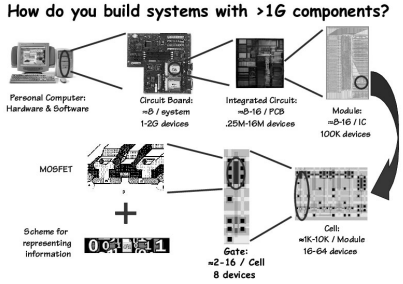
Wstęp do informatyki
budowa sprzętu

Ignacy Pardyka

Podstawy techniki cyfrowej cz. 1

Podstawy techniki cyfrowej

How do you build systems with >1G components?



Kodowanie

Fixed-length encodings

If all choices are equally likely (or we have no reason to expect otherwise), then a fixed-length code is often used. Such a code will use at least enough bits to represent the information content.

ex. Decimal digits 10 = {0,1,2,3,4,5,6,7,8,9}
4-bit BCD (binary code decimal)

$$\log_2(10) = 3.322 < 4\,bits$$

ex. ~84 English characters = {A-Z (26), a-z (26), 0-9 (11), punctuation (8), math (9), financial (4)}
7-bit ASCII (American Standard Code for Information Interchange)

$$\log_2(84) = 6.392 < 7\,bits$$

Kodowanie liczb

Encoding numbers

It is straightforward to encode positive integers as a sequence of bits. Each bit is assigned a weight. Ordered from right to left, these weights are increasing powers of 2. The value of an n-bit number encoded in this fashion is given by the following formula:

$$v = \sum_{i=0}^{n-1} 2^i b_i$$

2^{11}	2^{10}	2^9	2^8	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
0	1	1	1	1	1	0	1	0	0	0	0

$= 2000_{10}$

Oftentimes we will find it convenient to cluster groups of bits together for a more compact representation. Two popular groupings are clusters of 3 bits and 4 bits.

03720
Octal - base 8

000 - 0	
001 - 1	
010 - 2	
011 - 3	
100 - 4	
101 - 5	
110 - 6	
111 - 7	

0x7d0
Hexadecimal - base 16

0000 - 0	1000 - 8
0001 - 1	1001 - 9
0010 - 2	1010 - a
0011 - 3	1011 - b
0100 - 4	1100 - c
0101 - 5	1101 - d
0110 - 6	1110 - e
0111 - 7	1111 - f

Kodowanie liczb całkowitych

Signed integers: 2's complement

N bits

-2^{N-1}	2^{N-2}	•••	•••	•••	2^3	2^2	2^1	2^0

"sign bit" Range: -2^{N-1} to $2^{N-1}-1$ "decimal" point

8-bit 2's complement example:
$11010110 = -2^7 + 2^6 + 2^4 + 2^2 + 2^1 = -128 + 64 + 16 + 4 + 2 = -42$

If we use a two's complement representation for signed integers, the same binary addition mod 2^n procedure will work for adding positive and negative numbers (don't need separate subtraction rules). The same procedure will also handle unsigned numbers!

By moving the implicit location of "decimal" point, we can represent fractions too:
$1101.0110 = -2^3 + 2^2 + 2^0 + 2^{-2} + 2^{-3} = -8 + 4 + 1 + 0.25 + 0.125 = -2.625$

Kodowanie

When choices aren't equally probable

When the choices have different probabilities (p_i), you get more information when learning of a unlikely choice than when learning of a likely choice

Information from $choice_i = \log_2(1/p_i)$ bits
Average information from a choice $= \Sigma p_i \log_2(1/p_i)$

Example

$choice_i$	p_i	$\log_2(1/p_i)$
"A"	1/2	1 bit
"B "	1/6	2.58 bits
"C"	1/6	2.58 bits
"D"	1/6	2.58 bits

Average information
= (.5)(1) + (3)(.167)(2.58)
= 1.54 bits

Can we find an encoding where transmitting 1000 choices approaches 1540 bits on the average? Using two bits for each choice = 2000 bits

