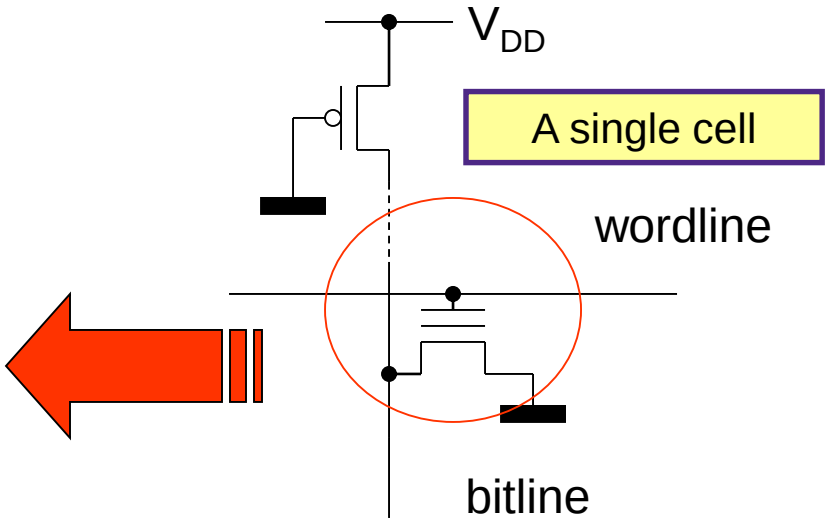
- The information is stored by the threshold voltage of a MOS transistor
  - The threshold voltage can be changed
  - During read process:
    - The transistor conducts/doesn't conduct (SLC – single level cell)
    - It can conduct a different amount of current on a given voltage (MLC – multi level cell (4 states), TLC - triple level cell (8 states))
- Threshold voltage: the gate-source voltage when the inversion channel will occur
- The threshold voltage depends on the charge stored in the insulator
  - E.g. in the case of n-channel, negative charge blocks the channel, thus the threshold voltage will be higher
  - The positive charge helps form the inversion layer. In extreme cases the transistor can conduct current when V_{GS}=0.
- If we can create a semiconductor construction which can store charge, then we can increase or decrease the threshold voltage of the memory transistor.
  - The device becomes programmable
  - Different constructions:
    - Floating (not connected polysilicon) gate, so called floating gate
    - Multilayer semiconductor which contains charge trapping sites able to hold an electrostatic charge (Silicon-Oxide-Nitride-Oxide-Silicon - SONOS)

EPROM

- Electrically programmable, can be erased by exposure to a strong ultraviolet light source.
- The information is stored in a Floating Gate Avalanche MOS (**FAMOS**) tr.

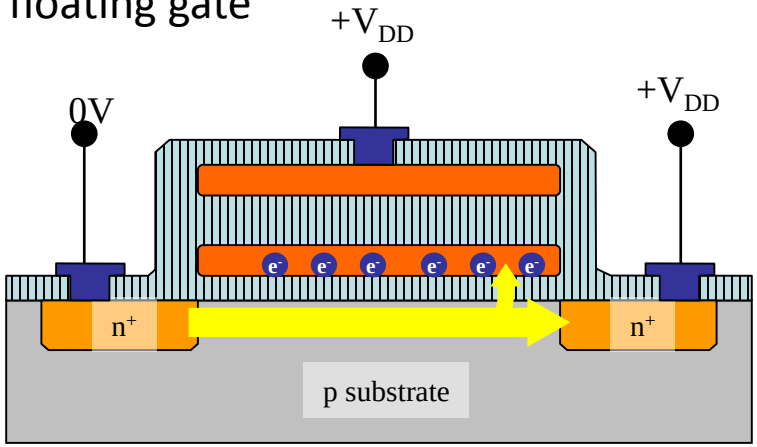






It has a **floating gate** where electrons can be transferred using **avalanche breakdown**.

- Programming
  - The source is connected to the ground, a large positive voltage (~12 V) is applied to the drain.
  - The transistor's channel breaks down – a large current of high energy electrons flows in the channel
  - Some of the high energy electrons get through the potential barrier of the oxide (3.2 eV) and get to the floating gate

