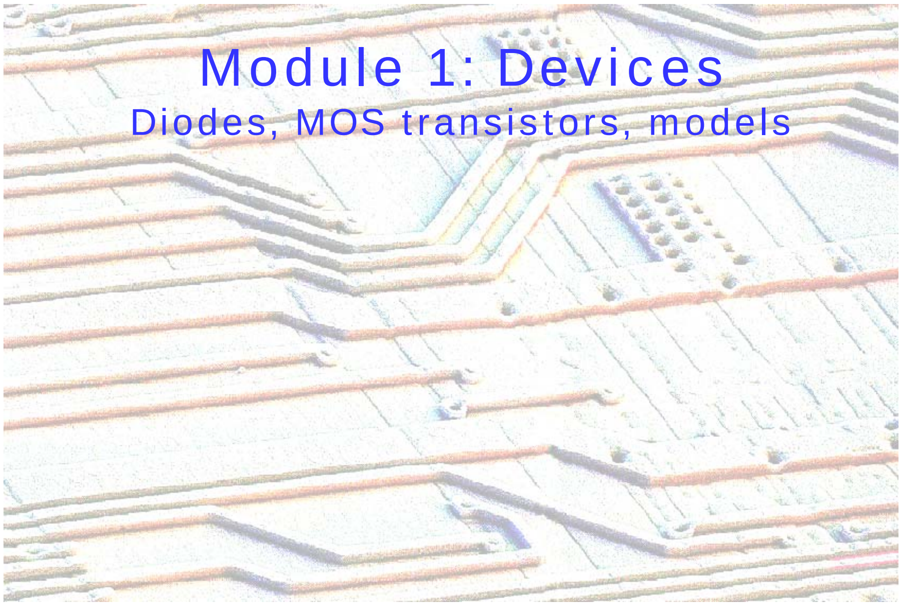
Module 1: Devices

Diodes, MOS transistors, models



Goal of this chapter

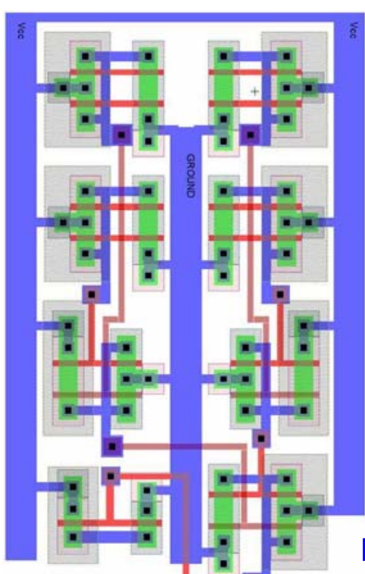
- **Present intuitive understanding of device operation**
- **Review of basic device equations**
- **Introduction of models for manual analysis**
- **Review of models for SPICE simulation**
- **Analysis of secondary and deep-sub-micron effects**
- **Future trends**

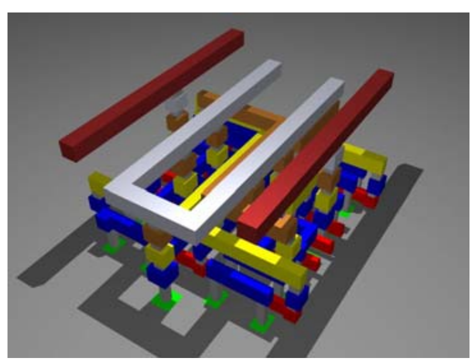
- **In depth:**
  - **ET4296 - Advanced Device Physics – 0/0/2/0 (René van Swaaij)**
  - **Neamen: *Semiconductor Physics and Devices, basic principles*, 2003, McGraw Hill**
  - **Taur and Ning: *Fundamentals of Modern VLSI Devices* 1998, Cambridge University Press**

Outline

- Semiconductor Physics
- The diode
  - Depletion, I-V relations, capacitance,
- The MOS transistor
  - First glance, threshold, I-V relations, models
  - Dynamic behavior (capacitances), resistances,
- Process variations

Chip Anatomy





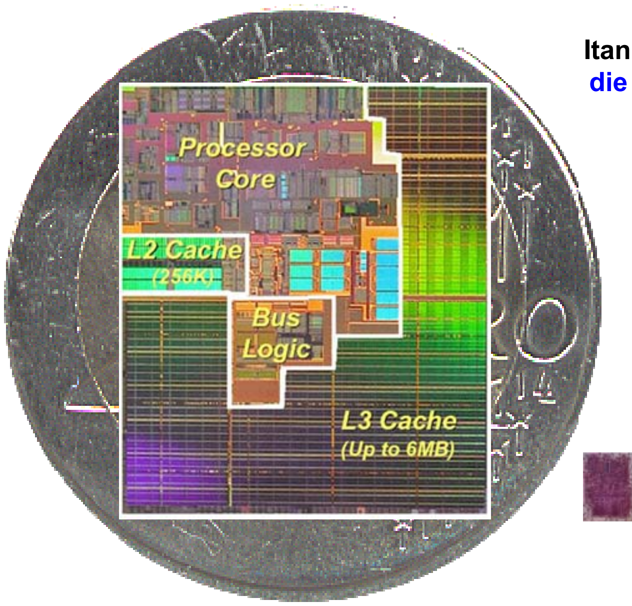
(Rendering of) result after fabrication

Not 1-to-1, sorry

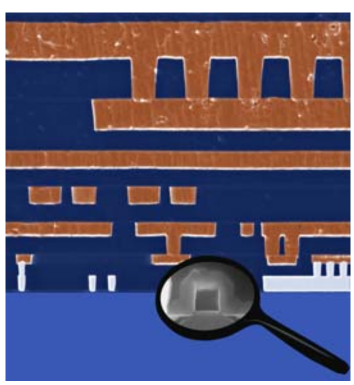
Layout of chip, final design result

Evolution

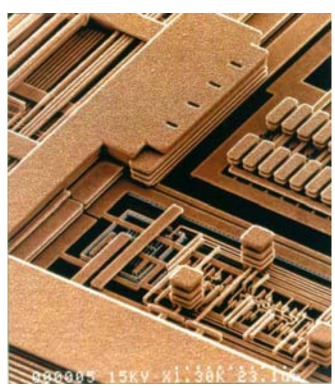
Intel 4004 (1971) and Itanium (McKinley, 2002) die compared to 2€ coin



