Chapter 2
Interconnect Analysis

Delay Modeling
S2P and Effective Capacitance

Wei Wu
Electrical Engineering Department
University of California, Los Angeles
Outline

- Delay models
  - Review of Elmore Delay
  - Second Order Analysis (S2P algorithm)
  - Gate delay
Review of Elmore Delay

50% Delay for lumped RC model

\[ V(t) = 1 - e^{-\frac{t}{RC}} \]

\[ e^{-\frac{t}{RC}} = 0.5 \]

\[ t = 0.69RC \]

How about more complex circuits?
Elmore delay at node $i$ is

$$
\tau_{Di} = 0.69(R_{i1}C_1 + R_{i2}C_2 + R_{i3}C_3 + R_{i4}C_4 + R_{ii}C_i)
$$

$$
= 0.69(R_1C_1 + R_1C_2 + (R_1 + R_3)C_3 + (R_1 + R_3)C_4 + (R_1 + R_3 + R_i)C_i)
$$

$$
= 0.69(R_1(C_1 + C_2 + C_3 + C_4 + C_i) + R_3(C_3 + C_4 + C_i) + R_iC_i)
$$
The Elmore delay is the metric of choice for performance-driven design applications due to its simple, explicit form and ease with which sensitivity information can be calculated.

However, for deep submicron technologies (DSM), the accuracy of the Elmore delay is insufficient.
Outline

- Delay models
  - Review of Elmore Delay
  - Second Order Analysis (S2P algorithm)
  - Gate delay
Alternative solution: S2P Algorithm

Second Order Analysis (Stable 2-Pole Delay Algorithm)

General Idea: MOR

Original RC network:

\[ H(s), Y(s) = \sum_{n=1}^{q} \frac{k_n}{s - p_n} + k_0 \]

2 poles Model: (Model Order Reduction: AWE)

\[ h(s) = \frac{b_0 + b_1 s}{1 + a_1 s + a_2 s^2} = \frac{k_1}{s - p_1} + \frac{k_2}{s - p_2} \]

Generate the time domain single from 2P model:

eg:

Impulse Response: \( \hat{h}(t) = (k_1 e^{p_1 t} + k_2 e^{p_2 t}) \times u(t) \)

Moments of \( H(s) \)

Moments of \( H(s) \) are coefficients of the Taylor’s Expansion of \( H(s) \) about \( s=0 \)

\[
H(s) = H_0 + s \frac{d}{ds} H(s) \bigg|_{s=0} + s^2 \frac{1}{2!} \frac{d^2}{ds^2} H(s) \bigg|_{s=0} + s^3 \frac{1}{3!} \frac{d^3}{ds^3} H(s) \bigg|_{s=0} + ... \\
= m^{(0)} + s m^{(1)} + s^2 m^{(2)} + s^3 m^{(3)} + ... \\
m^{(j)} = \frac{1}{j!} \frac{d^j}{ds^j} H(s) \bigg|_{s=0}
\]
Let $Y(s)$ be an driving point admittance function of a general RC circuit. Consider its representation in terms poles and residues.

$$Y(s) = \sum_{n=1}^{q} \frac{k_n}{s - p_n} + k_0$$  \text{ where } q \text{ is the exact order of the circuit}

Moments of $Y(s)$ can be written as:

$$m_i = \sum_{n=1}^{q} \frac{k_n}{p_n^{i+1}}$$  \text{ i > 0}
Compute $m_1$, $m_2$, $m_3$ and $m_4$ for $Y(s)$

Find the two poles at the driving point admittance as follows:

$$ a_1 = \frac{m_2m_3 - m_1m_4}{m_1m_3 - m_2^2}, \quad a_2 = \frac{m_2m_4 - m_3^2}{m_1m_3 - m_2^2} $$

$$ p_1 = \frac{-a_1 + \sqrt{a_1^2 - 4a_2}}{2a_2}, \quad p_2 = \frac{-a_1 - \sqrt{a_1^2 - 4a_2}}{2a_2} $$

To match the voltage moments at the response nodes, choose

$$ k_2 = \frac{(m_0^*/p_1 - m_1^*)}{1/p_2 - 1/p_1} p_2 $$

$$ k_1 = (-m_0^* - k_2/p_2)p_1 $$

and the S2P approximation is then expressed as:

$$ \hat{h}(s) = \frac{k_1}{s-p_1} + \frac{k_2}{s-p_2} $$

Note that $m_0^*$ and $m_1^*$ are the moments of $H(s)$. $m_0^*$ is the Elmore delay.