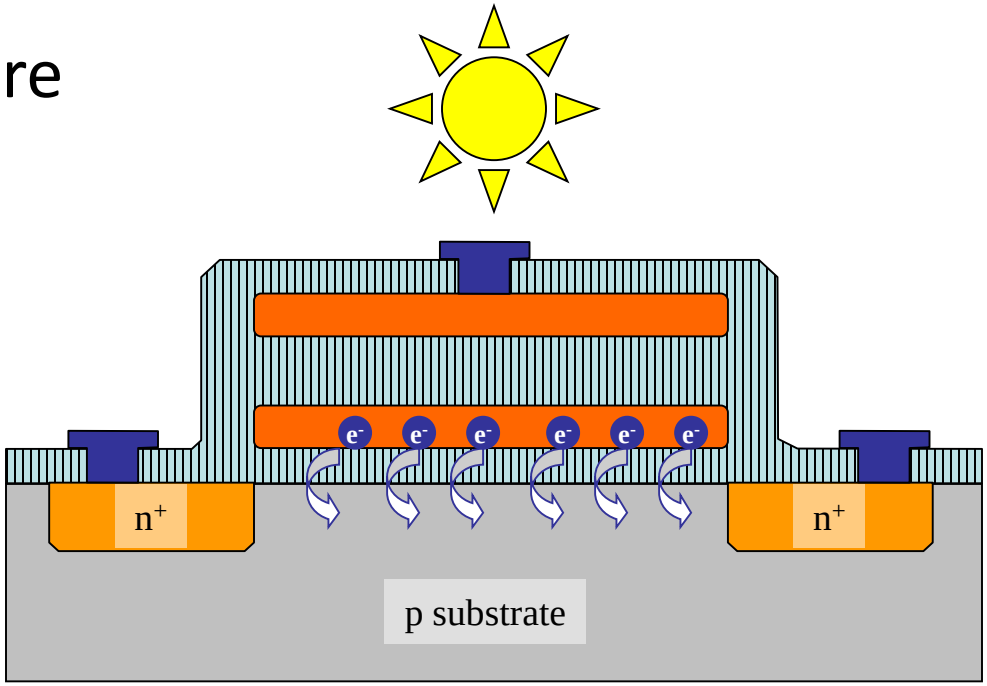


The charge accumulated on the floating gate stays there for years (10 years are guaranteed by the manufacturers)

The electric field of the charge **increases the threshold voltage** of the transistor, thus the channel will not be created when the supply voltage is applied to the gate.

When not programmed, the device operates as a normal MOS FET.

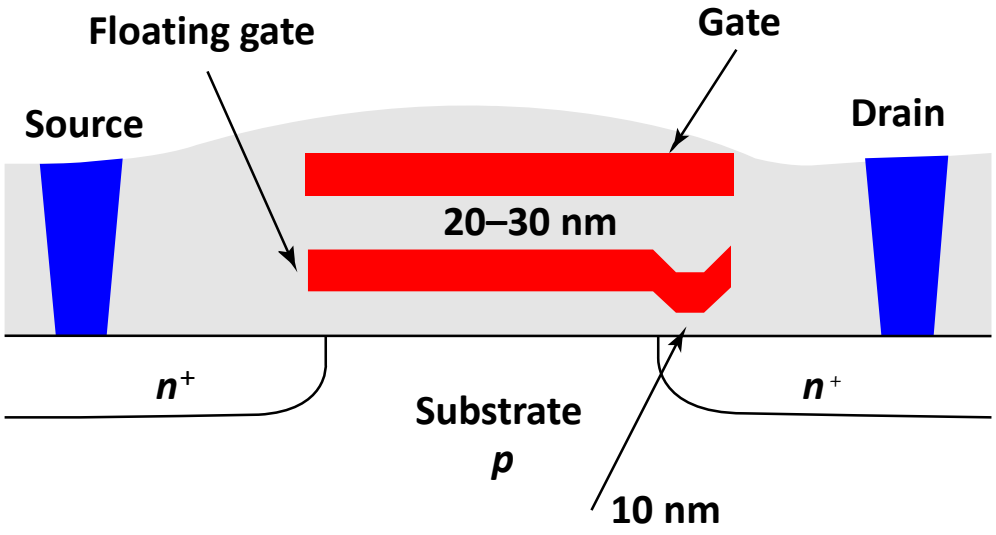
EPROM erasure

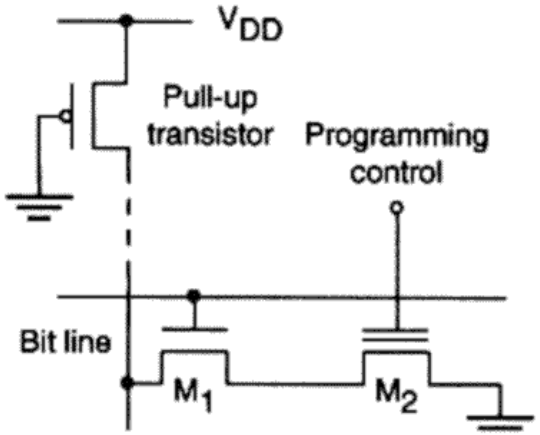


- A window is opened at the top of the EPROM where sunlight can get to the device.
- High energy UV light forces the electrons back to the semiconductor. An EEPROM can be cleared in ~20 minutes.