SEM Images

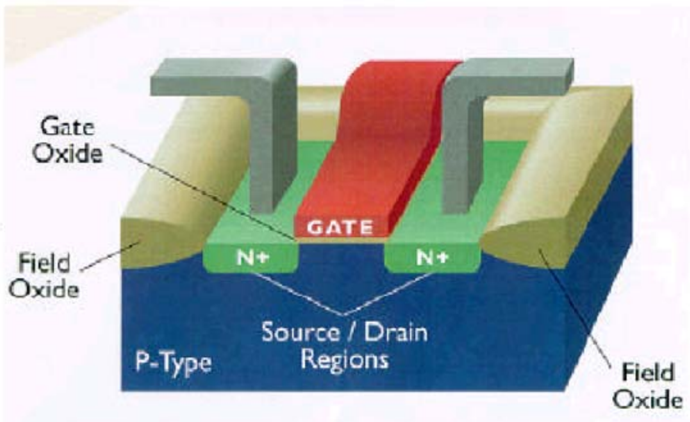


Cross-section

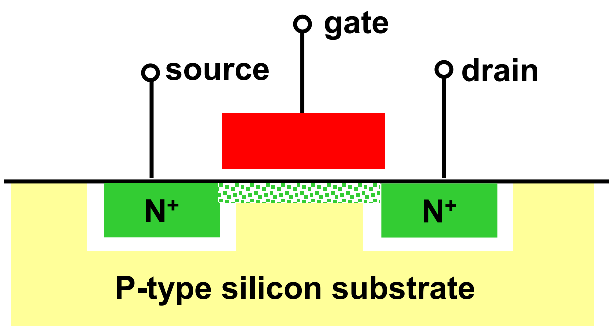


3d perspective

MOS Transistor 3D Structure



NMOS Cross-section



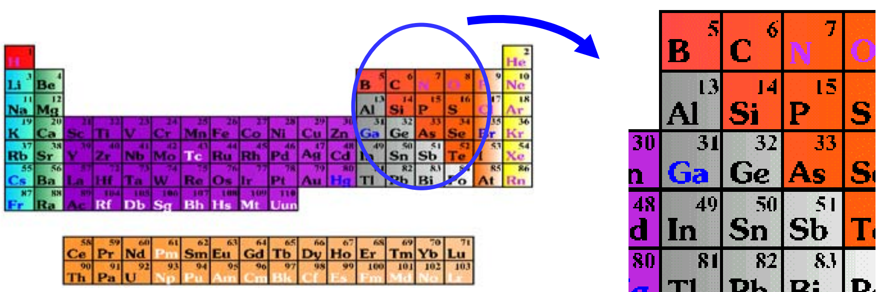
- 3 types of conductors
  - Good conductors (metallic)
  - N-type semi-conductors
  - P-type semi-conductors

What is a Semiconductor?

Material whose conductivity, due to charges of both signs, is normally in the range between that of metals and insulators and in which the electric charge carrier density can be changed by external means.

[http://goldbook.iupac.org/S05591.html]

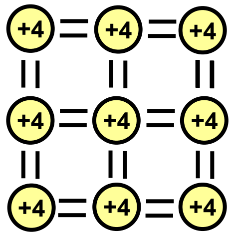
Periodic System



http://www.chemicool.com/

Name	Symbol	#	Valence
Silicon	Si	14	4
Boron	B	5	3
Phosphor	P	15	5
Arsenic	As	33	5
Germanium	Ge	32	4

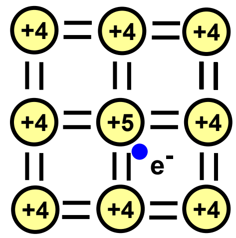
Semiconductor Physics



- Intrinsic Si
- Ideal crystal structure
- Valence 4
- almost no free carriers
- almost no conduction

$$[n] = [p] = n_i = 1{,}5.10^{10} / cm^3$$

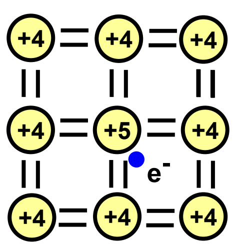
at 300 K for silicon



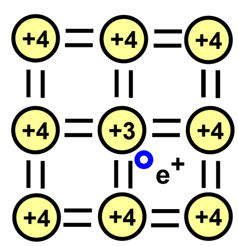
- doping with valence 5 atoms (Phosphor, Arsenic) introduces "loose electrons"
- electron donor
- conductivity depends on doping level

$$n.p = n_i^2 \text{ (in equilibrium)}$$

Semiconductor Physics



- doping with valence 5 atoms (Phosphor, Arsenic) introduces "loose electrons"
- electron donor
- conductivity depends on doping level



- doping with valence 3 atoms (Boron) introduces "loose holes"
- electron acceptors
- hole conductivity lower than electron conductivity

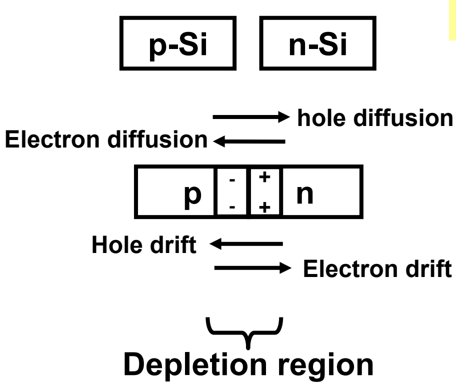
Semiconductor Physics

Si in equilibrium : $n.p = n_i^2 = 2.25 \times 10^{20}$ at 300K

Intrinsic Si : $n = p = n_i$

$N_D >> N_A$	$N_A >> N_D$
Electron donors: As, P	Electron acceptors: B
n-type Si	p-type Si
$n \approx N_D,\ p = n_i^2 / n$	$p \approx N_A,\ n = n_i^2 / p$
Electrons: majority carriers	Holes: majority carriers
Holes: minority carriers	Electrons: minority carriers
Resistive material	Hole conductivity lower than electron conductivity
Conductivity depends on N_D	

