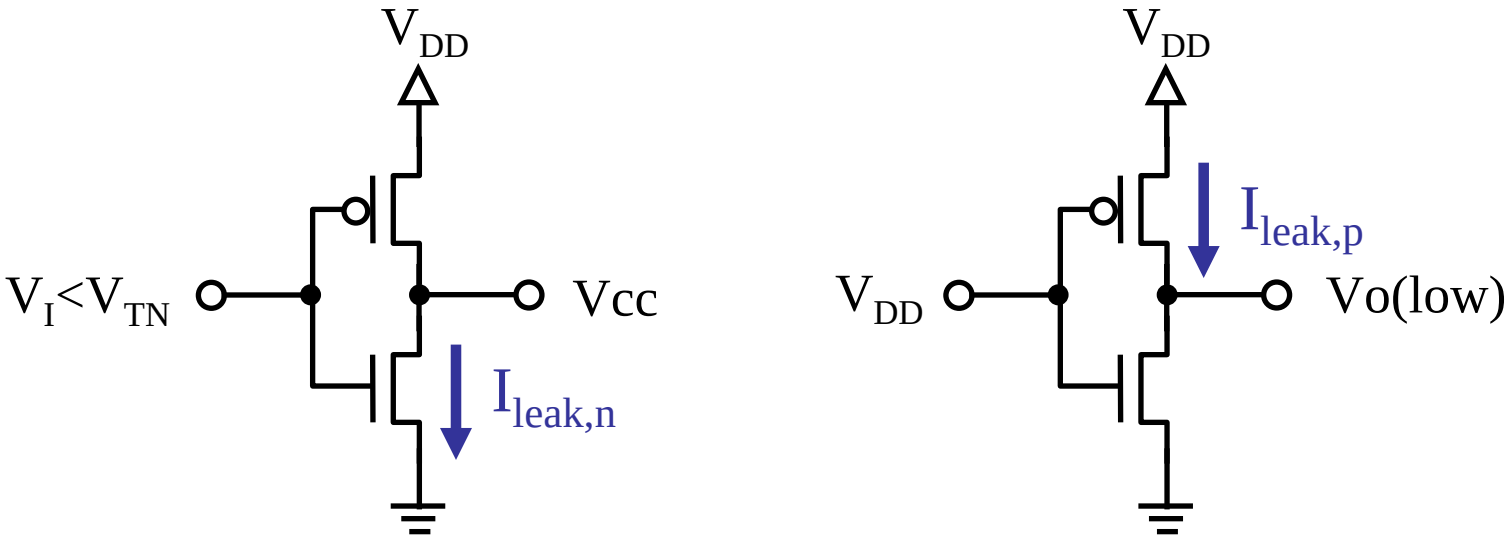
CMOS Inverter Power Dissipation

CMOS inverter power

- Power has three components
  - Static power: when input isn't switching
  - Dynamic capacitive power: due to charging and discharging of load capacitance
  - Dynamic short-circuit power: direct current from V_{DD} to G_{nd} when both transistors are on

