



DEPARTMENT OF COMPUTER SYSTEM ENGINEERING

Digital Integrated Circuits - ENCS333

Dr. Khader Mohammad

Lecture #11- Timing and power

Integrated-Circuit Devices and Modeling

Timing

- Intro (What is timing all about)
- Static timing versus dynamic timing
- Timing primer
- Flop based timing
- How clock uncertainty impacts timing analysis
- Latch based timing (transparency, time borrowing, time borrowing FFs etc)
- What is a design window?
- Gated clock & domino timing
- Phase paths, MCPs & cycle adjusted margins
- Understanding EVRs & BVRs
- Full chip rollups and slack allocation
- Advanced Topics (loop paths, zero-cycle paths)

What do we mean by “Timing” ?

- Timing Analysis is a method of analyzing and solving timing problems to ensure that the design meets the target frequency.
- A circuit meets target frequency when all of the signals:
 - Arrive just before they are needed.
 - Remain until no longer needed.
 - Be guaranteed stable in between.
- Critical paths are paths in the circuit along which the signal doesn't reach its destination on time; this causes the functionality of that part of the circuit to be impaired.

Static Timing

- Static timing analysis computes the worst case delays for paths in a given circuit.
- It is a non-simulation approach to timing analysis used to analyze the propagation of delays.
- Pathmill and Tango-XT are examples of static timing tools.

Static Timing

- Searches path by path. Performs static timing analysis (STA) of a circuit by analyzing the circuit in a non-simulated way (no input vectors.)
 - Traverses all possible paths between each source-sink pair.
 - Calculates the path delay and the arrival time at each signal along the path and internal sampling requirements.
 - Sorts the paths by the amount of violation at their sink node and prints them in a report.
- Report can be used to:
 - Determine how to fine tune the timing of the design to meet the target frequency
 - Analyze timing violations and performance bottlenecks in circuits.

Static vs. Dynamic Timing

- Static timing Analysis:
 - Non simulation approach used to analyze propagation of delays.
 - Computes worst case path delays by accumulating pre-characterized device delays along a given path.
 - Path determination is based on the possibility of an event on a device input causing an event on a device output.
- Dynamic Timing Analysis:
 - Circuit simulation approach that obtains a very accurate timing analysis of a path.
 - Uses input waveforms and path sensitization, generates output waveforms.

Comparison of Static vs. Dynamic Timing

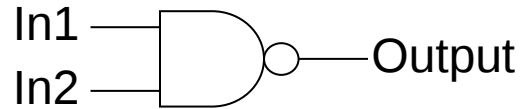
Static (path analysis)	Dynamic (simulation)
• No input vectors; full coverage	• Uses input vectors; coverage depends on vector selection
• Points to exact origin of timing problems (critical paths)	• Requires a debug process to find the origin of timing problems
• Ensures full coverage (Much faster than simulation)	• Requires a lot of CPU time for good coverage
• Relies on approximated models which are less accurate than simulation	• Highest degree of accuracy possible through software simulation
• Inherent problem: false paths	• Inherent problem: test vector generation
• Simulates only one input switching at a time	• Simulates multiple inputs switching simultaneously
• Gives worst case delays only	• Gives precise delay information

When do we use Static or Dynamic Timing?

- Static Timing is used more often due to the size of the designs.
- Dynamic timing is typically used where analog behavior is expected and not modeled correctly by static timing.

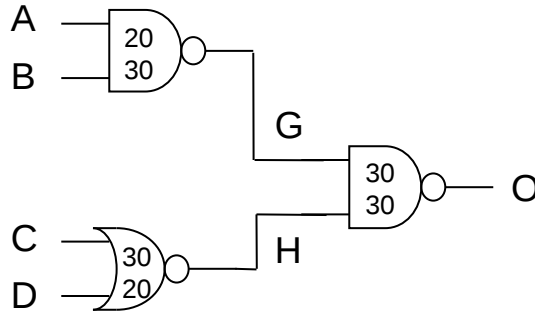
Static vs. Dynamic Timing:

Example 1



- Given a simple two-input NAND gate, static analysis will yield four paths analyzed:
 - In1 rising -> output falling
 - In1 falling -> output rising
 - In2 rising -> output falling
 - In2 falling -> output rising
- Given the same NAND gate, dynamic analysis can yield six main transitions. The above four, as well as
 - In1 rising & In2 rising -> output falling
 - In1 falling & In2 falling -> output rising
- Theoretically, Dynamic Timing has an infinite number of different paths, as the user determines the exact timing of the rise/fall of the inputs in relation to each other.

Static vs Dynamic Timing: Example2



Path #1 A -> O

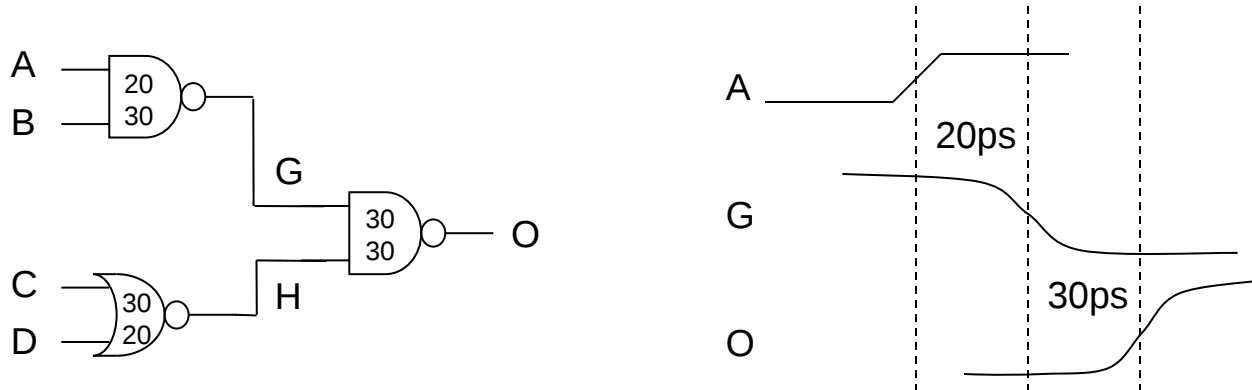
A Rising -> G Falling -> O Rising
20ps + 30ps = 50ps

Path #2 A -> O

A Falling -> G Rising -> O Falling
30ps + 30ps = 60ps

- Static timing analysis is based upon the possibility of a path existing.
- All paths are traversed and delays are calculated.
- These delays are used to check signal arrival times in order to determine setup and hold margin.

Static vs Dynamic Timing: Example2 (cont.)