PN junction



Join n-Si with p-Si

↓

Concentration gradient of free carriers

↓

Diffusion current

↓

Space charge (depletion) region

↓

Electric field

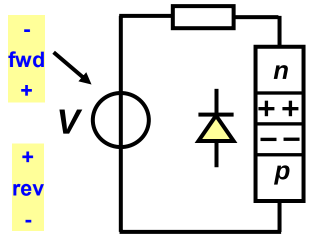
↓

Drift current opposite to diffusion

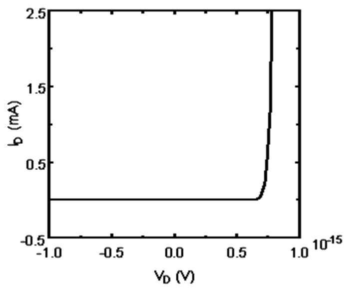
↓

equilibrium

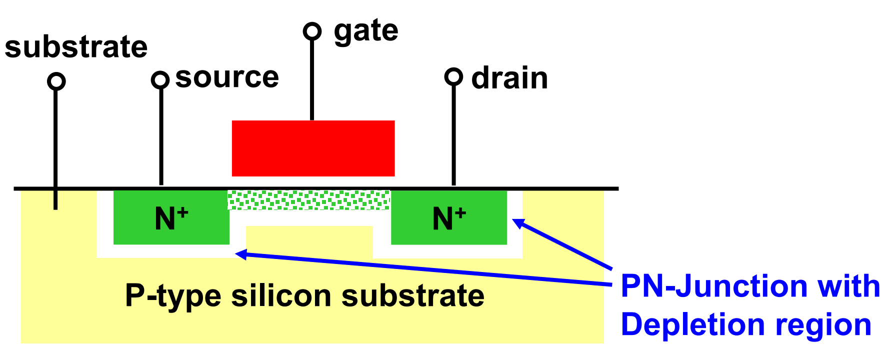
Diode Conduction



- By applying an external voltage, **width of depletion region** can be **changed**
- Forward: becomes smaller and smaller, finaly **conduction**
- Reverse: becomes wider and wider **=> no conduction**

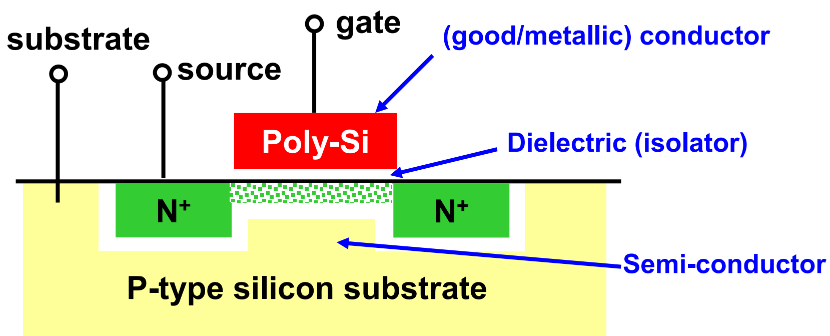


$$I_D = I_S\left(e^{V_D/\phi_T} - 1\right)$$



- Drain and source voltage relative to substrate voltage should maintain isolation!
  Substrate = Bulk = B, Source = S, Drain = D
  $V_{SB} > 0$, $V_{DB} > 0$
- Normally the case

MOS Capacitor

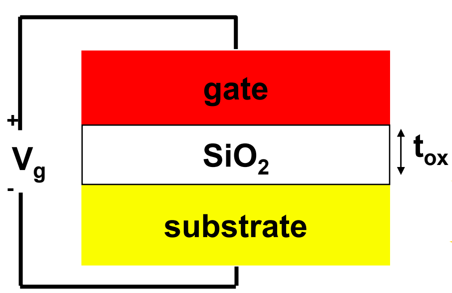


Inversion layer

- Presence of inversion layer depends on voltages relative to substrate
- If present: **conduction** between source and drain

MOSFET gate as capacitor

- Gate capacitance helps determine charge in channel which forms **inversion region**.
- Basic structure of gate is **parallel-plate capacitor**:



$C_{ox} = \varepsilon_{ox} / t_{ox}$

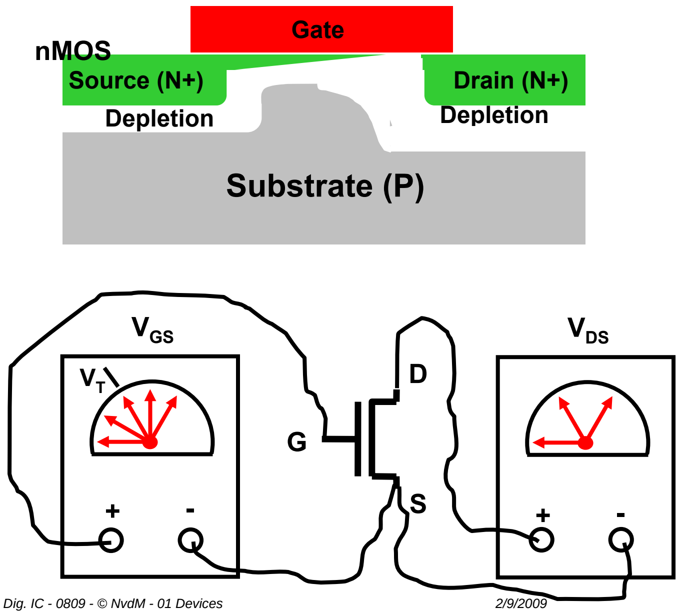
$\varepsilon_{ox} = \varepsilon_0 \varepsilon_r$

$\varepsilon_0 = 8.85 \times 10^{-12}$ F/m

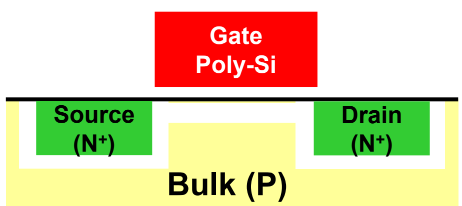
$\varepsilon_r = 3.9$ (SiO_2)

Note: [F] vs. [F/m²]

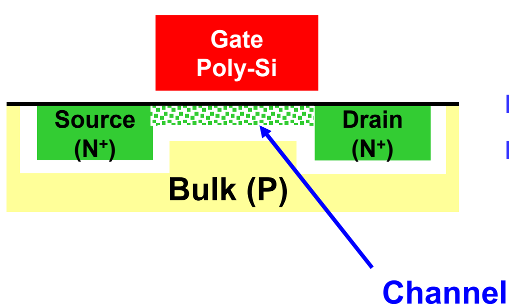
nMOS Transistor Operation



MOS Summary



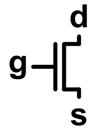
- Low gate voltage
- Transistor off



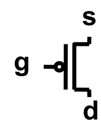
- High gate voltage
- Conduction between source and drain

CMOS – *Complementary* Metal Oxide Semiconductor Technology

2 Distinct Transistor Types



n-type



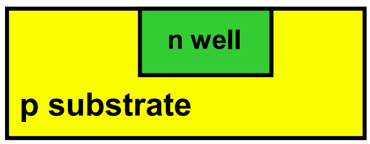
p-type

n-type:

- "on" when V_g is high
- With **n-type s/d**
- **Electrons** (n) as **carrier**
- Built in **p-type Si**

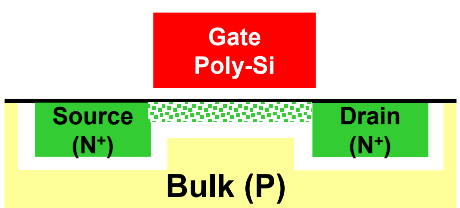
p-type:

- "on" when V_g is low
- With **p-type s/d**
- **Holes** (p) as **carrier**
- Built in **n-type Si**



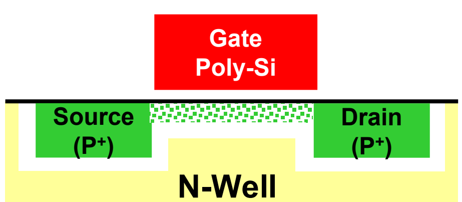
n-well (for PMOS) in p-type substrate (for NMOS)

NMOS vs PMOS



NMOS

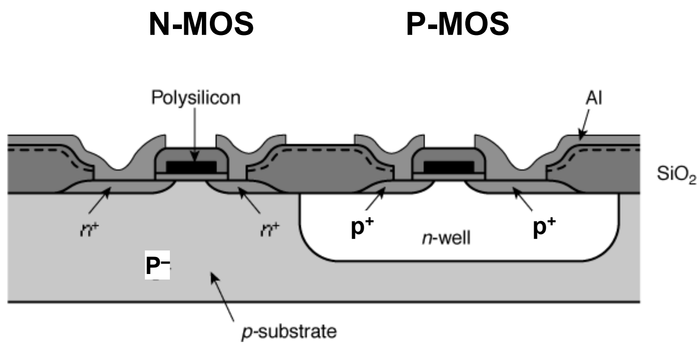
- On when gate voltage is high
- Off when gate voltage is low



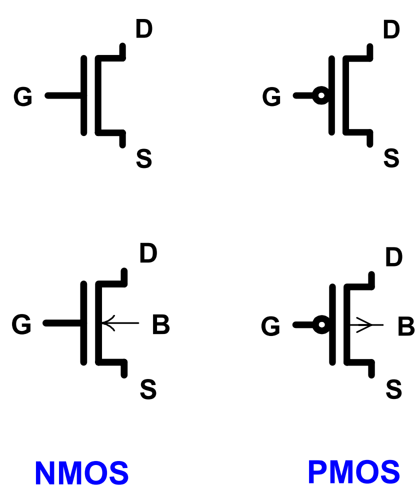
PMOS

- N and P regions opposite of NMOS
- On when gate voltage is low
- Off when gate voltage is high

Cross-Section of CMOS Technology



MOS Transistors



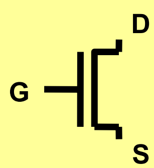
3-terminal model

bulk assumed to be connected to appropriate supply

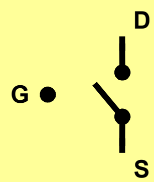
4-terminal model

B = bulk (substrate)

MOS Transistor
Switch Level Models



NMOS



Simplest possible useful model

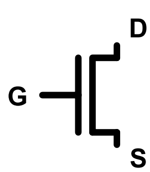
Position of switch depends on gate voltage

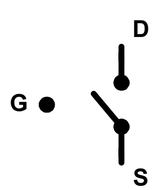
V_G	NMOS	PMOS
hi	closed	open
lo	open	closed

- **Connection between source and drain depends on gate voltage, current can flow from source to drain and vice versa if closed**
- **No static current flows into gate terminal**

MOS Transistor Treshold Voltage

Treshold voltage V_T: point at which transistor turns on

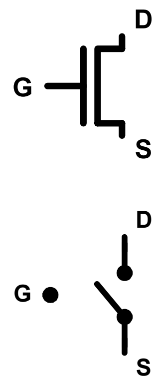




Position of switch depends on gate voltage (relative to source)

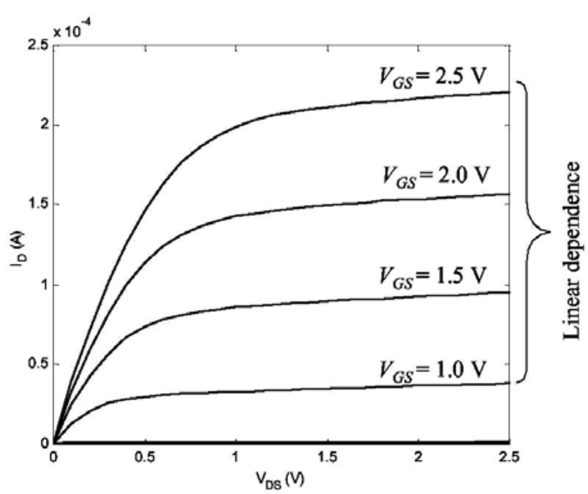
V_{GS}	NMOS	PMOS
$V_{GS}>V_T$	closed	open
$V_{GS}<V_T$	open	closed

Is this all there is?



- You don't believe that (CMOS) life can be so **simple**, do you?
- {TPS} some of the things that you would expect to be **non-idealities** of CMOS as a switch
- Since we want to design CMOS circuits, we need a **deeper understanding** of CMOS circuit

Drain Current I_D

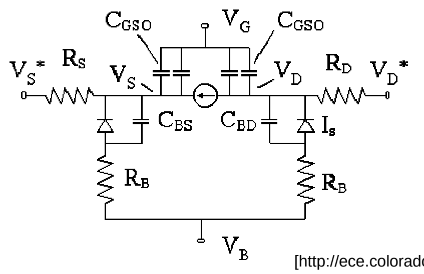


(b) Short-channel transistor (L_d = 0.25 μm)

- You might remember quadratic dependence
- Not true anymore for short-channel devices (velocity saturation)

SPICE Model

- Some very advanced models to describe MOS devices over all combinations of terminal voltage
  - Includes dynamic behavior, thermal, …
  - BSIM3, BSIM4, HICUPS, PSP, MEXTRAM, …
  - Can be used for Circuit Simulation
  - >10k lines of C-code

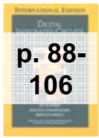


[http://ece.colorado.edu/~bart/book/book/index.html]

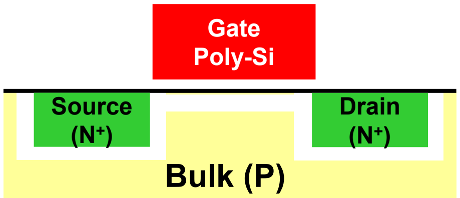
MOS Operating Regimes

- Off
- Saturation
- Linear – Triode – Resistive
- Velocity Saturation
- (Sub-threshold)

- Different formulas for drain current I_D in each region
- You need to understand these principles



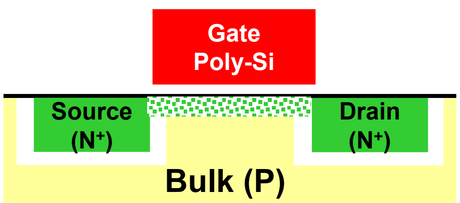
Off-regime (NMOS)