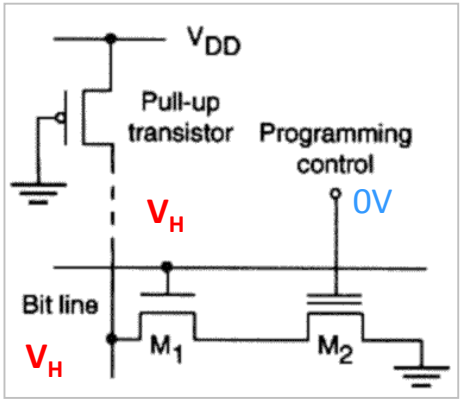
EEPROM

- Electrically erasable programmable read-only memory
- EEPROMs (or E^2PROMs) can be written and erased electronically.
- Each cell consists of a selector and a floating gate transistor.
  - FLOTOX – floating gate tunneling oxide
  - The floating gate transistor is special in that the gate reaches above the drain.
  - The floating gate is separated from the drain by a very thin oxide (5-10 nm)
  - The electrons can get through the oxide by tunneling.

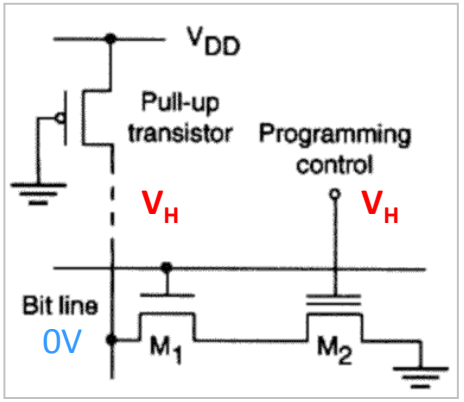




- Programming
  - Gate is grounded, large positive voltage is connected to drain
  - Electrons tunnel from the floating gate towards the drain
  - The floating gate is charged to a positive potential, that decreases the threshold voltage of the transistor to Vth = 0 V.

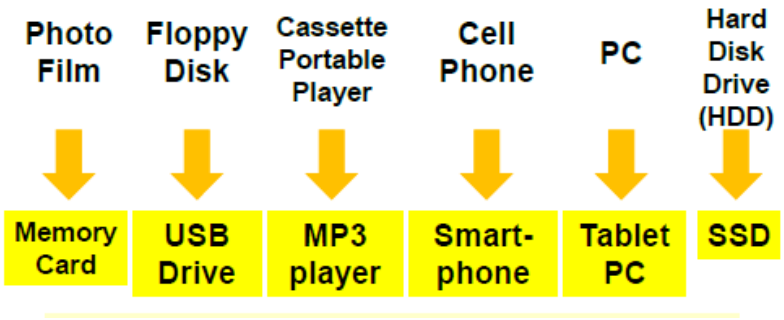


- Erasure
  - Positive voltage is applied to the gate and the drain is grounded
  - A tunneling current of the opposite direction flows, the floating gate is charged to a negative potential, which turns the transistor off.
  - The programming control input is only used for programming – for normal readout, the cell's other transistor is used

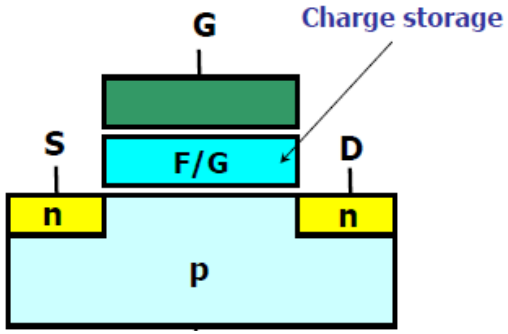


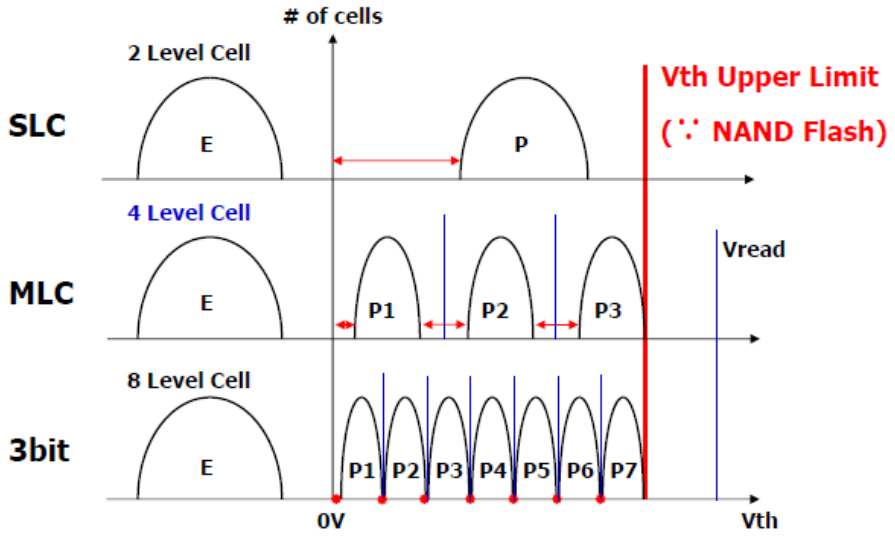
FLASH EEPROM

- The most successful type of memory
- Flash memory (both NOR and NAND types) was invented by Fujio Masuoka while working for Toshiba (~1980)
- The name "flash" was suggested by Masuoka's colleague, Shōji Ariizumi, because the erasure process of the memory contents reminded him of the flash of a camera.
- 1992 – 700nm, 16Mbit
- It has replaced many data storing technologies