Kodowanie Huffmana

Variable-length encodings

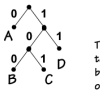
(David Huffman, 1950)

Use shorter bit sequences for high probability choices, longer sequences for less probable choices

$choice_i$	p_i	encoding
"A"	1/2	0
"B "	1/6	100
"C"	1/6	101
"D"	1/6	11

A B D A D C

010011011101



Huffman Decoding Tree

Average information
= (.5)(1) + (2)(.167)(3) + (.167)(2)
= 1.836 bits

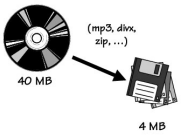
Transmitting 1000 choices takes an average of 1836 bits… better but not optimal

To get a more efficient encoding (closer to information content) we need to encode sequences of choices, not just each choice individually. This is the approach taken by most file compression algorithms…

Kompresja danych

Data Compression

Key: *re-encoding to remove redundant information: match data rate to actual information content.*



"Outside of a dog, a book is man's best friend. Inside of a dog, its too dark to read…"

-Groucho Marx

AB4bl*m9@+M(p

Ideal: No redundant info – Only unpredictable bits transmitted. Result appears *random!*

Kompresja danych

Does *recompression* work?

If ZIP compression of a 40MB Bible yields a 4MB ZIP file, what happens if we compress *that?*

- Do we get a 400 KB file?

ZIP

- Can we compress that, to get a 40K file??

ZIP

- Can we compress the whole Bible down to a single bit???

ZIP

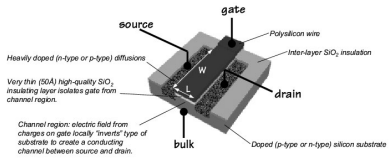
- Is it a 0 or a 1????

"?"

HINT: if the initial compression works perfectly, the result has *no* redundancy!

Technologia MOSFET

MOSFETS: Gain & non-linearity

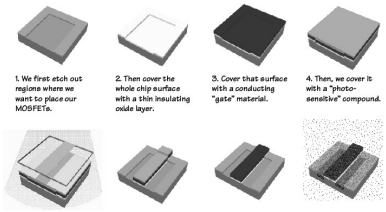


MOSFETs (metal-oxide-semiconductor field-effect transistors) are four-terminal voltage-controlled switches. Current flows between the diffusion terminals if the voltage on the gate terminal is large enough to create a conducting "channel", otherwise the mosfet is off and the diffusion terminals are not connected.

Why are MOS devices King?

Technologia produkcji układów

#1: Self-Aligning Gates



1. We first etch out regions where we want to place our MOSFETs.
2. Then cover the whole chip surface with a thin insulating oxide layer.
3. Cover that surface with a conducting "gate" material.
4. Then, we cover it with a "photo-sensitive" compound.
5. We expose the chip to an optical "mask".
6. We etch away the exposed materials.
7. Remove the photoresist.
8. Bombard the substrate with ions. The gate protects the region beneath it from implantation.

Technologia cd.

Let's build a MOSFET

Start with a 500µ slice of a silicon ingot that has been *doped* with an acceptor (typically boron) to increase the concentration of holes to $10^{14}/cm^3$ - $10^{18}/cm^3$. At room temperature, all the dopants in this *p-type* material are ionized, turning the silicon into a *semiconductor*.

We'll build many copies of the same circuit onto a single wafer. Only a certain percentage of the chips will work; those that work will run at different speeds. The yield decreases as the size of the chips increases and the feature size decreases.

Wafers are processed by automated fabrication lines. To minimize the chance of contaminants ruining a process step, great care is taken to maintain a meticulously clean environment. So put on your bunny suits and let's begin…

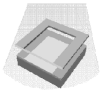
Technologia – wzory

Creating patterns on the wafer

Images from: http://www.intel.com/education/chips/index.htm

A "thick" (0.4u) layer of SiO_2 is formed by oxidizing the surface of the wafer with wet oxygen (we rust it!). The SiO_2 will erve as insulation between the conductive substrate and subsequent conductive layers we'll build on top of the oxide.