$V_{GS} < V_T$

- Easiest – current I_D is essentially zero
- When V_{GS} approaches V_T, a small current starts to flow (**sub-threshold current**)
  - Important phenomenon for small, low-voltage devices
  - Important opportunity for ultra-low power circuits

Triode-regime (NMOS)



$V_{GS} > V_T, V_{GD} > V_T$

$\Leftrightarrow$

$V_{GS} > V_T, V_{DS} < V_{GS} - V_T$

- Inversion both at source-side and drain-side of channel
- I_D depends on V_{DS}: triode behavior
- I_D depends on $V_{GS} - V_T$
- Inversion not completely symmetric if $V_{DS} > 0$

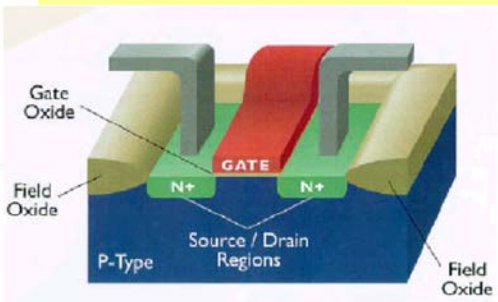
$$I_D = k\left[(V_{GS} - V_T)V_{DS} - \frac{1}{2}V_{DS}^2\right]$$

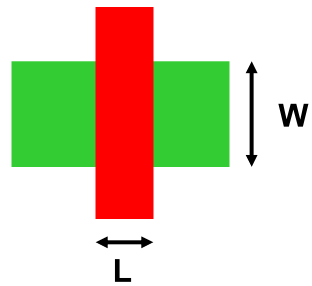
k, V_T: device parameters

Device Sizing

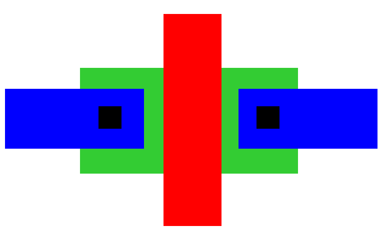
$$I_D = k\left[(V_{GS} - V_T)V_{DS} - \frac{1}{2}V_{DS}^2\right] \qquad k = k'\frac{W}{L}$$

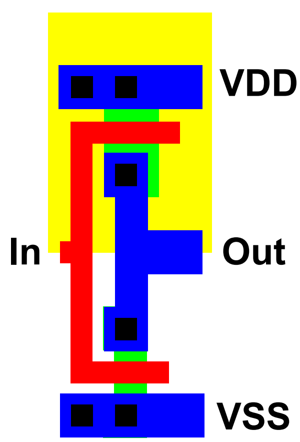
k	**Device transconductance**
k'	**Process transconductance**
W	**Device width**
L	**Device length**

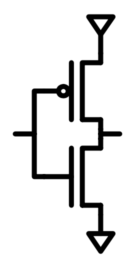




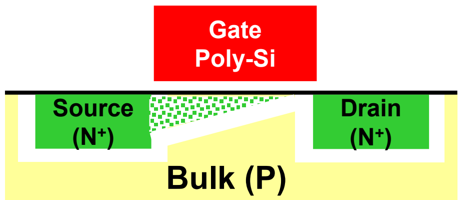
Device Layout







Saturation (NMOS)



$V_{GS} > V_T, V_{GD} < V_T$

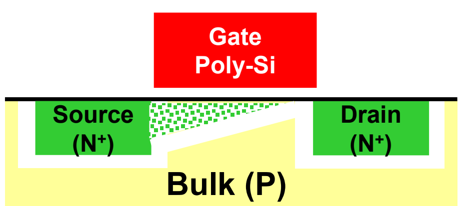
$\Leftrightarrow$

$V_{GS} > V_T, V_{DS} > V_{GS}-V_T$

- **Inversion only at source-side of channel**
- **Still, current will be flowing between S and D**
- **Current does not (strongly) depend on V_{DS}: current source behavior**

$$I_D = \frac{1}{2}k(V_{GS} - V_T)^2$$

Velocity Saturation



$V_{DS} > V_{DSAT}$

- **When V_{DS} large enough – current doesn't increase anymore since carrier velocity is limited by scattering to lattice**
- **Visible in (advanced) short-channel devices**
- **Can happen in (otherwise) saturation conditions, but also in triode conditions**

$$I_D = k\left[(V_{GS} - V_T)V_{DSAT} - \frac{1}{2}V_{DSAT}^2\right]$$

(*k*, V_T,V_{DSAT}: device data)

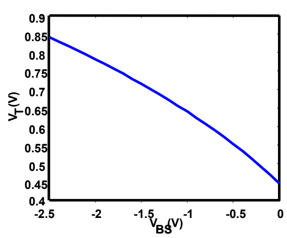
More Effects

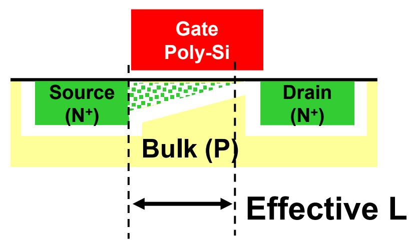
- **Body Effect: V_T depends on V_{SB}**

$$V_T = V_{T0} + \gamma\left(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{|-2\phi_F|}\right)$$

- **Channel length Modulation:**
  I_D depends on V_{DS} also in saturation

$$I_D = I_D \times (1 + \lambda V_{DS})$$





Summary

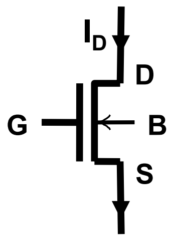
$$I_D = \begin{cases} k\left[(V_{GS} - V_T)V_{DS} - \frac{1}{2}V_{DS}^2\right] & V_{GS} > V_T,\ V_{DS} < V_{GS}-V_T \\ \frac{1}{2}k(V_{GS} - V_T)^2 & V_{GS} > V_T,\ V_{DS} > V_{GS}-V_T \\ k\left[(V_{GS} - V_T)V_{DSAT} - \frac{1}{2}V_{DSAT}^2\right] & V_{DS} > V_{DSAT} \end{cases}$$

$$I_D = I_D \times (1 + \lambda V_{DS})$$

$$V_T = V_{T0} + \gamma\left(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{|-2\phi_F|}\right)$$