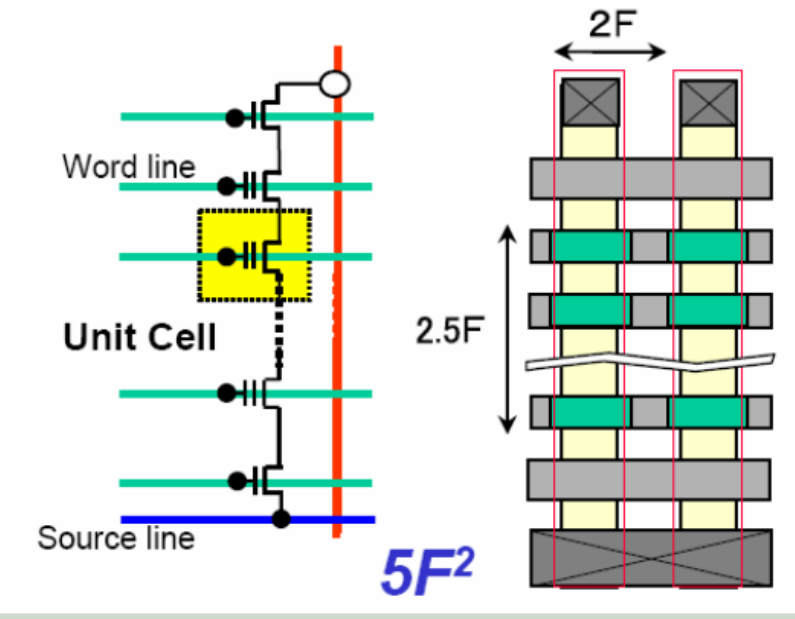
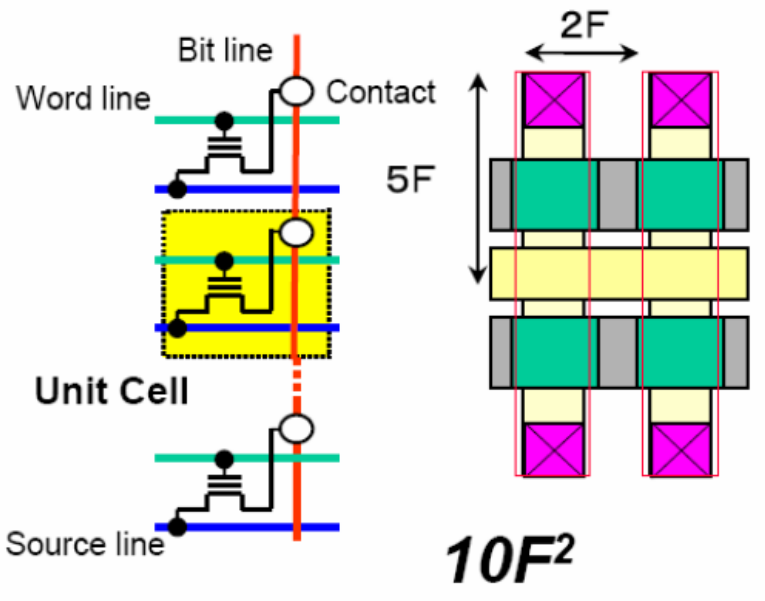
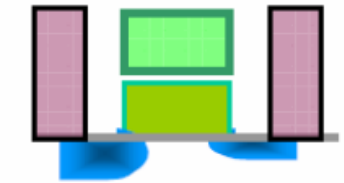