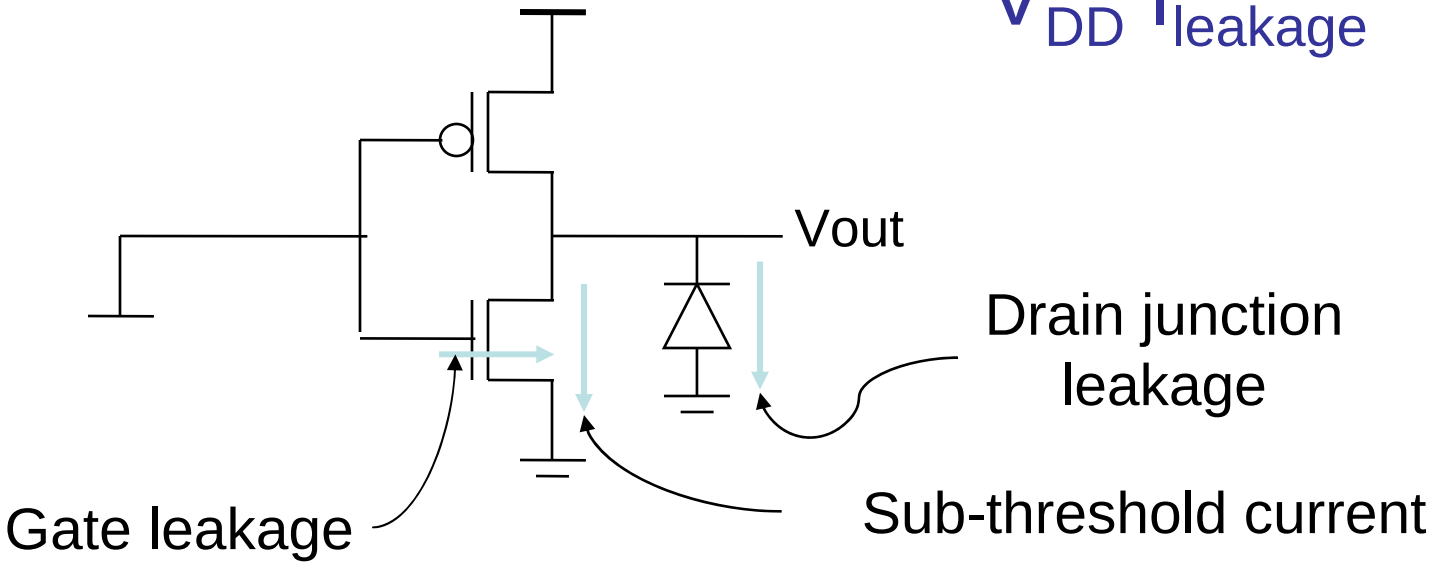
CMOS inverter static power

- Static power consumption:
  - Static current: in CMOS there is no static current as long as $V_{in} < V_{TN}$ or $V_{in} > V_{DD}+V_{TP}$
  - Leakage current: determined by “off” transistor
  - Influenced by transistor width, supply voltage, transistor threshold voltages



Leakage (Static) Power Consumption

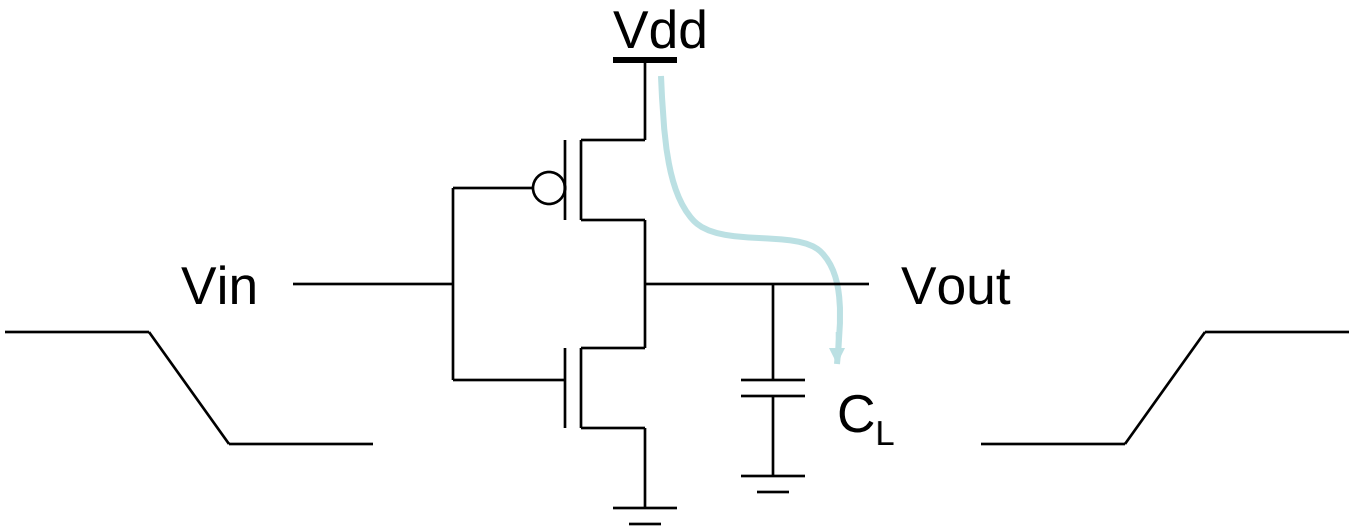
V_{DD} $I_{leakage}$



Sub-threshold current is the dominant factor.

