Complementary Metal Oxide Semiconductor (CMOS)

Franssila: Chapters 26,28

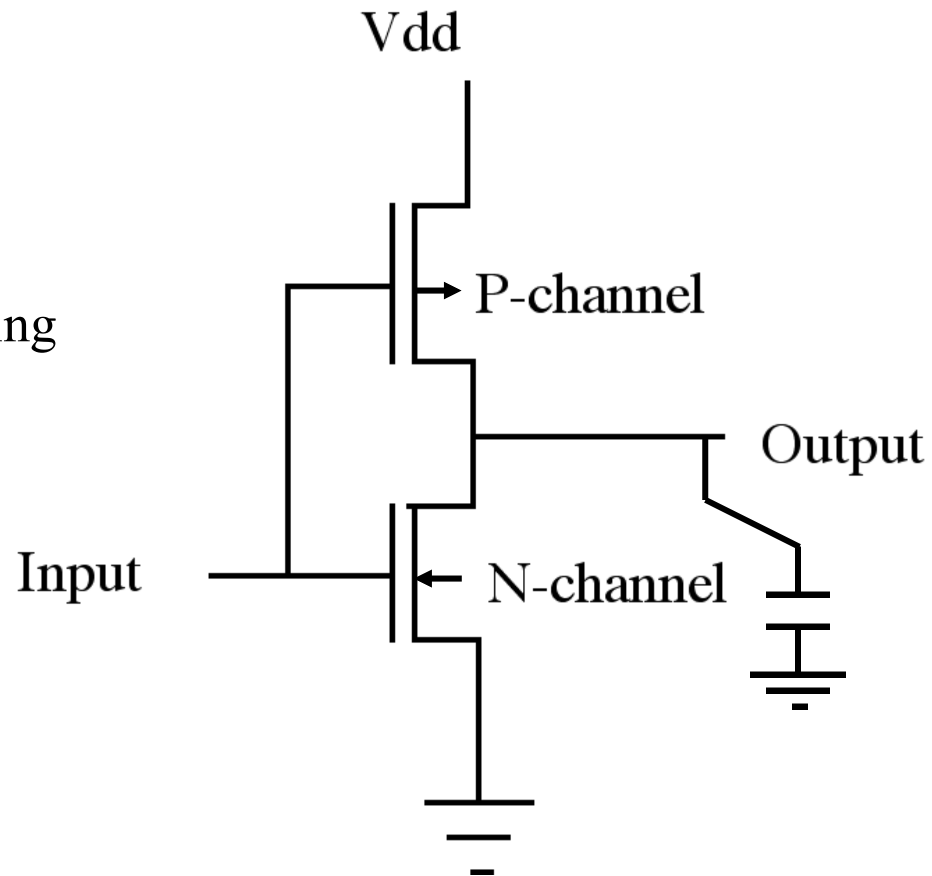
Complementary Metal Oxide Semiconductor (CMOS)

The dominant technology for microprocessors

Low power dissipation through the use of n-type and p-type MOSFETs

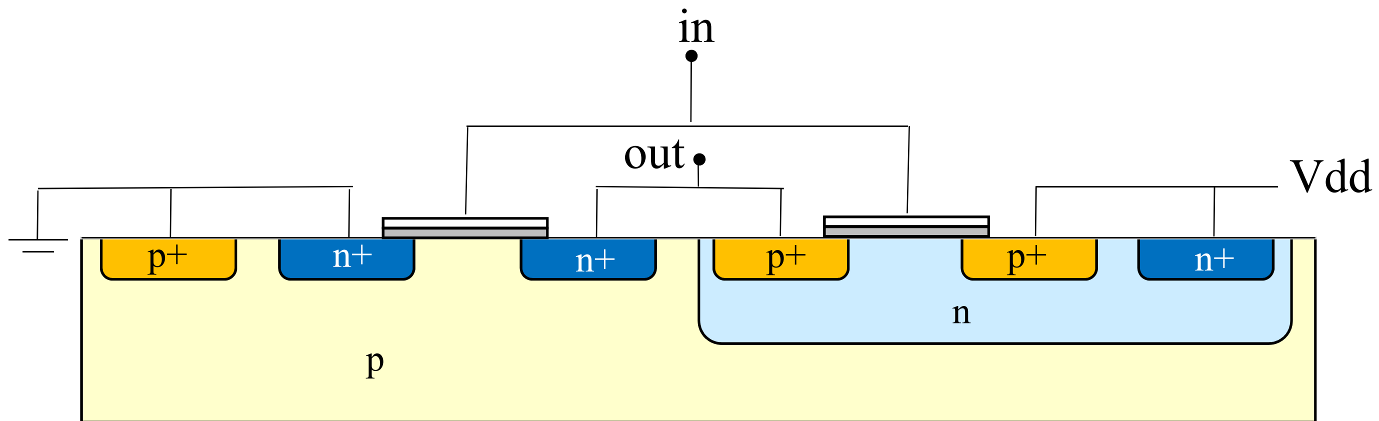
CMOS inverter

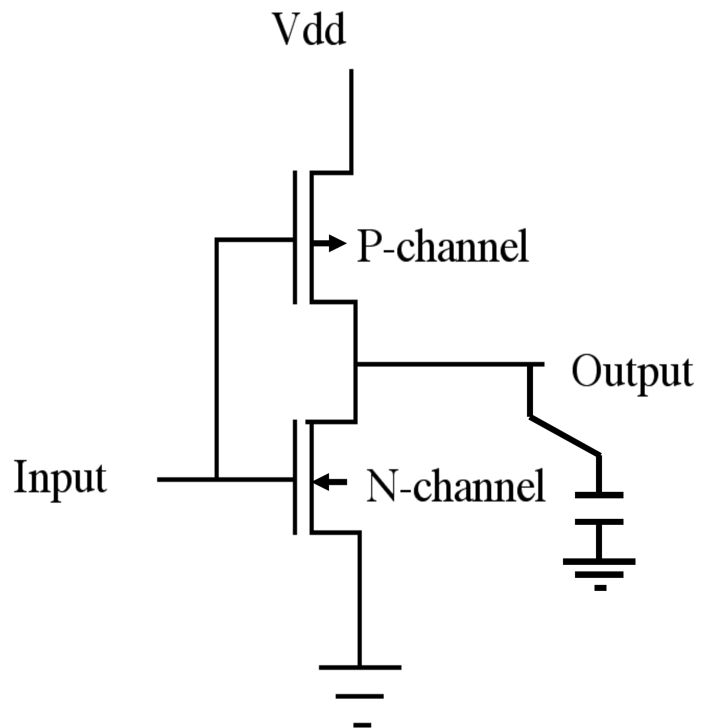
Dissipates little power
except when it is switching

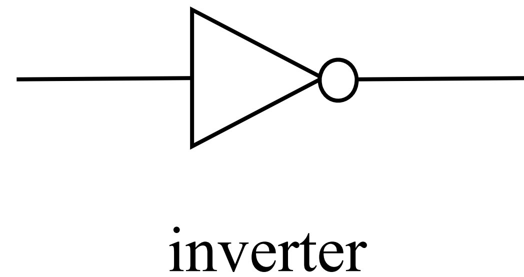


$$E = QV_{dd} = CV_{dd}^2$$

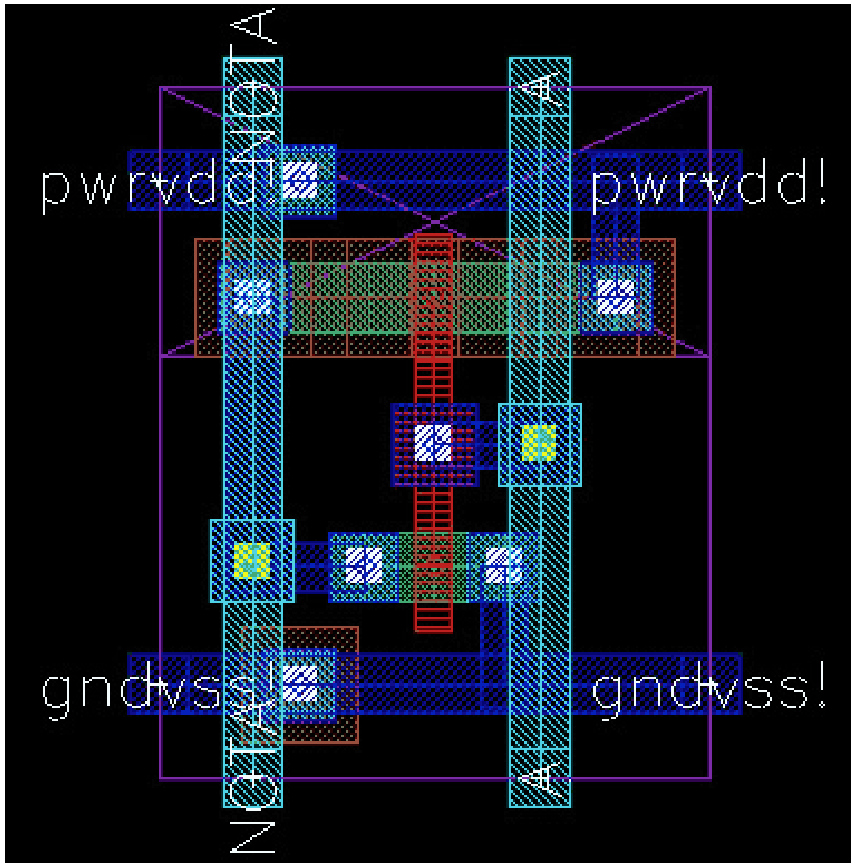
CMOS inverter

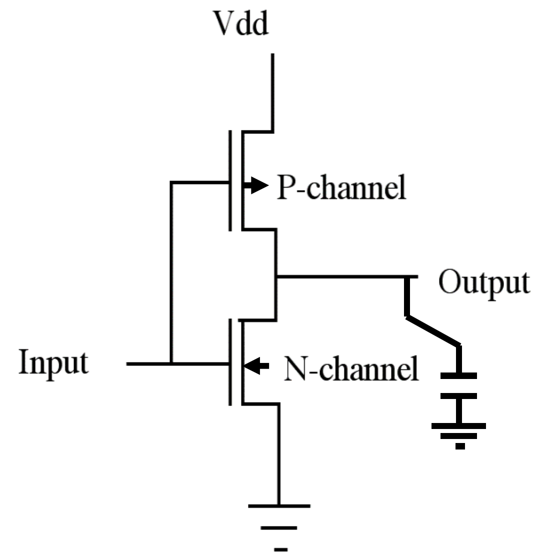


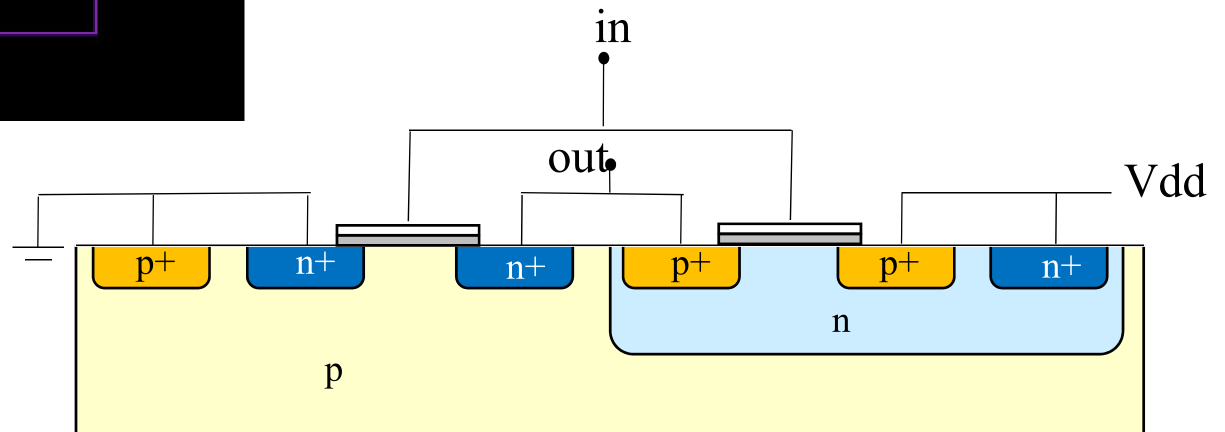




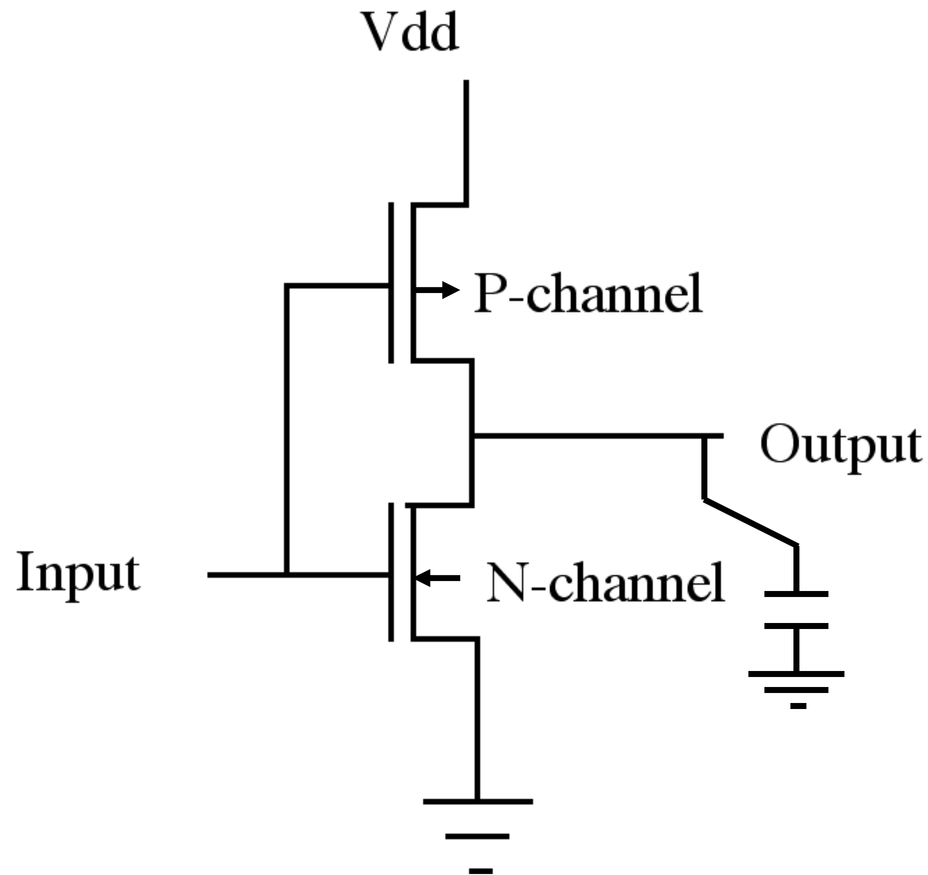
CMOS inverter



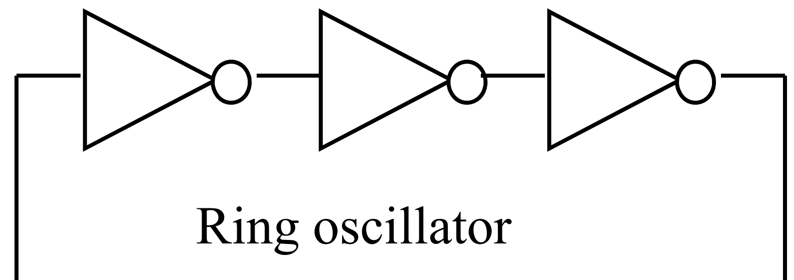




Gate delay

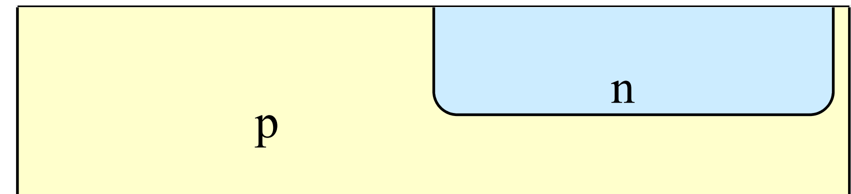


Gate delay is limited by $C_{gate}V_{dd}/I$.



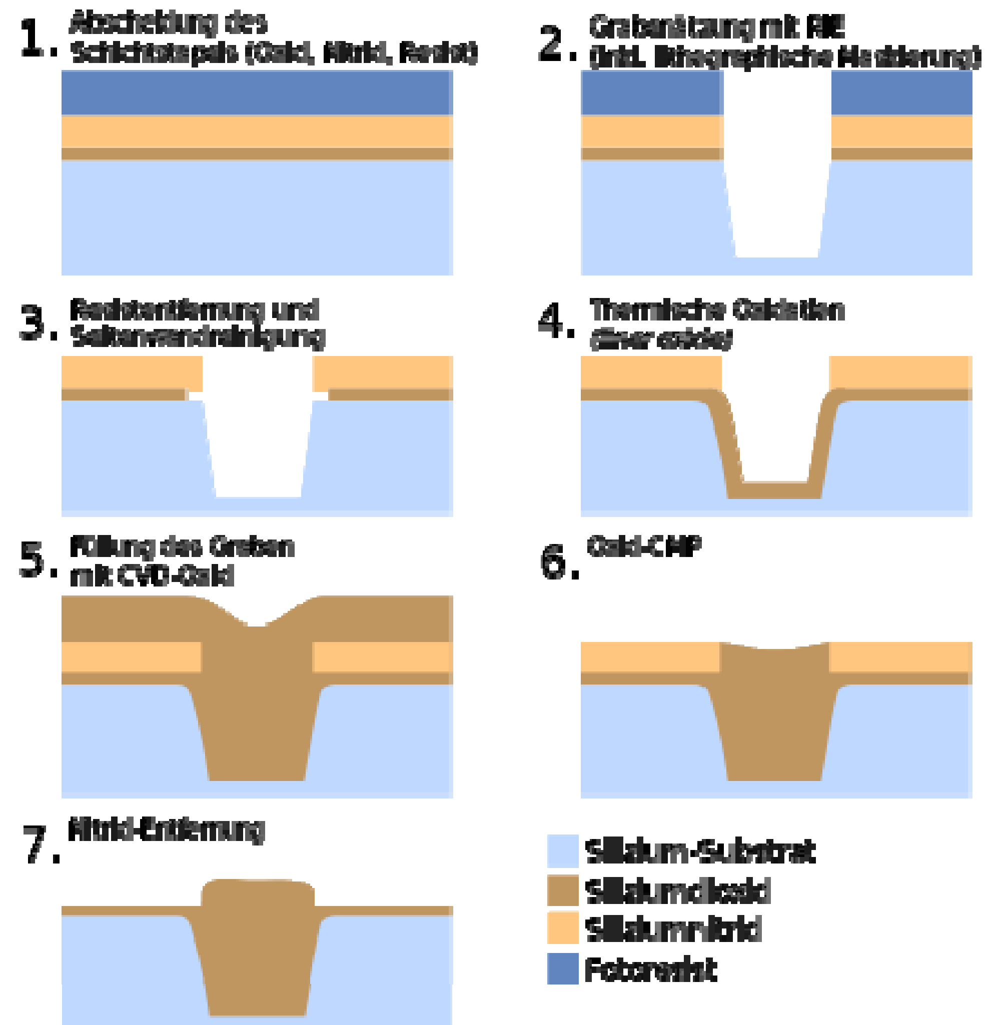
Well formation

Mask oxide
Lithography
Etching oxide
Resist strip and cleaning
Dopant predeposition
(for instance 10^{13} cm^{-2}
@ 40keV)
Drive-in
(500 min. 1000 C)
Oxide etching