FLASH EEPROM

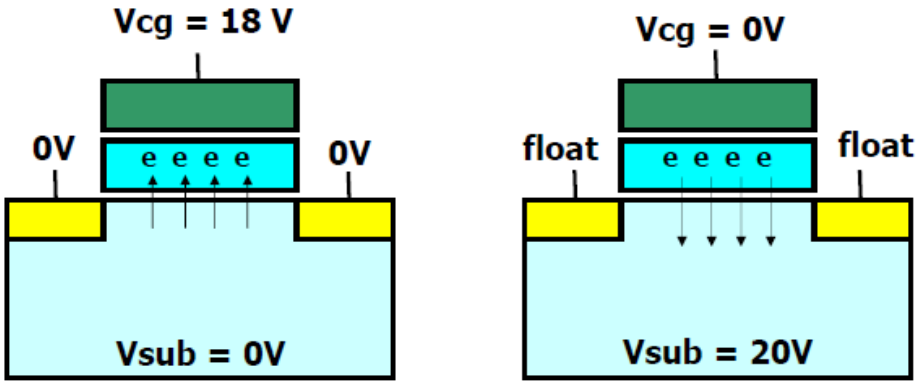




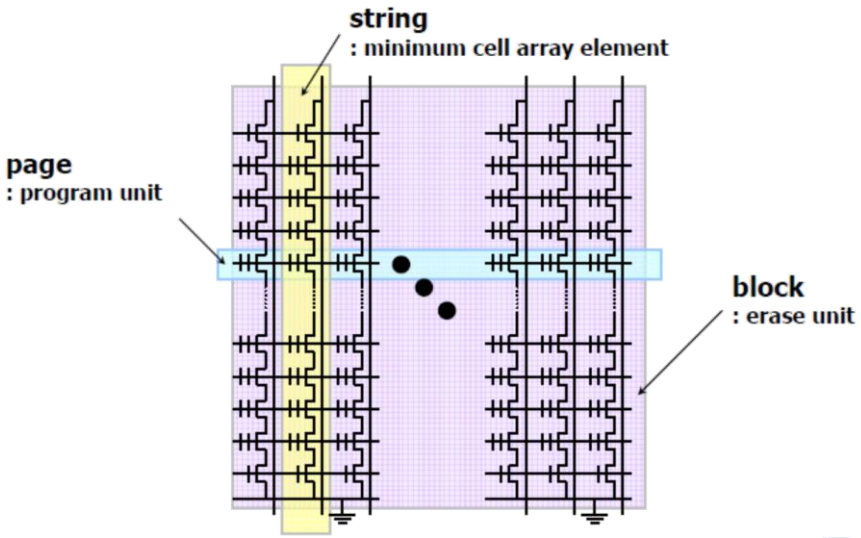
- The two-transistor EEPROM cell is substituted with one special transistor, so the cell size is approximately the same as that of EPROMs.
  - **Single Level Cell (SLC):** two distinct threshold levels
  - **Multi Level Cell (MLC):** several (4 to 8) distinct threshold levels
  - Marketing names:
    - MLC: 2 bits per transistor
    - TLC: (triple-level cell): 3 bits per transistor

	NAND	NOR
Cell matrix		
Cross-section view		
Proper-ties	Small cells, high storage capacity, lower active power, higher write speed (but lower read speed) File storage	Bigger cells, low storage capacity, lower standby power, higher read speed (but lower write speed) Code storage

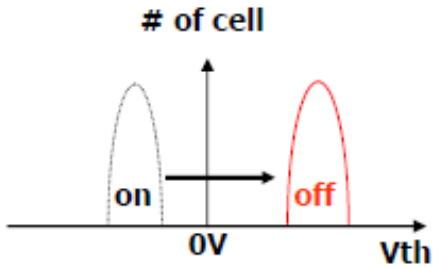
- NOR FLASH
  - The readout is equivalent to that of ROMs (by avalanche breakdown).
  - Erasure (by tunneling)
  - The structure is similar to NOR ROM/EPROM
- NAND FLASH
  - Programming and erasure is done by tunneling



Terminology

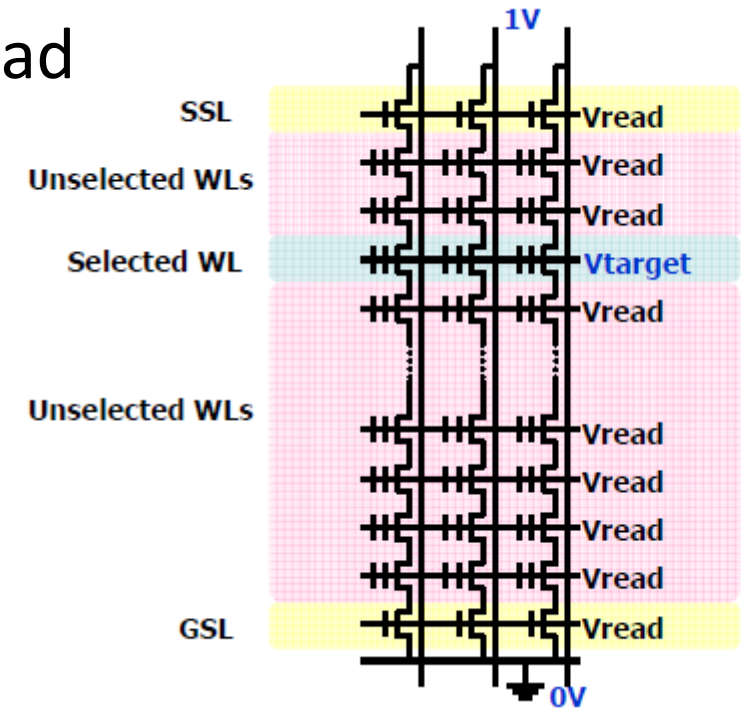


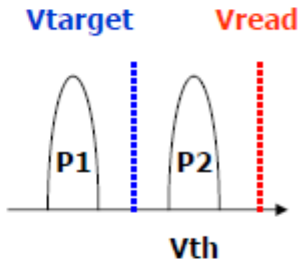
- Page
  - ~256-512 bytes, program unit
- String
  - Storing transistors + two access transistors
- Block
  - One erase unit: 64-128 pages