All increase exponentially with temperature!

Dynamic Power Consumption



Dynamic Capacitive Power and energy stored in the PMOS device

Case I: When the input is at logic 0: Under this condition the PMOS is conducting and NMOS is in cutoff mode and the load capacitor must be charged through the PMOS device.

Power dissipation in the PMOS transistor is given by,

$P_P = i_L V_{SD} = i_L(V_{DD} - V_O)$

The current and output voltages are related by,

$i_L = C_L dv_O/dt$

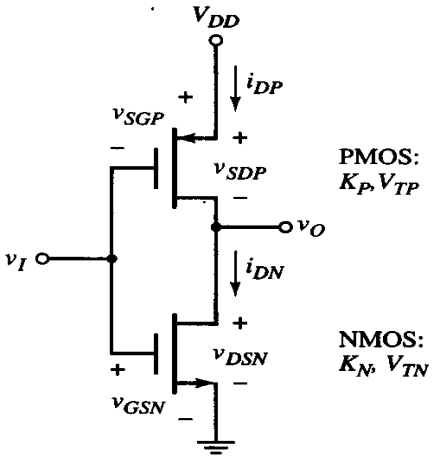
Similarly the energy dissipation in the PMOS device can be written as the output switches from low to high ,

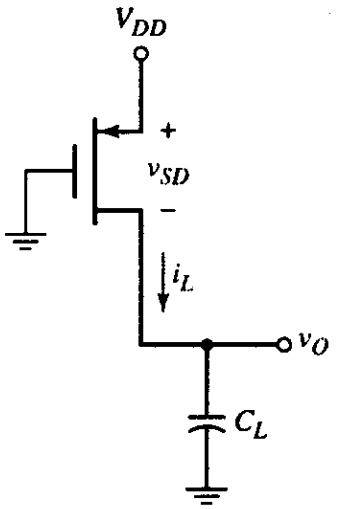
$$E_P = \int_0^\infty P_P dt = \int_0^\infty C_L (V_{DD} - v_O)\frac{dv_O}{dt}\, dt, \; E_P = C_L V_{DD}\int_0^{V_{DD}} dv_O - C_L \int_0^{V_{DD}} v_O dv_O$$

$$E_P = C_L V_{DD} v_O\Big|_0^{V_{DD}} - C_L \frac{v_O^2}{2}\Bigg|_0^{V_{DD}}, \; E_P = (C_L V_{DD} V_{DD} - 0) - (C_L \frac{V_{DD}^2}{2} - 0)$$

$$E_P = \frac{1}{2} C_L V_{DD}^2$$

Above equation showed the energy stored in the capacitor C_L when the output is high.



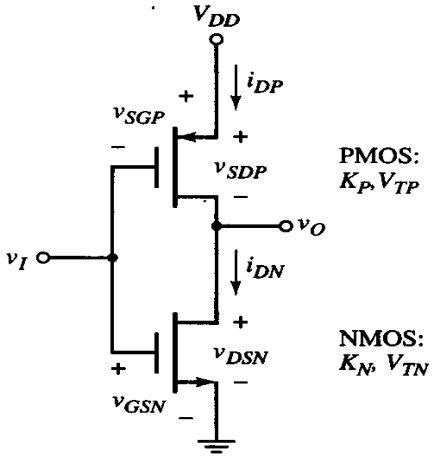


Power Dissipation and Total Energy Stored in the CMOS Device

Case II: when the input is high and out put is low:

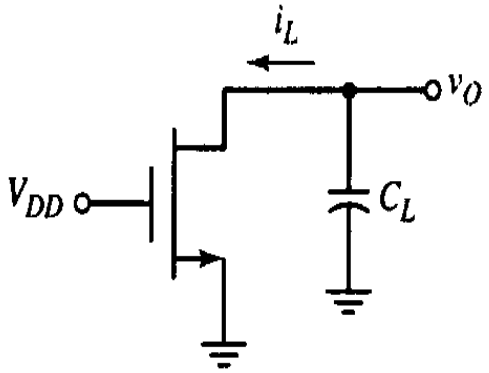
During switching all the energy stored in the load capacitor is dissipated in the NMOS device because NMOS is conducting and PMOS is in cutoff mode. The energy dissipated in the NMOS inverter can be written as,

$$E_N = \frac{1}{2} C_L V_{DD}^2$$



The total energy dissipated during one switching

$$E_T = E_P + E_N = \frac{1}{2} C_L V_{DD}^2 + \frac{1}{2} C_L V_{DD}^2 = C_L V_{DD}^2$$