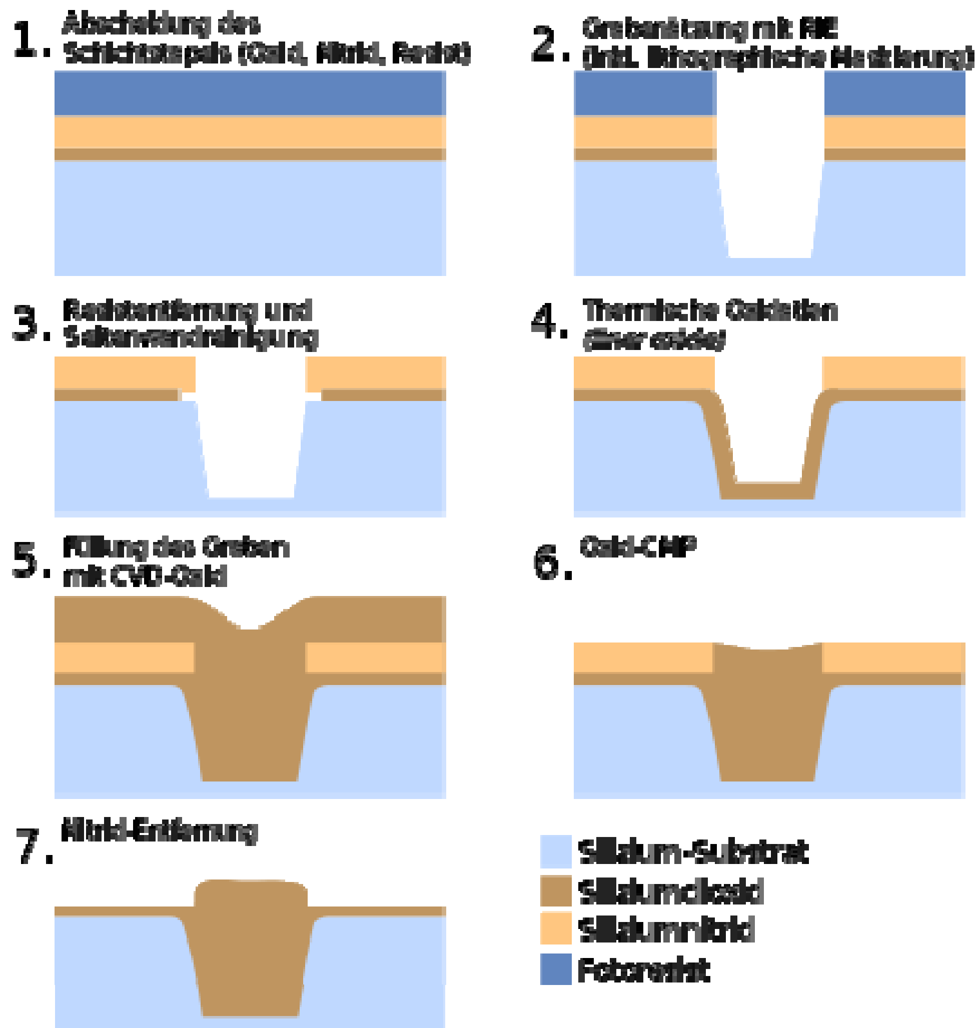
Shallow Trench Isolation (STI)

Pad oxide (thermal)
Mask nitride (LPCVD)
Lithography
Etching nitride/oxide/silicon
Resist strip and cleaning
Liner oxide (thermal)
CVD oxide deposition
CMP planarization of the oxide
Nitride etching
Oxide etching

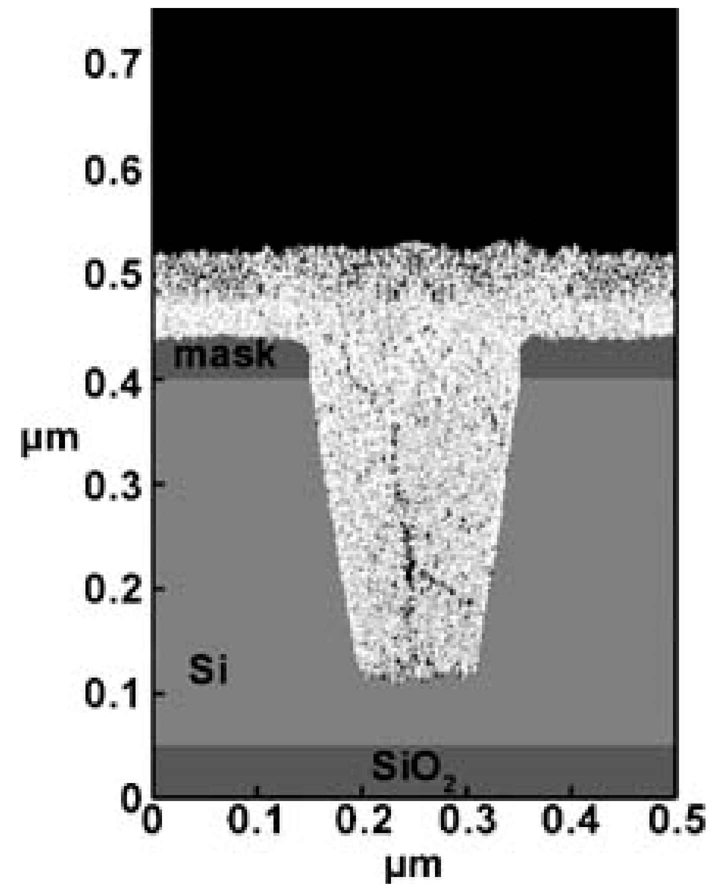


http://de.wikipedia.org/wiki/Grabenisolation

Shallow Trench Isolation (STI)



http://de.wikipedia.org/wiki/Grabenisolation



http://onlinelibrary.wiley.com/doi/10.1002/ppap.201100093/abstract

Self-aligned fabrication

p-Si 100 wafer

Dry oxidation

SiO_2 gate oxide

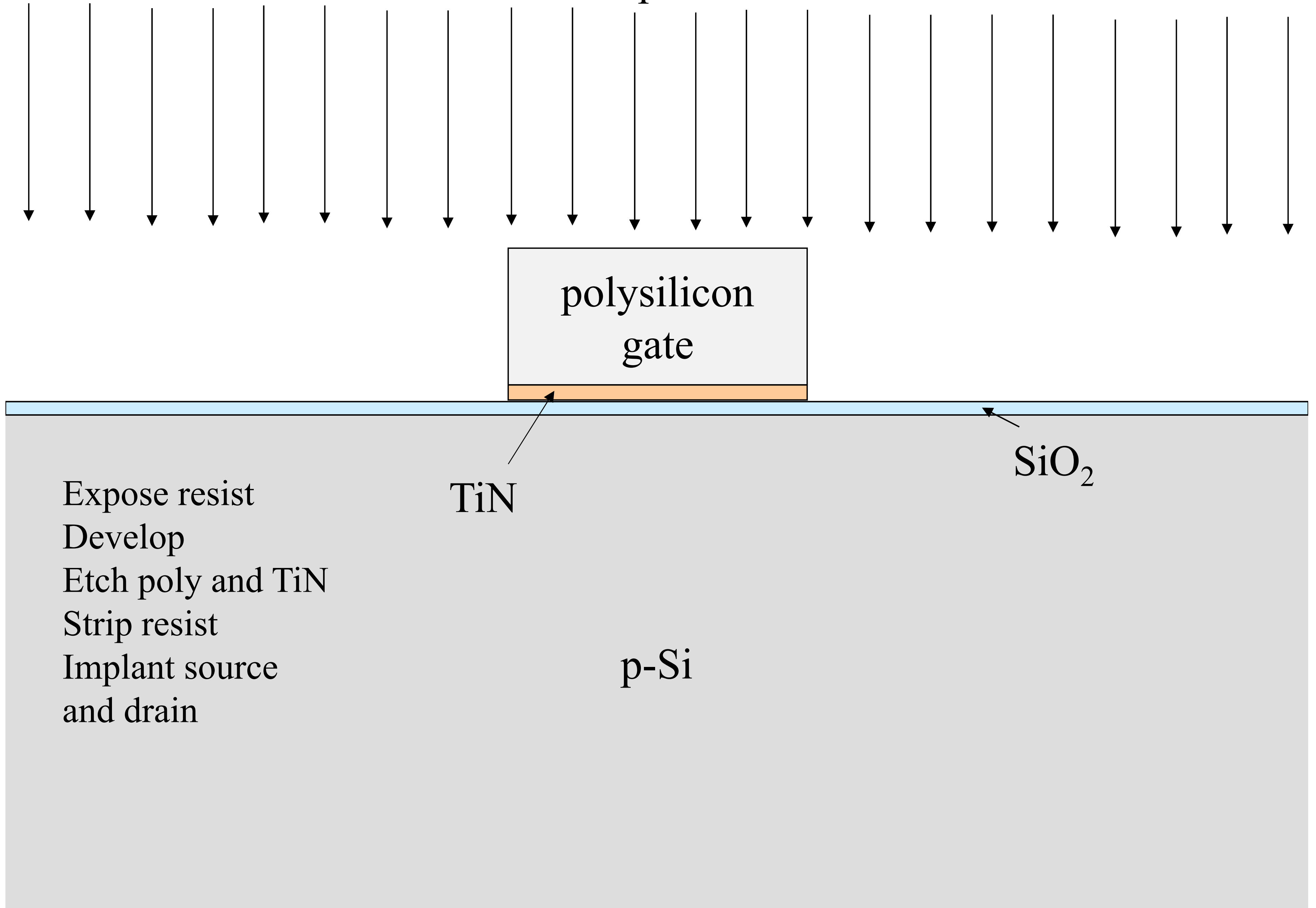
photoresist

polysilicon CVD: SiH_4 @ 580 to 650 °C

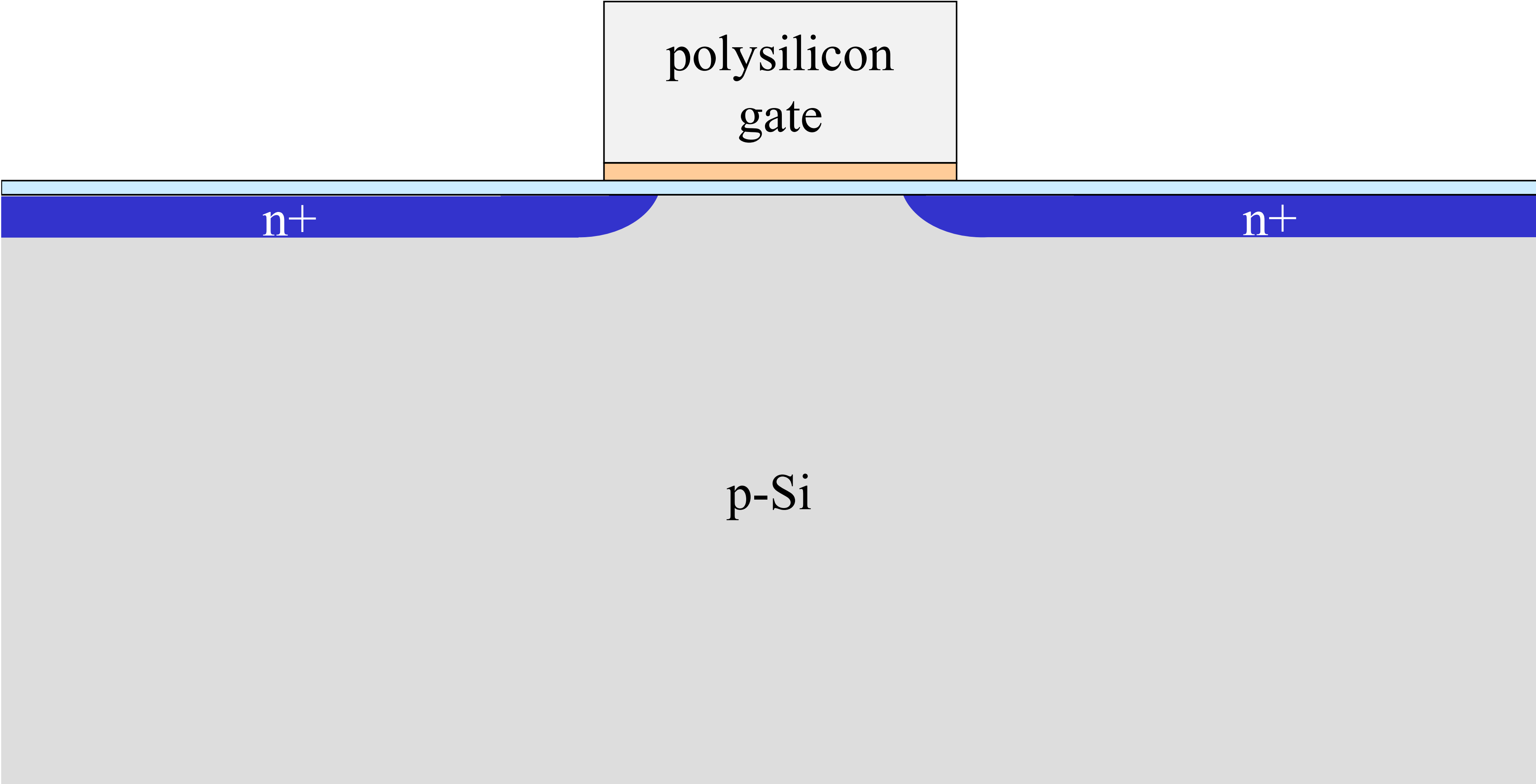
SiO_2

TiN (CVD)
30–70 μΩ·cm Conductive diffusion barrier

p-Si

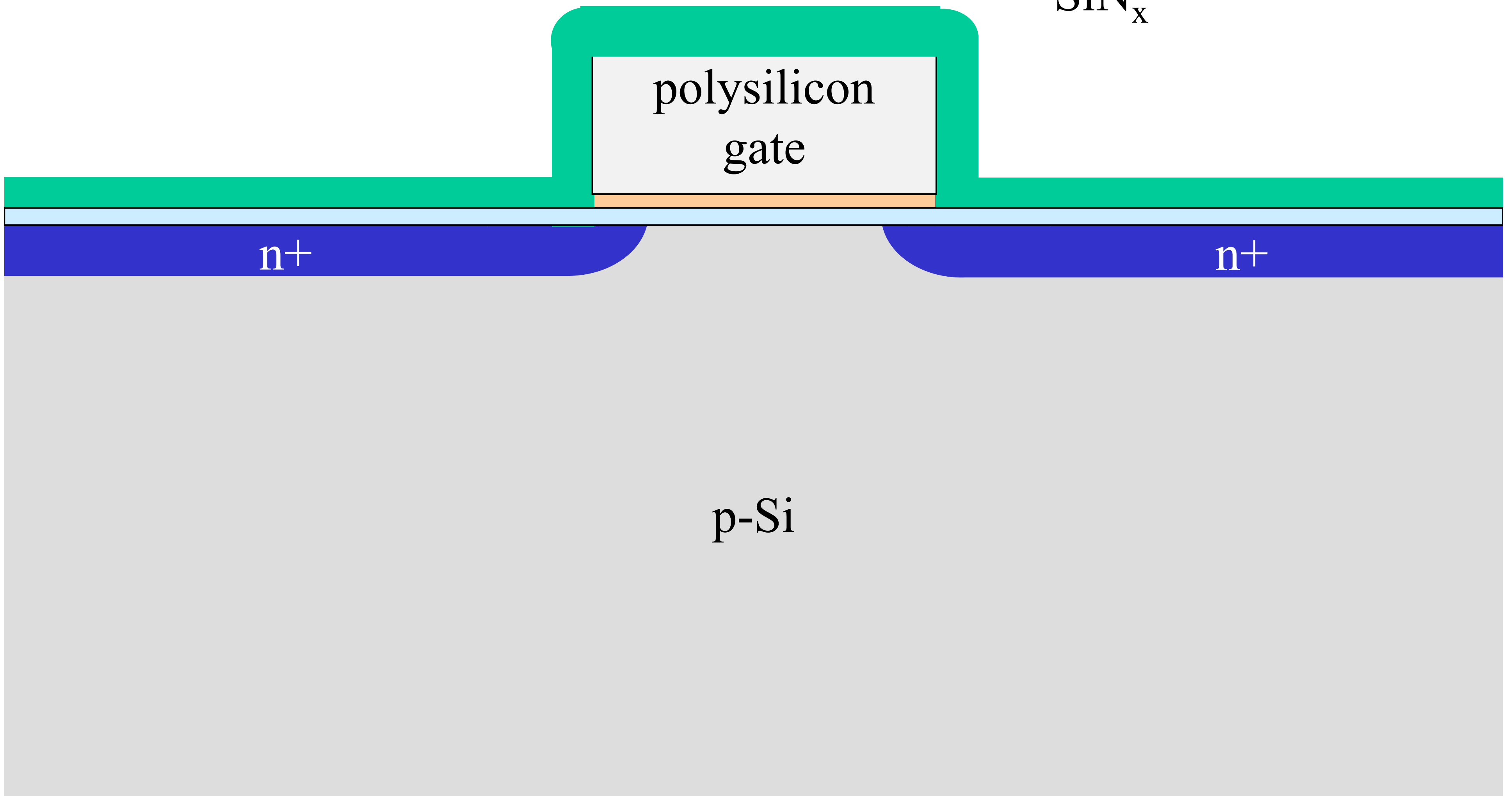

Implant
polysilicon gate
TiN
SiO$_2$
Expose resist
Develop
Etch poly and TiN
Strip resist
Implant source
and drain
p-Si