- In the case of SLC there are two states
  - ON: the transistor has negative threshold voltage, so it always conducts
  - OFF: the transistor has positive threshold voltage, so on the readout voltage it doesn't conduct

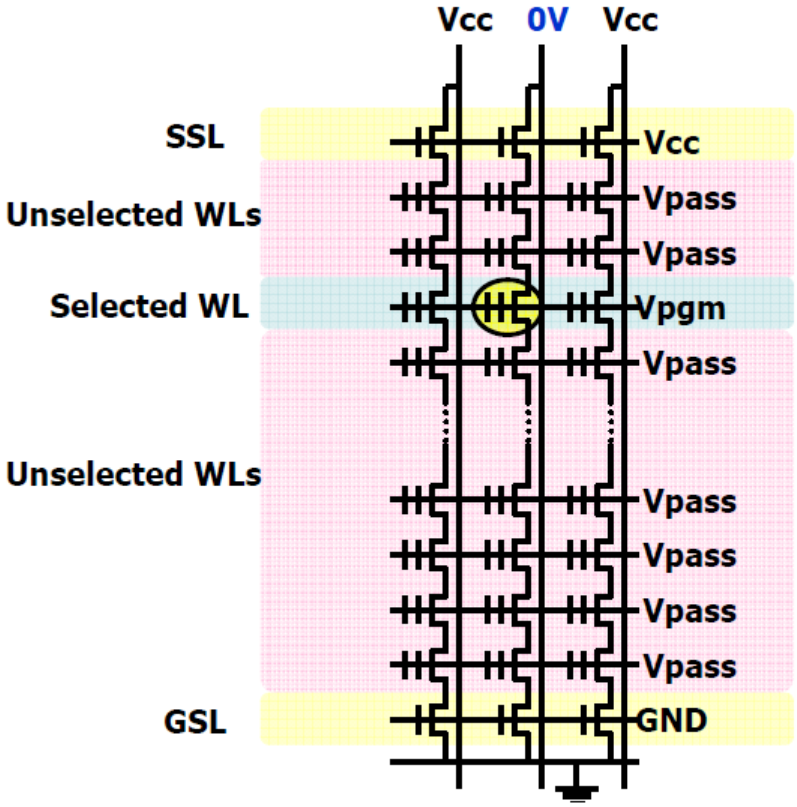
NAND Flash read

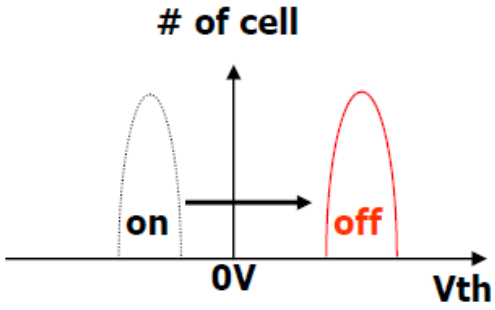




- Every column starts with a string select line (SSL) transistor, and ends with a ground select line transistor
- Read process
  - Source line is grounded
  - The unselected wordlines are connected to the high voltage (reading voltage), the activated wordline is connected to a small voltage (a bit higher than 0)
  - If the transistor conducts, it stores logic 1, otherwise it stores logic 0
  - In the case of MLC we have to try it on different target voltage levels

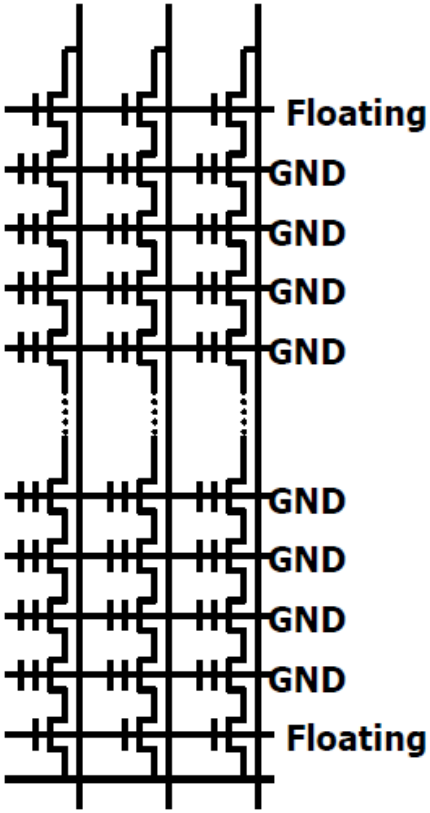
NAND FLASH programming

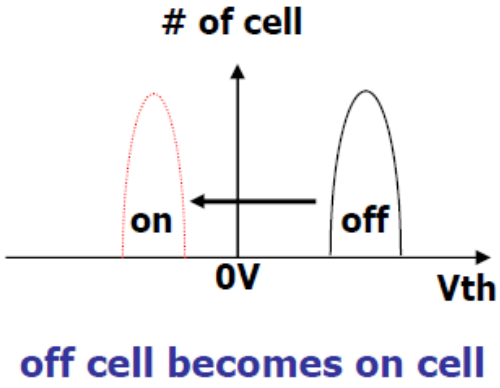




- The selected WL is connected to the program voltage. ~20V
- The unselected WLs are connected to the pass voltage (lower voltage)
- The threshold voltage of the selected transistor becomes higher, and the device turns off