The power dissipated in terms of frequency can be

$$E_T = P\,t \Rightarrow \quad P = \frac{E_T}{t} \Rightarrow \quad P = fE_T \Rightarrow \quad fC_L V_{DD}^2$$

This implied that the power dissipation in the CMOS inverter is directly proportional to switching frequency and V_{DD}^2

Dynamic capacitive power

- Formula for dynamic power:

$$P_{dyn} = C_L V_{DD}^2 f$$

- Observations
  - Does not (directly) depend on device sizes
  - Does not depend on switching delay
  - Applies to general CMOS gate in which:
    - Switched capacitances are lumped into C_L
    - Output swings from Gnd to V_{DD}
    - Input signal approximated as step function
    - Gate switches with frequency f

Not a function of transistor sizes!

Data dependent - a function of switching activity!

Lowering Dynamic Power

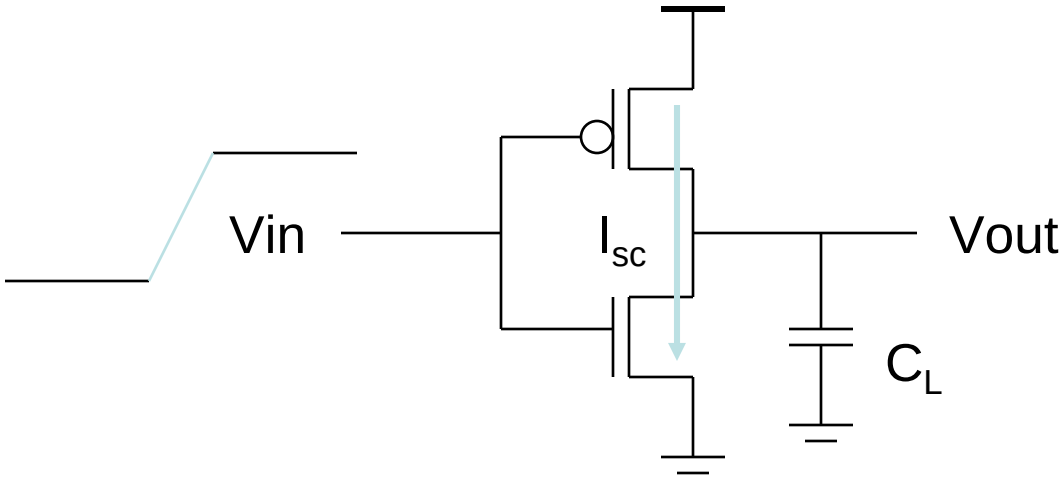
Capacitance:
Function of fan-out, wire length, transistor sizes

Supply Voltage:
Has been dropping with successive generations

$$P_{dyn} = C_L \, V_{DD}^2 \, f$$

Clock frequency:
Increasing...