Now we'll form a pattern in the SiO_2 using a mask & etch process. First the wafer is coated with a layer of photoresist. Photoresist becomes soluable when exposed to ultraviolet light…



Using a mask to protect parts of the wafer, we'll expose those portions of the wafer where we want to remove the photoresist. We'll use different masks when creating each of the different structures on the wafer.

Technologia - trawienie

The etching process

The exposed photoresist is removed with a solvent. The unexposed photoresist remains, masking portions of the underlying SiO_2 layer.

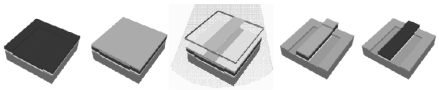
A chemical etch is then used to remove the revealed silicon dioxide.

Finally, the remaining photoresist is removed with a different solvent and we're left with pattern of insulating SiO_2 on top of exposed p-type substrate.

Technologia - utlenianie

Gate oxide & polysilicon

Now a "thin" (100 Å) layer of SiO_2, called gate oxide, is grown on the surface. The gate oxide needs to be of high quality: uniform thickness, no defects! The thinner the oxide, the more oomph the FET will have (we'll see why soon) but the harder it is to make it defect-free. Coming soon to a fab near you: 50 Å gate oxide…



On top of the thin oxide a 0.7u thick layer of polycrystalline silicon, called *polysilicon* or *poly* for short, is deposited by CVD. The poly layer is patterned and plasma etched (thin ox not covered by poly is etched away too!) exposing the surface where the source and drain junctions will be formed. Poly has a high sheet resistance of 20 Ω/sq which can be reduced by adding a layer of a silicided refractory metal such titanium ($TiSi_2$), tantalum ($TaSi_2$) or molybdenum ($MoSi_2$) => 1, 3 or 5 Ω/sq.

Technologia - dyfuzja

Source/drain diffusions

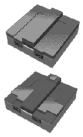
Donor implants are used to create self-aligned MOSFET source/drain diffusions and substrate contacts. Usually As is preferred to obtain shallow N-type diffusions and minimal lateral diffusion. High doses are needed to make low resistance (25 Ω/sq) diffusion wires. Afterwards a short thermal annealing step is performed to repair surface damage caused by the implantation.

This completes the construction of the MOSFET itself. Now we'll add the metal wiring layers…

Technologia – połączenia metaliczne

Metal interconnect

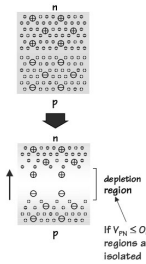
Each level of metal interconnect proceeds in two major steps. The first, shown in the sequence above, deposits an insulating layer of SiO_2 and forms vias (holes) in the oxide to the layer(s) below. In this example, vias are made to the source/drain diffusions and the polysilicon gate.



In the second major step, shown to the left, the aluminum is deposited, patterned, then etched to form low-resistance (.07 Ω/sq) interconnect. With planarization (a mechanical polishing step that creates a flat surface), multiple levels of metal interconnect are possible -- 3 to 6 layers are common in today's processes.

Technologia - połączenia

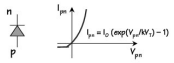
#2: Junctions as Insulation



Regions "doped" with P-type and N-type materials each conduct. At a PN boundary, the mobile carriers move away from the boundary, forming a depletion region with a voltage across it due to induced field.

At room temp, with doping concentrations of $10^{15}/cm^3$, this voltage is 0.6v. No current flows unless sufficient potential is applied to overcome this bias.

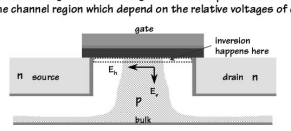
The net result is a *diode*:



Technologia – przełącznik FET

#3: FETs as switches

The four terminals of a Field Effect Transistor (gate, source, drain and bulk) connect to conducting surfaces that generate a complicated set of electric fields in the channel region which depend on the relative voltages of each terminal.

