VLSI Design Issues

Scaling/Moore’s Law has limits due to the physics of material.
- Now L (L=20nm??) affects tx delays (speed), noise, heat (power consumption)
- Scaling increases density of txs and requires “more” interconnect (highways & buses)-more delays (lowering speed) and heat.

Possible Solutions:
- New fabrication solutions/material. E.g., Interconnect layers, new material (copper & low k-material)
- **Improve physical Designs at the transistor level. Create better cell libraries (min Power, min-delay, max speed)**
- Exploit transister analog/physics characteristics
- Invent new transistors
- Invent new architectures
CMOS Inverter: DC Analysis

- Analyze DC Characteristics of CMOS Gates by studying an Inverter

- DC Analysis
  - DC value of a signal in static conditions

- DC Analysis of CMOS Inverter
  - Vin, input voltage
  - Vout, output voltage
  - single power supply, VDD
  - Ground reference
  - find \( V_{\text{out}} = f(V_{\text{in}}) \)

- Voltage Transfer Characteristic (VTC)
  - plot of Vout as a function of Vin
  - vary Vin from 0 to VDD (and in reverse!)
  - find Vout at each value of Vin

\[
\text{pFET: } V_{T_p} < 0 \\
\beta_p = k_p \left( \frac{W}{L} \right)_p
\]

\[
\text{nFET: } V_{T_n} > 0 \\
\beta_n = k_n \left( \frac{W}{L} \right)_n
\]
Inverter Voltage Transfer Characteristics

- **Output High Voltage, \( V_{OH} \)**
  - maximum output voltage
  - occurs when input is low (\( Vin = 0V \))
  - pMOS is ON, nMOS is OFF
  - pMOS pulls \( V_{out} \) to \( V_{DD} \)
  - \( V_{OH} = V_{DD} \)

- **Output Low Voltage, \( V_{OL} \)**
  - minimum output voltage
  - occurs when input is high (\( Vin = V_{DD} \))
  - pMOS is OFF, nMOS is ON
  - nMOS pulls \( V_{out} \) to Ground
  - \( V_{OL} = 0 \) V

- **Logic Swing**
  - Max swing of output signal
  - \( V_L = V_{OH} - V_{OL} \)
  - \( V_L = V_{DD} \)
Inverter Voltage Transfer Characteristics

- **Gate Voltage, \( f(V_{in}) \)**
  - \( V_{GSn} = V_{in} \), \( V_{SGp} = V_{DD} - V_{in} \)

- **Drain Voltage, \( f(V_{out}) \)**
  - \( V_{DSn} = V_{out} \), \( V_{SDp} = V_{DD} - V_{out} \)

**Transition Region (between \( V_{OH} \) and \( V_{OL} \))**

- **Vin low**
  - \( V_{in} < V_{tn} \)
    - \( M_n \) in Cutoff, OFF
    - \( M_p \) in Triode, V_{out} pulled to V_{DD}
  - \( V_{in} > V_{tn} < V_{out} \)
    - \( M_n \) in Saturation, strong current
    - \( M_p \) in Triode, \( V_{SG} \) & current reducing
    - V_{out} decreases via current through \( M_n \)
  - **Vin = V_{out} (mid point) \approx \frac{1}{2} V_{DD}**
    - \( M_n \) and \( M_p \) both in Saturation
    - maximum current at \( V_{in} = V_{out} \)

- **Vin high**
  - \( V_{in} > V_{out} \), \( V_{in} < V_{DD} - |V_{tp}| \)
    - \( M_n \) in Triode, \( M_p \) in Saturation
  - \( V_{in} > V_{DD} - |V_{tp}| \)
    - \( M_n \) in Triode, \( M_p \) in Cutoff

ECE 410, Prof. F. Salem/Prof. A. Mason notes update
Noise Margin

- Input Low Voltage, $V_{IL}$
  - $V_{in} < V_{IL} =$ logic 0
  - point 'a' on the plot
    - where slope, $\frac{\partial V_{in}}{\partial V_{out}} = -1$

- Input High Voltage, $V_{IH}$
  - $V_{in} > V_{IH} =$ logic 1
  - point 'b' on the plot
    - where slope = -1