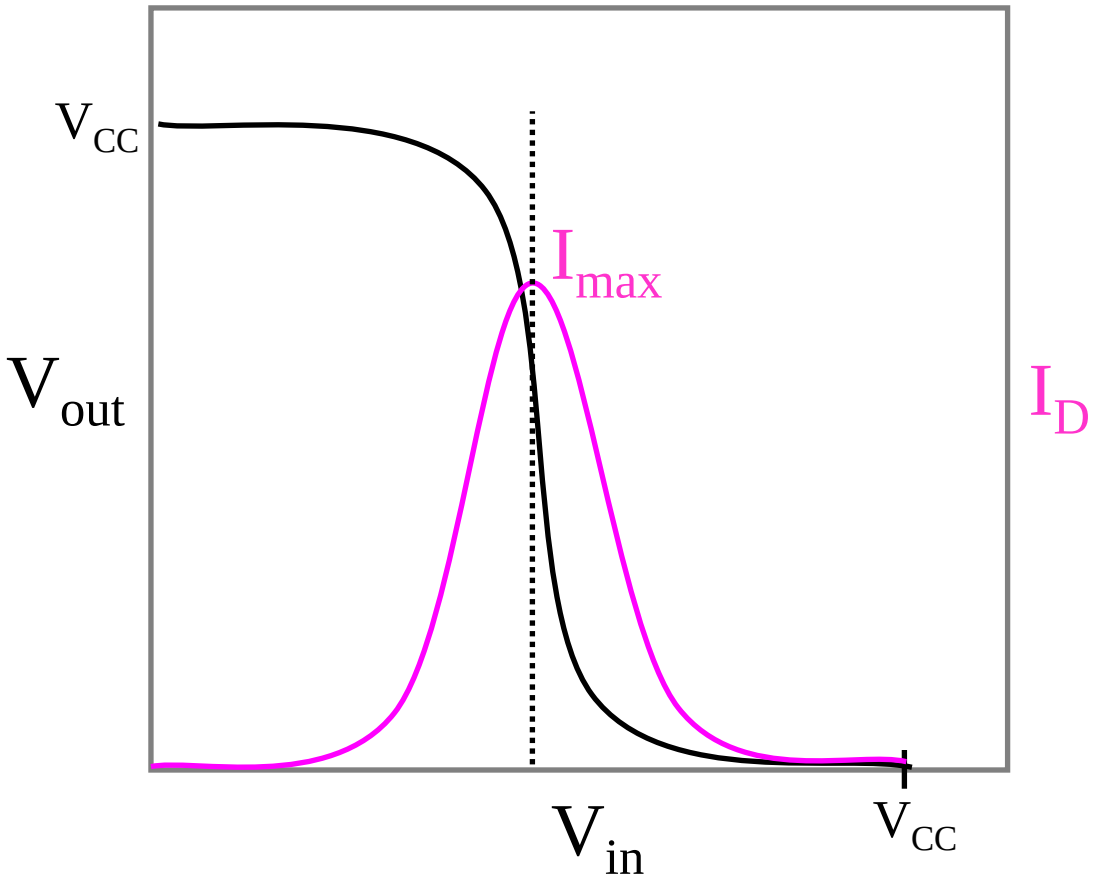
Short Circuit Power Consumption



Finite slope of the input signal causes a direct current path between V_{DD} and GND for a short period of time during switching when both the NMOS and PMOS transistors are conducting.