Self-aligned fabrication



Spacer

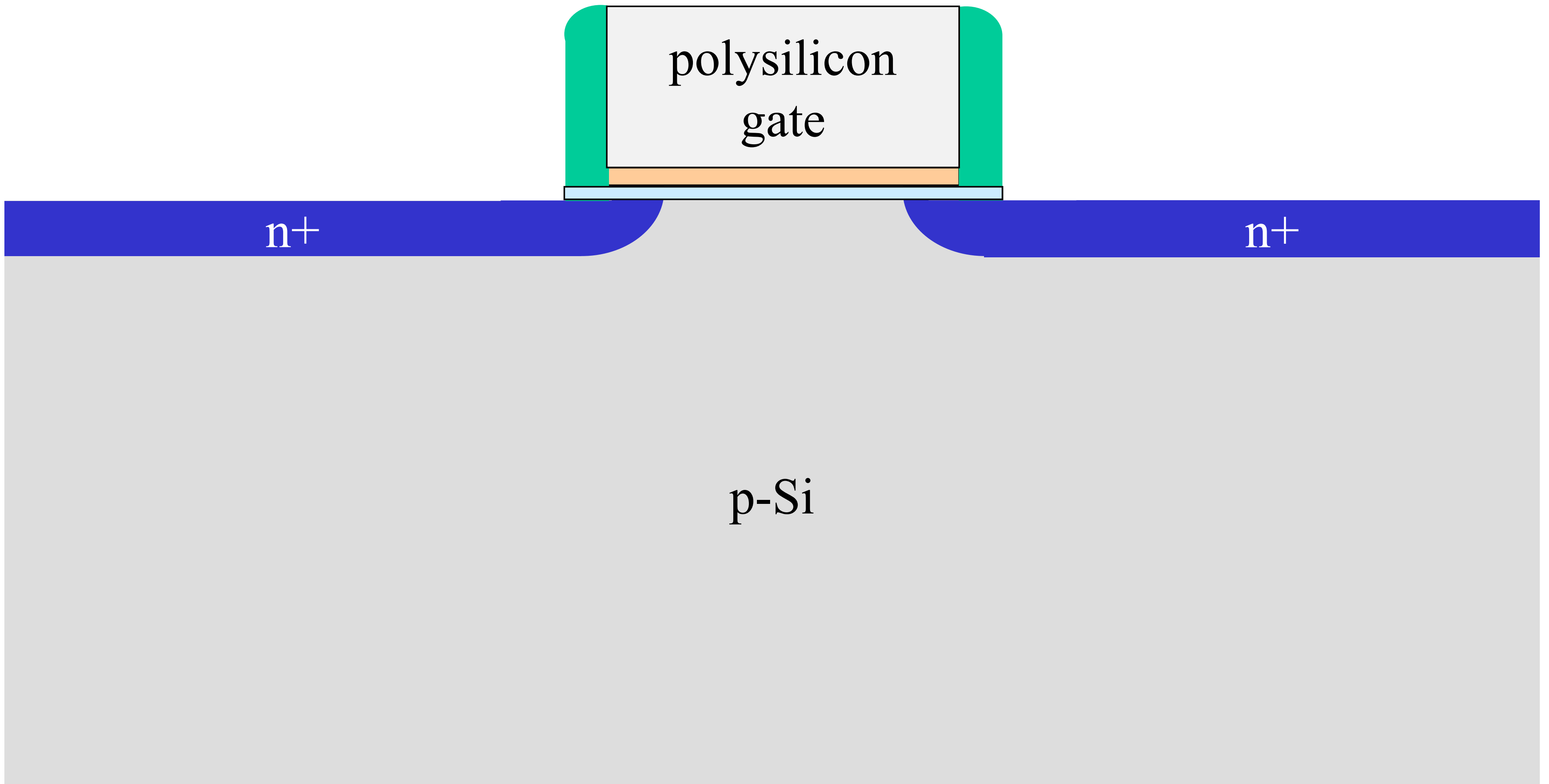
PECVD SiN_x

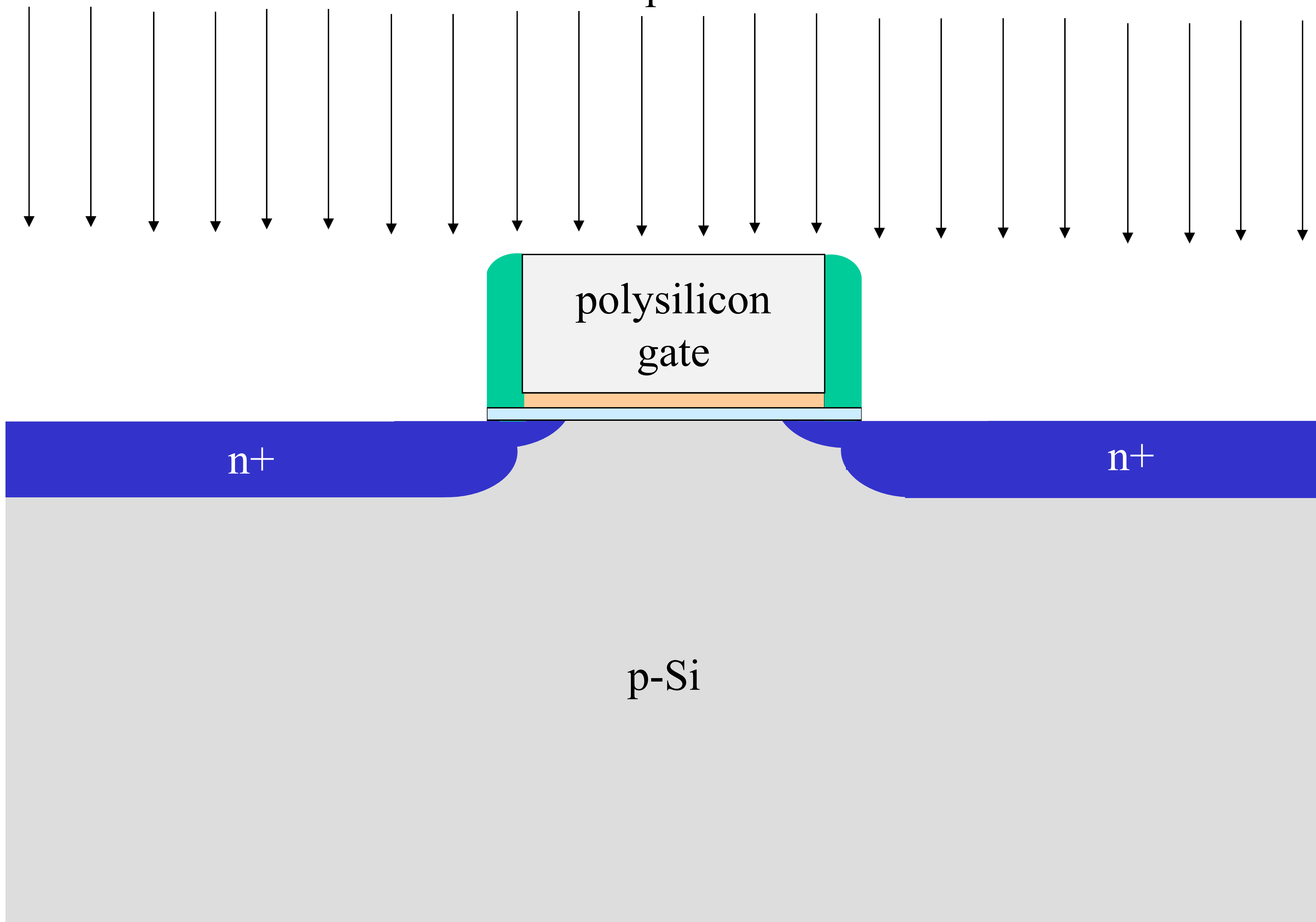


Spacer

Etch back to leave only sidewalls

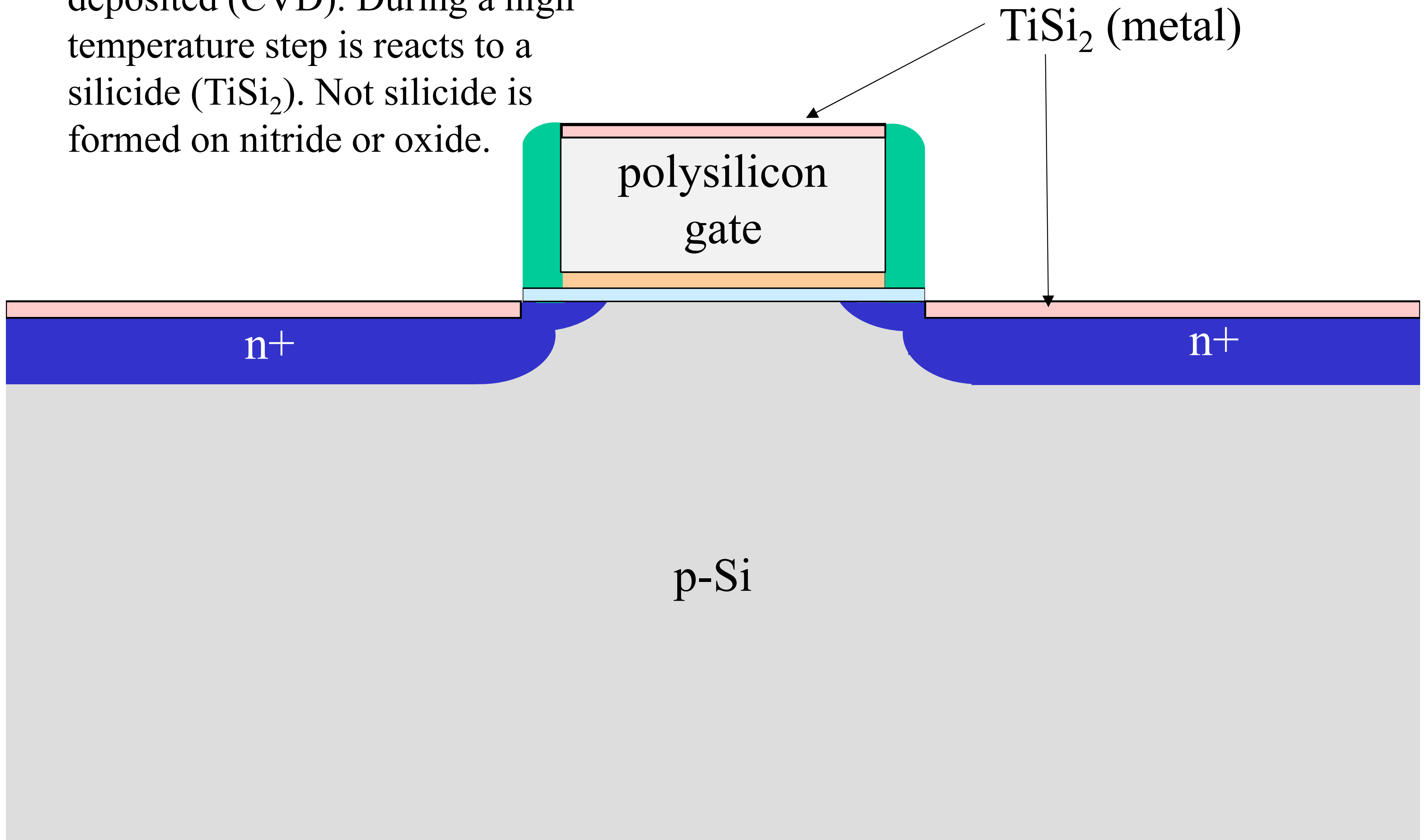
SiN_x




Implant
polysilicon
gate
n+
n+
p-Si

Salicide (Self-aligned silicide)

Transition metal (Ti, Co,W) is deposited (CVD). During a high temperature step is reacts to a silicide ($TiSi_2$). Not silicide is formed on nitride or oxide.



CMOS

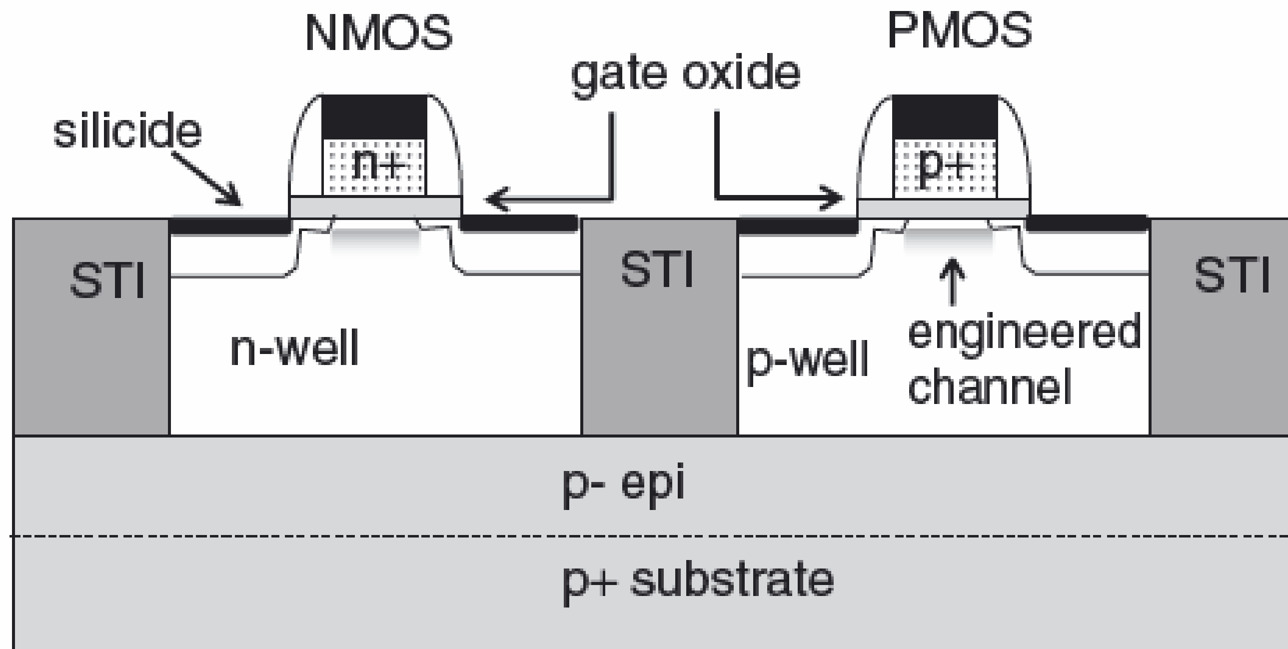
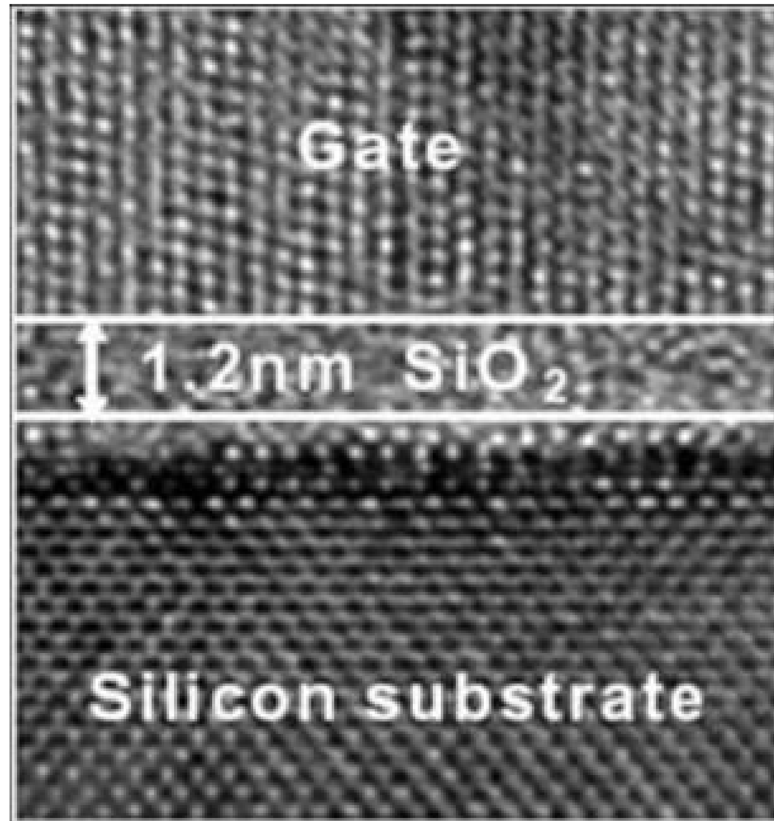


Figure 26.11 Deep submicron CMOS: 200 nm gate length, 5 nm gate oxide, 70 nm junction depth; n^+ poly for NMOS and p^+ poly for PMOS. Shallow trench isolation on epitaxial n^+/p^+ wafer

Fransila

Gate dielectric



Thinner than 1 nm: electrons tunnel

Large dielectric constant desirable
$\varepsilon_r(SiO_2) \sim 4$

$\varepsilon_r(Si_3N_4) \sim 7$

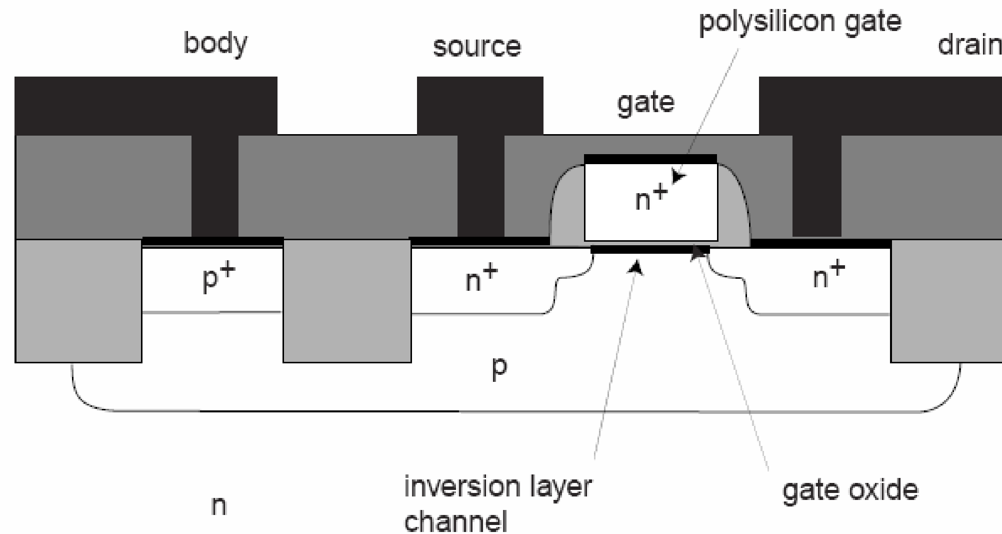
The heat dissipation problem

Microprocessors are hot ~ 100 C
Hotter operation will cause dopants to diffuse

When more transistors are put on a chip they must dissipate less power.

Power per transistor decreases like L^2.

Constant E-field Scaling



Gate length L, transistor width Z, oxide thickness t_{ox} are scaled down.

V_{ds}, V_{gs}, and V_T are reduced to keep the electric field constant.

Power density remains constant.

$$L \sim 45\ t_{ox}$$

1975 - 1990: "Days of happy scaling"

Constant E-field scaling

$$I_{sat} = \frac{Z}{2L}\mu_n \frac{\varepsilon_{ox}}{t_{ox}}\left(V_G - V_T\right)^2$$

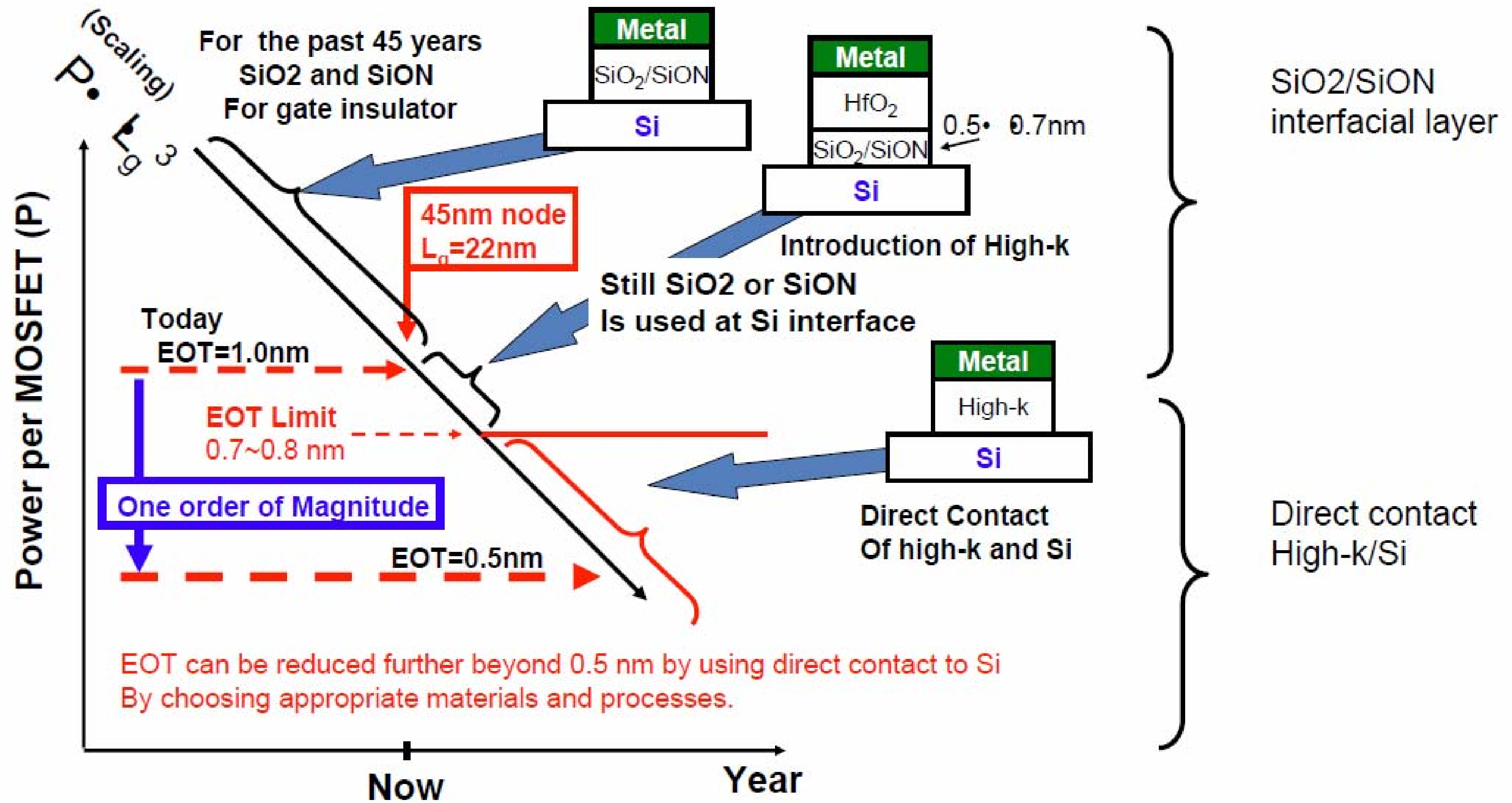
$$L \Rightarrow sL, \qquad Z \Rightarrow sZ, \qquad t_{ox} \Rightarrow st_{ox}, \qquad V_{th} \Rightarrow sV_{th}$$

$$I_{sat} \Rightarrow sI_{sat}$$ ← I_{sat} gets smaller

$$g_m = \frac{dI_D}{dV_G} = \frac{Z}{L}\mu_n \frac{\varepsilon_{ox}}{t_{ox}}\left(V_G - V_T\right)$$ ← Transconductance stays the same.