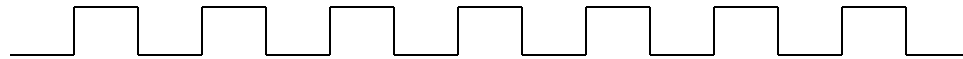


- Using Dynamic Timing on the same example, inputs B, C, and D need to be sensitized to exercise the path: A rising -> G falling -> O rising
 - B must be set high
 - C and D must be set low

Timing Primer

- Terms we need to learn very very well:
 - Reference Clock
 - Lead edge signals
 - Trail edge signals
 - Valid Windows
 - Required Windows
 - Margin
 - Margin Violations

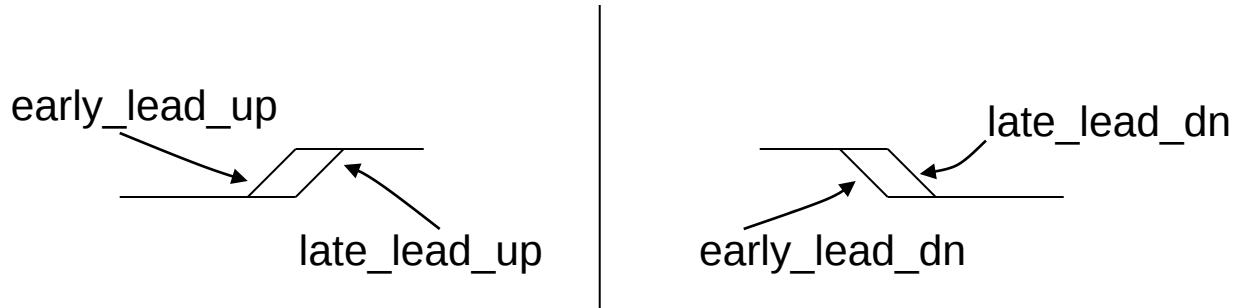
What is a Reference Clock?



- An ideal clock that does not have jitter or skew
- Generally does not exist anywhere in the design
- All timing is analyzed and specified against a reference clock
- There may be more than one per design

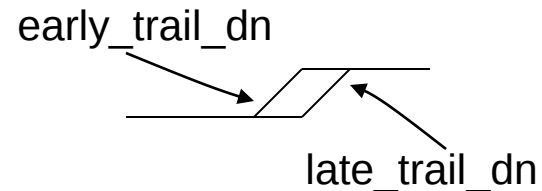
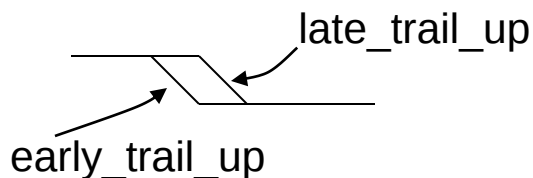
What are Lead Edge Signals ?

- The lead edge is the time that the signal become stable/active.
- If a signal rises and becomes active, this is **lead_up**.
- If the signal falls and becomes active, it is a **lead_dn** edge.
- A signal may have multiple lead up edges at different times. For this case we can describe them using the terms: **early_lead** and **late_lead**



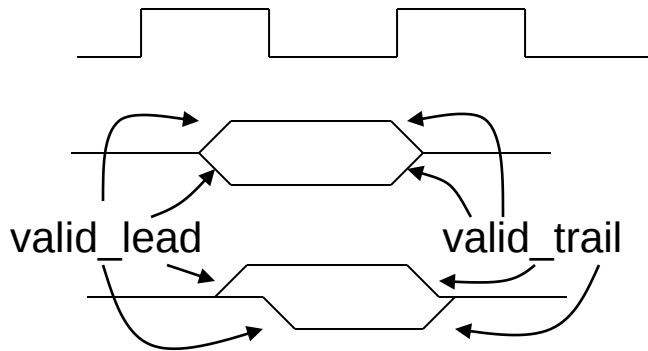
What are Trail Edge Signals ?

- The trail edge is the time that the signal become inactive (or goes away).
- The **trail_up** edge (falling edge) refers to the trail edge of the “up” (logical 1) window.
- The **trail_dn** edge (rising edge) refers to the trail of the “dn” (logical 0) window.



What are Valid Windows?

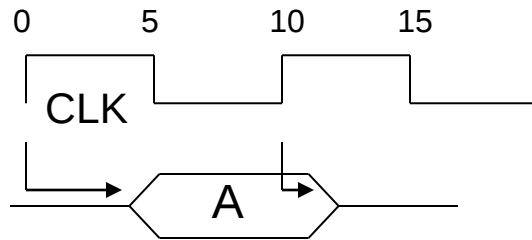
- Valid windows are referenced to a clock event and indicate the time when a signal is stable.
- The `valid_lead_up` edge and the `valid_lead_dn` edge need not have the same timing. These edges may come from entirely different paths of logic.



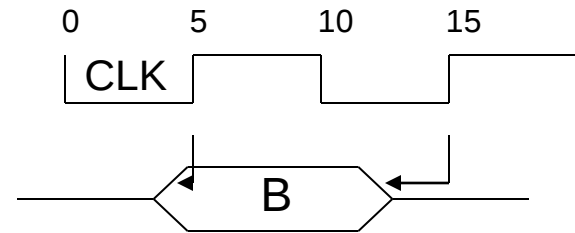
- `valid_lead` will correspond to `late_lead`
- `valid_trail` will correspond to `early_trail`

Valid Window Examples

- Valid times are conventionally specified as
AR or AF (after rise or after fall)



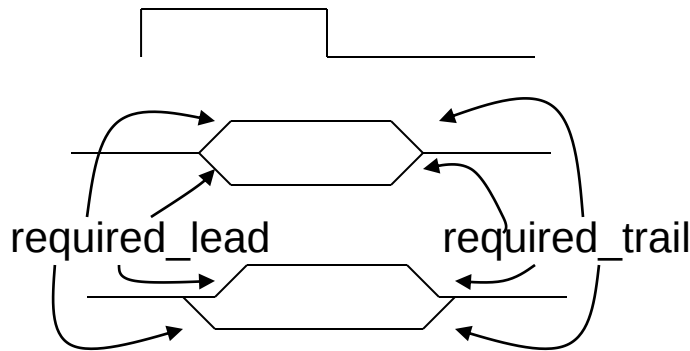
A: $\text{valid_lead} = 4$ ar CLK
 $\text{valid_trail} = 2$ ar CLK



B: $\text{valid_lead} = -1$ ar CLK
 $\text{valid_trail} = -3$ ar CLK

What are Required Windows?

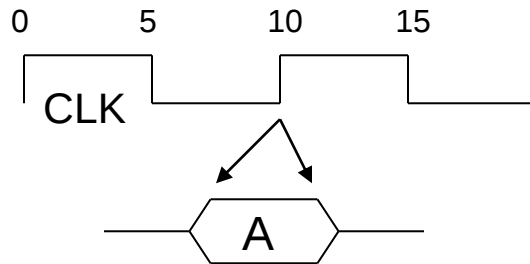
- Required windows indicate the time when a signal must be valid or active to meet timing. This signal must be valid for the whole time of the defined window.
- Required windows come from circuit timing requirements for setup times and hold times.