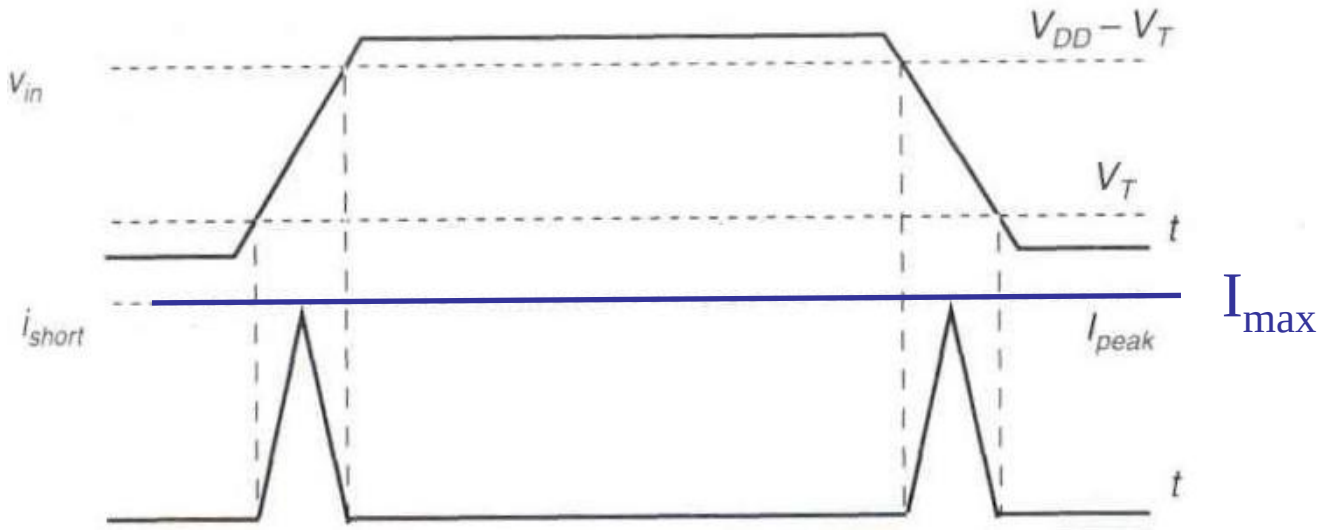
Dynamic short-circuit power

- Short-circuit current flows from V_{DD} to Gnd when both transistors are on
- Plot on VTC curve:



I_{max}: depends on saturation current of devices

Dynamic short-circuit power



- Approximate short-circuit current as a triangular wave
- Energy per cycle:

$$E_{sc} = V_{CC}\frac{I_{\max} t_r}{2} + V_{CC}\frac{I_{\max} t_f}{2} = \frac{t_r + t_f}{2} V_{CC} I_{\max}$$

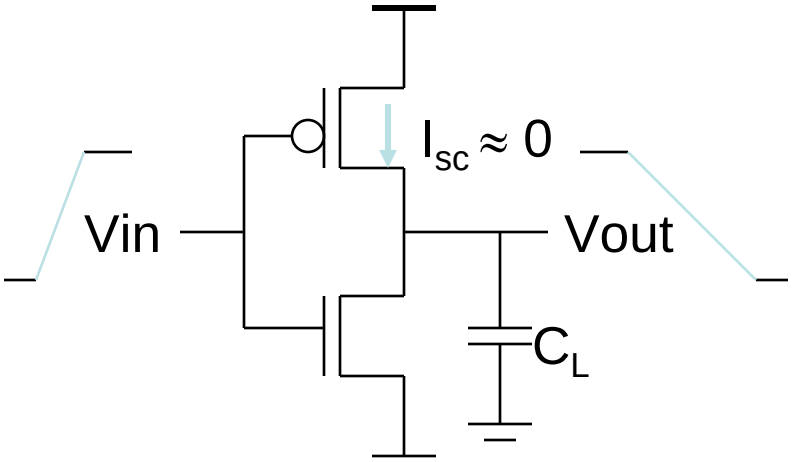
$$P_{sc} = \frac{t_r + t_f}{2} V_{CC} I_{\max} f$$

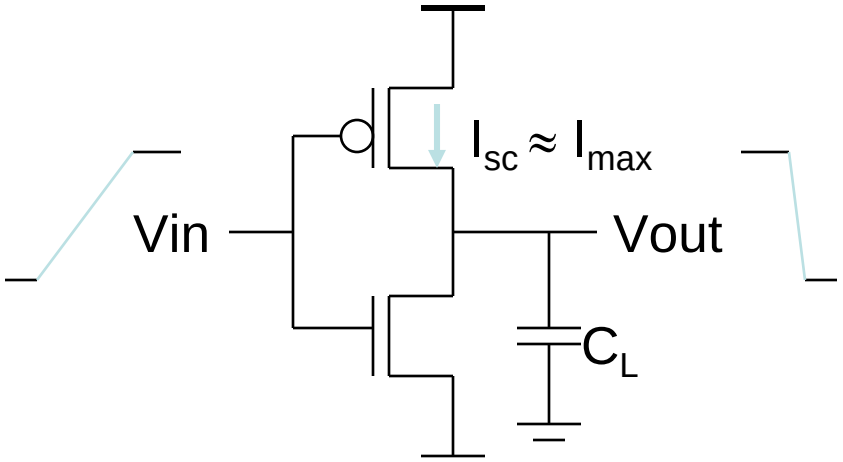
Short Circuit Currents Determinates

$$P_{sc} = t_{sc} V_{DD} I_{peak} f_{0\rightarrow1}$$

- Duration and slope of the input signal, t_{sc}
- I_{peak} determined by
  - the saturation current of the P and N transistors which depend on their sizes, process technology, temperature, etc.
  - strong function of the ratio between input and output slopes
    - a function of C_L