Power per transistor decreases like L^2. Power per unit area remains constant. Maximum operating frequency remains constant.

Direct contact technology of high-k to Si



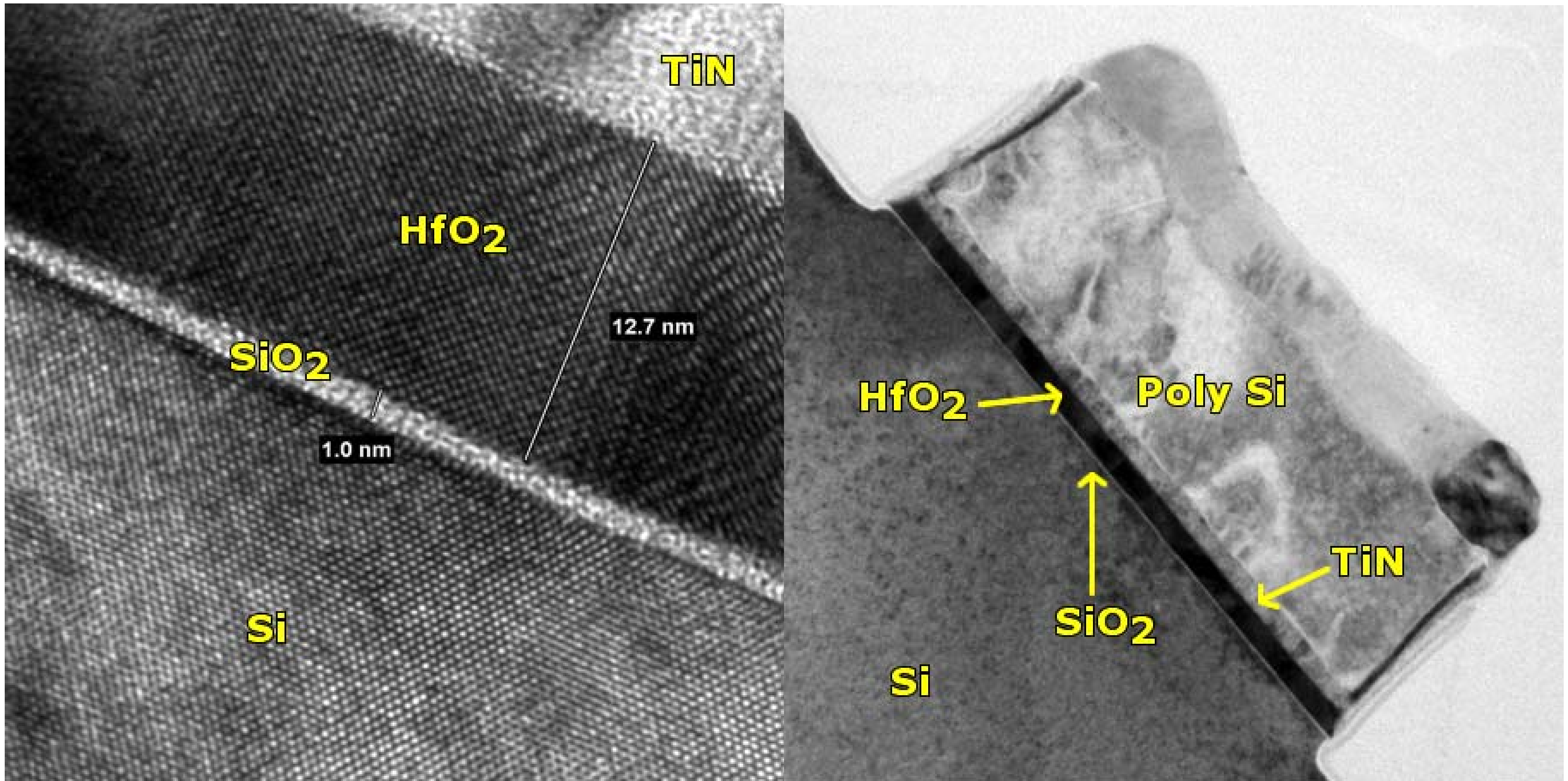
www.iwailab.ep.titech.ac.jp

Equivalent oxide thickness (EOT)

$$\mathrm{EOT} = \frac{\varepsilon_{\mathrm{Si0_2}}}{\varepsilon} t_{\text{high-k}} + t_{\mathrm{Si0_2}}$$

Scaling can continue using oxides too thick to tunnel if they have a higher dielectric constant.

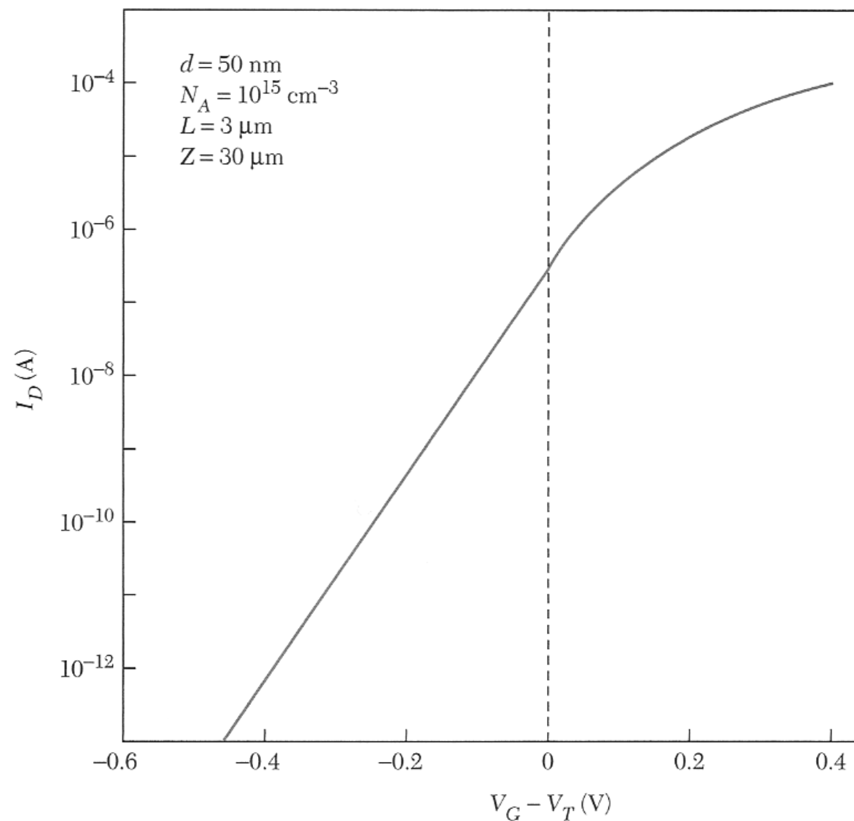
High-k dielectrics



http://nano.boisestate.edu/research-areas/gate-oxide-studies/

Subthreshold current

For $V_G < V_T$ the transistor should switch off but there is a diffusion current. The current is not really off until ~ 0.5 V below the threshold voltage.

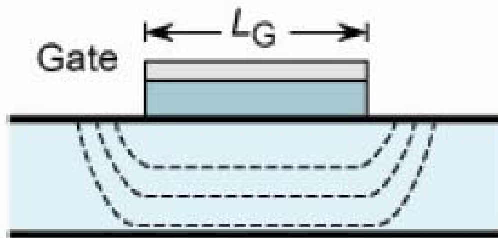


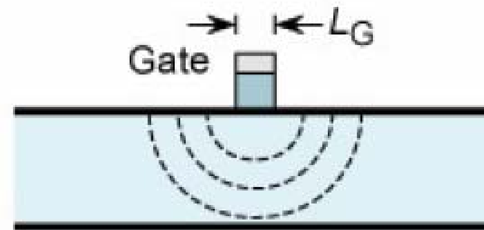
Weak inversion

$$I_D \propto \exp\left(\frac{e(V_G - V_T)}{k_B T}\right)$$

Subthreshold swing: 70-100 mV/decade

Short channel effects





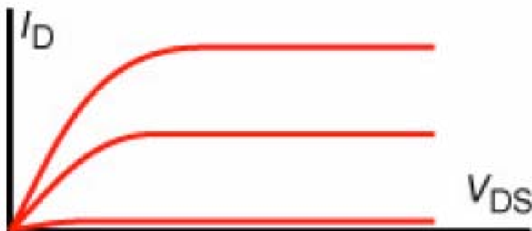
Short-channel effects:

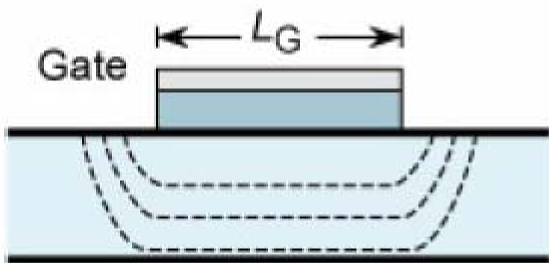
Threshold-voltage shift

Lack of pinch-off

Increased leakage current

Increase of output conductance

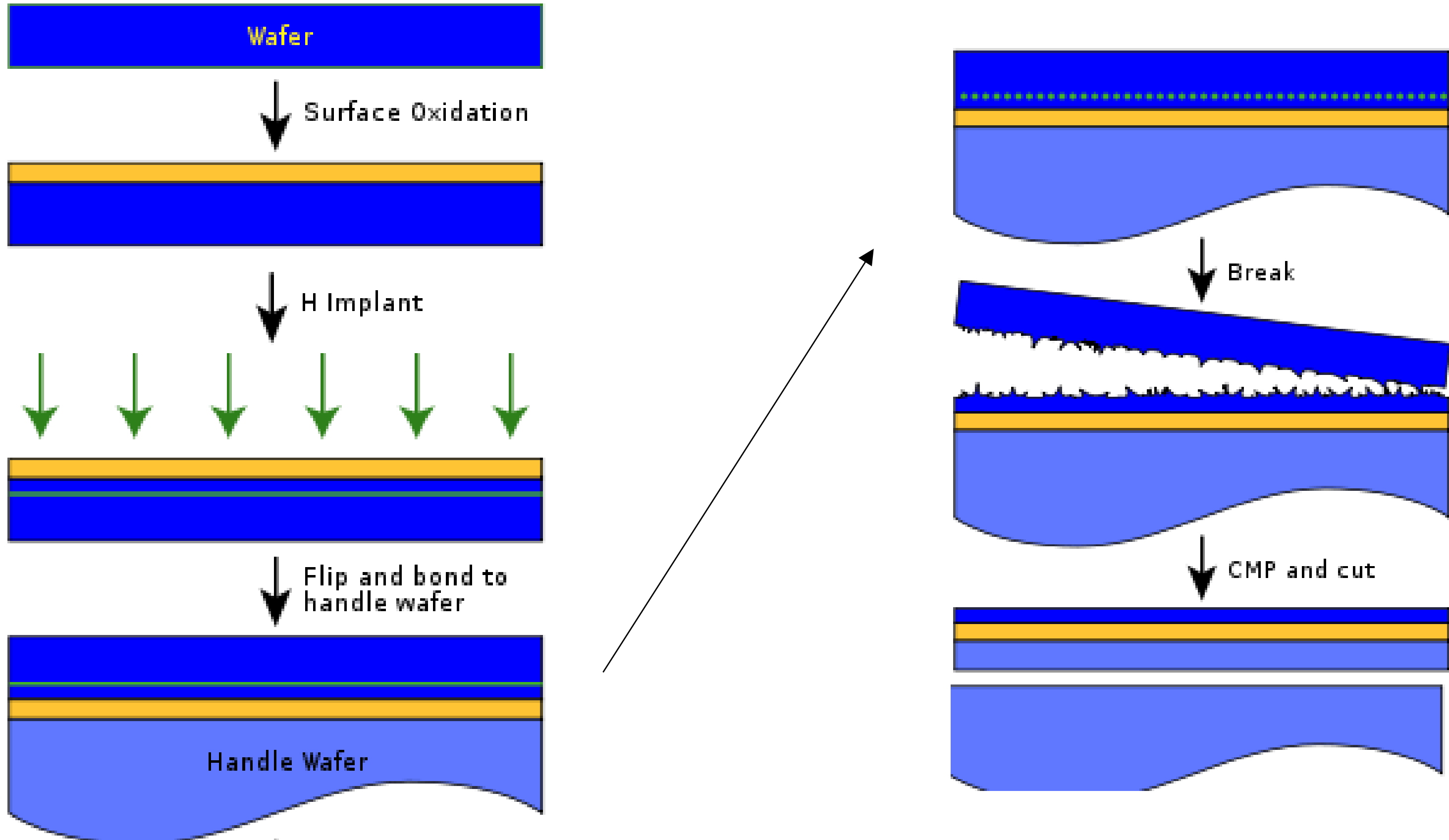




SOI: silicon on insulator

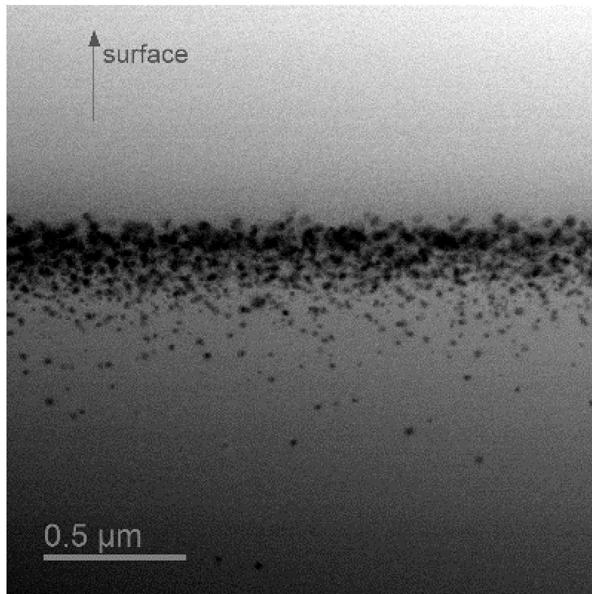
© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

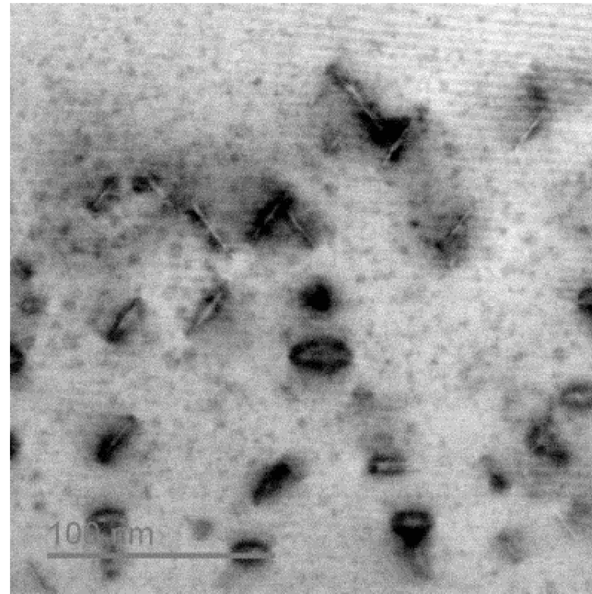
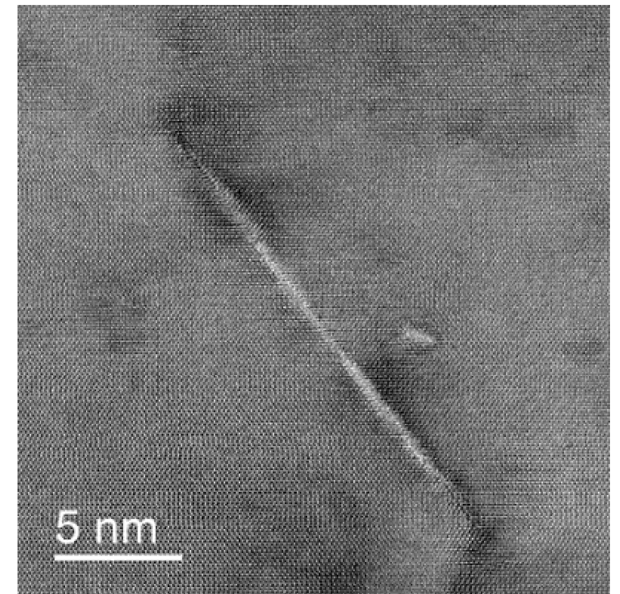
Smart Cut



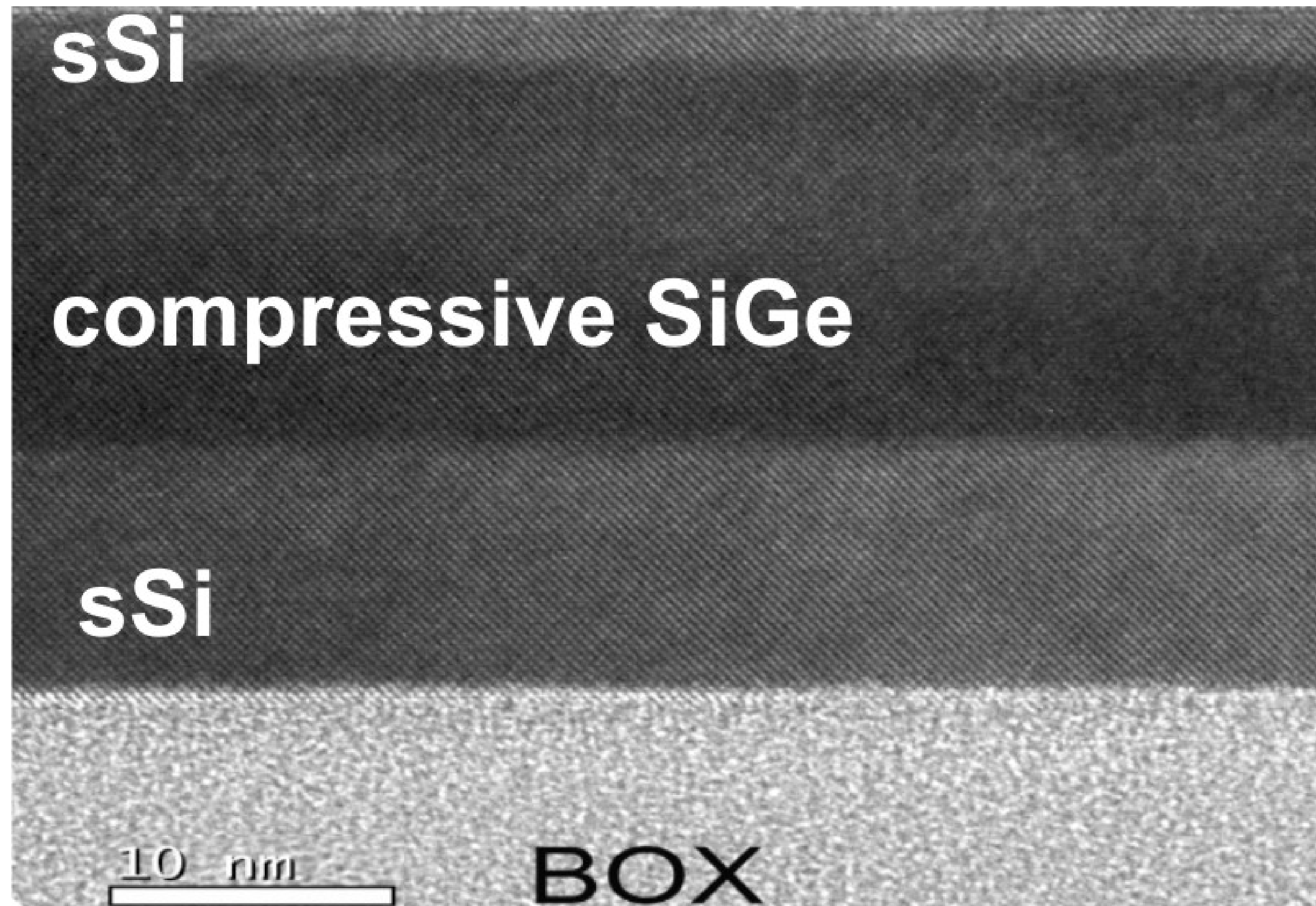
http://en.wikipedia.org/wiki/Silicon_on_insulator