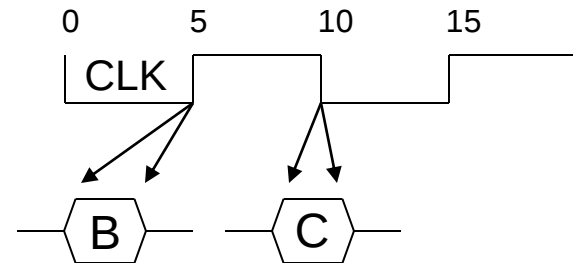
- `required_lead` will correspond to `early_lead`
- `required_trail` will correspond to `late_trail`

Required Window Examples

- Required times are conventionally specified as
BR or BF (before rise or before fall)



A: required_lead = 4 br CLK
required_trail = -3 br CLK

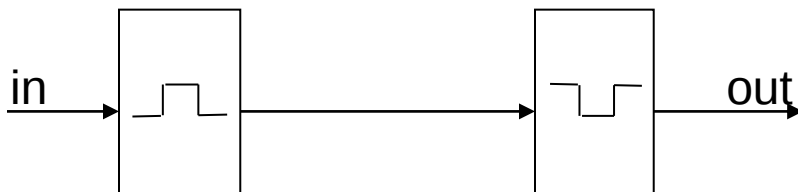


B: required_lead = 5 br CLK
required_trail = 2 br CLK

C: required_lead = 2 bf CLK
required_trail = -1 bf CLK

A Note about Coloring Conventions

- The color (af/ar) of the **VALID** lead is always referenced to the “opening” edge of the closest sequential element to the output along a path.
- The color (bf/br) of the **REQUIRED** lead edge is always referenced to the “closing” edge of the closest sequential element to the input along a path.



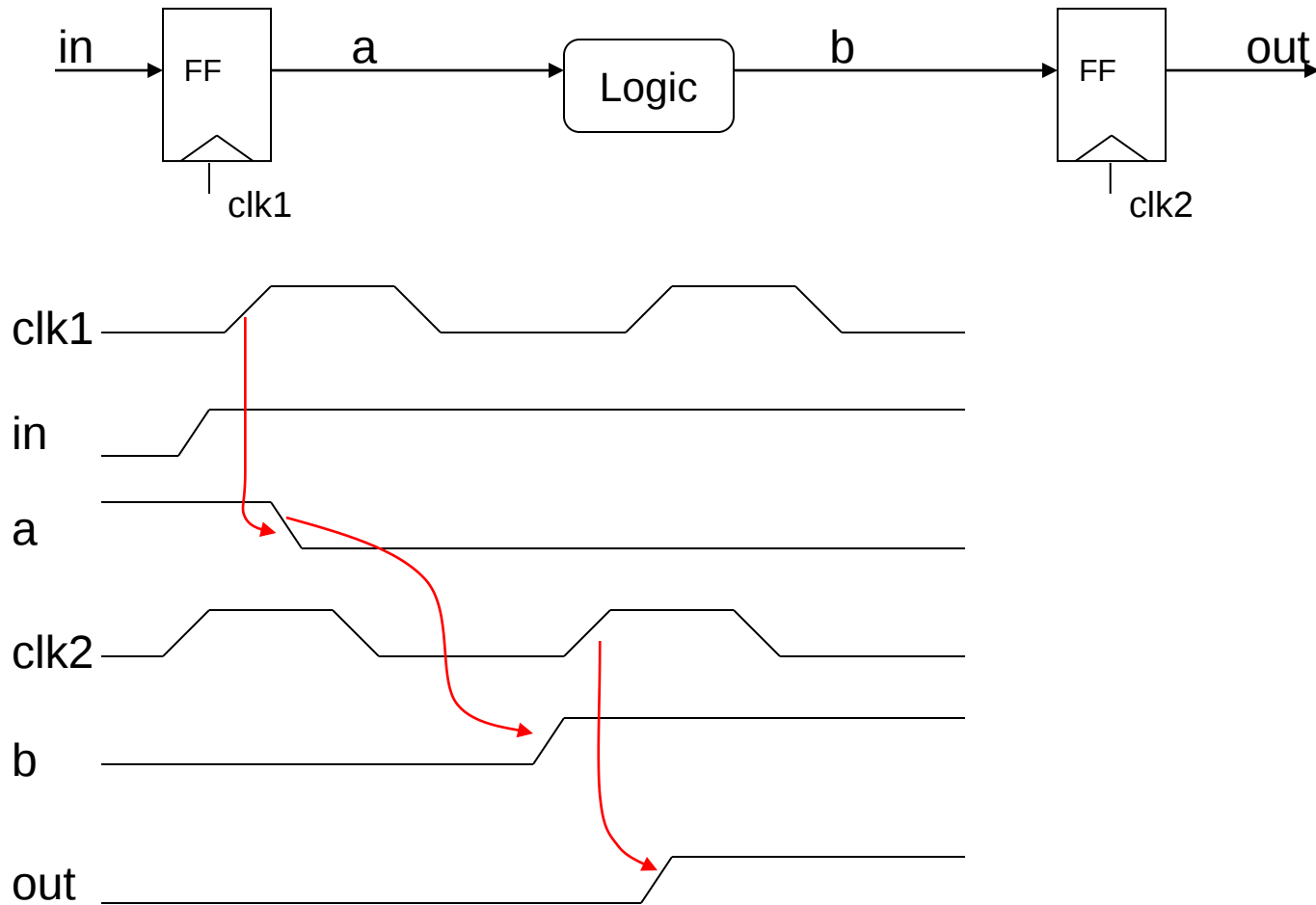
For the example:

VALID = af

REQUIRED = bf

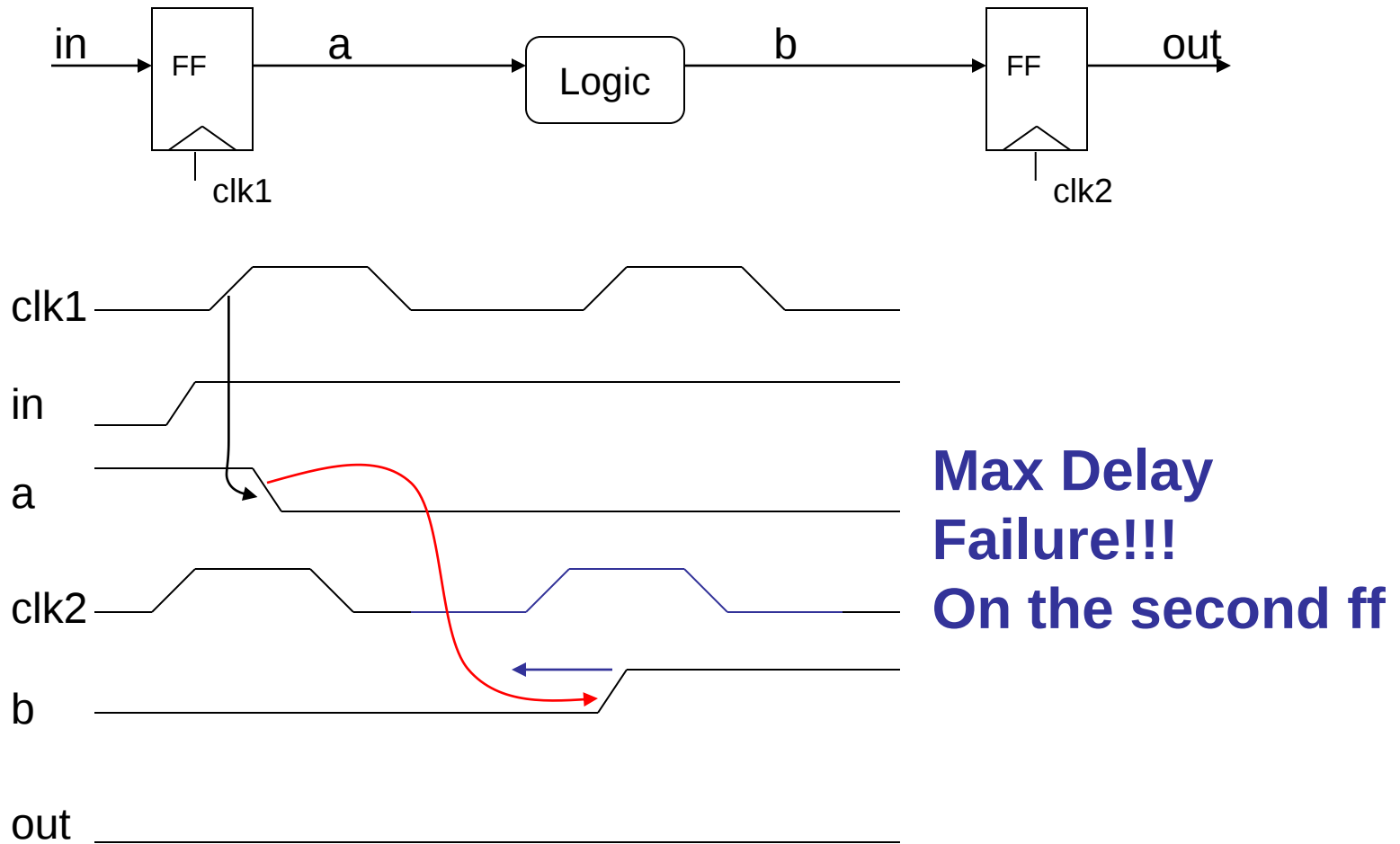
Max/Min Delay Margins

Example 3



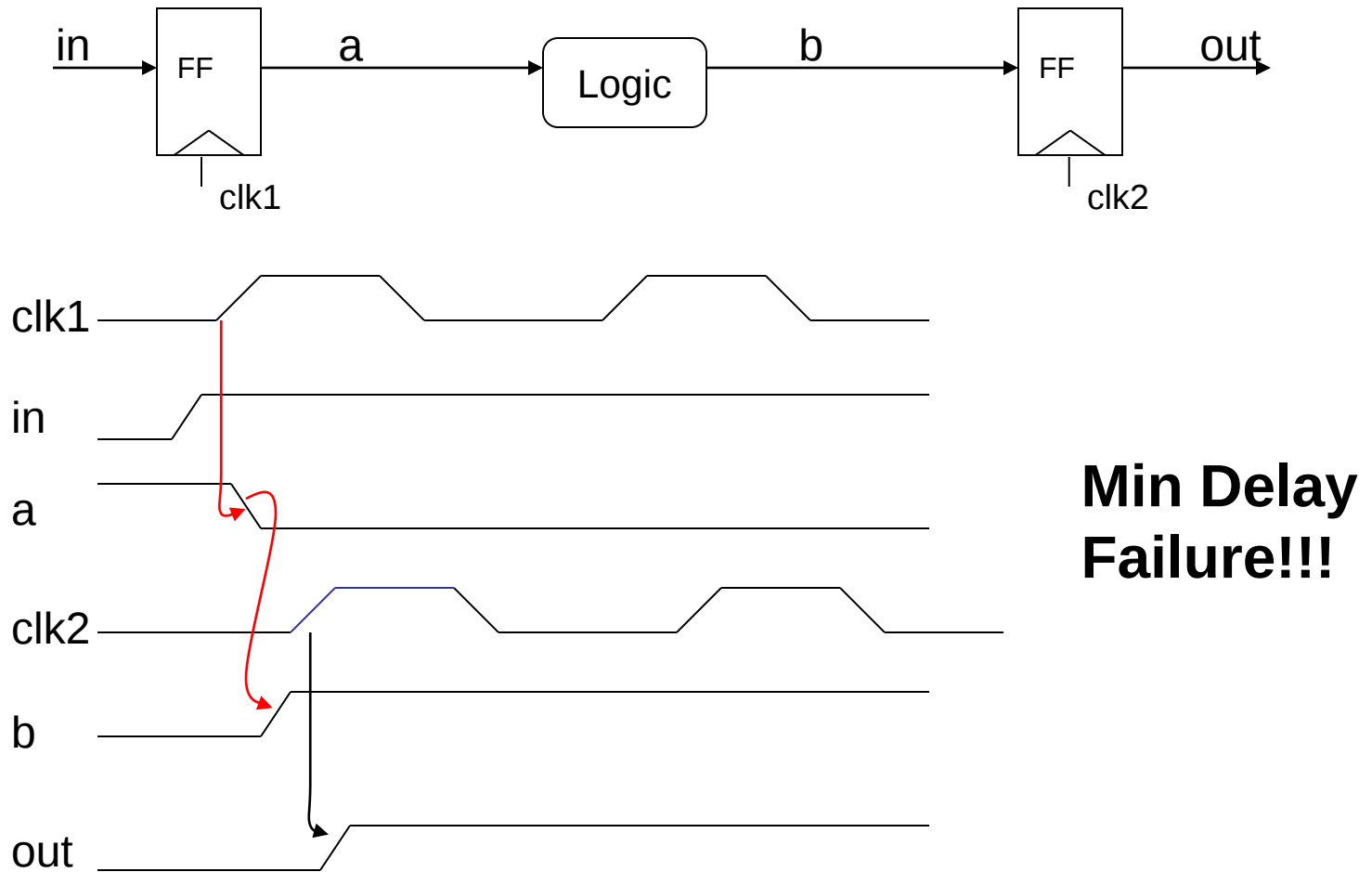
Max/Min Delay Margins

Example 3 (cont.)



Max/Min Delay Margins

Example 3 (cont.)



Sources of Clock Uncertainty

- Skew – Unintentional mismatch in the arrival of two local clock edges produced by the same global clock edge.