Impact of C_L on P_{sc}





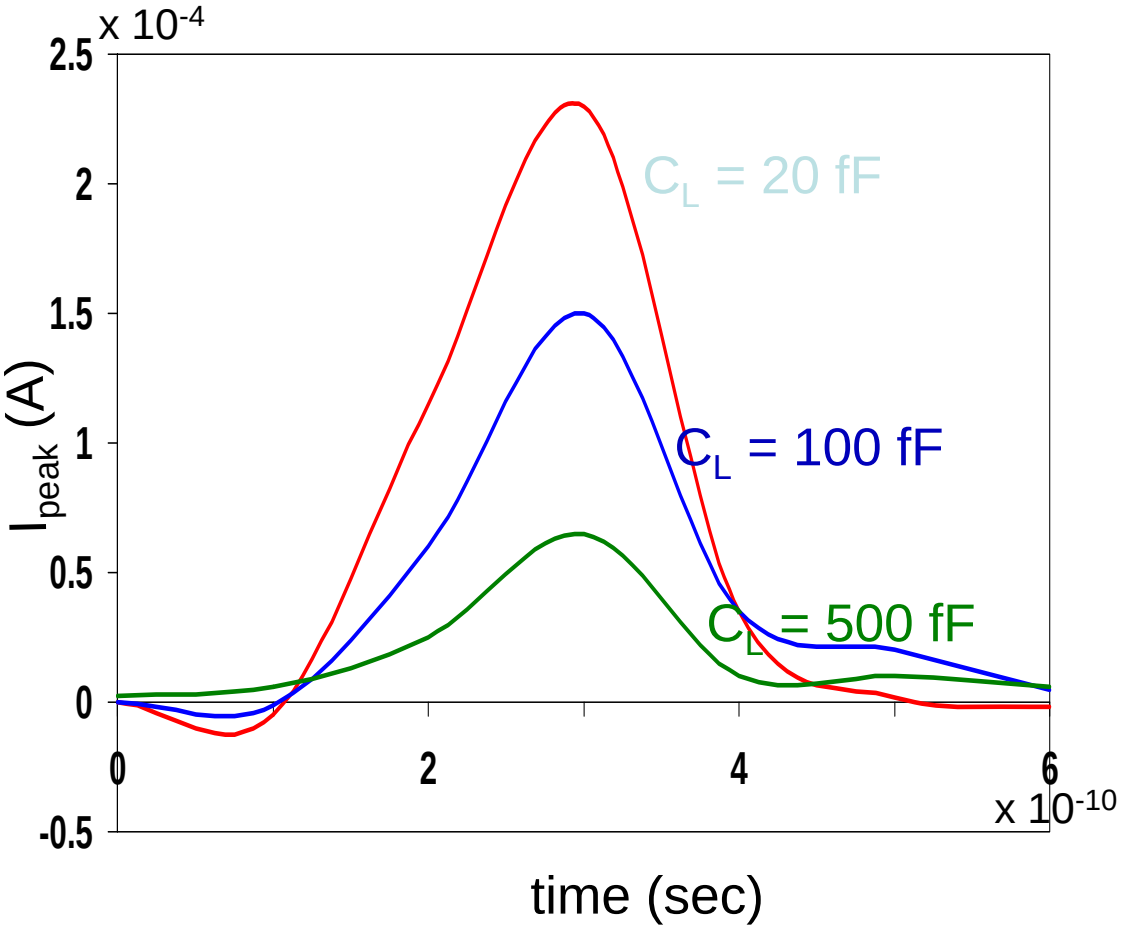
Large capacitive load

Small capacitive load

Output fall time significantly larger than input rise time.

Output fall time substantially smaller than the input rise time.

I_{peak} as a Function of C_L



When load capacitance is small, I_{peak} is large.

Short circuit dissipation is minimized by matching the rise/fall times of the input and output signals - slope engineering.