- Voltage Noise Margins
  - measure of how stable inputs are with respect to signal interference
  - $V_{NM_H} = V_{OH} - V_{IH} = V_{DD} - V_{IH}$
  - $V_{NM_L} = V_{IL} - V_{OL} = V_{IL}$
  - desire large $V_{NM_H}$ and $V_{NM_L}$ for best noise immunity
Switching Threshold

- **Switching threshold** = point on VTC where $V_{out} = V_{in}$
  - also called midpoint voltage, $V_M$
  - here, $V_{in} = V_{out} = V_M$

- **Calculating $V_M$**
  - at $V_M$, both nMOS and pMOS in Saturation
  - in an inverter, $I_{Dn} = I_{Dp}$, always!
  - solve equation for $V_M$

$$I_{Dn} = \frac{\mu_n C_{OX} W}{2L} (V_{GSn} - V_{in})^2 = \frac{\beta_n}{2} (V_{GSn} - V_{in})^2 = \frac{\beta_p}{2} (V_{SGp} - |V_{tp}|)^2 = I_{Dp}$$

- express in terms of $V_M$

$$\frac{\beta_n}{2} (V_M - V_{in})^2 = \frac{\beta_p}{2} (V_{DD} - V_M - |V_{tp}|)^2 \Rightarrow \sqrt{\frac{\beta_n}{\beta_p}} (V_M - V_{in}) = V_{DD} - V_M - |V_{tp}|$$

- solve for $V_M$

$$V_M = \frac{V_{DD} - |V_{tp}| + V_{in} \sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}}$$
Effect of Transistor Size on VTC

- Recall
  \[ \beta_n = k_n' \frac{W}{L} \]
  \[ \beta_p = \frac{k_p'}{\beta_p} \left( \frac{W}{L} \right)_p \]

- If nMOS and pMOS are same size
  - \((W/L)_n = (W/L)_p\)
  - \(C_{oxn} = C_{oxp}\) (always)
  \[ \beta_n = \frac{\mu_n C_{oxn} \left( \frac{W}{L} \right)_n}{\beta_p} = \frac{\mu_n}{\mu_p} \approx 2\text{or}3 \]

- If \(\left( \frac{W}{L} \right)_p = \frac{\mu_n}{\mu_p} \), then \(\beta_n = 1\)
  since \(L\) normally min. size for all tx, can get betas equal by making \(W_p\) larger than \(W_n\)

- Effect on switching threshold
  - if \(\beta_n \approx \beta_p\) and \(V_{tn} = |V_{tp}|\), \(V_M = VDD/2\), exactly in the middle

- Effect on noise margin
  - if \(\beta_n \approx \beta_p\), \(V_{IH}\) and \(V_{IL}\) both close to \(V_M\) and noise margin is good

\[ V_M = \frac{VDD - |V_{tp}| + V_m \sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}} \]
Example

- **Given**
  - \( k'n = 140\mu A/V^2, \ Vtn = 0.7V, \ VDD = 3V \)
  - \( k'p = 60\mu A/V^2, \ Vtp = -0.7V \)

- **Find**
  - a) tx size ratio so that \( V_M = 1.5V \)
  - b) \( V_M \) if tx are same size

Transition pushed lower as beta ratio increases
CMOS Inverter: Transient Analysis

- Analyze Transient Characteristics of CMOS Gates by studying an Inverter

- Transient Analysis
  - signal value as a function of time

- Transient Analysis of CMOS Inverter
  - $V_{in}(t)$, input voltage, function of time
  - $V_{out}(t)$, output voltage, function of time
  - $V_{DD}$ and Ground, DC (not function of time)
  - find $V_{out}(t) = f(V_{in}(t))$

- Transient Parameters
  - output signal rise and fall time
  - propagation delay
Transient Response