Main causes are:

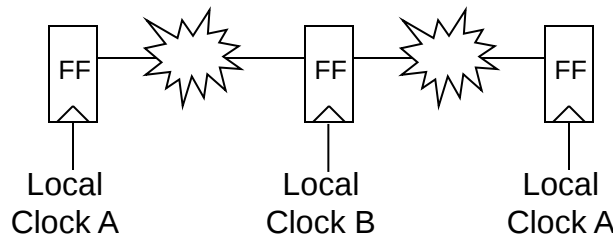
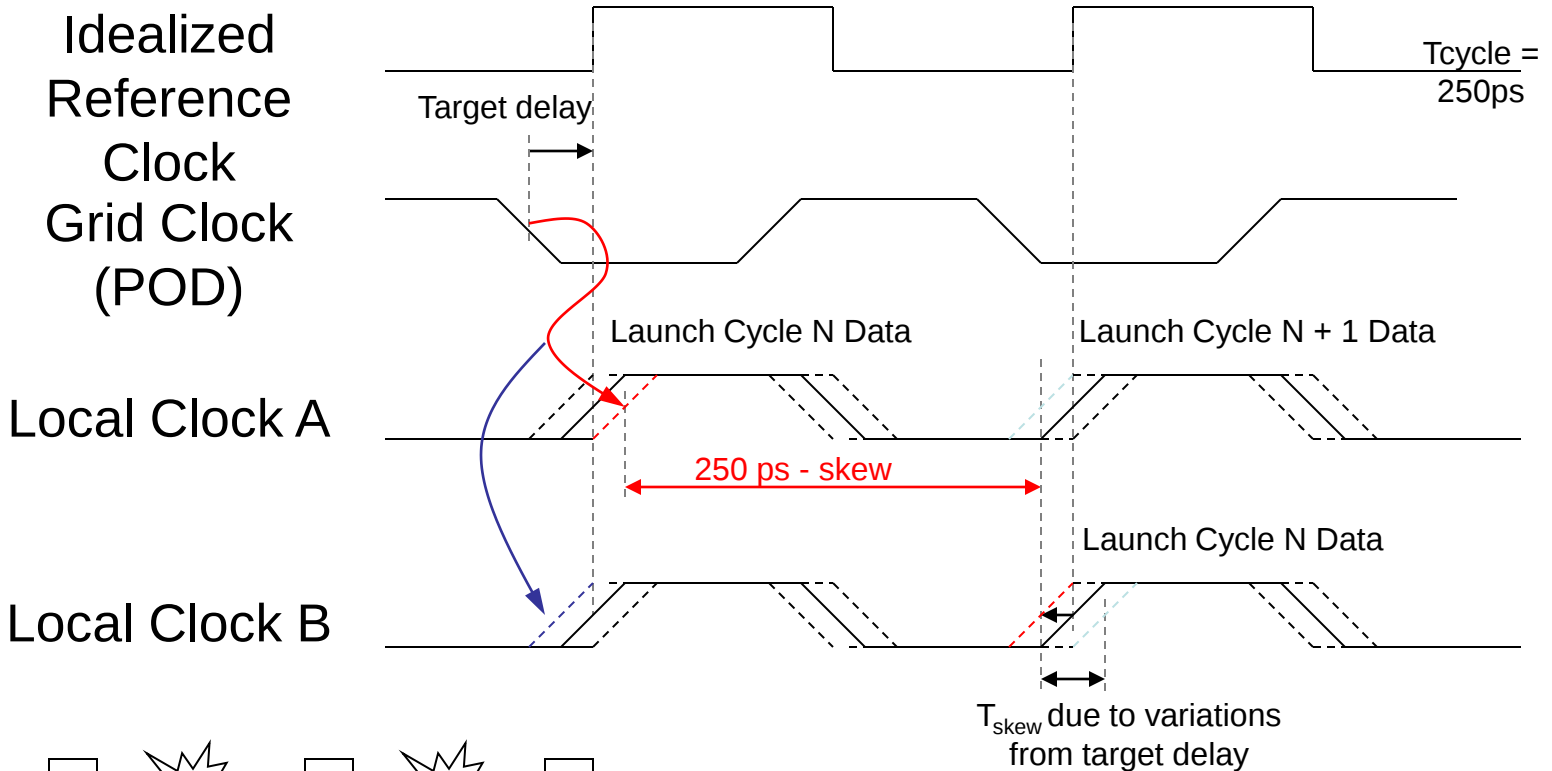
- RC mismatch in the clock network
 - Cross-coupling in the clock network
 - (P)rocess, (V)oltage, (T)emperature mismatch in the clock distribution network
- Jitter* – Variation in the cycle time of the global clock.

Main causes are:

- PLL Jitter
- Coupling in the global clock distribution network

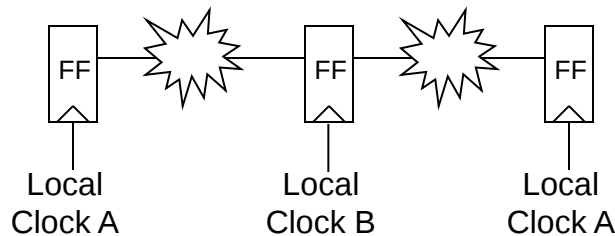
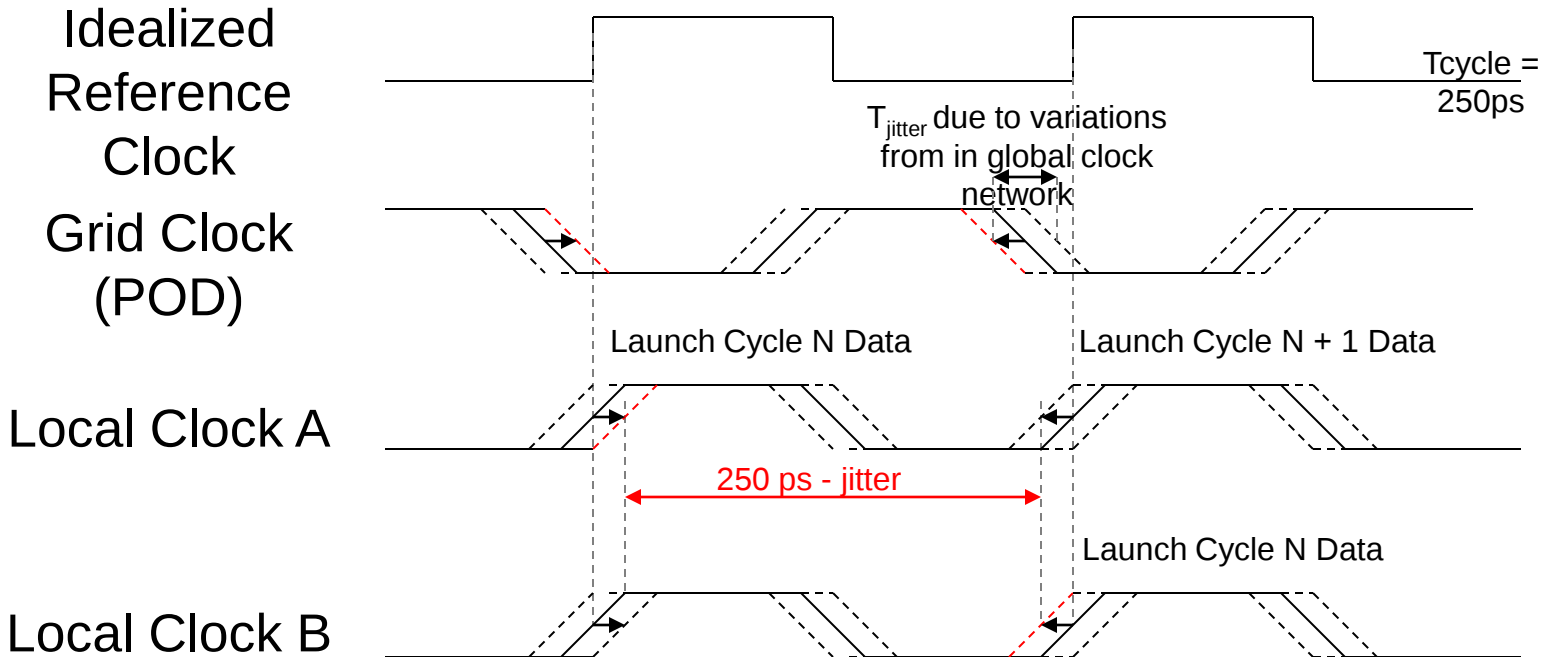
* My definition of jitter might vary from what other people call jitter. For our conversation we'll use my definition.