500 psec input slope

Inverter power consumption

- Total power consumption

$$P_{tot} = P_{dyn} + P_{sc} + P_{stat}$$

$$P_{tot} = C_L V_{CC}^2 f + V_{CC} I_{\max} \left(\frac{t_r + t_f}{2} \right) f + V_{CC} I_{leak}$$

Power reduction

- Reducing dynamic capacitive power:
  - Lower the voltage (Vdd)!
    - Quadratic effect on dynamic power
  - Reduce capacitance
    - Short interconnect lengths
    - Drive small gate load (small gates, small fan-out)
  - Reduce frequency
    - Lower clock frequency -
    - Lower signal activity

$$P_{dyn} = C_L V_{DD}^2 f$$

Examples

f=500MHz

C_L=15fF/gate

V_{DD}=2.5V

P_{dyn}=50μW

For a design with I million gate

P_{dyn}=50W!

Is this possible in reality? If not why?

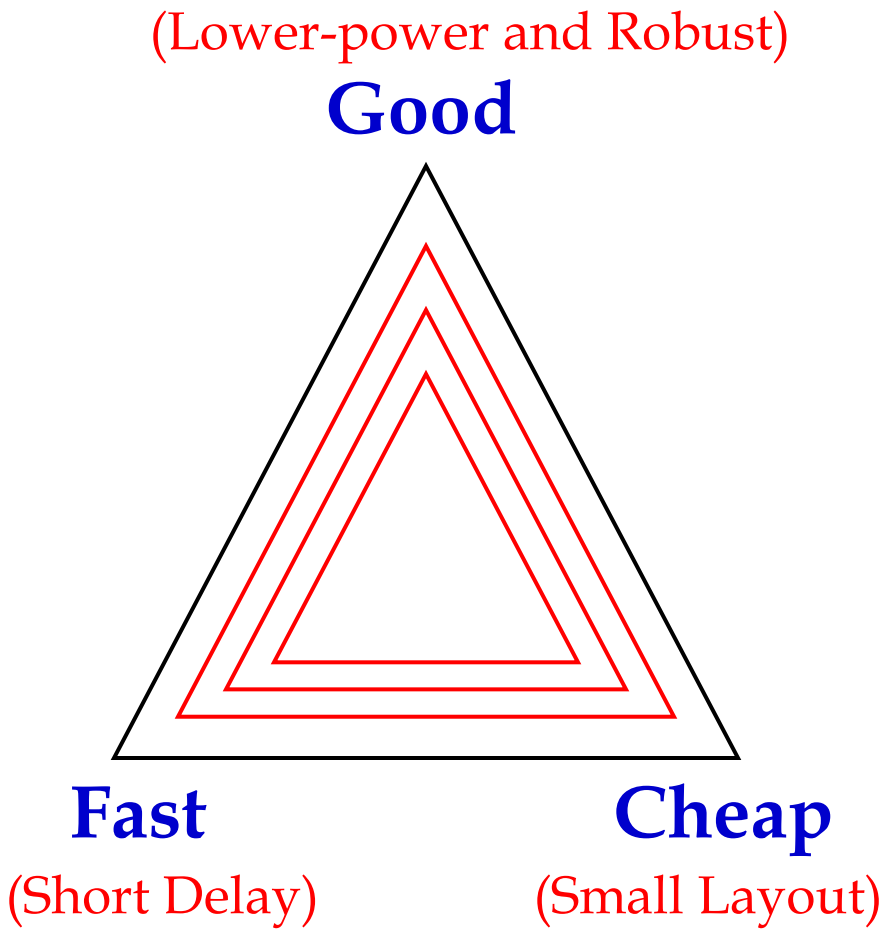
$P_{dyn}=E_{dyn}/2t_p$=580μW for t_p=32.5ps

$P_{dyn}=E_{dyn}/2t_p$=155μW for f=4GHz(250ps)

Power reduction

- Reducing short-circuit current:
  - Fast rise/fall times on input signal
  - Reduce input capacitance
  - Insert small buffers to “clean up” slow input signals before sending to large gate
- Reducing leakage current:
  - Small transistors (leakage proportional to width)
  - Lower voltage

Retrospect on Design Trade-offs



- **Design trade-offs dance around the triangle, but still important**
- **Fundamental improvement that shrinks the triangle:**
  - **Scaling in technology (lithography improvement)**
  - **New functionality**
  - **New architecture**
  - **New algorithms**