Smart Cut



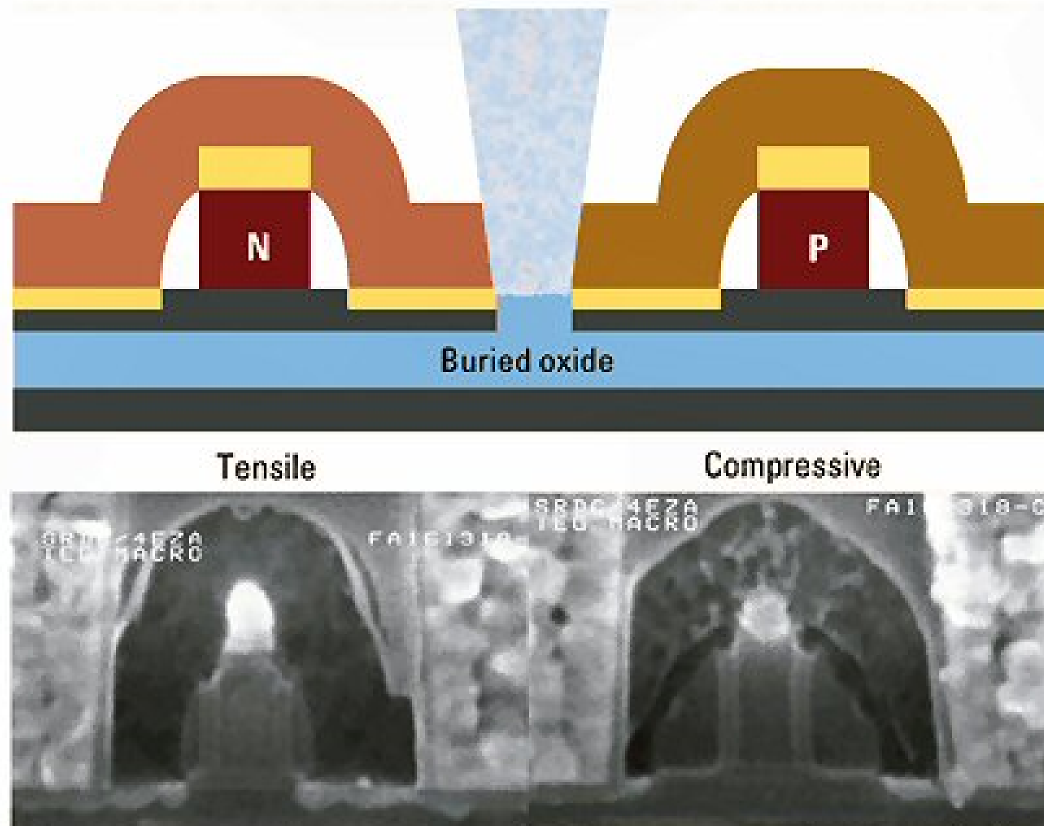
STEM images of FZ-silicon implanted with 400 keV protons at a dose of 10^{16} cm^{-2}.



TEM image of a 3 nm Si cap/ 15 nm SiGe 50% /10 nm strained SOI structure grown at CEA-Leti and used for p-SiGe MOSFET fabrication.

http://www.fz-juelich.de/pgi/pgi-9/EN/Forschung/08-strained%20silicon/04_Biaxially%20strained%20Si_SiGe_%28S%29SOI%20heterostructure/_node.html

Dual stress liners



Tensile silicon nitride film over the NMOS and a compressive silicon nitride film over the PMOS improves the mobility.

http://www.xbitlabs.com/articles/cpu/display/athlon64-venice_2.html

CMOS SOI

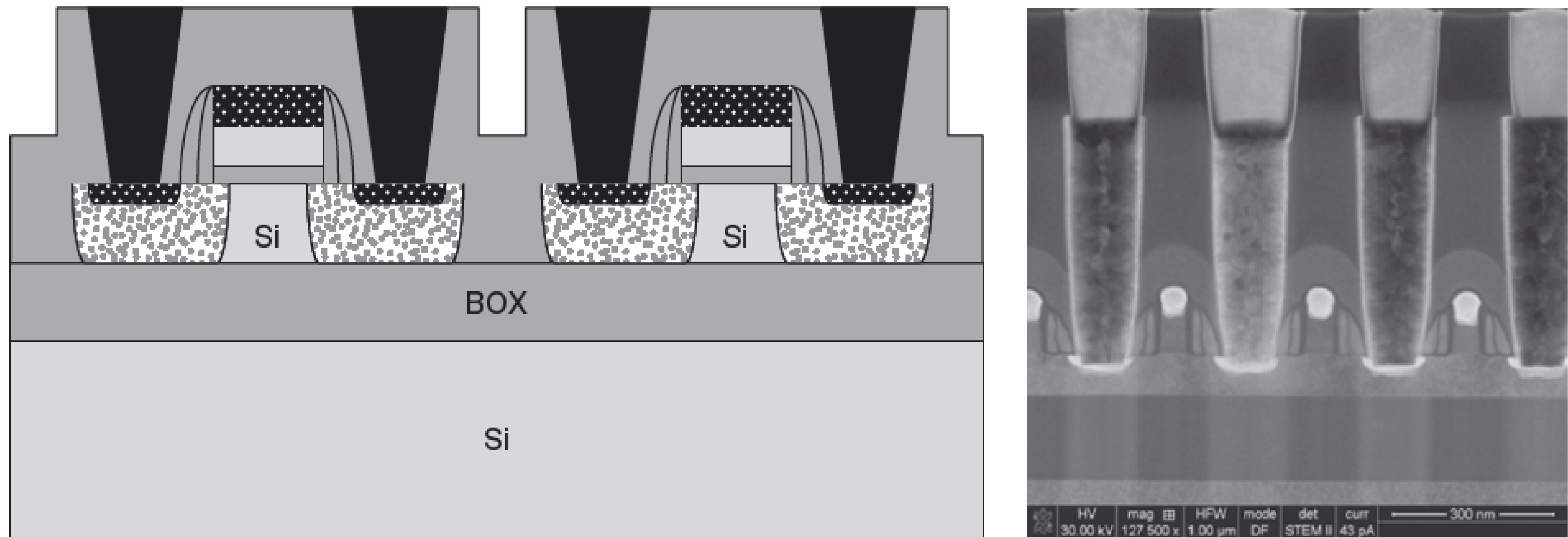
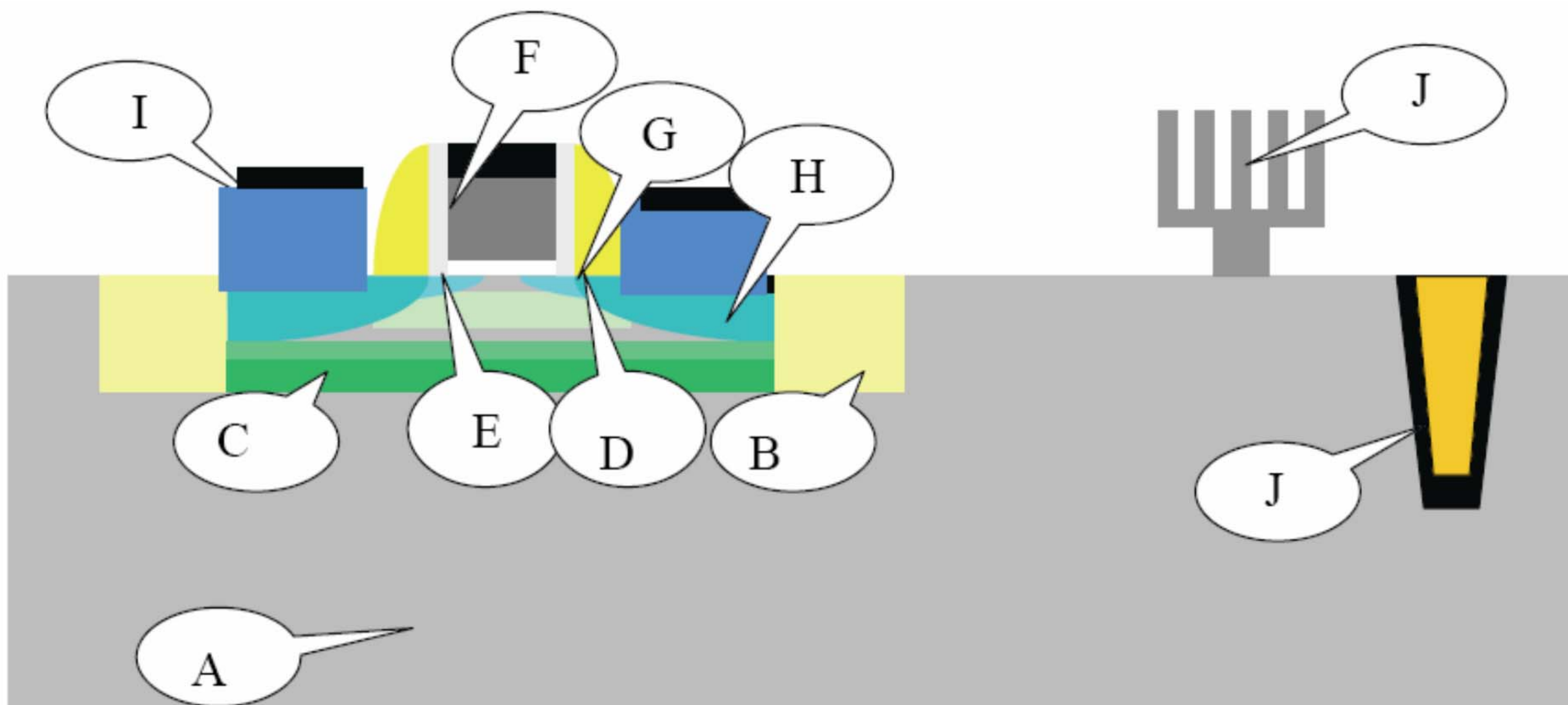


Figure 26.16 SOI MOSFET with first-level metal, schematic and TEM. Courtesy Brandon Van Leer, FEI Company4

Fransila

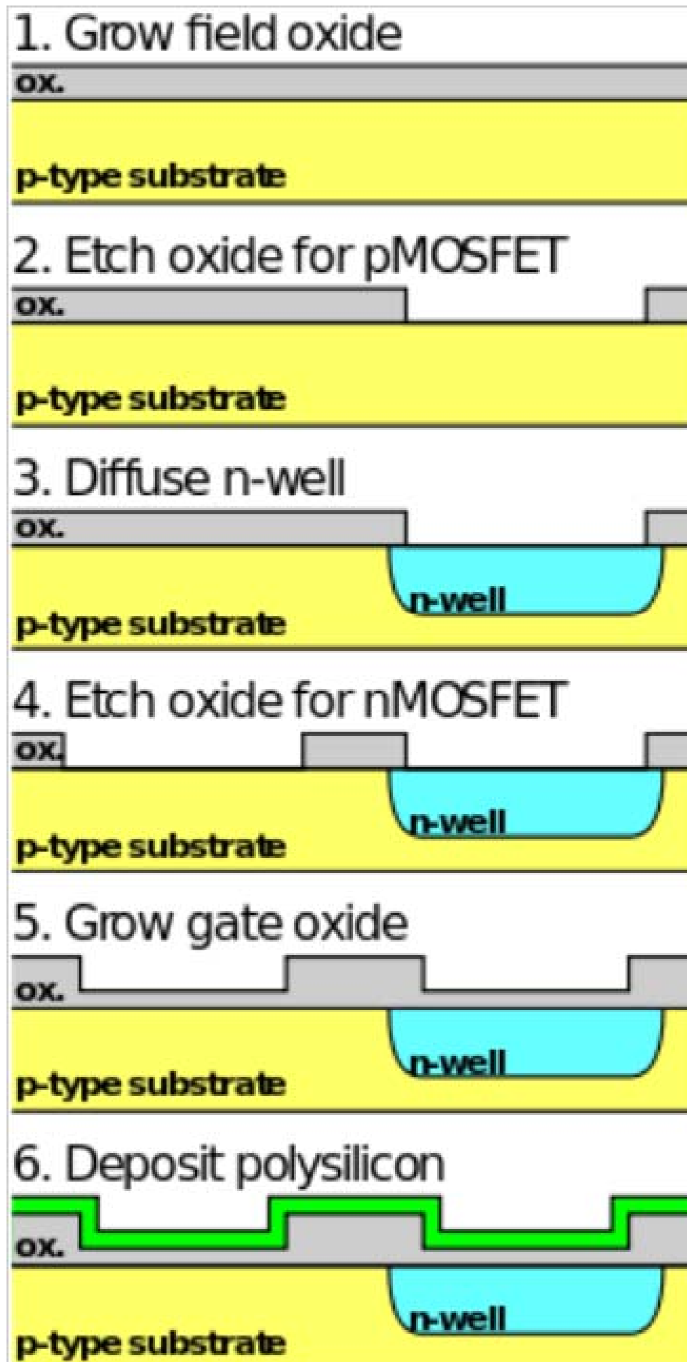
Table PIDS2a High-performance (HP) Logic Technology Requirements - TCAD																
Year of Production	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
Logic Industry "Node Range" Labeling (nm) [based on 0.71x reduction per "Node Range" ("Node" = 2x Mx)]	"16/14"		"11/10"		"8/7"		"6/5"		"4/3"		"3/2.5"		"2/1.5"		"1/0.75"	
MPU/ASIC Metal 1 (M1) ½ Pitch (nm) (contacted)	40	32	32	28.3	25.3	22.5	20.0	17.9	15.9	14.2	12.6	11.3	10.0	8.9	8	7.1
L_g : Physical Gate Length for HP Logic (nm)	20	18	16.7	15.2	13.9	12.7	11.6	10.6	9.7	8.8	8.0	7.3	6.7	6.1	5.6	5.1
L_{ch} : Effective Channel Length (nm) [3]	16.0	14.4	13.4	12.2	11.1	10.2	9.3	8.5	7.8	7.0	6.4	5.8	5.4	4.9	4.5	4.1
V_{dd} : Power Supply Voltage (V)																
Bulk/SOI/MG	0.86	0.85	0.83	0.81	0.80	0.78	0.77	0.75	0.74	0.72	0.71	0.69	0.68	0.66	0.65	0.64
EOT: Equivalent Oxide Thickness																
Bulk/SOI/MG (nm)	0.80	0.77	0.73	0.70	0.67	0.64	0.61	0.59	0.56	0.54	0.51	0.49	0.47	0.45	0.43	0.41
Dielectric constant (K) of gate dielectrics	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0	16.5	17.0	17.5	18.0	18.5	19.0	19.5	20.0
Physical gate oxide thickness (nm)	2.56	2.57	2.53	2.51	2.49	2.46	2.42	2.42	2.37	2.35	2.29	2.26	2.23	2.19	2.15	2.10
Channel Doping (10^{18} /cm^3) [4]																
Bulk	6.0	7.0	7.7	8.4	9.0											
SOI/MG	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Body Thickness (nm) [5]																
SOI																
MG	6.4	5.8	5.3	4.9	4.4	4.1	3.7	3.4	3.1	2.8	2.6	2.3	2.1	2.0	1.8	1.6
T_{BOX} : Buried Oxide Thickness for SOI (nm) [6]																
SOI																
CET: Capacitance Equivalent Thickness (nm) [7]																
Bulk/SOI/MG	1.10	1.07	1.03	1.00	0.97	0.94	0.91	0.89	0.86	0.84	0.81	0.79	0.77	0.75	0.73	0.71
C_{ch} intrinsic (fF/μm) [8]																
Bulk/SOI/MG	0.502	0.465	0.448	0.420	0.396	0.373	0.352	0.329	0.311	0.289	0.273	0.255	0.240	0.225	0.212	0.198
Mobility (cm^2 /V-s)																
Bulk	400	400	400	400	400											
SOI																
MG	250	250	250	250	250	250	200	200	200	200	200	150	150	150	150	150
I_{off} (nA/μm) [9]																
Bulk/SOI/MG	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
$I_{d,sat}$: NMOS Drive Current (μA/μm) [10]																
Bulk	1,348	1,355	1,340	1,295	1,267											
SOI																
MG	1670	1,680	1,700	1,660	1,660	1,610	1,600	1,480	1,450	1,350	1,330	1,170	1,100	1,030	970	900
$V_{t,lin}$ (V) [11]																
Bulk	0.306	0.327	0.324	0.357	0.378											

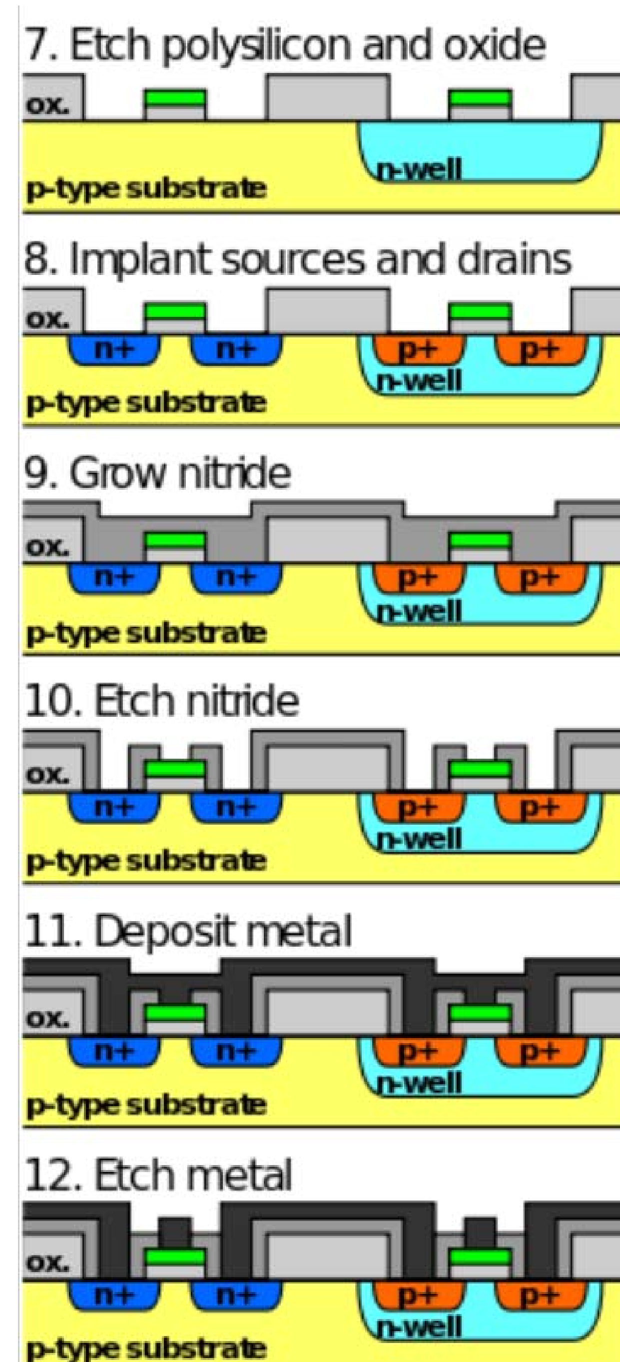


A: Starting Material
C: Well Doping
E: Channel Doping and Channel Strain
G: Extension Junction and Halo
I: Elevated Junction and Contacts