Clock Skew Illustrated



- Variations in distribution cause unexpected delay differences from the POD (point of divergence) to the local clock @ the sequential
- Skew impacts both setup and hold checks

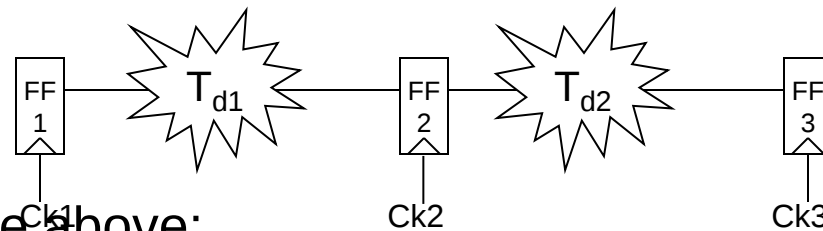
Clock Jitter Illustrated



- Variations in PLL & global distribution network cause jitter
- Jitter only impacts setup checks

How Skew & Jitter Fit in the Setup Margin Calculation

- Without Skew/Jitter
 - $\text{Margin}_{\text{setup}} = \text{Required}_{\text{abs}} - \text{Valid}_{\text{abs}}$
- With Skew/Jitter
 - $\text{Margin}_{\text{setup}} = \text{Required}_{\text{abs}} - \text{Valid}_{\text{abs}} - \text{MaxSkew}$



- For the example above:

$$\text{Margin}_{\text{setup}} = T_{\text{cycle}} - \text{Clk} \rightarrow Q_{\text{FF1}} - T_{d1} - \text{Setup}_{\text{FF2}} - (\text{MaxSkew}_{\text{ck1-ck2}} + \text{Jitter})$$

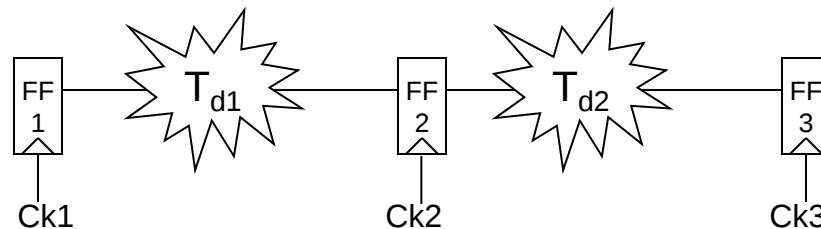
$$\text{Margin}_{\text{setup}} = T_{\text{cycle}} - \text{Clk} \rightarrow Q_{\text{FF2}} - T_{d2} - \text{Setup}_{\text{FF3}} - (\text{MaxSkew}_{\text{ck2-ck3}} + \text{Jitter})$$

Each cycle pays setup, skew & jitter

Setup Margin Examples

$$\begin{aligned}\text{Margin}_{\text{setup FF1} \rightarrow \text{FF2}} &= T_{\text{cycle}} - \text{Clk} \rightarrow Q_{\text{FF1}} - T_{d1} - \text{Setup}_{\text{FF2}} - (\text{MaxSkew}_{\text{FF1-FF2}} + \text{Jitter}) \\ &= 250 - (23 + 185 + 37 + 15 + 16) \\ &= 250 - 276 = -26\text{ps}\end{aligned}$$

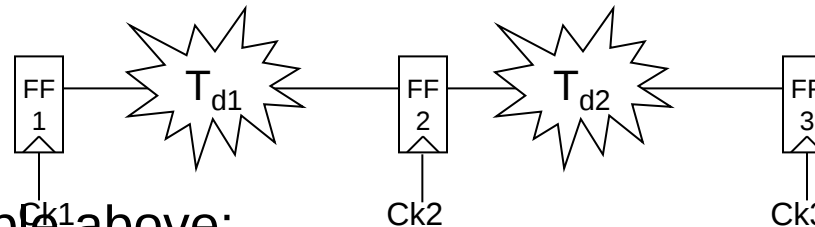
$$\begin{aligned}\text{Margin}_{\text{setup FF2} \rightarrow \text{FF3}} &= T_{\text{cycle}} - \text{Clk} \rightarrow Q_{\text{FF2}} - T_{d2} - \text{Setup}_{\text{FF3}} - (\text{MaxSkew}_{\text{FF2-FF3}} + \text{Jitter}) \\ &= 250 - (25 + 155 + 41 + 10 + 16) \\ &= 250 - 247 = 3\text{ps}\end{aligned}$$



T _{d1}	0.185		T _{d2}	0.155
Clk → Q _{FF1}	0.023		Setup _{FF1}	0.038
Clk → Q _{FF2}	0.025		Setup _{FF2}	0.037
Clk → Q _{FF3}	0.024		Setup _{FF3}	0.041
MaxSkew _{Ck1-Ck2}	0.015		MaxSkew _{Ck2-Ck3}	0.010
T _{cycle}	0.250		Jitter	0.016

How Skew Fits in the Hold Margin Calculation

- Without Skew
 - $\text{Margin}_{\text{hold}} = \text{Valid}_{\text{abs}} - \text{Required}_{\text{abs}}$
- With Skew
 - $\text{Margin}_{\text{hold}} = \text{Valid}_{\text{abs}} - \text{Required}_{\text{abs}} - \text{MinSkew}$



- For the example above:

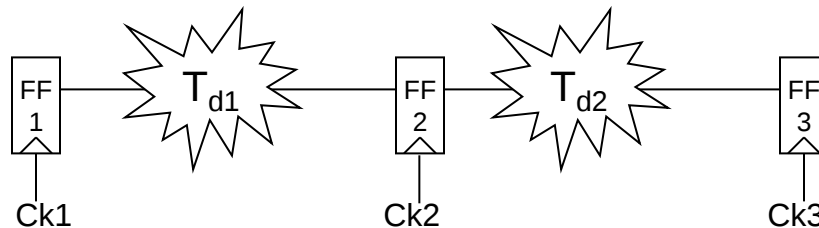
$$\text{Margin}_{\text{hold}} = \text{Clk} \rightarrow Q_{\text{FF1}} + T_{d1} - \text{Hold}_{\text{FF2}} - \text{MinSkew}_{\text{ck1-ck2}}$$

$$\text{Margin}_{\text{hold}} = \text{Clk} \rightarrow Q_{\text{FF2}} + T_{d2} - \text{Hold}_{\text{FF3}} - \text{MinSkew}_{\text{ck2-ck3}}$$

Hold Margin Examples

$$\begin{aligned}\text{Margin}_{\text{hold FF1} \rightarrow \text{FF2}} &= \text{Clk} \rightarrow Q_{???} + T_{d??} - \text{Hold}_{???} - \text{MinSkew}_{???} \\ &= 18 + 55 - 25 - 45 \\ &= 3\end{aligned}$$

$$\begin{aligned}\text{Margin}_{\text{hold FF2} \rightarrow \text{FF3}} &= \text{Clk} \rightarrow Q_{???} + T_{d??} - \text{Hold}_{???} - \text{MinSkew}_{???} \\ &= 20 + 40 - 28 - 40 \\ &= -8\end{aligned}$$



T_{d1}	0.055		T_{d2}	0.040
$\text{Clk} \rightarrow Q_{\text{FF1}}$	0.018		Hold_{FF1}	0.026
$\text{Clk} \rightarrow Q_{\text{FF2}}$	0.020		Hold_{FF2}	0.025
$\text{Clk} \rightarrow Q_{\text{FF3}}$	0.022		Hold_{FF3}	0.028
$\text{MinSkew}_{\text{Ck1-Ck2}}$	0.045		$\text{MinSkew}_{\text{Ck2-Ck3}}$	0.040

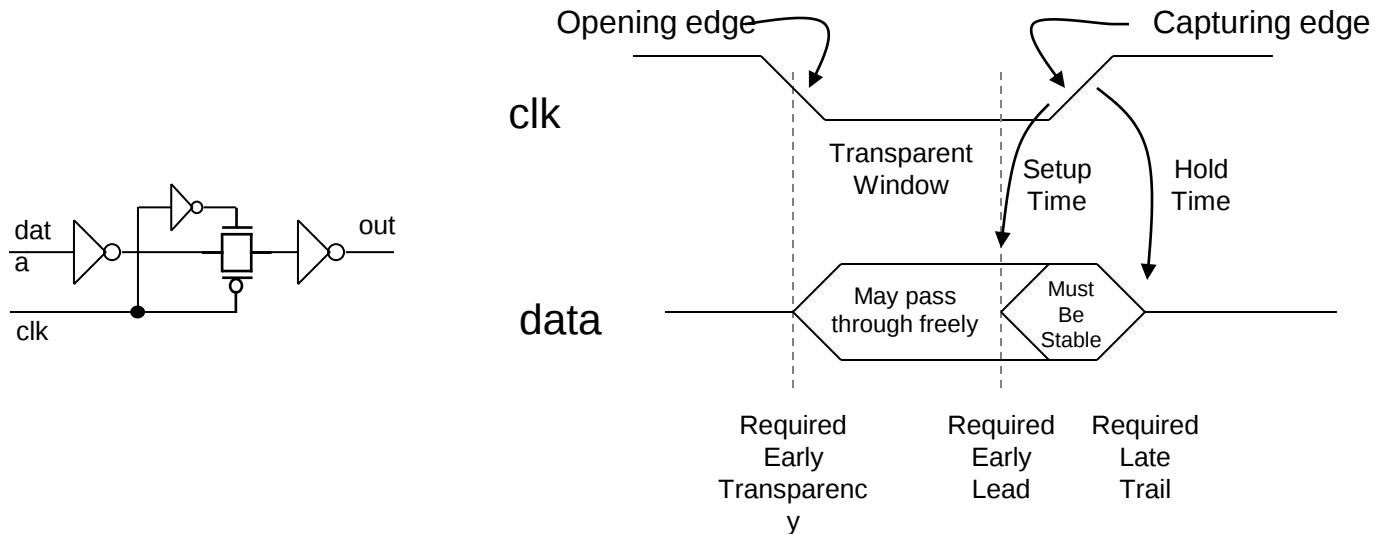
Clock Uncertainty Key Notes

- Always assume worst case skew
 - Launching clocks are slow in setup analysis, capturing clocks are fast
 - Launching clocks are fast in hold analysis, capturing clocks are slow
- Skew can be minimized by the designer, jitter cannot
 - Since skew is related to local clock variation, you can control this by keeping the POD as close as possible
 - Jitter is mainly caused by the PLL & the global grid, you can't impact this as a fub designer
- Max skew and Min skew are different
 - Min skew is bigger because it must encompass more of the natural variation because a hold failure is fatal
 - Max skew is smaller because a setup failure only impacts binsplit

What is a Transparency Window?

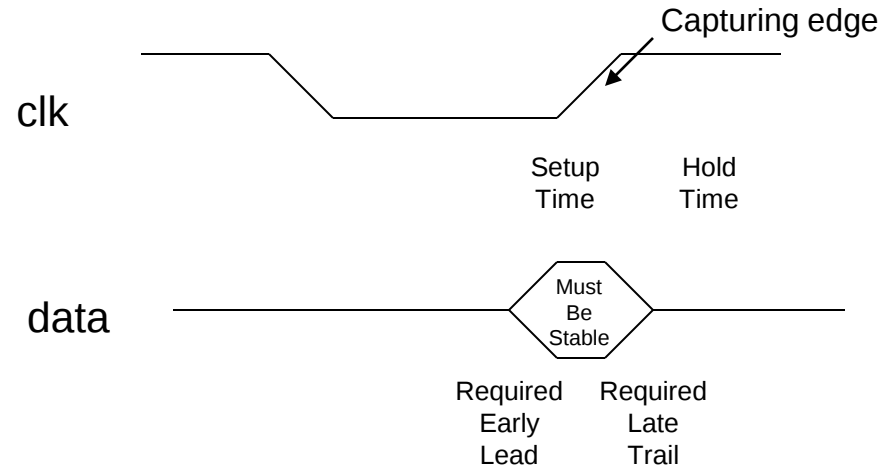
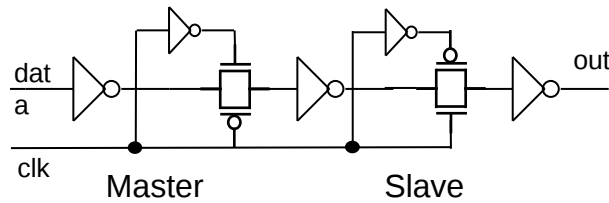