Next slide presents alternative formulation

MOS Models for Manual Analysis



determined by circuit
V_{DS}, V_{GS}, V_{SB}

determined by designer
W, L

determined by technology
$k', \lambda, V_{DSAT}, V_{T0}, \gamma, \phi_F$

MOS model for manual analysis

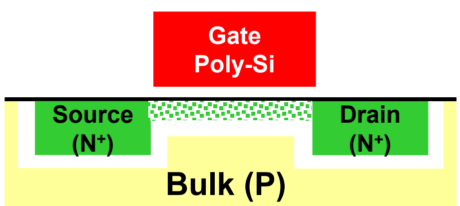
$$I_D = k\left(V_{GT}V_{MIN} - 0.5V_{MIN}^2\right)\left(1 + \lambda V_{DS}\right) \quad \text{for } V_{GT} \geq 0$$

$$= 0 \quad \text{for } V_{GT} \leq 0$$

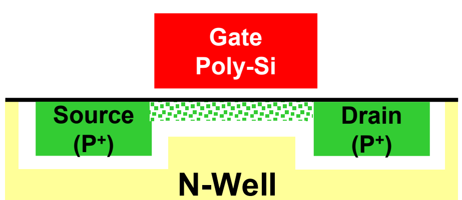
$$V_{MIN} = MIN\left(V_{DS}, V_{GT}, V_{DSAT}\right) \qquad k = k'\frac{W}{L}$$

$$V_{GT} = V_{GS} - V_T, \qquad V_T = V_{T0} + \gamma\left(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{|-2\phi_F|}\right)$$

NMOS vs PMOS



NMOS
- On when gate voltage is high
- Off when gate voltage is low

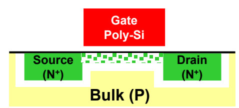


PMOS
- N and P regions opposite of NMOS
- On when gate voltage is low
- Off when gate voltage is high

NMOS vs PMOS

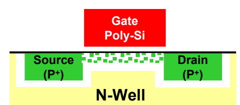
- **PMOS: all polarities reversed for same operation region**
- **Hole-mobility < Electron-mobility**
- **$|k'_{PMOS}| < k'_{NMOS}$**

NMOS vs PMOS



NMOS

- On when gate voltage is high
- Off when gate voltage is low



PMOS

- N and P regions opposite of NMOS
- On when gate voltage is low
- Off when gate voltage is high

NMOS Regions of Operation

Triode	⇔	$V_{gs} > V_t$ and $V_{ds} < V_{gs} - V_t$
Saturation	⇔	$V_{gs} > V_t$ and $V_{ds} > V_{gs} - V_t$
Cut-off	⇔	$V_{gs} \leq V_t$

PMOS Regions of Operation

Triode	⇔	$V_{gs} < V_t$ and $V_{ds} > V_{gs} - V_t$
Saturation	⇔	$V_{gs} < V_t$ and $V_{ds} < V_{gs} - V_t$
Cut-off	⇔	$V_{gs} \geq V_t$

Universal

Triode	⇔	$\|V_{gs}\| > \|V_t\|$ and $\|V_{ds}\| < \|V_{gs}\| - \|V_t\|$
Saturation	⇔	$\|V_{gs}\| > \|V_t\|$ and $\|V_{ds}\| > \|V_{gs}\| - \|V_t\|$
Cut-off	⇔	$\|V_{gs}\| \leq \|V_t\|$

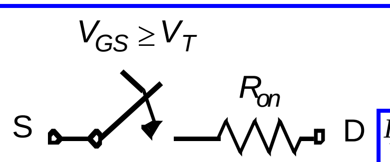
Unified Model Parameters

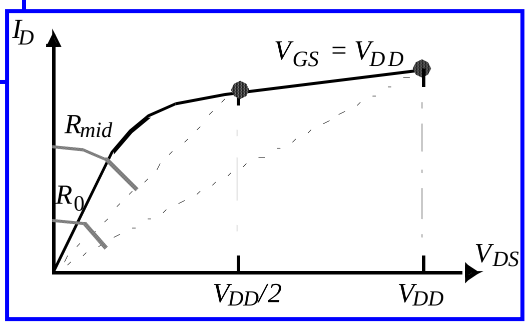
Table 3.2 Parameters for manual model of generic 0.25 μm CMOS process (minimum length device).

	V_{T0} (V)	γ ($V^{0.5}$)	V_{DSAT} (V)	k' (A/V^2)	λ (V^{-1})
NMOS	0.43	0.4	0.63	115×10^{-6}	0.06
PMOS	−0.4	-0.4	-1	-30×10^{-6}	-0.1

- **Parameters depend on technology**
- **See tables in front and back cover of book**

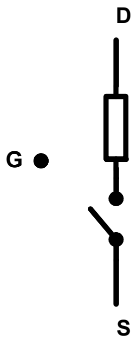
The Transistor as a Switch





$$R_{eq} = \frac{1}{2}\left(\frac{V_{DD}}{I_{DSAT}(1+\lambda V_{DD})} + \frac{V_{DD}/2}{I_{DSAT}(1+\lambda V_{DD}/2)}\right) \approx \frac{3}{4}\frac{V_{DD}}{I_{DSAT}}\left(1 - \frac{5}{6}\lambda V_{DD}\right)$$

MOS Transistor
Switch Level Model (Empirical).



$$R_{on} = \frac{L}{W} R_{eq}$$

Position of switch depends on gate to source voltage

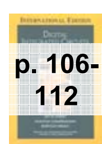
V_{GS}	NMOS	PMOS
hi	closed	open
lo	open	closed

R_{eq}: Practice!

R_{eq} \ V_{dd} (V)	1	1.5	2	2.5
NMOS (kΩ)	35	19	15	13
PMOS (kΩ)	115	55	38	31

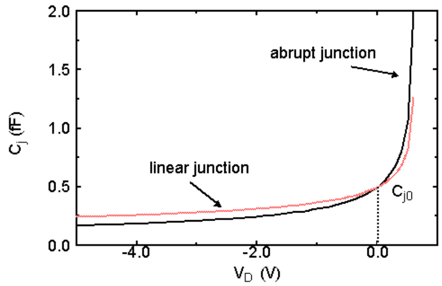
Dynamic Behavior

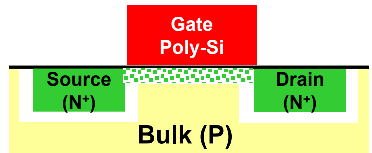
- (Solely) governed by **time needed to (dis)charge** (intrinsic, parasitic) **capacitances** associated with device and interconnect
- **Essential knowledge** for designing high-quality ckts.
- Many caps are **non-linear**
- **Spice models** can accurately take C into account
- Need **simplification** and insight for *design*
  - Linearization
  - Lumping/merging