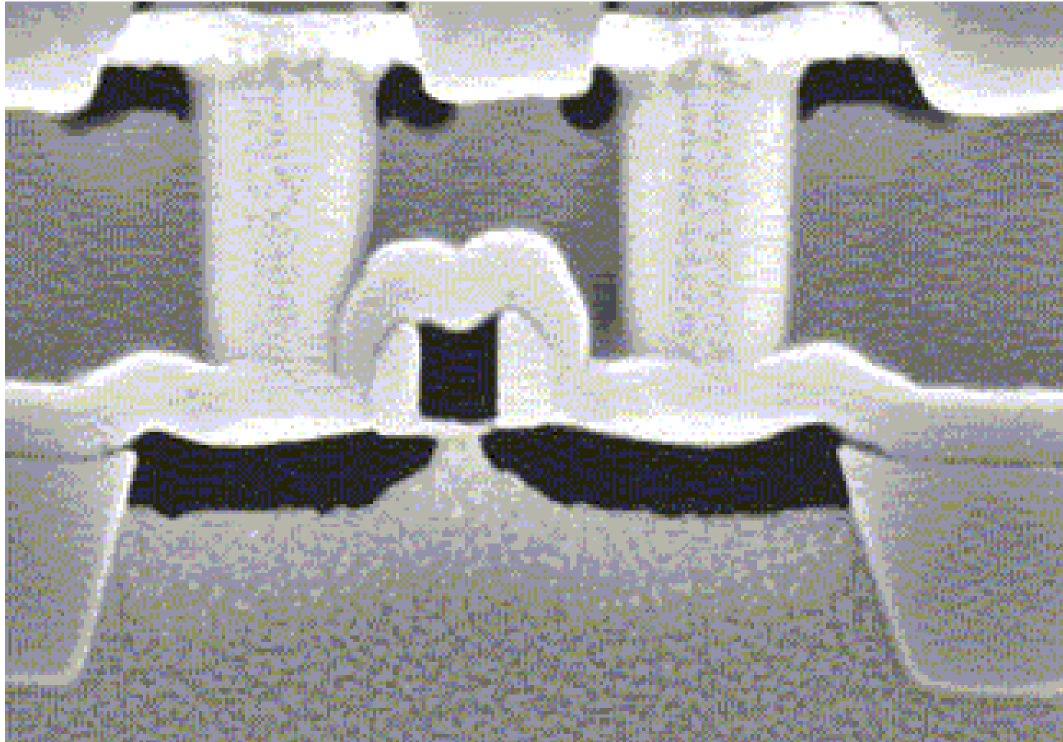
B: Isolation
D: Channel Surface (Preparation)
F: Gate Stack (Including Flash) and Spacer
H: Contacting Source/Drain Junction
J: DRAM Stack/Trench Cap. & FeRAM Storage

http://en.wikipedia.org/wiki/CMOS





Contact holes

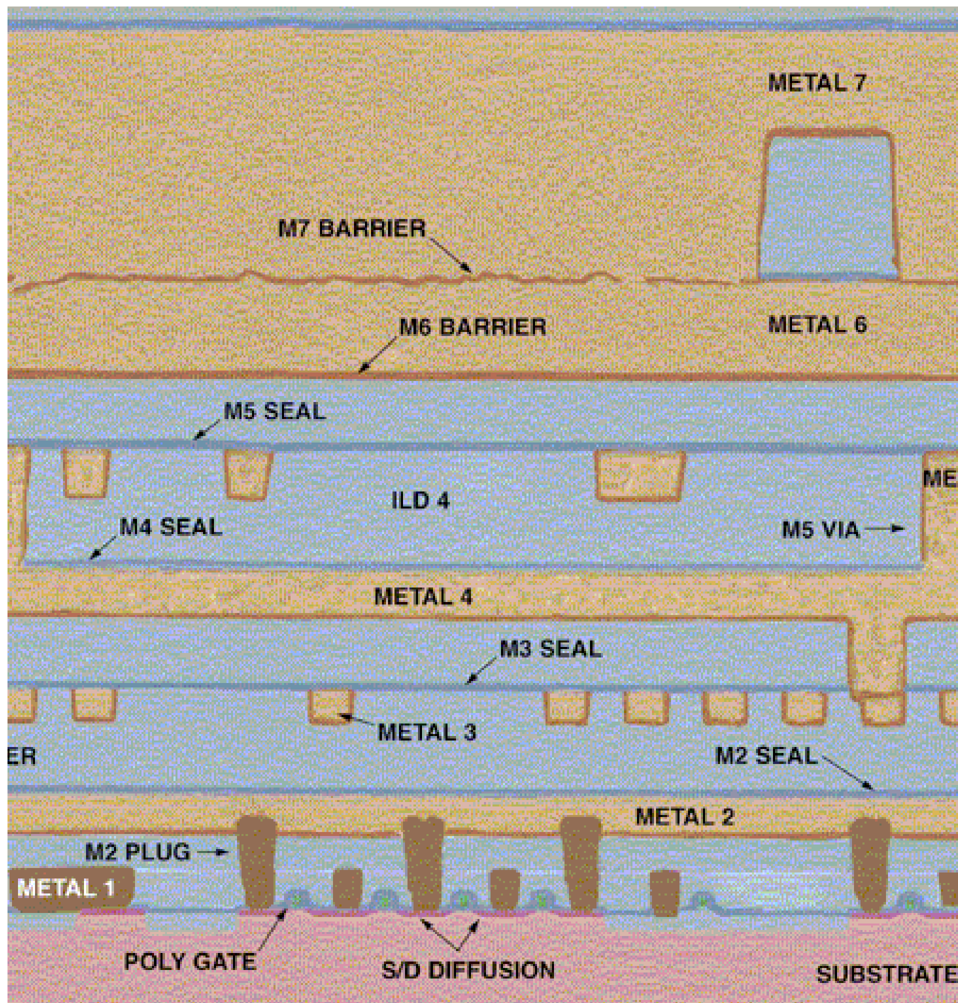


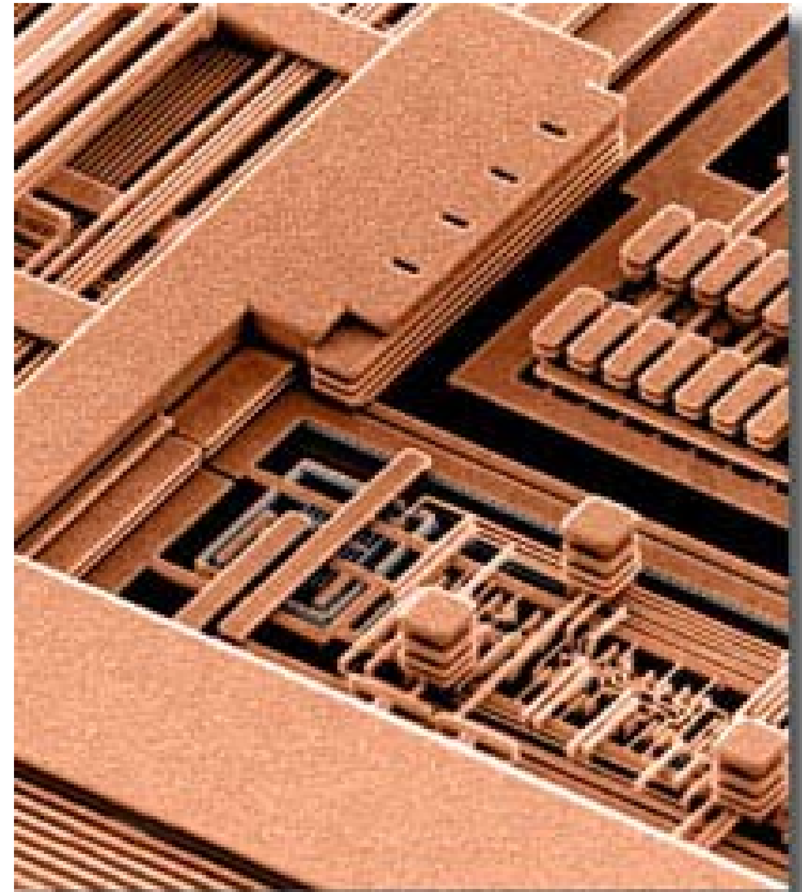
Tungsten CVD using the WF_6. Exceptionally good conformality.

Adhesion/barrier layer such as Ti/TiN (CVD). Protects Si from attack by fluorine, ensures adhesion of W to the silicon dioxide.

Marks the start of back end processes. The temperature cannot go higher than 450 C.

Back-end Metallization





Cu (ρ = 1 μΩ cm) and Al (ρ = 3 μΩ cm) are used for wiring layers.

Multi-level Metallization

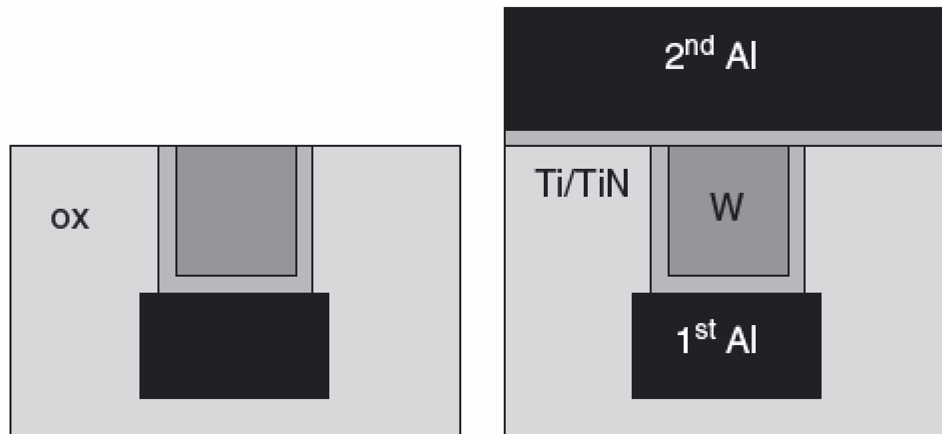


Figure 28.5 Aluminum bottom metal with Ti/TiN/W contact plug after etchback (left) and with second Ti/TiN/aluminum metal layer (right)

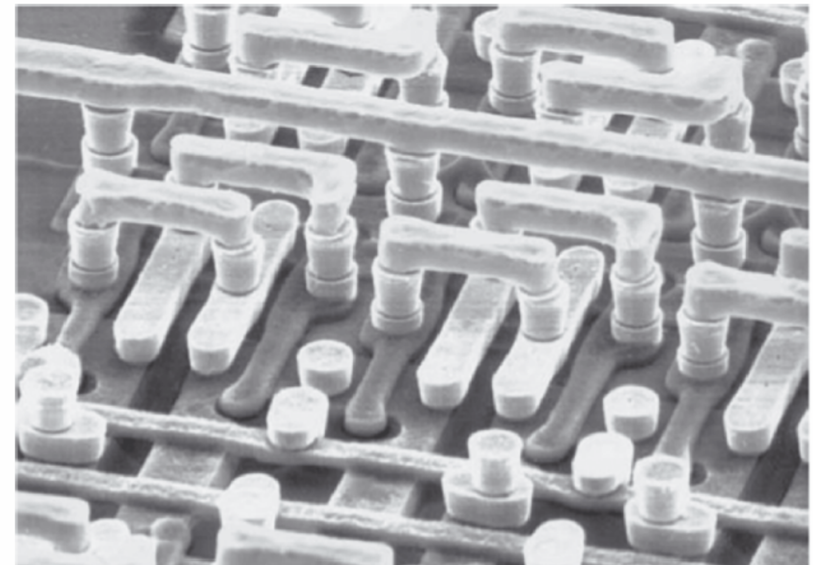
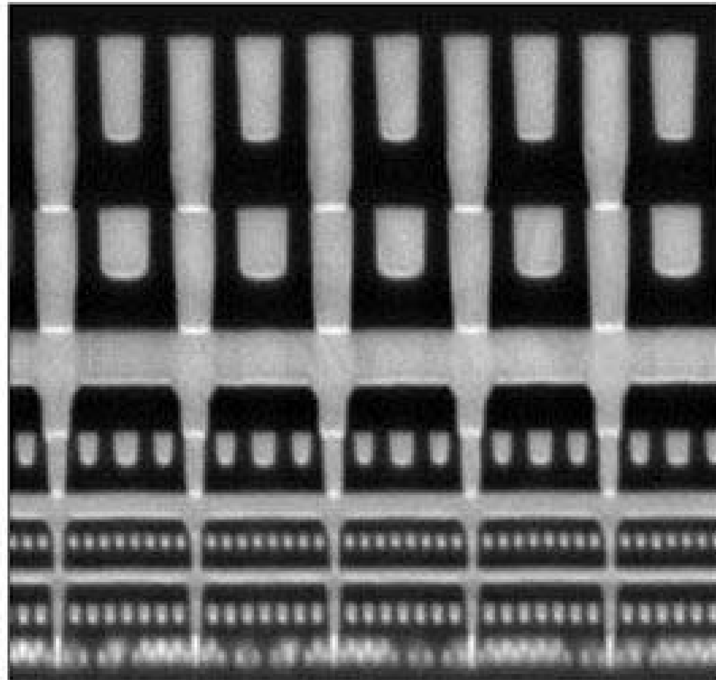


Figure 28.6 Multilevel metallization with all dielectric layers etched away. $TiSi_2$/poly gates, tungsten plugs and local wires, Al global wires. Reproduced from Mann *et al.* (1995) by permission of IBM

Fransila

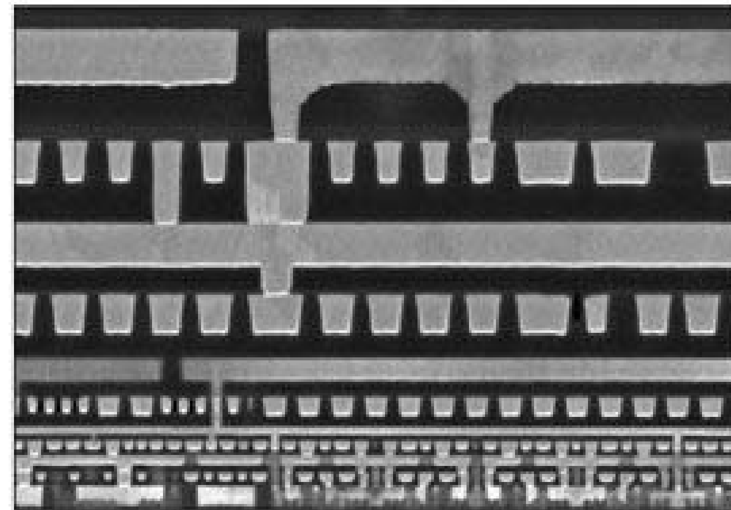
Interconnects

22 nm Process



80 nm minimum pitch

14 nm Process

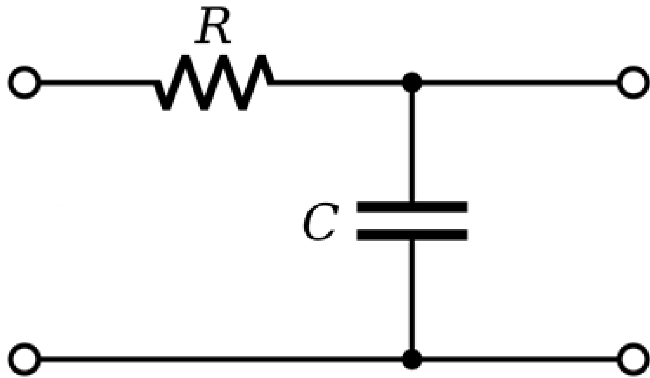


52 nm (0.65x) minimum pitch

Copper must not diffuse into the silicon. Separate equipment is used for the back end of line.

http://www.tomshardware.com/reviews/intel-14nm-broadwell-y-core-m,3904.html

Interconnect RC delay



Low resistivity metal
Low-k dielectrics
  low polarizability
  high porosity

Smaller circuits:
  lower gate delay
  larger RC delay

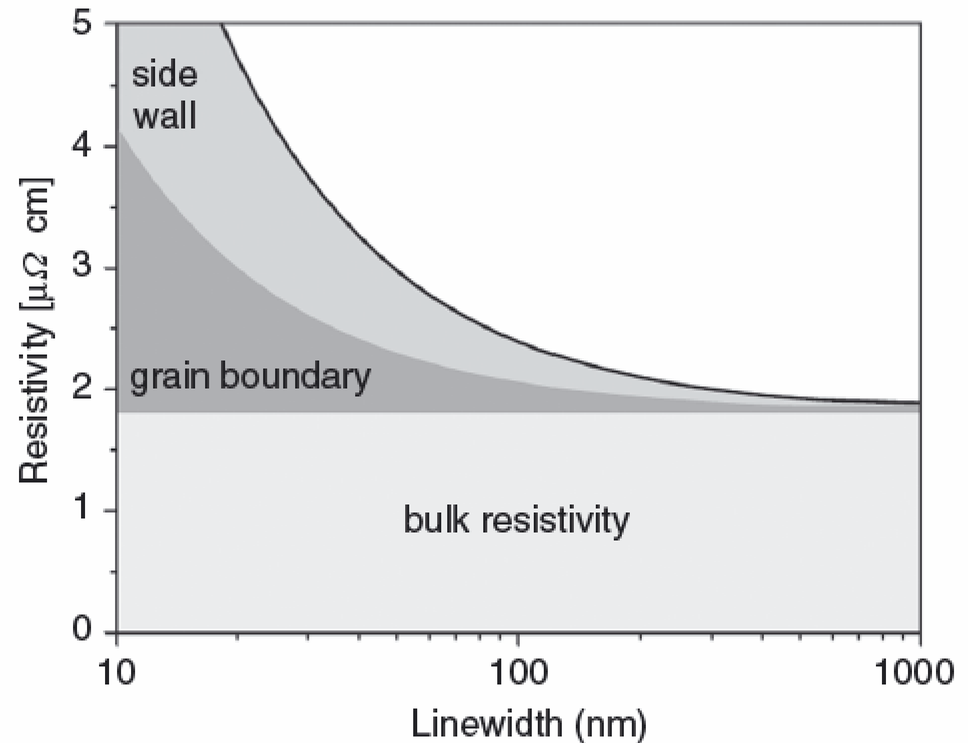
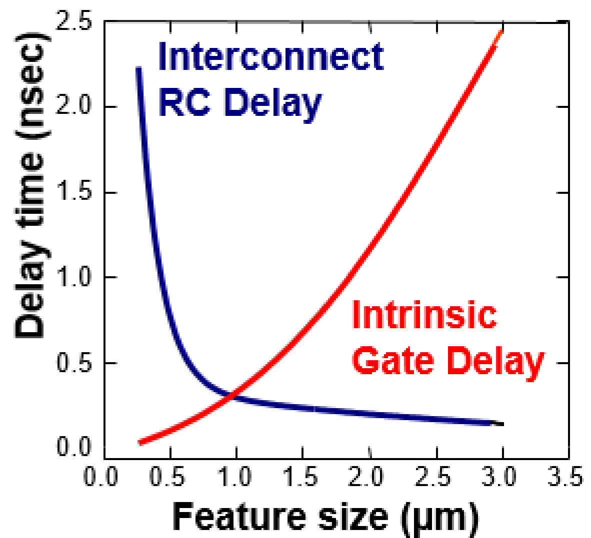


Figure 28.14 Copper resistivity as a function of linewidth. Courtesy of The Semiconductor Industry Association. The International Technology Roadmap for Semiconductors, 2007 Edition. International SEMATECH:Austin, TX, 2007