- Recall: the RC nMOS Transistor Model

![Diagram of an nMOS transistor and an RC circuit model]
Transient Response

- Response to step change in input
  - delays in output due to parasitic R & C
- Inverter RC Model
  - Resistances (linear model)
    - $R_n = 1/\left[\beta_n(V_{DD} - V_{tn})\right]$
    - $R_p = 1/\left[\beta_p(V_{DD} - |V_{tp}|)\right]$
  - Output Cap. (only output is important)
    - $C_{Dn}$ (nMOS drain capacitance)
      - $C_{Dn} = \frac{1}{2} Cox W_n L + C_j A_{Dnbot} + C_{jsw} P_{Dnsw}$
    - $C_{Dp}$ (pMOS drain capacitance)
      - $C_{Dp} = \frac{1}{2} Cox W_p L + C_j A_{Dpbot} + C_{jsw} P_{Dpsw}$
    - Load capacitance, due to gates attached at the output
      - $C_L = 3 C_{in} = 3 (C_{Gn} + C_{Gp})$, 3 is a “typical” load
    - Total Output Capacitance
      - $C_{out} = C_{Dn} + C_{Dp} + C_L$

\[V_in\] \[\Rightarrow\] \[\downarrow\] \[\downarrow\] \[\Rightarrow\]
\[\downarrow\] \[\downarrow\] \[\downarrow\]
\[C_{L}\] \[\downarrow\] \[\downarrow\] \[\downarrow\]
\[C_{in}\] \[\downarrow\] \[\downarrow\] \[\downarrow\]
\[\downarrow\] \[\downarrow\] \[\downarrow\]
Fall Time

• Fall Time, $t_f$
  - time for output to fall from '1' to '0'
  - derivation:
    \[
    i = -C_{out} \frac{\partial V_{out}}{\partial t} = \frac{V_{out}}{R_n}
    \]
    
    - initial condition, $V_{out}(0) = V_{DD}$
    - solution
    \[
    V_{out}(t) = V_{DD}e^{-\frac{t}{\tau_n}}
    \]
    \[
    t = \tau_n \ln\left(\frac{V_{DD}}{V_{out}}\right)
    \]
  - definition
    - $t_f$ is time to fall from 90% value $[V_1,t_x]$ to 10% value $[V_0,t_y]$
    \[
    t = \tau_n \left[\ln\left(\frac{V_{DD}}{0.1V_{DD}}\right) - \ln\left(\frac{V_{DD}}{0.9V_{DD}}\right)\right]
    \]
    - $t_f = 2.2 \, \tau_n$
Rise Time

- **Rise Time,** $t_r$
  - time for output to rise from '0' to '1'
  - derivation:
    $$ i = C_{out} \frac{\partial V_{out}}{\partial t} = \frac{V_{DD} - V_{out}}{R_p} $$
    - initial condition, $V_{out}(0) = 0V$
    - solution
      $$ V_{out}(t) = V_{DD} \left[ 1 - e^{-t/t_p} \right] $$
    - time constant
      $$ \tau_p = R_p C_{out} $$

- **definition**
  - $t_f$ is time to rise from 10% value $[V_0, t_u]$ to 90% value $[V_1, t_v]$
  - $t_r = 2.2 \tau_p$

- **Maximum Signal Frequency**
  - $f_{max} = 1/(t_r + t_f)$
    - faster than this and the output can't settle
Propagation Delay

- **Propagation Delay,** $t_p$
  - measures speed of output reaction to input change
  - $t_p = \frac{1}{2} (t_{pf} + t_{pr})$
- **Fall propagation delay,** $t_{pf}$
  - time for output to fall by 50%
  - reference to input switch
- **Rise propagation delay,** $t_{pr}$
  - time for output to rise by 50%
  - reference to input switch
- **Ideal expression** (if input is step change)
  - $t_{pf} = \ln(2) \tau_n$
  - $t_{pr} = \ln(2) \tau_p$
- **Total Propagation Delay**
  - $t_p = 0.35(\tau_n + \tau_p)$

Propagation delay measurement:
- from time input reaches 50% value
- to time output reaches 50% value

Add rise and fall propagation delays for total value
Switching Speed - Resistance

• **Rise & Fall Time**
  - \( t_f = 2.2 \tau_n, \ t_r = 2.2 \tau_p \)

• **Propagation Delay**
  - \( t_p = 0.35(\tau_n + \tau_p) \)

• **In General**
  - delay \( \propto \tau_n + \tau_p \)
  - \( \tau_n + \tau_p = Cout (R_n+R_p) \)

• **Define delay in terms of design parameters**
  - \( R_n+R_p = \frac{(V_{DD}-Vt)(\beta_n + \beta_p)}{\beta_n \beta_p (V_{DD}-Vt)^2} \)
  - \( R_n+R_p = \frac{\beta_n + \beta_p}{\beta_n \beta_p (V_{DD}-Vt)} \)

