In this lecture you will learn:

- Digital Logic
- The CMOS Inverter
- Charge and Discharge Dynamics
- Power Dissipation
- Digital Levels and Noise

### A NFET Inverter

![NFET Inverter Circuit Diagram]

<table>
<thead>
<tr>
<th>A</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Digital Signal Levels

Valid logic levels:

- $V_{IL} =$ Maximum valid logical LOW input
- $V_{IH} =$ Minimum valid logical HIGH input
- $V_{OL} =$ Maximum valid logical LOW output
- $V_{OH} =$ Minimum valid logical HIGH output

$V_{IL} , V_{IH} \text{ and } V_{OL} , V_{OH}$ are determined by the unity gain points on the transfer curve (Otherwise amplification can corrupt the logic levels as they propagate in a chain)

Gain is necessary to realize logic gates …!!!

Digital Signal Levels and Noise

Noise is reduced at the output when the input is within the valid range
Noise can be amplified at the output when the input is outside the valid range
Noise Margins

And when noise is present……

\[ V_{OL} + \text{noise} < V_{IL} \]
\[ V_{OH} - \text{noise} > V_{IH} \]

One must have:

Noise margins:

\[ NM_L = V_{IL} - V_{OL} \]
\[ NM_H = V_{OH} - V_{IH} \]

Noise and device variations sets the minimum \( V_{DD} \) one can use

The Ideal Inverter Transfer Curve

A perfectly symmetric curve with a near-vertical transition is an ideal transfer curve because:

- Noise margins can be made very large
- Logical HIGH voltage can be made very small (because the noise margins are so large) resulting smaller power dissipation
The input/output characteristics are not symmetric.

Noise margins are good but not excellent.

The Load Capacitance

$C_L$ is the capacitance of the subsequent CMOS stage(s) as well as of the interconnects.
A NFET Inverter: Charging Dynamics

Problem:

When the output is LOW, charging of the output to HIGH is slow because charging current is not uniform.

\[
I_{OUT} = \frac{V_{DD} - V_{OUT}}{R} = C_L \frac{dV_{OUT}}{dt}
\]

\[
\Rightarrow V_{OUT}(t) = V_{DD} \left[1 - e^{-\frac{t}{RC_L}}\right]
\]
**Problem:**

When the input is HIGH, and the output is LOW, current keeps flowing through the FET and the resistor forever!!

This is an example of static power dissipation – extremely bad!

---

**A NFET Inverter: Static Power Dissipation**

---

**A CMOS Inverter: Noise Margins**

If $V_{TN}$ and $V_{DD} + V_{TP}$ are close to each other, the transition region can be made narrow and sharp

→ The noise margins can be very wide!!
A CMOS Inverter: Charging and Discharging Dynamics

When the input goes HIGH or LOW, power is only dissipated during the time when the output makes the transition – after this period, there is no power dissipation.

There is no static power dissipation (ideally!), only dynamic power dissipation!!

A CMOS Inverter: No Static Power Dissipation

When the input goes HIGH or LOW, power is only dissipated during the time when the output makes the transition – after this period, there is no power dissipation.

There is no static power dissipation (ideally!), only dynamic power dissipation!!
**Rise Times, Fall Times and Propagation Delays**

- \( t_r \): Rise time b/w 10% to 90% of the total upward swing
- \( t_f \): Fall time b/w 90% to 10% of the total downward swing
- \( t_{PHL} \): Propagation delay for \( H \rightarrow L \) b/w 50% points
- \( t_{PLH} \): Propagation delay for \( L \rightarrow H \) b/w 50% points

**A CMOS Inverter: Charging Dynamics**

When the output is LOW, initial charging of the output to HIGH is done with a uniform current supplied by the PFET in saturation:

\[
I_{OUT} = C_L \frac{dV_{OUT}}{dt}
\]

\[
\Rightarrow \frac{k_p}{2} (V_{IN} - V_{DD} - V_{TP})^2 = C_L \frac{dV_{OUT}}{dt}
\]

\[
\Rightarrow \frac{k_p}{2} (V_{DD} + V_{TP})^2 = C_L \frac{dV_{OUT}}{dt}
\]

\[
\Rightarrow V_{OUT}(t) = \frac{k_p}{2C_L} (V_{DD} + V_{TP})^2 t
\]