Fransila

Interconnects



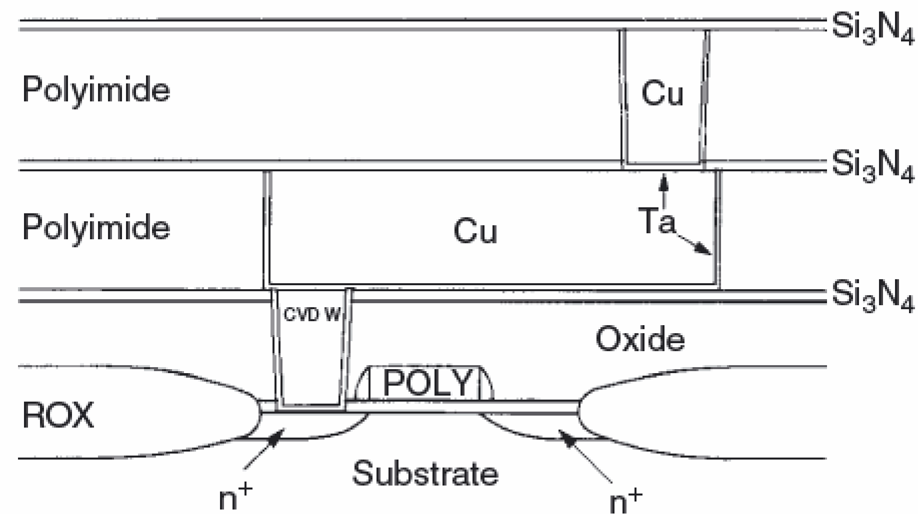
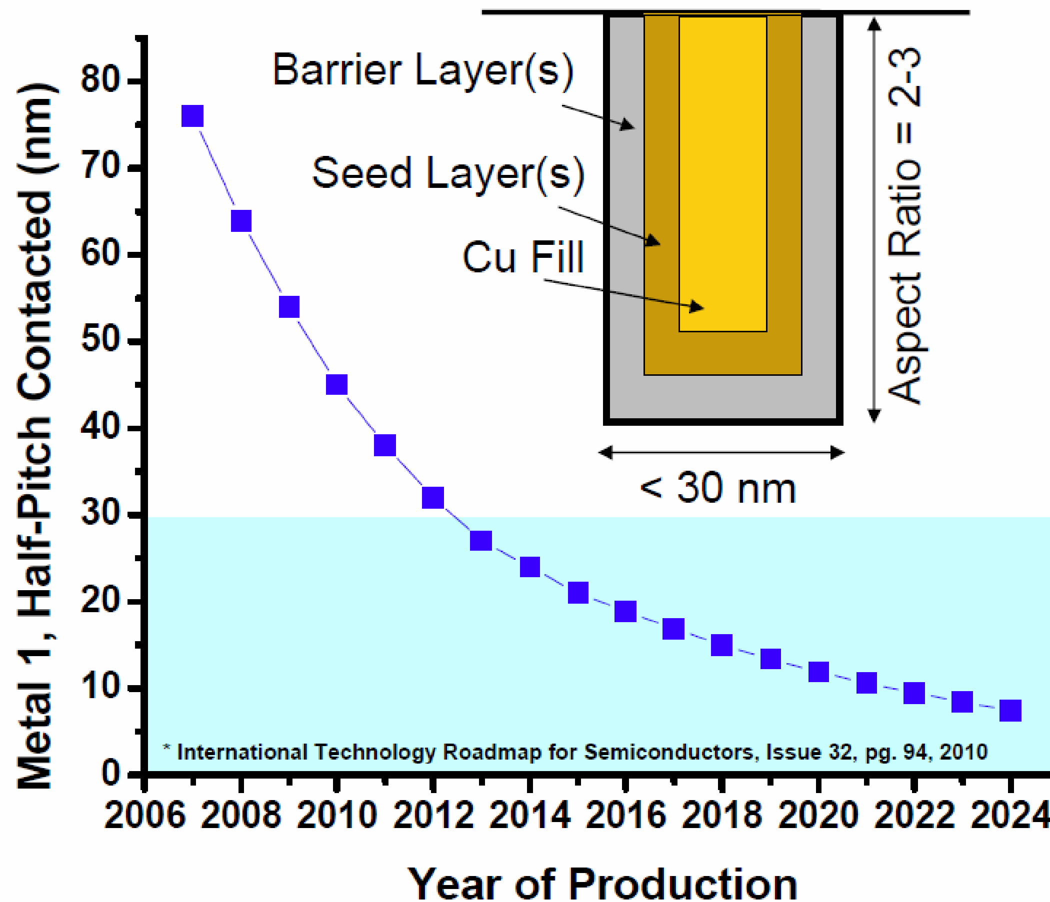


Figure 28.9 Cu/polyimide multilevel metallization with Ta barriers, W plugs and silicon nitride polish stop layers. Reproduced from Small and Pearson (1990) by permission of IBM

Conventional Damascene Copper Extendibility?



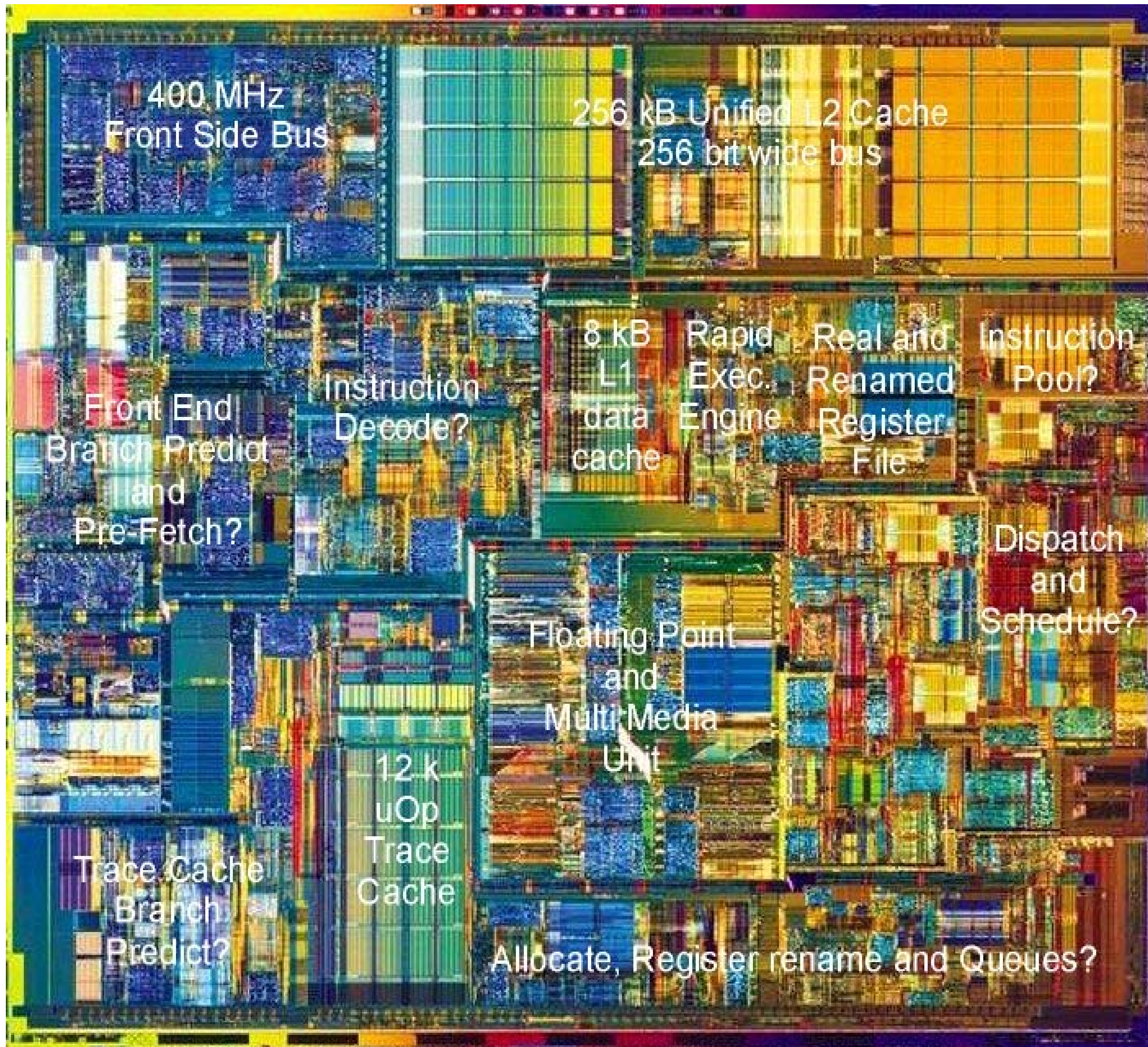
Barrier Layer(s)
TaN/Ta, AlOx, MnOx, other

Seed Layer(s)
PVD Cu, CVD/ALD Co, CVD/ALD Ru, other

Fill Metallurgy
Plated Cu, CVD/ALD Cu, PVD Cu, other

IBM

Microprocessor



Intel® Pentium® 4 90 nm

Intel® Pentium® D 65 nm

Intel® Core™2 Duo 45 nm
Intel® Atom™ Z6xx Series 45 nm
Intel® Core™2 Celeron 45 nm

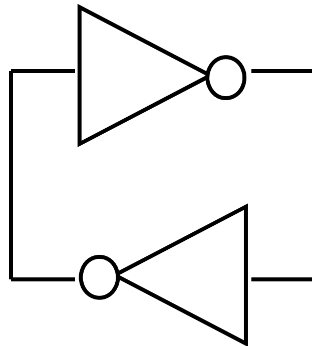
Intel® Core™ i7-900 32 nm
Intel® Xeon® 5600 Series 32 nm

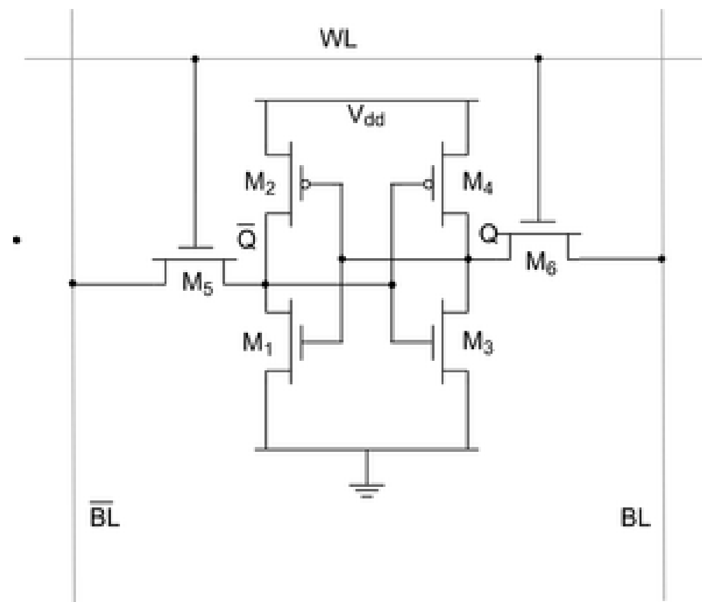
Intel® Ivy bridge tri-gate 22 nm

Intel® Haswell FinFET 16 nm

SRAM

Static random access memory

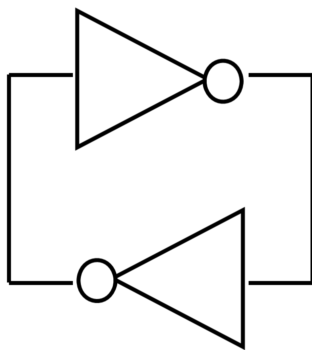


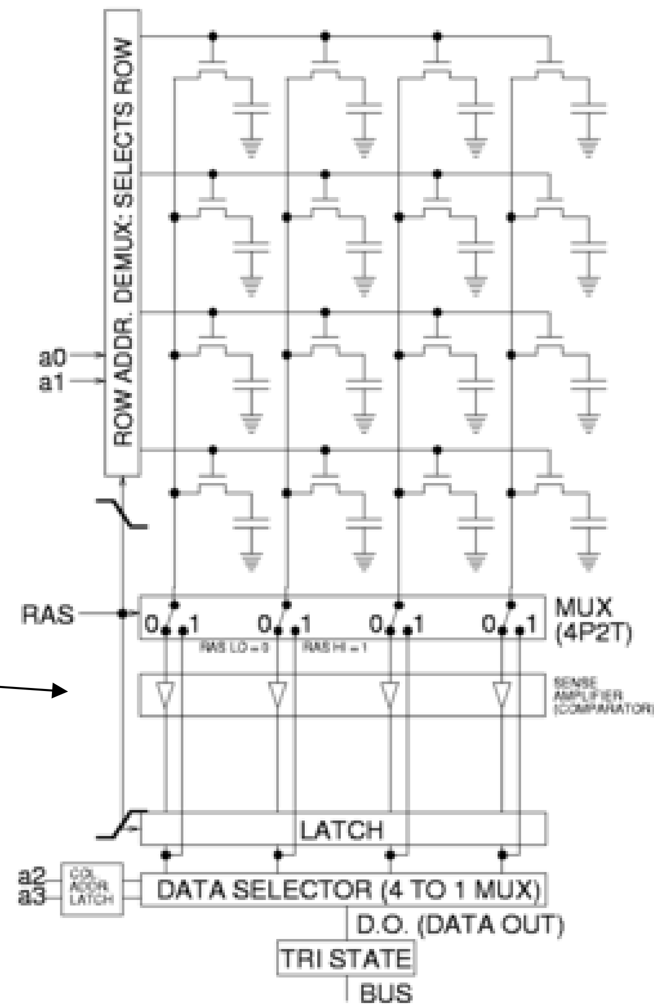


No refresh circuitry needed.

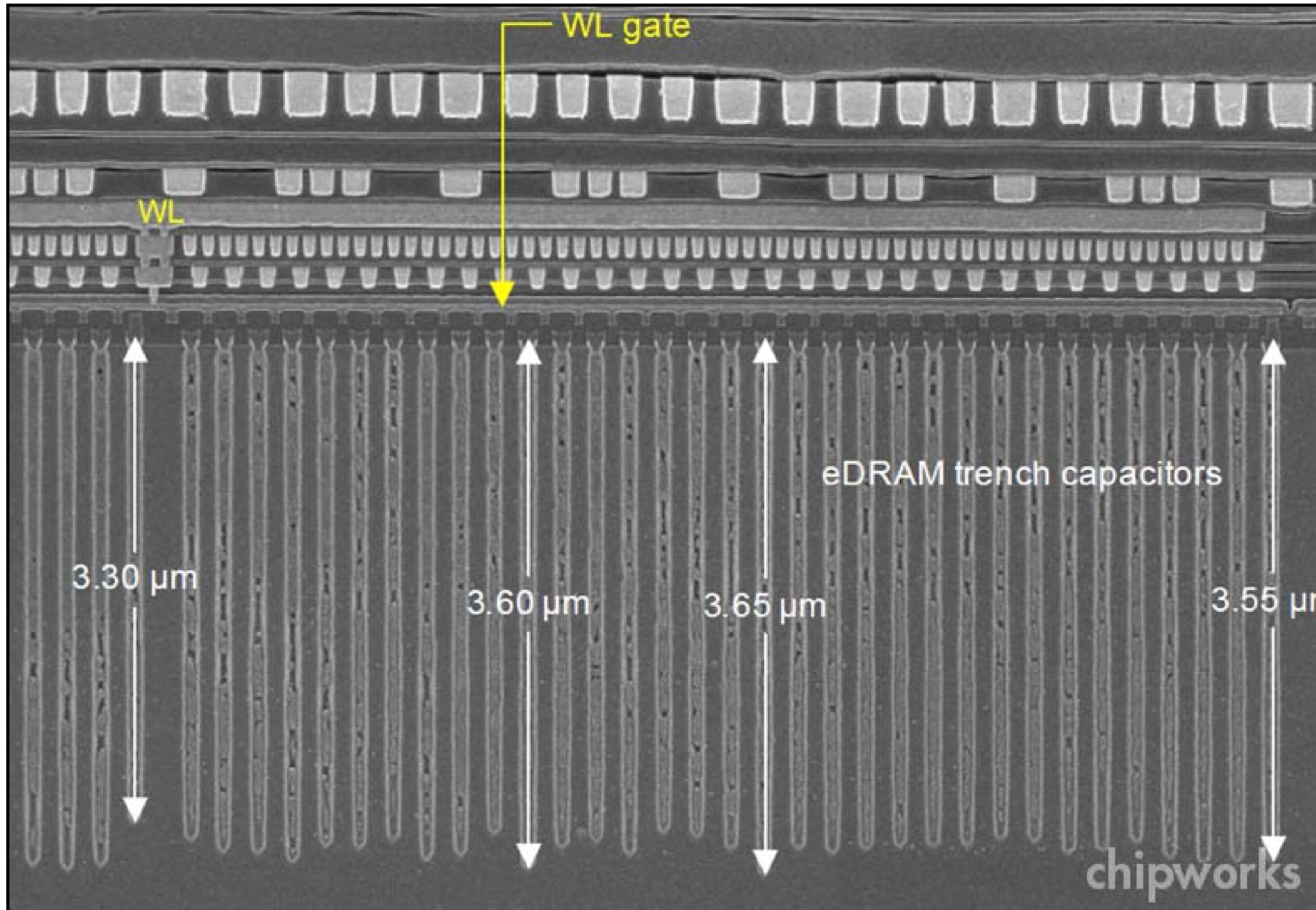
Dynamic random access memory (DRAM)

Read and refresh DRAM with a SRAM cell





DRAM

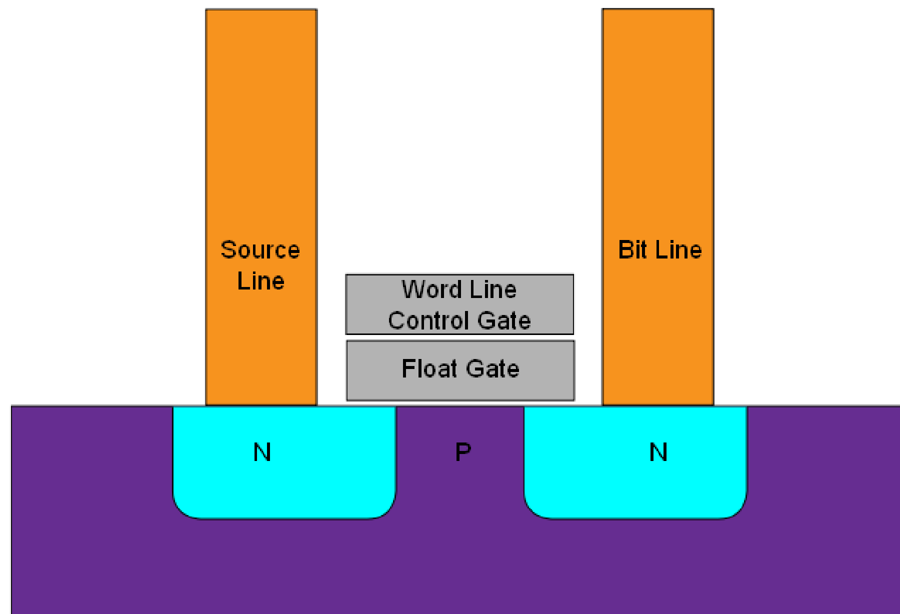


75:1

Silicon oxynitride SiO_xN_y dielectric

http://electroiq.com/chipworks_real_chips_blog/

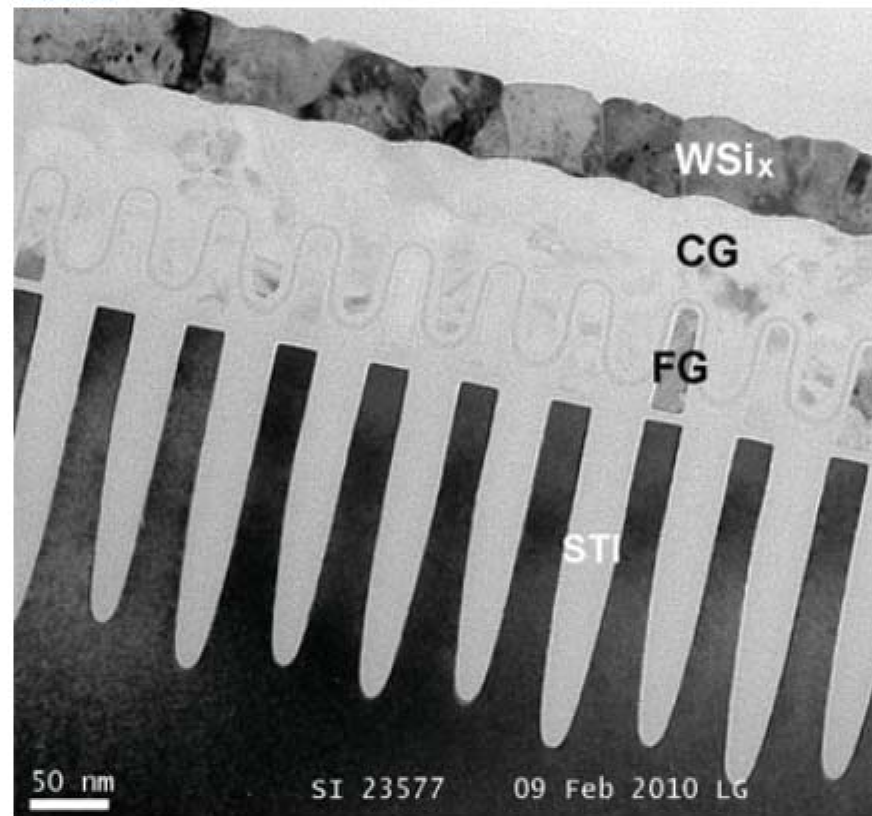
Flash memory



Charge is stored on a floating gate

nonvolatile

Intel Micron Flash Technologies (IMFT)
Shallow Trench Isolation (STI)
Control Gate (CG)
Floating Gate (FG)
Self-Aligned Doubled Patterning (SADP)