• if \( Vt = Vtn = |Vtp| \)

\[ \tau_n = R_n C_{out} \quad \tau_p = R_p C_{out} \]

\[ R_n = \frac{1}{[\beta_n(V_{DD}-Vtn)]} \quad \beta = \mu_{Cox} (W/L) \]

\[ R_p = \frac{1}{[\beta_p(V_{DD}-|Vtp|)]} \]

\[ C_{out} = C_{Dn} + C_{Dp} + C_L \]

**Beta Matched** if \( \beta_n = \beta_p = \beta \),

\[ R_n+R_p = \frac{2}{\beta (V_{DD}-Vt) \mu_{Cox} W (V_{DD}-Vt)} = \frac{2L}{\beta (V_{DD}-Vt) \mu_{Cox} W (V_{DD}-Vt)} \]

**Width Matched** if \( W_n = W_p = W \), and \( L = L_n = L_p \)

\[ R_n+R_p = \frac{L (\mu_n + \mu_p)}{(\mu_n \mu_p) \mu_{Cox} W (V_{DD}-Vt)} \]

**To decrease R's,** \( \downarrow L, \uparrow W, \uparrow VDD, \ (\uparrow \mu_p, \uparrow \mu_{Cox}) \)
Switching Speed - Capacitance

- From Resistance we have
  - $\downarrow L$, $\uparrow W$, $\uparrow VDD$, ($\uparrow \mu_p, \uparrow Cox$)
  - but $\uparrow VDD$ increases power
  - $\uparrow W$ increases $Cout$

- $Cout = C_{Dn} + C_{Dp} + C_L$

  - if $L=L_n=L_p$

  - $C_L = 3 (C_{Gn} + C_{Gp}) = 3 Cox (W_nL+W_pL)$

  - $C_{Dn} = \frac{1}{2} Cox W_n L + C_j A_{Dnbot} + C_{jsw} P_{Dnsw}$

  - $C_{Dp} = \frac{1}{2} Cox W_p L + C_j A_{Dpbot} + C_{jsw} P_{Dpsw}$

\[\text{To decrease } Cout, \downarrow L, \downarrow W, (\uparrow C_j, \downarrow Cox)\]

- $Cout \approx L (W_n+W_p) [3 \frac{1}{2} Cox +2 C_j]$

- $Cout \propto L (W_n+W_p)$

- Delay $\propto Cout (R_n+R_p) \propto L W \frac{L}{W VDD} = \frac{L^2}{VDD}$

Decreasing $L$ (reducing feature size) is best way to improve speed!
Switching Speed - Local Modification

- Previous analysis applies to the overall design
  - shows that reducing feature size is critical for higher speed
  - general result useful for creating cell libraries

- How do you improve speed within a specific gate?
  - increasing W in one gate will not increase \( C_G \) of the load gates
    - \( Cout = C_{Dn} + C_{Dp} + C_L \)
    - increasing W in one logic gate will increase \( C_{Dn/p} \) but not \( C_L \)
      - \( C_L \) depends on the size of the tx gates at the output
      - as long as they keep minimum W, \( C_L \) will be constant
  - thus, increasing W is a good way to improve the speed within a local point
  - But, increasing W increases chip area needed, which is bad
    - fast circuits need more chip area (chip "real estate")

- Increasing VDD is not a good choice because it increases power consumption
CMOS Power Consumption

- \( P = P_{DC} + P_{dyn} \)
  - \( P_{DC} \): DC (static) term
  - \( P_{dyn} \): dynamic (signal changing) term

- \( P_{DC} \)
  - \( P = I_{DD} V_{DD} \)
    - \( I_{DD} \): DC current from power supply
    - ideally, \( I_{DD} = 0 \) in CMOS: ideally only current during switching action
    - leakage currents cause \( I_{DD} > 0 \), define quiescent leakage current, \( I_{DDQ} \) (due largely to leakage at substrate junctions)
  - \( P_{DC} = I_{DDQ} V_{DD} \)

- \( P_{dyn} \), power required to switch the state of a gate
  - charge transferred during transition, \( Q_e = Cout V_{DD} \)
  - assume each gate must transfer this charge 1x/clock cycle
  - \( P_{average} = V_{DD} Q_e f = Cout V_{DD}^2 f \), \( f \) = frequency of signal change