Condition for the PFET to be in saturation:

\[
V_{DS} < V_{GS} - V_{TP}
\]

\[
\Rightarrow V_{OUT} - V_{DD} < V_{IN} - V_{DD} - V_{TP}
\]

\[
\Rightarrow V_{OUT} < V_{IN} - V_{TP} \approx -V_{TP}
\]

When \( V_{OUT} \) becomes larger than \(-V_{TP}\), then the PFET goes into the linear region...
A CMOS Inverter: Charging Dynamics

For times $0 < t < t_1$ when the PFET (M2) is in saturation:

$$V_{OUT}(t) = \frac{k_p}{2C_L} (V_{DD} + V_{TP})^2 t$$

$$\Rightarrow t_1 = -\frac{V_{TP}}{k_p (V_{DD} + V_{TP})} 2C_L = \frac{-2V_{TP}}{(V_{DD} + V_{TP})} \tau$$

$$\tau = \frac{C_L}{k_p (V_{DD} + V_{TP})}$$

A CMOS Inverter: Charging Dynamics

When $V_{OUT}$ becomes larger than $-V_{TP}$ then the PFET goes into the linear region, and from then onwards:

$$I_{OUT} = C_L \frac{dV_{OUT}}{dt}$$

$$\Rightarrow k_p (V_{IN} - V_{DD} - V_{TP} - \frac{(V_{OUT} - V_{DD})}{2})(V_{OUT} - V_{DD}) = C_L \frac{dV_{OUT}}{dt}$$

$$\Rightarrow -k_p (V_{DD} + V_{TP} + \frac{(V_{OUT} - V_{DD})}{2})(V_{OUT} - V_{DD}) = C_L \frac{dV_{OUT}}{dt}$$

One can obtain faster charging compared to a resistor in place of a PFET!
A CMOS Inverter: Charging Dynamics

One can obtain faster charging compared to a resistor:

\[
-t_r \approx \frac{2C_L}{k_p(V_{DD}+V_{TP})} \left\{ \frac{-V_{TP}}{V_{DD}+V_{TP}} - \frac{1}{2} \ln \left[ \frac{0.1V_{DD}}{2(V_{DD}+V_{TP}) - 0.1V_{DD}} \right] \right\}
\]

\[
\alpha \approx \frac{L^2}{\mu_p V_{DD}} \ll t_r \quad \text{if } C_L \ll C_{gs}
\]

In reality, the load capacitance is not just due to the next FET gate - it also includes the interconnect capacitances.
**A CMOS Inverter: Discharging Dynamics**

One can obtain faster discharging compared to a resistor:

\[
T_{\text{DD}} \sim \frac{2C_L}{k_n(V_{DD} - V_{TN})} \left( \frac{V_{TN}}{V_{DD} - V_{TN}} - \frac{1}{2} \ln \left[ \frac{0.1V_{DD}}{2(V_{DD} - V_{TN}) - 0.1V_{DD}} \right] \right)
\]

\[
\sim \frac{L^2}{\mu_n V_{DD}} \quad \text{FET transit time}
\]

Assume the input changes abruptly.

In reality, the load capacitance is not just due to the next FET gate - it also includes the interconnect capacitances.

**A CMOS Inverter: Charging and Discharging Dynamics**

How to charge and discharge faster?

- Decrease \( L \)
- Increase charging current
  - Increase supply voltage \( V_{DD} \)

At what price?