NAND FLASH erasure



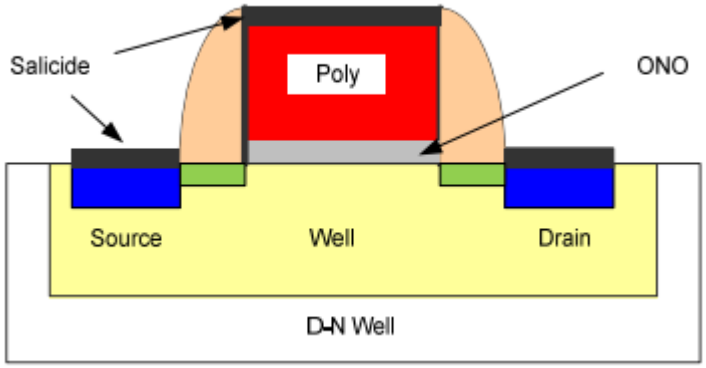


off cell becomes on cell

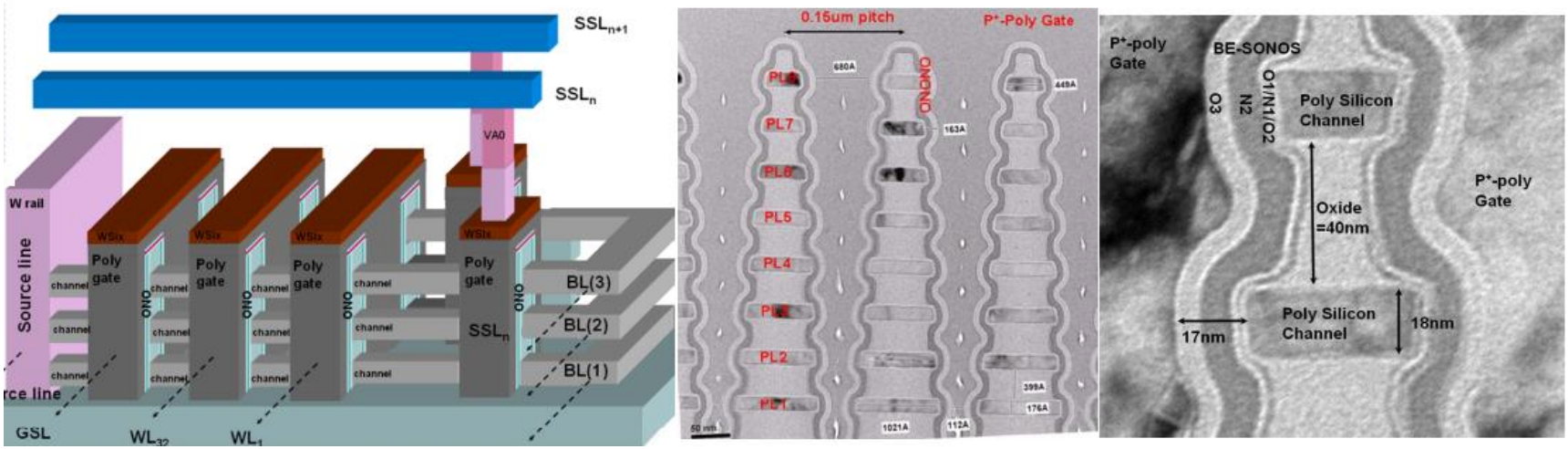
- The bulk of the whole block is connected to the erase voltage. All the gates are connected to ground potential. The threshold voltage becomes negative (all transistors are turned on)

Modern FLASH – SONOS transistor

- Silicon-Oxide-Nitride-Oxide-Silicon
- A standard polysilicon N-channel MOSFET transistor with the addition of a small sliver of silicon nitride inserted into the transistor's gate oxide
- The sliver of nitride is non-conductive but contains a large number of charge trapping sites able to hold an electrostatic charge
  - Devices require much lower write voltages, typically 5–8 V (instead of 12-20V)
  - It has higher reliability than the floating gate one (~100 million write cycles)