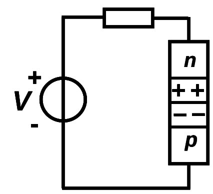


p. 106-112

Junction Capacitance





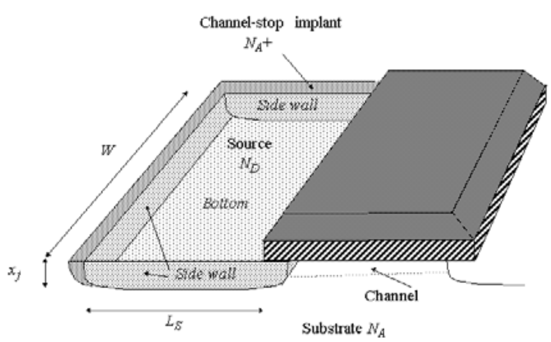


$$C_j = \frac{C_{j0}}{(1 - V_D/\phi_0)^m}$$

m = 0.5: abrupt junction
m = 0.33: linear junction

- Space-charge / depletion region creates electric field
- Electric field energy works like capacitor (energy is ½CV^2)
- Width of depletion region depends on voltage: non-linear C

MOS S/D Capacitance



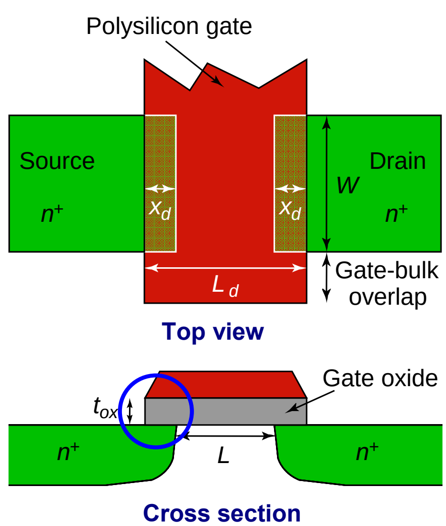
$$\begin{aligned} C_{diff} &= C_{bottom} + C_{sidewall} \\ &= C_j \times AREA + C_{jsw} \times PERIMETER \\ &= C_j L_s W + C_{jsw}(2L_s + W) \end{aligned}$$

Diode capacitances ⇒ use linearized values

Gate Capacitance

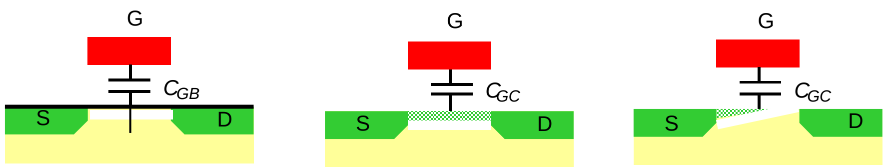
- **Gate-to-channel**
  - **Complex function of operating voltages**
  - **Because channel charge depends on voltages**
- **Gate-source and gate-drain overlap capacitance**
  - **Just (almost) linear capacitances depending only on geometry**

Gate-Overlap Capacitance



$$C_{gate} = \frac{\varepsilon_{ox}}{t_{ox}} WL$$

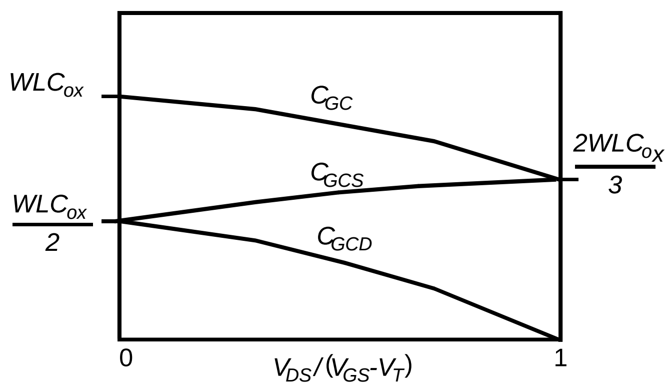
Gate-Channel Capacitance



Operation Region	C_{gb}	C_{gs}	C_{gd}
Cutoff	$C_{ox}WL_{eff}$	0	0
Triode	0	$C_{ox}WL_{eff}/2$	$C_{ox}WL_{eff}/2$
Saturation	0	$(2/3)C_{ox}WL_{eff}$	0

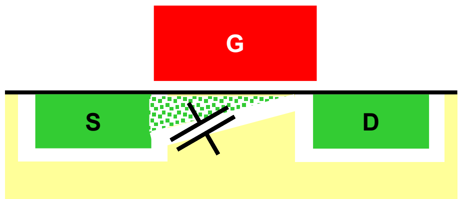
Most important regions in digital design: saturation and cut-off

Gate-Channel Capacitance



Degree of saturation

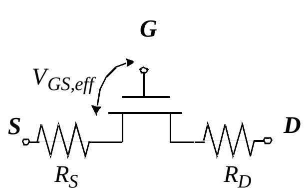
Channel-Bulk Capacitance

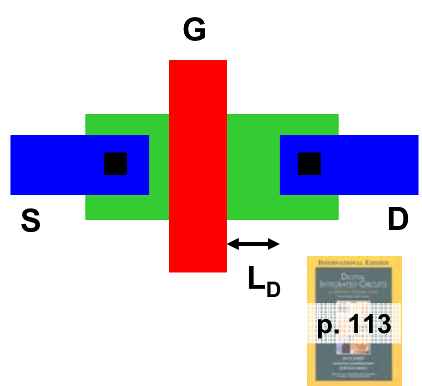


- Channel-bulk cap C_{CB}
- Only when transistor is on
- Parallel to C_{SB}

TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 01 Devices 2/9/2009 53

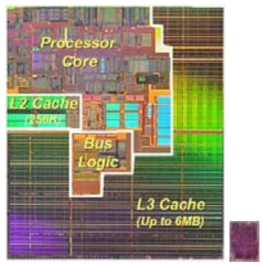
Device Parasitic Resistance





TUD/EE ET4293 - Dig. IC - 0809 - © NvdM - 01 Devices 2/9/2009 54

Technology Scaling



McKinley 4004

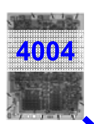
Processor	4004	McKinley	Montecito
Year	1971	2002	2005
Feat. size	10μm	0.18μm	90nm
Die size	$12mm^2$	$421mm^2$	$596mm^2$
Transistors	2300	221×10^6	1.7×10^9 91% in L3 cache
Clock	108 kHz	1GHz	1.8GHz
Perform. (spec2000)	0.01	810	~1600

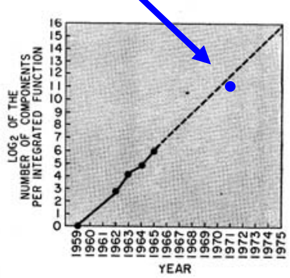
4004 →

Comparative interconnect dimensions

▯▯▯▯▯ ← McKinley

Moore's Law





The number of transistors that can be integrated on a single chip will double every 18 months