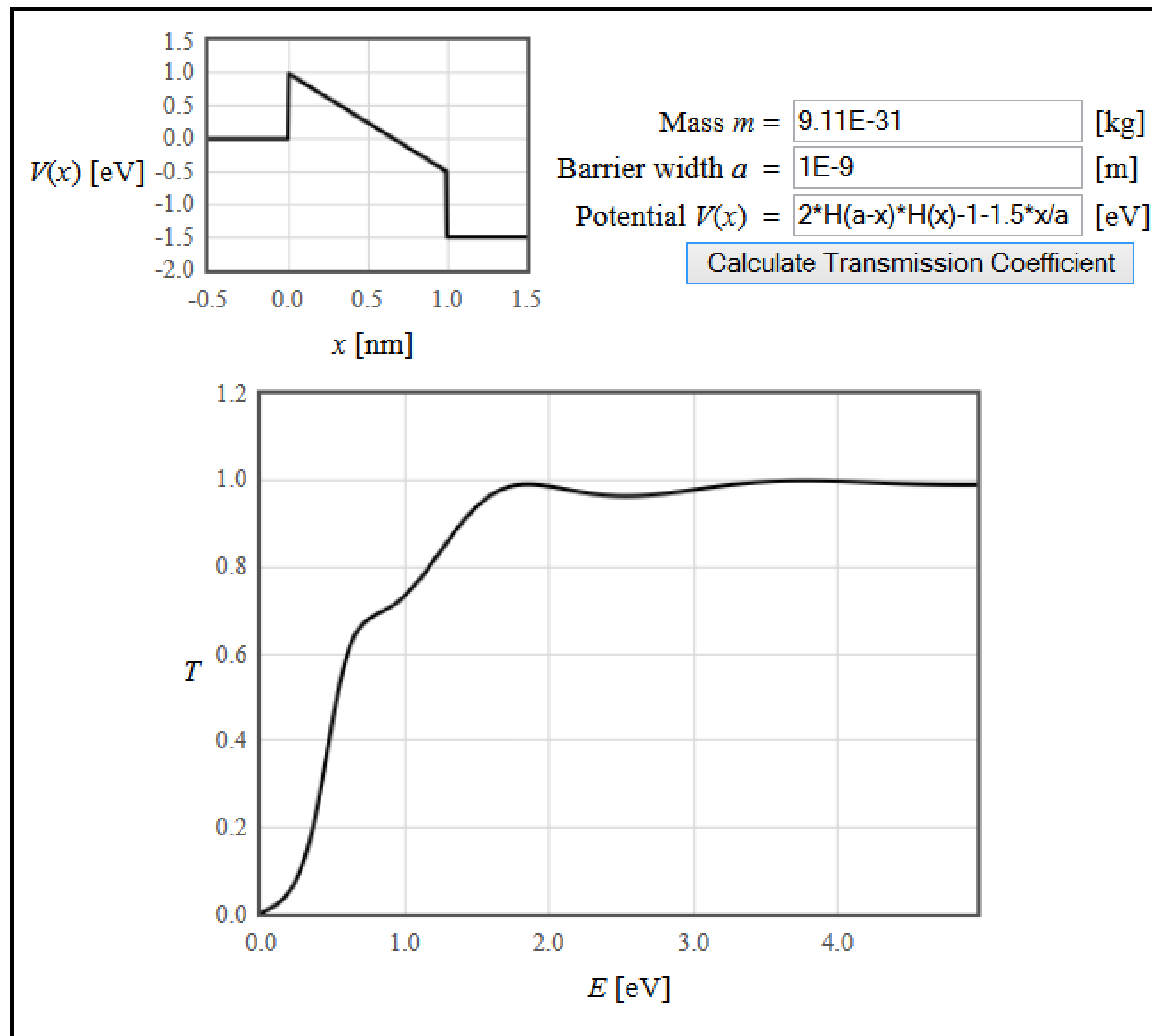
Topographical SEM image of the IMFT 25-nm flash memory array at gate level: array trench depth is shallower to allow a denser flash array

The extent of immersion-lithography tool usage cannot be known, but our end-of-the-wordline analysis and STI pattern analysis of the IMFT device has shown some interesting spacing patterns that could give useful insight into the lithography and SADP processes. Technically, immersion lithography is the mainstream technology for NAND flash integration for sub-50 nm and is used along with SADP to shrink line widths and avoid overlay issues. Strongly enhanced DP (two exposures + spacer approach) could extend immersion to 21 nm and beyond. Since the extreme ultraviolet lithography (EUVL) tool is not going to be ready till 2012, immersion would continue to fill the gap up to 2x-nm node and beyond.

http://www.dongyitech.com/en/newsdetail.asp?newsid=128

Tunneling through an arbitrarily shaped potential barrier

In quantum mechanics, there is some probability that a particle of mass m will tunnel through a potential barrier even if the energy of the particle is less than the energy of the barrier. During a direct tunneling process, the energy of the electron remains contant. The form below calculates the transmission coefficient for tunneling. The shape of the tunnel barrier can be arbitrarily defined in the interval between $x = 0$ and $x = a$. The potential is assumed to be constant to the left of the tunnel barrier at the value $V(x=0)$ and constant to the right of the barrier at the value $V(x=a)$.



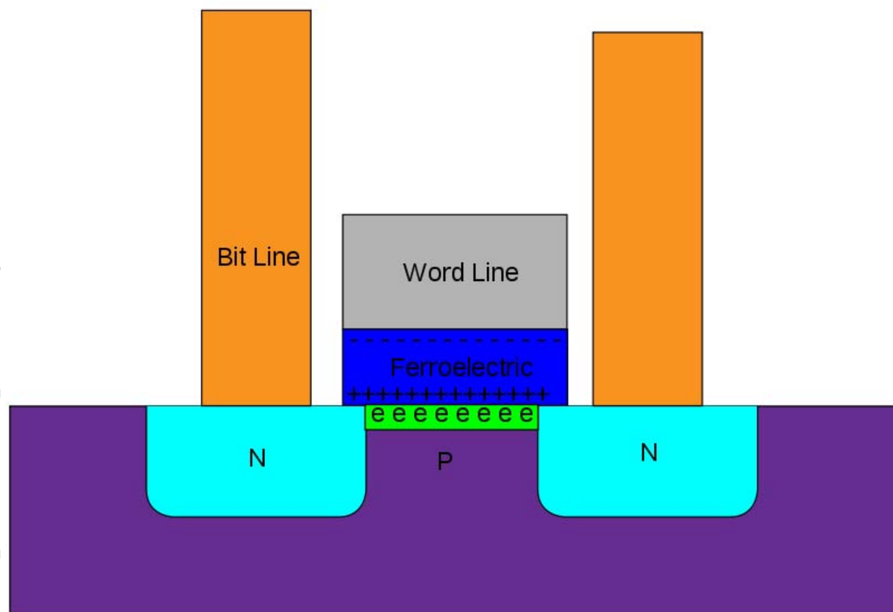
http://lamp.tu-graz.ac.at/~hadley/ss1/appendix/tunnel/tunneltrans.php

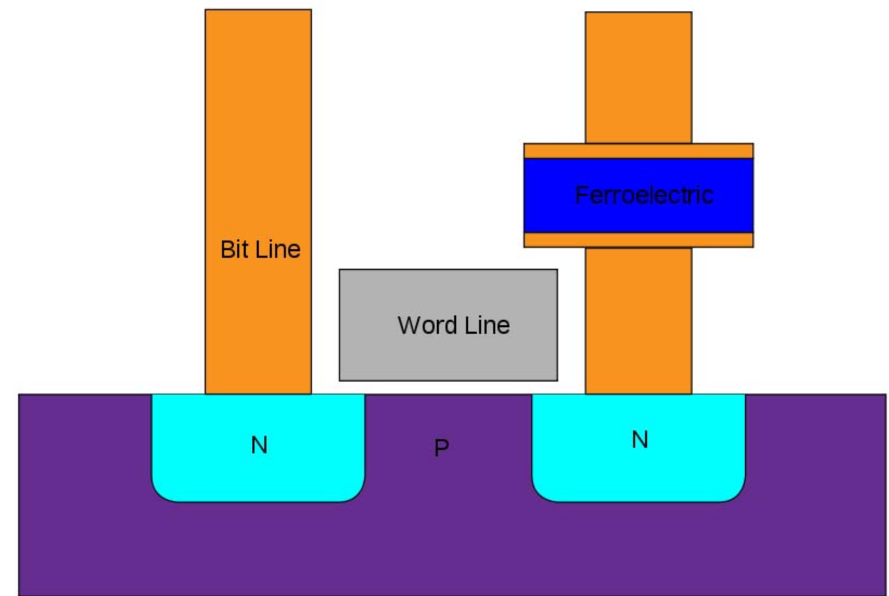
Ferroelectric RAM

FeRAM uses a Ferroelectric material like PZT to store information.

Sometimes used in smart cards.

http://en.wikipedia.org/wiki/Ferroelectric_RAM

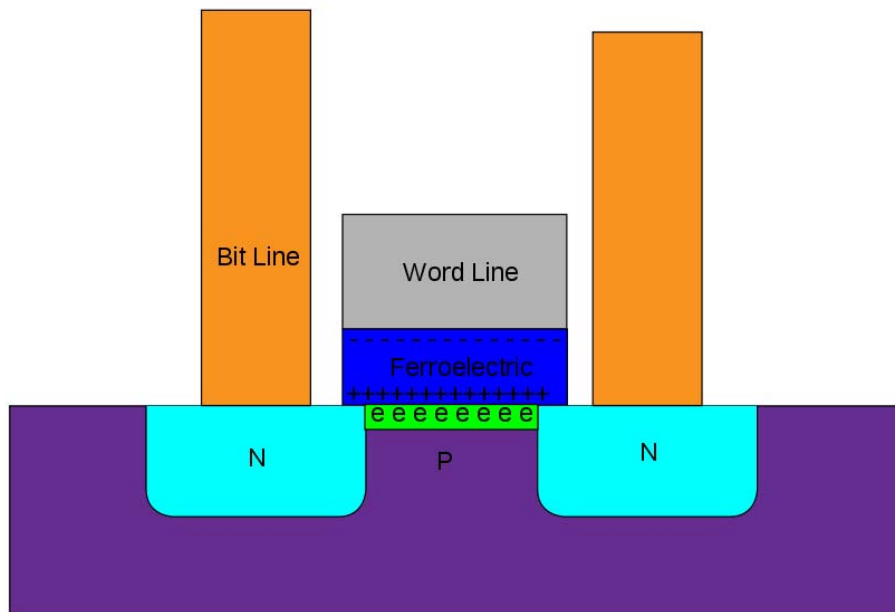


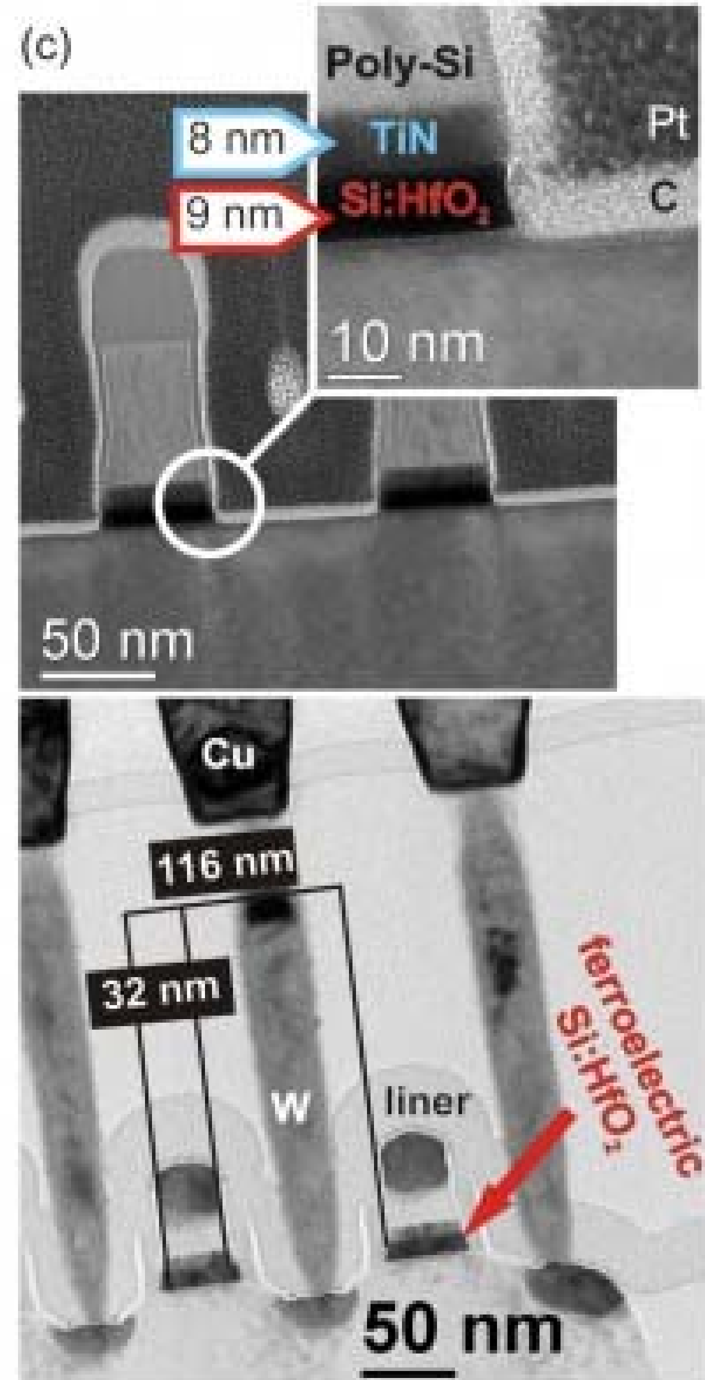


nonvolatile

To read, try to write a 0,
if a current flows, it was a 1.

Ferroelectric RAM



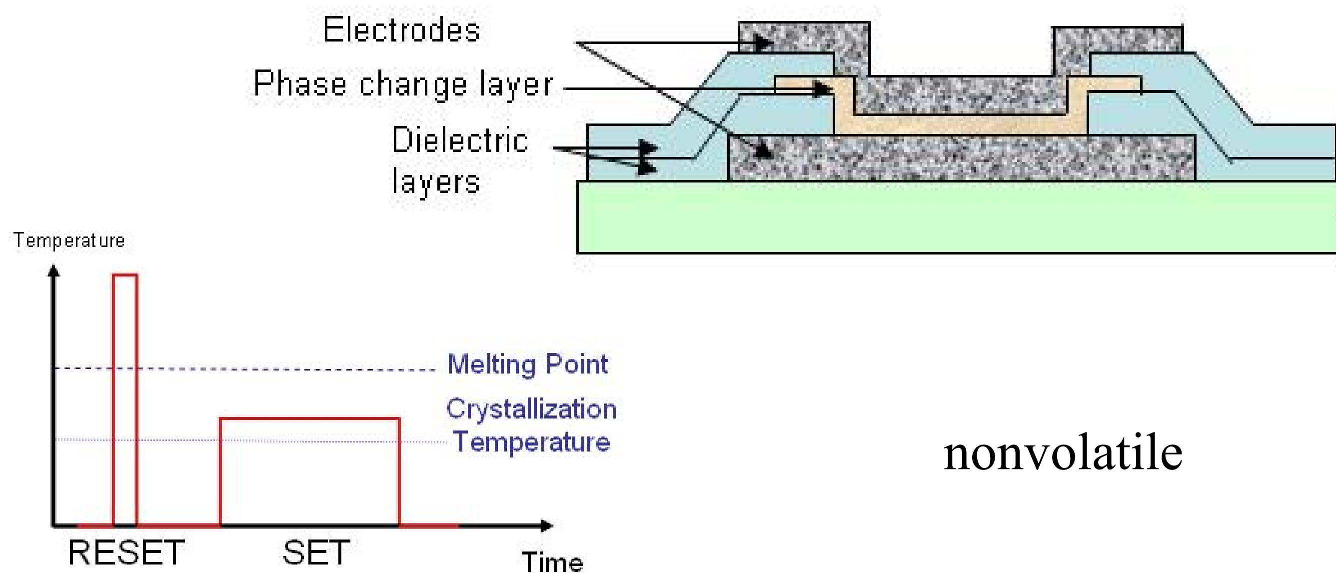


http://baldengineering.blogspot.co.at/2013_12_01_archive.html

Phase change memory

Phase-change memory (PRAM) uses chalcogenide materials. These can be switched between a low resistance crystalline state and a high resistance amorphous state.

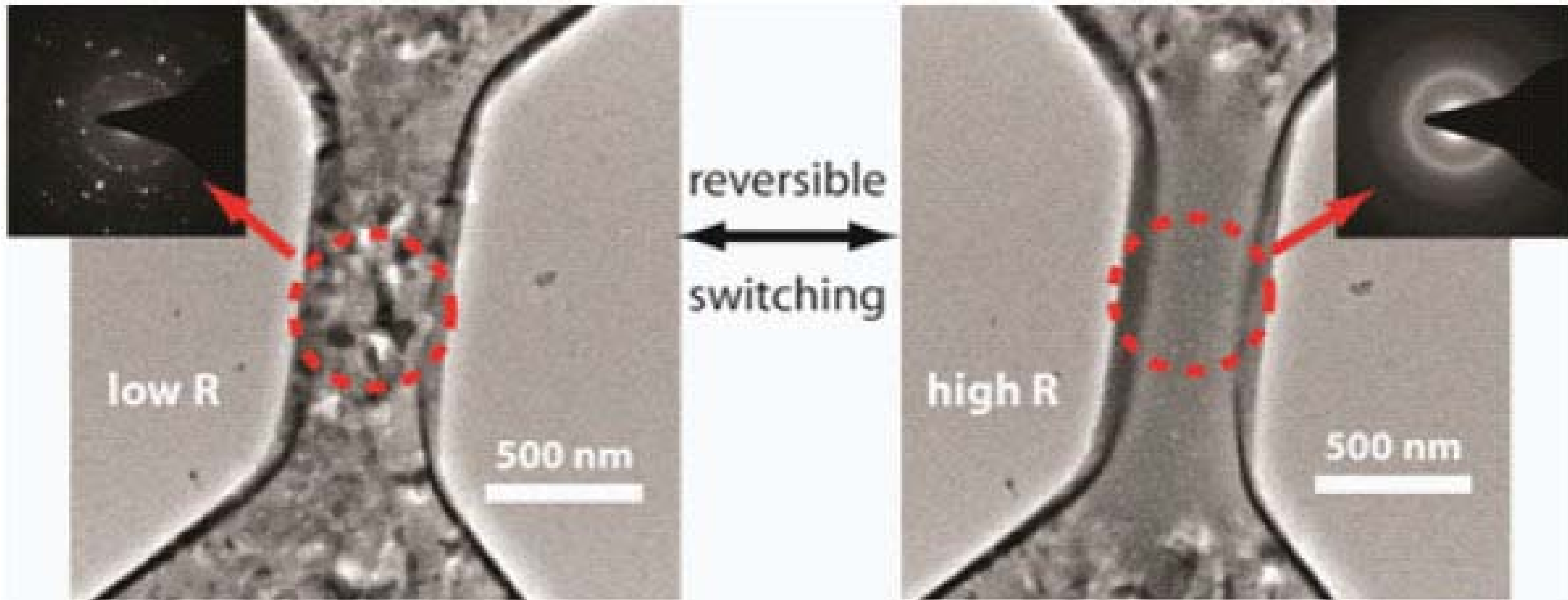
GeSbTe is melted by a laser in rewritable DVDs and by a current in PRAM.



nonvolatile

Phase change material

Electron diffraction in a TEM of a GeSbTe alloy.



http://web.stanford.edu/group/cui_group/research.htm