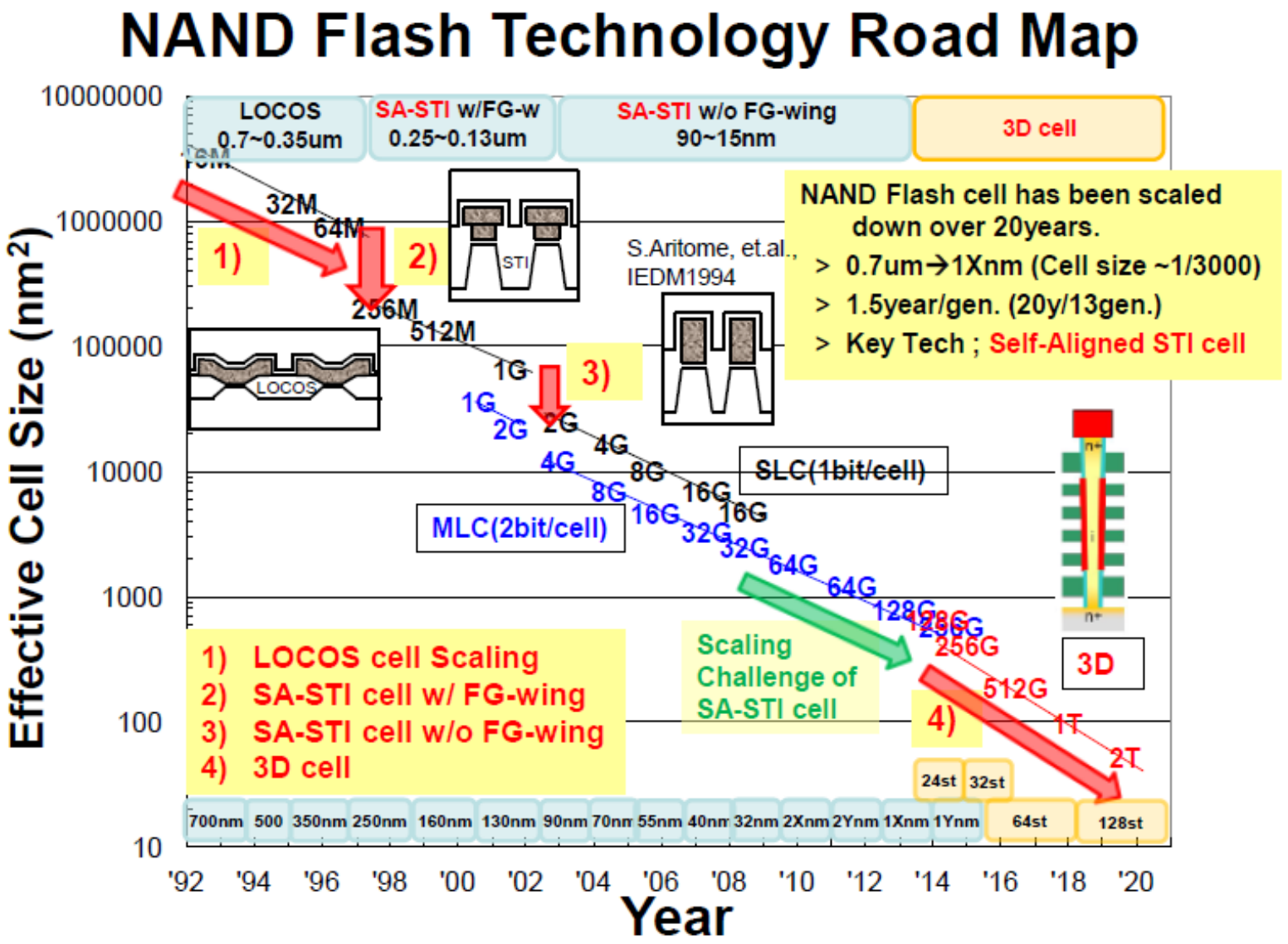


VNAND – vertical NAND



- The transistor is replaced with a thin-film transistor (TFT)
- The channel material is polysilicon
- The charge storing layer is an ONO structure
- 24-32 layers can be produced

Source: http://www.samsung.com/semiconductor/products/flash-storage/v-nand/


NAND Flash Technology Road Map
LOCOS
0.7~0.35um
SA-STI w/FG-w
0.25~0.13um
SA-STI w/o FG-wing
90~15nm
3D cell
NAND Flash cell has been scaled down over 20years.
> 0.7um→1Xnm (Cell size ~1/3000)
> 1.5year/gen. (20y/13gen.)
> Key Tech ; Self-Aligned STI cell
S.Aritome, et.al., IEDM1994
32M
64M
256M
512M
1G
2G
4G
8G
16G
32G
64G
128G
256G
512G
1T
2T
1)
2)
3)
4)
SLC(1bit/cell)
MLC(2bit/cell)
1) LOCOS cell Scaling
2) SA-STI cell w/ FG-wing
3) SA-STI cell w/o FG-wing
4) 3D cell
Scaling Challenge of SA-STI cell
3D
24st
32st
64st
128st
700nm
500
350nm
250nm
160nm
130nm
90nm
70nm
55nm
40nm
32nm
2Xnm
2Ynm
1Xnm
1Ynm
Effective Cell Size (nm²)
10000000
1000000
100000
10000
1000
100
10
'92
'94
'96
'98
'00
'02
'04
'06
'08
'10
'12
'14
'16
'18
'20
Year