Assuming \( C_L \) is not dominated by interconnect capacitance (not generally true), the only way to increase \( k_n \) and \( k_p \) of FETs and at the same time decrease \( C_{gs} \) is to decrease the FET length \( L \):

\[
t_f \approx \frac{L^2}{\mu V_{DD}}
\]

\[
t_f \approx \frac{L^2}{\mu V_{DD}}
\]
ECE 315 – Spring 2005 – Farhan Rana – Cornell University

Intel FET Gate length Trends

Smaller transistor provides:
- Higher performance
- Lower power
- Lower cost per FET

Mark Bohr, Intel (2014)

A CMOS Inverter: Dynamic Power Dissipation

Q: How much energy is dissipated (in the PFET and the wires) in charging the capacitor to HIGH from LOW?

A: Irrespective of how it is charged, the net energy dissipation in charging a capacitor equals the energy stored in the capacitor after charging!

\[ E_D = \frac{1}{2} C_L V_{DD}^2 \]
Q: How much energy is dissipated (in the NFET and the wires) in discharging the capacitor from HIGH to LOW?

A: Irrespective of how it is discharged, the net energy dissipation in discharging a capacitor equals the energy stored in the capacitor before discharging!

\[ E_D = \frac{1}{2} C_L V_{DD}^2 \]

Total energy dissipation in one charge and discharge cycle:

\[ E_D = C_L V_{DD}^2 \]

---

Thermodynamics, Entropy, Information, and Computation

Question: How much energy does it require to compute or process one bit of information?

The question was answered by Rolf W. Landauer (1927-1999) (IBM)

Any thermodynamically irreversible operation that manipulates information increases entropy, and an associated amount of energy is unavoidably dissipated as heat.

The minimum amount of energy needed to process or compute one bit of information equals:

\[ KT \log(2) \]

\( KT \log(2) \approx 17.9 \text{ meV} \) at room temperature

For the smallest CMOS inverter intel has:

\[ E_D = C_L V_{DD}^2 \approx 62.5 \text{ eV} \]

\( C_L \approx C_{gs} = \frac{2 \varepsilon_{ox} W L}{3 t_{ox}} = 10^{-17} \text{ Farads} \)

\( V_{DD} = 1 \text{ V} \)

This is almost ~3500 times larger than the fundamental thermodynamic limit ...!!!

This is almost ~5000 times smaller than where CMOS was in the 80’s

There is plenty of room for improvement...!!!
Total energy dissipation in one charge and discharge cycle per FET:
\[ E_D = C_{gs} V_{DD}^2 \]

Ignoring interconnect capacitance

Total energy dissipation in one charge and discharge cycle if \( N_{FET} \) FETs in the chip are active:
\[ E_D = N_{FET} C_{gs} V_{DD}^2 \]

Total power dissipation (energy dissipation per second) if \( N_{FET} \) FETs are active:
\[ P_D = N_{FET} f_{CLK} C_{gs} V_{DD}^2 \]

Number of cycles per second ~ \( f_{CLK} \)

E8500 45 nm Chipset

Clock speed = 3.16 GHz
Gate length: 45 nm = .045 \( \mu \)m
Number of FETs in the chip = 410 \( \times \) 10^6
Power supply voltage ~ 2 V

\[ P_D = N_{FET} f_{CLK} C_{gs} V_{DD}^2 \]

Our power estimate:
\[ P_D = N_{FET} f_{CLK} C_{gs} V_{DD}^2 \]
\[ = 42 \text{ Watts!!} \]

Actual published number:
\[ P_T = 65 \text{ Watts!!} \]
\[ = P_D + P_S \]
Smaller transistor provides:
- Higher performance
- Lower power
- Lower cost per FET

**Intel FET Gate length Trends**

- $V_{DD} \sim 2$ V
- $V_{DD} < 1$ V

**CMOS Trends**

- Number of 1000s of FETs per cm²
- Clock speed (MHz)
- Power Dissipated (W/cm²)

**Intel CPU Trends**

- Dual-Core Itanium 2
- Pentium 4
- 386

*Source: Intel, Wikipedia, K. Ousterhout*