S2P Vs. Elmore Delay

![Graph showing the comparison between S2P and Elmore delay over time](graph.png)
Outline

- Delay models
  - Review of Elmore Delay
  - Second Order Analysis (S2P algorithm)
  - Gate delay
Gate Delay

Overall flow of delay calculation

\[ \text{Delay}_{AC} = \text{Delay}_{AB} + \text{Delay}_{BC} \]

\( \text{Delay}_{AB} \): Gate delay
\( \text{Delay}_{BC} \): RC network delay
General Model of a Gate
Definitions

Output Transition Time

Gate Delay

Vin

Vout

10%

90%

Time

Gate/Cell

C_{load}

T_{in}

\tau_{in}

Definitions
Gate Delay and Output Transition Time

\[
\text{Gate Delay} = f(T_{in}, C_{load})
\]

\[
\text{Output Transition Time} = f(T_{in}, C_{load})
\]

The gate delay and the output transition time are functions of both input slew and the output load.
Output Response for Different Loads

Output Response with different Load Capacitance

CL = 0.1 pF, 0.5 pF, 1 pF, 1.5 pF
Three approaches for gate propagation delay computation are based on:
- Delay look-up tables
- K-factor approximation
  - Effective capacitance

Delay look-up table is currently in wide use especially in the ASIC design flow.
Effective capacitance promises to be more accurate when the load is not purely capacitive.
What is the delay when $C_{load}$ is $505 \text{fF}$ and $T_{in}$ is $90 \text{pS}$?
We can fit the output transition time v.s. input transition time and output load as a polynomial function, e.g.

\[ T_{output} = k_1 + k_2 C_{load} + T_{in}(k_3 + k_4 C_{load}) - k_5 C_{load}^2 \]

A similar equation gives the gate delay
How to calculate the $C_{\text{load}}$?

The gate delay and the output transition time are functions of both input slew and the output load.

Gate Delay $= f(T_{\text{in}}, C_{\text{load}})$

Output Transition Time $= f(T_{\text{in}}, C_{\text{load}})$
Second-order RC-π Model

- Using Taylor Expansion around $s = 0$

\[ Y_{in}(s) = A_1 s + A_2 s^2 + A_3 s^3 + ..... \]

\[ \hat{Y}_{in}(s) = (C_1 + C_2) s - R_\pi C_2^2 s^2 + R_\pi^2 C_2^3 s^3 + ..... \]

\[ C_1 = A_1 - \frac{A_2^2}{A_3} \quad R_\pi = -\frac{A_3^2}{A_2^3} \quad C_2 = \frac{A_2^2}{A_3} \]
Second-order RC-$$\pi$$ Model (Cont’d)

This equation requires creation of a four-dimensional table to achieve high accuracy.

This is however costly in terms of memory space and computational requirements.
The “Effective Capacitance” approach attempts to find a single capacitance value that can be replaced instead of the RC-\(\pi\) load such that both circuits behave similarly during transition.

What's the value of \(C_{eff}\), \(C_1+C_2??\)
Output Response for Effective Capacitance

Effective Capacitance

Voltage (Volt)

Time (s)

Wn/Wp = 40/80
Ln/Lp = 0.25 μm
trin = 100 ps
C1 = 0.1 pF
C2 = 0.5 pF
R = 500 ohm
Ceff = 0.134 pF
50% delay error = 0.9%
20%–80% rise time error = 7%
Effective Capacitance (Cont’d)

Effective Capacitance 2

Voltage (volt)

Time (s)

Wn/Wp = 40/80
Ln = Lp = 0.25 um
trin = 100 ps
C1 = 0.1 pF
C2 = 0.5 pF
R = 200 ohm

Ceff = 0.19 pF
50% delay error = 4%
20%–80% rise time error = 22%
Effective Capacitance (Cont’d)

Because of the shielding effect of the interconnect resistance, the driver will only “see” a portion of the far-end capacitance $C_2$

$C_{\text{eff}} = C_1 + kC_2$

$0 < k < 1$

- $R_{\pi} \rightarrow 0$  
  $k = 1$

- $R_{\pi} \rightarrow \infty$  
  $k = 0$
Effective Capacitance for Different Resistive Shielding

\[ Wn/Wp=40/80 \]
\[ Ln=Lp=0.25\mu m \]
\[ trin=100\,ps \]
\[ C1=1e^{-13}F \]
\[ C2=5e^{-13}F \]
\[ R=0\sim2000\,ohm \]
\[ C_{eff}=1.34e^{-13}F \]
Assumption: If two circuits have the same loads and output transition times, then their effective capacitances are the same

=> the effective capacitance is only a function of the output transition time and the load

Macy’s Iterative Solution

1. Compute $\alpha$ from $C_1$ and $C_2$
2. Choose an initial value for $C_{\text{eff}}$
3. Compute $T_{\text{out}}$ for the given $C_{\text{eff}}$ and $T_{\text{in}}$
4. Compute $\beta$
5. Compute $\gamma$ from $\alpha$ and $\beta$
6. Find new $C_{\text{eff}}$
7. Go to step 3 until $C_{\text{eff}}$ converges

$\alpha = \frac{C_1}{C_1 + C_2}$

$\beta = \frac{T_{\text{out}}}{R\pi C_2}$

$\gamma = \frac{C_{\text{eff}}}{C_1 + C_2}$

Delay model
- Elmore Delay model
- S2P model
  - Reduce the RC network to a 2-pole model
  - AWE is adopted for MOR
- Gate delay:
  - look-up table, k-factor approximation,
  - effective capacitance
References


- Jason Cong, Lei He, Cheng-Kok Koh, Patrick H. Madden, Performance optimization of VLSI interconnect layout, Integration, the VLSI Journal, v.21 n.1-2, p.1-94, Nov. 1996 (Section 2.1-2.2)


- Jorge Rubinstein, and etc. “Signal Delay in RC Tree Networks”, TCAD'83