- **Total Power**, \( P = I_{DDQ} V_{DD} + Cout V_{DD}^2 f \)
  - Power increases with \( Cout \) and frequency, and strongly with \( VDD \) (second order).
Multi-Input Gate Signal Transitions

- In multi-input gates multiple signal transitions produce output changes

- What signal transitions need to be analyzed?
  - for a general N-input gate with $M_0$ low output states and $M_1$ high output states
    - # high-to-low output transitions $= M_0 \cdot M_1$
    - # low-to-high output transitions $= M_1 \cdot M_0$
    - total transitions to be characterized $= 2 \cdot M_0 \cdot M_1$
  - example: NAND has $M_0 = 1, M_1 = 3$
  - don't test/characterize cases without output transitions

- Worst-case delay is the slowest of all possible cases
  - worst-case high-to-low
  - worst-case low-to-high
  - often different input transitions for each of these cases
Series/Parallel Equivalent Circuits

- Scale both \( W \) and \( L \)
  - no effective change in \( W/L \)
  - increases gate capacitance

inputs must be at same value/voltage

- Series Transistors
  - increases effective \( L \)

- Parallel Transistors
  - increases effective \( W \)
NAND: DC Analysis

- Multiple Inputs
- Multiple Transitions
- Multiple VTCs
  - VTC varies with transition
    - transition from 0,0 to 1,1 pushed right of others
    - why?
  - $V_M$ varies with transition
    - assume all tx have same $L$
    - $V_M = V_A = V_B = V_{out}$
      - can merge transistors at this point
    - if $W_{pA} = W_{pB}$ and $W_{nA} = W_{nB}$
      - series nMOS, $\beta_n \Rightarrow \frac{1}{2} \beta_n$
      - parallel pMOS, $\beta_p \Rightarrow 2 \beta_p$
    - can now calculate the NAND $V_M$
NAND Switching Point

- **Calculate VM for NAND**
  - 0,0 to 1,1 transition
    - all tx change states (on, off)
    - in other transitions, only 2 change
  - \( V_M = V_A = V_B = V_{out} \)
  - set \( I_{Dn} = I_{Dp} \), solve for \( V_M \)
  
  \[
  V_M = \frac{V_{DD} - \left| V_p \right| + V_m \frac{1}{2} \sqrt{\frac{\beta_n}{\beta_p}}} {1 + \frac{1}{2} \sqrt{\frac{\beta_n}{\beta_p}}}
  \]

- denominator reduced more
  - VTC shifts right

- **For NAND with N inputs**
  
  \[
  V_M = \frac{V_{DD} - \left| V_p \right| + V_m \frac{1}{N} \sqrt{\frac{\beta_n}{\beta_p}}} {1 + \frac{1}{N} \sqrt{\frac{\beta_n}{\beta_p}}}
  \]

- series nMOS means more resistance to output **falling**, shifts VTC to right
  to balance this effect and set \( V_M \) to \( V_{DD}/2 \)
  can increase \( \beta \) by increasing \( W_n \)

but, since \( \mu_n > \mu_p \), \( V_M \approx V_{DD}/2 \) when \( W_n = W_p \)
**NOR: DC Analysis**

- **Similar Analysis to NAND**
- **Critical Transition**
  - 0,0 to 1,1
  - when all transistors change
- **$V_M$ for NOR2 critical transition**
  - if $W_{pA} = W_{pB}$ and $W_{nA} = W_{nB}$
    - parallel nMOS, $\beta_n \Rightarrow 2 \beta_n$
    - series pMOS, $\beta_p \Rightarrow \frac{1}{2} \beta_p$

\[
V_M = \frac{V_{DD} - |V_p| + 2V_m \sqrt{\frac{\beta_n}{\beta_p}}}{1 + 2 \sqrt{\frac{\beta_n}{\beta_p}}}
\text{ for NOR2}
\]
\[
V_M = \frac{V_{DD} - |V_p| + NV_m \sqrt{\frac{\beta_n}{\beta_p}}}{1 + N \sqrt{\frac{\beta_n}{\beta_p}}}
\text{ for NOR-N}
\]