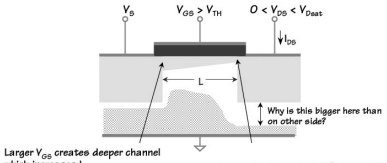


INVERSION:
A sufficiently strong ($> V_{TH}$) positive voltage applied to the will attract enough electrons to the surface to create a conducting n-type channel between the source and drain.

CONDUCTION:
If a channel exists, a horizontal field will cause a drift current from the drain to the source.

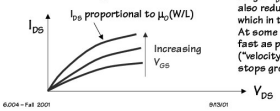
Technologia - obszary

"Linear" operating region



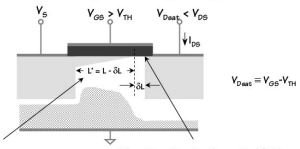
Larger V_{GS} creates deeper channel which increases I_{DS}

Larger V_{DS} increases drift current but also reduces vertical field component which in turn makes channel less deep. At some point, electrons are traveling as fast as possible through the channel ("velocity saturation") and the current stops growing linearly.



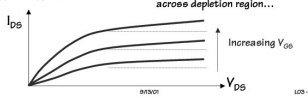
Technologia - obszary

Saturated operating region



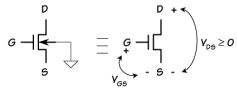
This looks just like a fet with a channel length of $L' < L$. Shorter L' implies greater I_{DS}. As V_{DS} increases, δL gets larger.

When $V_{DS} = V_{GS}-V_{TH}$, the vertical field component is reduced and the channel is pinched-off. Electrons just keep traveling across depletion region…



Technologia - tranzystor

NFET Summary

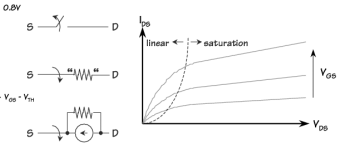


Operating regions:

cut-off: $V_{GS} < V_{TH}$ (0.8V)

linear: $V_{GS} \geq V_{TH}$, $V_{DS} < V_{Dsat}$ ($V_{GS} - V_{TH}$)

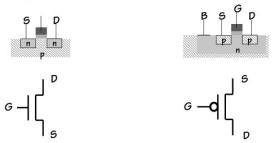
saturation: $V_{GS} \geq V_{TH}$, $V_{DS} \geq V_{Dsat}$



Technologia – tranzystor

#4: FETs come in two flavors

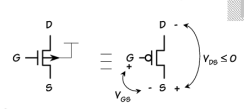
By embedding p-type source and drain in a n-type substrate, we can fabricate a complement to the N-FET:



The use of both NFETs and PFETs – complimentary transistor types – is a key to CMOS (complementary MOS) logic families.

Technologia

PFET Summary

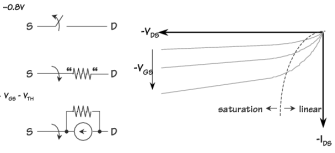


Operating regions:

cut-off: $V_{GS} > V_{TH}$ (-0.8V)

linear: $V_{GS} \leq V_{TH}$, $V_{DS} > V_{Dsat}$ ($V_{GS} - V_{TH}$)

saturation: $V_{GS} \leq V_{TH}$, $V_{DS} \leq V_{Dsat}$



Technologia

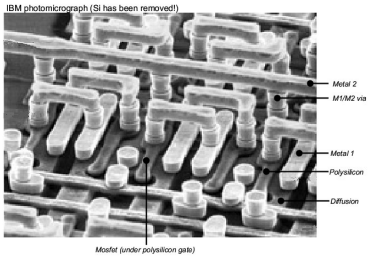
Wires: metal interconnect

Deposit SiO_2 insulation | Etch openings for vias/contacts | Deposit Al/Cu conductor | Etch away Al/Cu leaving wires

After a layer of SiO_2 insulation has been deposited, aluminum or copper is deposited, patterned, then etched to form low-resistance (.07 Ω/sq) interconnect. With planarization of the SiO_2 (a mechanical polishing step that creates a flat surface), multiple levels of metal interconnect are possible -- 3 to 6 layers are common in today's processes.