Gordon Moore, co-founder of Intel [Electronics, Vol 38, No. 8, 1965]

Chip Size Trends – 2007 ITRS Functions/Chip Model

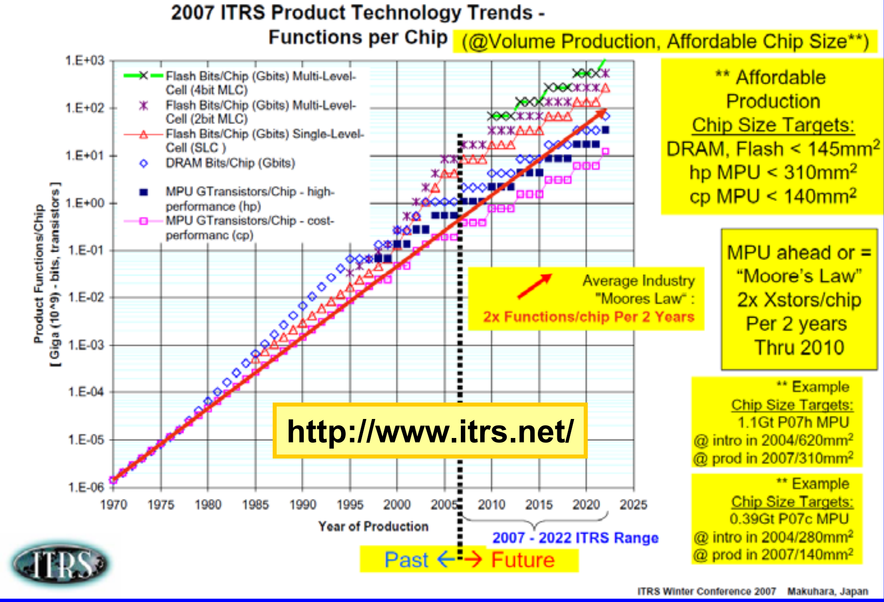
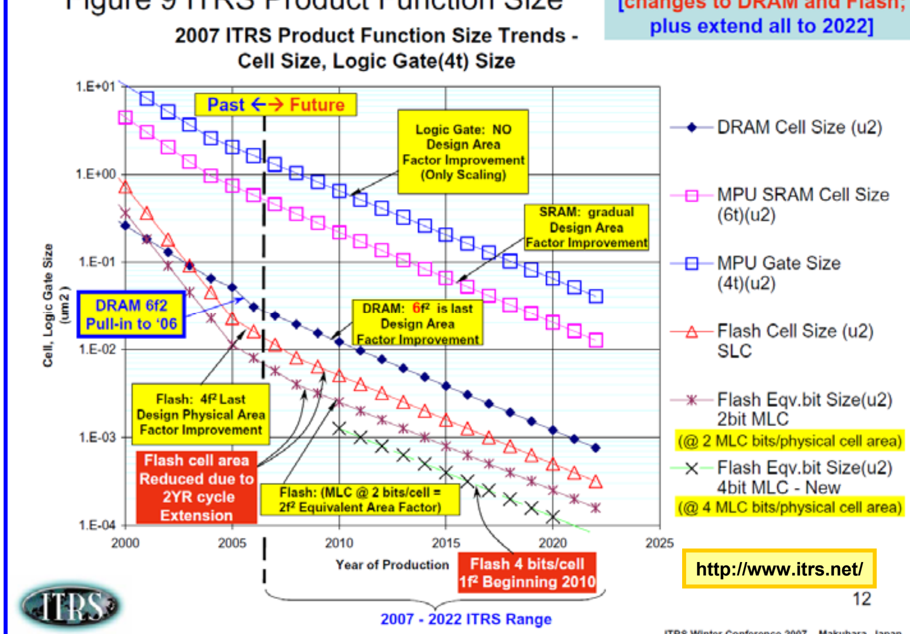
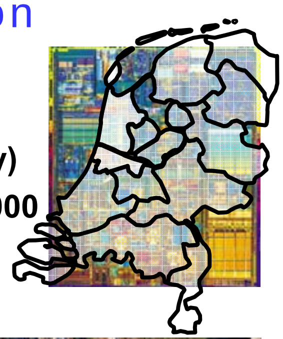


Figure 9 ITRS Product Function Size



IC Technology—Comparison



Chip: 4 cm^2

Netherlands: 40,000 km^2 (approximately)

Scale: 2 cm / 200 km = 1:10,000,000

A 90nm chip compares to Netherlands full of roads:

- 0.9 meters wide
- 0.9 meters apart
- 8 layers

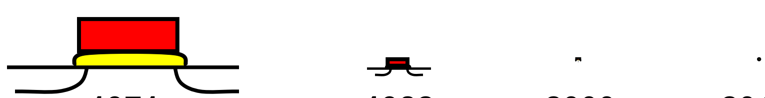
IC Technology Scaling

Scaling improves density and performance

- First order **scaling theory**

		2000 / 1971
dimensions, voltage	1/S	0.02
intrinsic delay	1/S	0.02
power per transistor	$1/S^2$	0.0004
power-delay product	$1/S^3$	0.000008

- Scaling **trend**

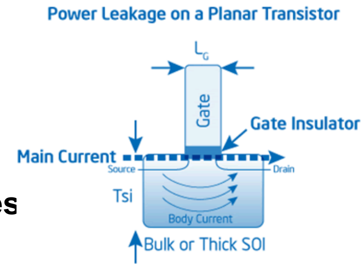


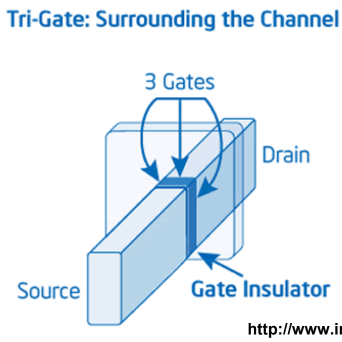
1971 S=1 | 1982 S≈5 | 2000 S≈50 | 2010 S≈200

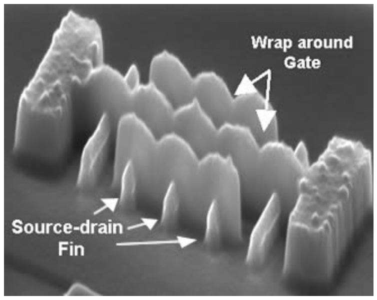
first µproc

Advanced Issues

- Variability, manufacturing tolerances
- Scaling
- Reliability
- Advanced device architectures







http://www.intel.com/technology/silicon/tri-gate-demonstrated.htm