- series pMOS resistance means slower rise
- VTC shifted to the left
- to set $V_M$ to $V_{DD}/2$, increase $W_p$
  - this will increase $\beta_p$

<table>
<thead>
<tr>
<th>$V_A$</th>
<th>$V_B$</th>
<th>$V_{out}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>$V_{DD}$</td>
</tr>
<tr>
<td>(i)</td>
<td>$V_{DD}$</td>
<td>0</td>
</tr>
<tr>
<td>(ii)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Transition table

(b) VTC family

---

ECE 410, Prof. F. Salem/Prof. A. Mason notes update

Lecture Notes 7.23
NAND: Transient Analysis

- NAND RC Circuit
  - R: standard channel resistance
  - C: \( C_{\text{out}} = C_L + C_{Dn} + 2C_{Dp} \)
- Rise Time, \( t_r \)
  - Worst case charge circuit
    - 1 pMOS ON
    - \( t_r = 2.2 \, \tau_p \)
    - \( \tau_p = R_p \, C_{\text{out}} \)
  - best case charge circuit
    - 2 pMOS ON, \( R_p \Rightarrow R_p/2 \)
- Fall Time, \( t_f \)
  - Discharge Circuit
    - 2 series nMOS, \( R_n \Rightarrow 2R_n \)
    - must account for internal cap, \( C_x \)
    - \( t_f = 2.2 \, \tau_n \)
    - \( \tau_n = C_{\text{out}} \, (2 \, R_n) + C_x \, R_n \)
  - \( C_x = C_{Sn} + C_{Dn} \)
NOR: Transient Analysis

- **NAND RC Circuit**
  - R: standard channel resistance
  - C: $C_{out} = C_L + 2C_{Dn} + C_{Dp}$

- **Fall Time, $t_f$**
  - Worst case discharge circuit
    - 1 nMOS ON
      - $t_f = 2.2 \tau_n$
        - $\tau_n = R_n C_{out}$
    - best case discharge circuit
      - 2 nMOS ON, $R_n \Rightarrow R_n/2$

- **Rise Time, $t_r$**
  - Charge Circuit
    - 2 series pMOS, $R_p \Rightarrow 2R_p$
      - must account for internal cap, $C_y$
    - $t_r = 2.2 \tau_p$
      - $\tau_p = C_{out} (2R_p) + C_y R_p$
NAND/NOR Performance

- Inverter: symmetry ($V_M=V_{DD}/2$), $\beta_n = \beta_p$
  - $(W/L)_p = \mu_n/\mu_p (W/L)_n$
- Match INV performance with NAND
  - pMOS, $\beta_p = \beta_p$, same as inverter
  - nMOS, $\beta_N = 2\beta_n$, to balance for 2 series nMOS
- Match INV performance with NOR
  - pMOS, $\beta_p = 2\beta_p$, to balance for 2 series pMOS
  - nMOS, $\beta_N = \beta_n$, same as inverter
- NAND and NOR will still be slower due to larger $C_{out}$
- This can be extended to 3, 4, ... $N$ input NAND/NOR gates

$\beta$ is adjusted by changing transistor size (width)
NAND/NOR Transient Summary

- **Critical Delay Path**
  - paths through series transistors will be slower
  - more series transistors means worse delays

- **Tx Sizing Considerations**
  - increase $W$ in series transistors
  - balance $\beta_n/\beta_p$ for each cell

- **Worst Case Transition**
  - when all series transistor go from OFF to ON
  - and all internal caps have to be
    - charged (NOR)
    - discharged (NAND)
Performance Considerations

- Speed based on $\beta_n$, $\beta_p$ and parasitic caps
- DC performance ($V_M$, noise) based on $\beta_n/\beta_p$
- Design for speed not necessarily provide good DC performance
- Generally set tx size to optimize speed and then test DC characteristics to ensure adequate noise immunity

• Review Inverter: Our performance reference point
  - for symmetry ($V_M = V_{DD}/2$), $\beta_n = \beta_p$
    - which requires $(W/L)_p = \mu_n/\mu_p (W/L)_n$
  • Use inverter as reference point for more complex gates

• Apply slowest arriving inputs to series node closest to output
  - let faster signals begin to charge/discharge nodes closer to VDD and Ground
Timing in Complex Logic Gates