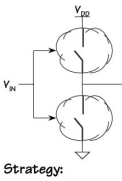
Technologia – połączenia między warstwami

Multiple interconnect layers



Technologia – bramka logiczna

Finally… logic gates!



Inverter recipe:

pullup: make this connection when V_{IN} near 0 so that $V_{OUT} = V_{DD}$

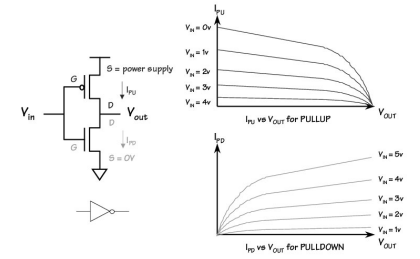
pulldown: make this connection when V_{IN} near V_{DD} so that $V_{OUT} = 0$

General Strategy:

- one power supply => low impedance source for 2 levels
- At equilibrium, with valid inputs,
  - Either pulldown or pullup is cut off
  - output connected to either V_{DD} or ground
  - NO current flow or power consumption
- Power dissipation only!

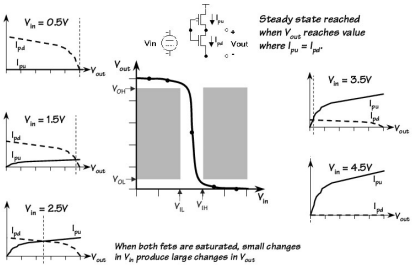
Technologia

CMOS Inverter



Technologia

CMOS Inverter VTC

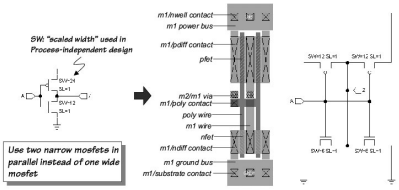


Steady state reached when V_{out} reaches value where $I_{pu} = I_{pd}$.

When both fets are saturated, small changes in V_{in} produce large changes in V_{out}

Technologia

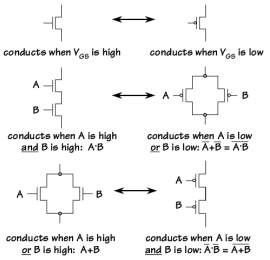
Standard Cell Layout for Inverter



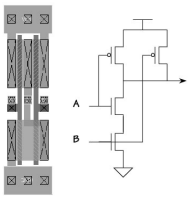
Use two narrow mosfets in parallel instead of one wide mosfet

Physical design of a CMOS gate is represented by a mask layout showing where material on each layer (ndiff, pdiff, poly, m1, m2, …) should be placed on the silicon wafer. Each manufacturing process has a set of *design rules* that determine minimum widths, spacings, overlaps, etc.

CMOS complements



A pop quiz!

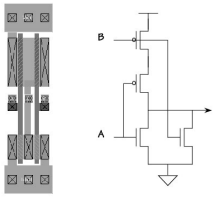


What function does this gate compute?

A	B	C
0	0	1
0	1	1
1	0	1
1	1	0

NAND: $C = \overline{AB}$

Here's another…



What function does this gate compute?

A	B	C
0	0	1
0	1	0
1	0	0
1	1	0

NOR: $C = \overline{A+B}$

General CMOS gate recipe

Step 1. Figure out NFET pulldown network that does what you want, e.g.,

$F = \overline{A*(B+C)}$

Step 2. Walk the hierarchy replacing nfets with pfets, series subnets with parallel subnets, and parallel subnets with series subnets

Step 3. Combine pfet pullup network from Step 2 with nfet pulldown network from Step 1 to form fully-complementary CMOS gate.