- Critical delay path is due to series-connected transistors
- Example: $f = x (y+z)
  - assume all tx are same size
- Fall time critical delay
  - worst case, $x$ ON, and $y$ or $z$ ON
  - $t_f = 2.2 \tau_n$
    - $\tau_n = R_n C_n + 2 R_n C_{out}$
    - $C_{out} = 2C_{Dp} + C_{Dn} + C_L$
    - $C_n = 2C_{Dn} + C_{Sn}$
- Rise time critical delay
  - worst case, $y$ and $z$ ON, $x$ OFF
  - $t_r = 2.2 \tau_p$
    - $\tau_p = R_p C_p + 2 R_p C_{out}$
    - $C_{out} = 2C_{Dp} + C_{Dn} + C_L$
    - $C_p = C_{Dp} + C_{Sp}$

Size vs. tx speed considerations

\[ W_{nx} \Rightarrow R_n \text{ but } C_{out} \text{ and } C_n \]
\[ W_{ny} \Rightarrow C_n \text{ but } R_n \]
\[ W_{pz} \Rightarrow R_p \text{ but } C_{out} \text{ and } C_p \]
\[ W_{px} \Rightarrow \text{no effect on critical path!} \]
Sizing in Complex Logic Gates

- Improving speed within a single logic gate
- An Example: \( f = (a \cdot b + c \cdot d) \times \)
- nMOS
  - discharge through 3 series nMOS
  - set \( \beta_N = 3 \beta_n \)
- pMOS
  - charge through 2 series pMOS
  - set \( \beta_P = 2 \beta_p \)
  - but, \( M_{xp} \) is alone so \( \beta_{P1} = \beta_p \)
    - but setting \( \beta_{P1} = 2 \beta_p \) might make layout easier
- These large transistors will increase capacitance and layout area and may only give a small increase in speed
- Advanced logic structures are best way to improve speed
Timing in Multi-Gate Circuits

- What is the worst-case delay in multi-gate circuits?
  - too many transitions to test manually

- Critical Path
  - longest delay through a circuit block
  - largest sum of delays, from input to output
  - intuitive analysis: signal that passes through most gates
    - not always true, can be slower path through fewer gates

path through most gates

critical path if delay at D input is very slow
Power in Multi-Input Logic Gates

- **Inverter Power Consumption**
  - \( P = P_{DC} + P_{dyn} = V_{DD}I_{DDQ} + C_{out}V^2_{DD}f \)
  - Assumes gates switch output state once per clock cycle, \( f \)

- **Multi-Input Gates**
  - Same DC component as inverter, \( P_{DC} = V_{DD}I_{DDQ} \)
  - For dynamic power, need to estimate “activity” of the gate, how often will the output be switching
  - \( P_{dyn} = aC_{out}V^2_{DD}f \), \( a \) = activity coefficient
  - Estimate activity from truth table

<table>
<thead>
<tr>
<th>( A )</th>
<th>( B )</th>
<th>( A + B )</th>
<th>( A \cdot B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- \( a = p_0p_1 \)
  - \( p_0 \) = prob. output is at 0
  - \( p_1 \) = prob. of transition to 1

NOR: \( p_0=0.75 \), \( p_1=0.25 \), \( a=3/16 \)
NAND: \( p_0=0.25 \), \( p_1=0.75 \), \( a=3/16 \)
Timing Analysis of Transmission Gates

- **TG** = parallel nMOS and pMOS

- **RC Model**
  - In general, only one tx active at same time
    - nMOS pulls output low
    - pMOS pushes output high
  - \( R_{TG} = \max (R_n, R_p) \)
  - \( C_{in} = C_{Sn} + C_{Dp} \)
    - If output at higher voltage than input
  - Larger W will decrease R but increase Cin

- Note: no connections to VDD-Ground. Input signal, Vin, must drive TG output; TG just adds extra delay
Pass Transistor

- Single nMOS or pMOS tx
- Often used in place of TGs
  - less area and wiring
  - can’t pull to both VDD and Ground
  - typically use nMOS for better speed
- Rise and Fall Times
  - $\tau_n = R_n C_{out}$
  - $t_f = 2.94 \tau_n$
  - $t_r = 18 \tau_n$
    - much slower than fall time
- nMOS can’t pull output to VDD
  - rise time suffers from threshold loss in nMOS