- Transparency Window
 - The time when a data input flows unblocked through a sequential



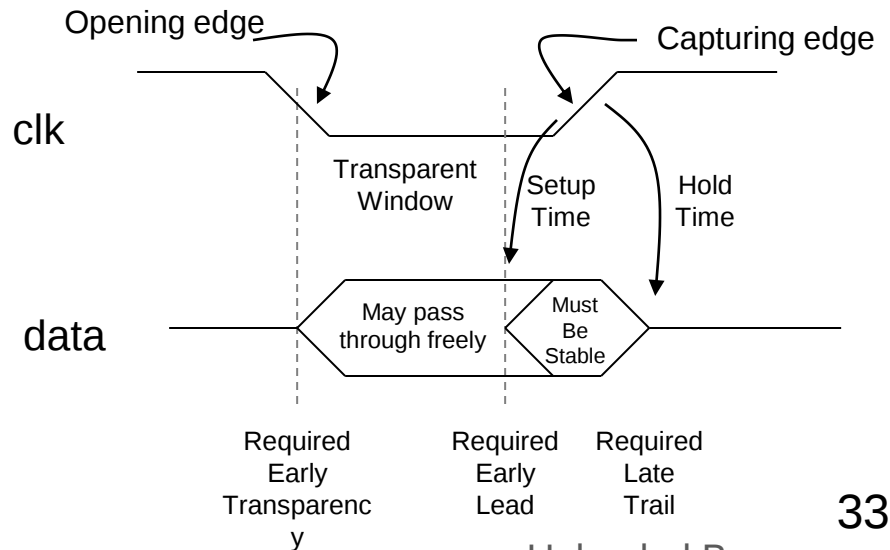
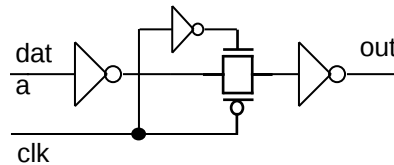
- Inputs may switch anytime in this window, so long as they are valid before the capturing edge of the clock, and remain valid until the hold time (+ skew)

Flop vs. Latch Timing Requirements



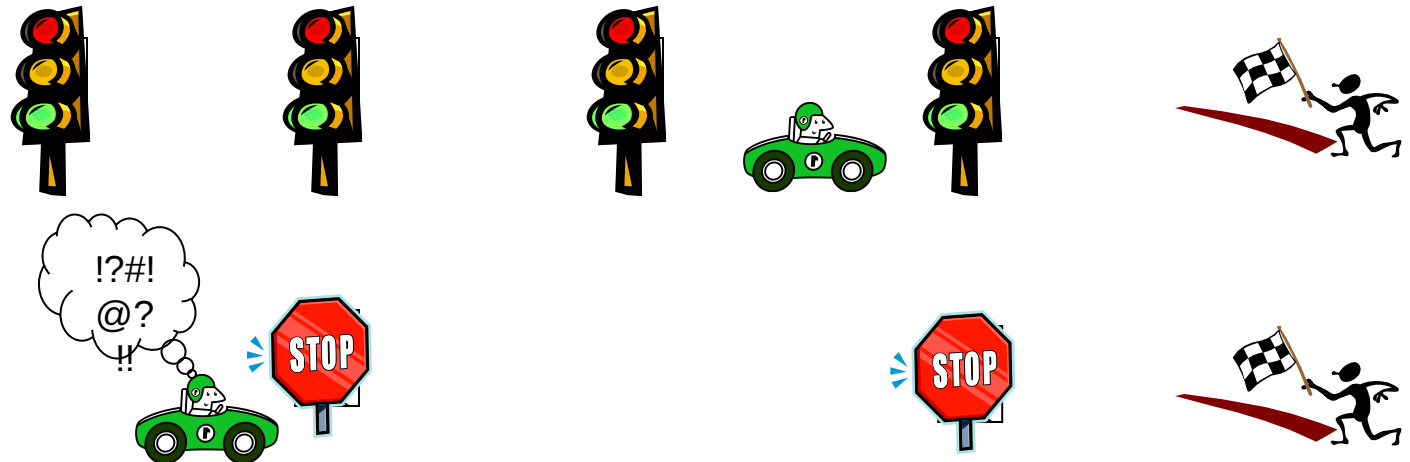
FLOP Timings

Latch Timings

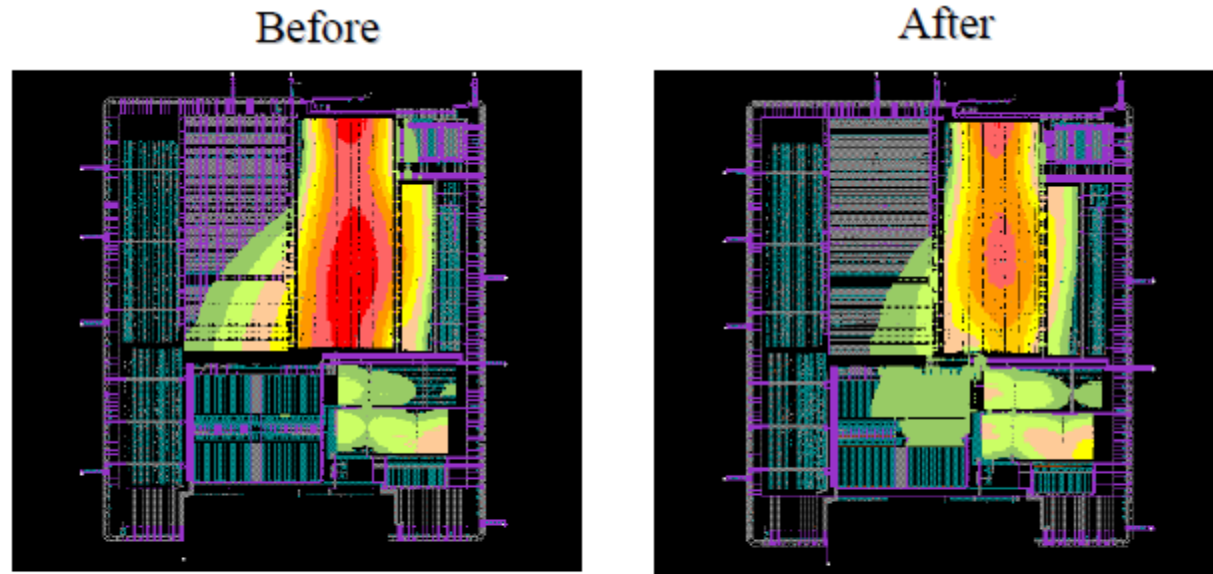


A Real World Example of Transparency

- Level sensitive latches allow transparency
 - When the latch opens, data may arrive and pass freely from input to output *without waiting for clock*
 - Data only needs to be stable **when the latch closes**
- Think of a latch like a street light and a flop light a stop sign. Catch the green and it's a free ride



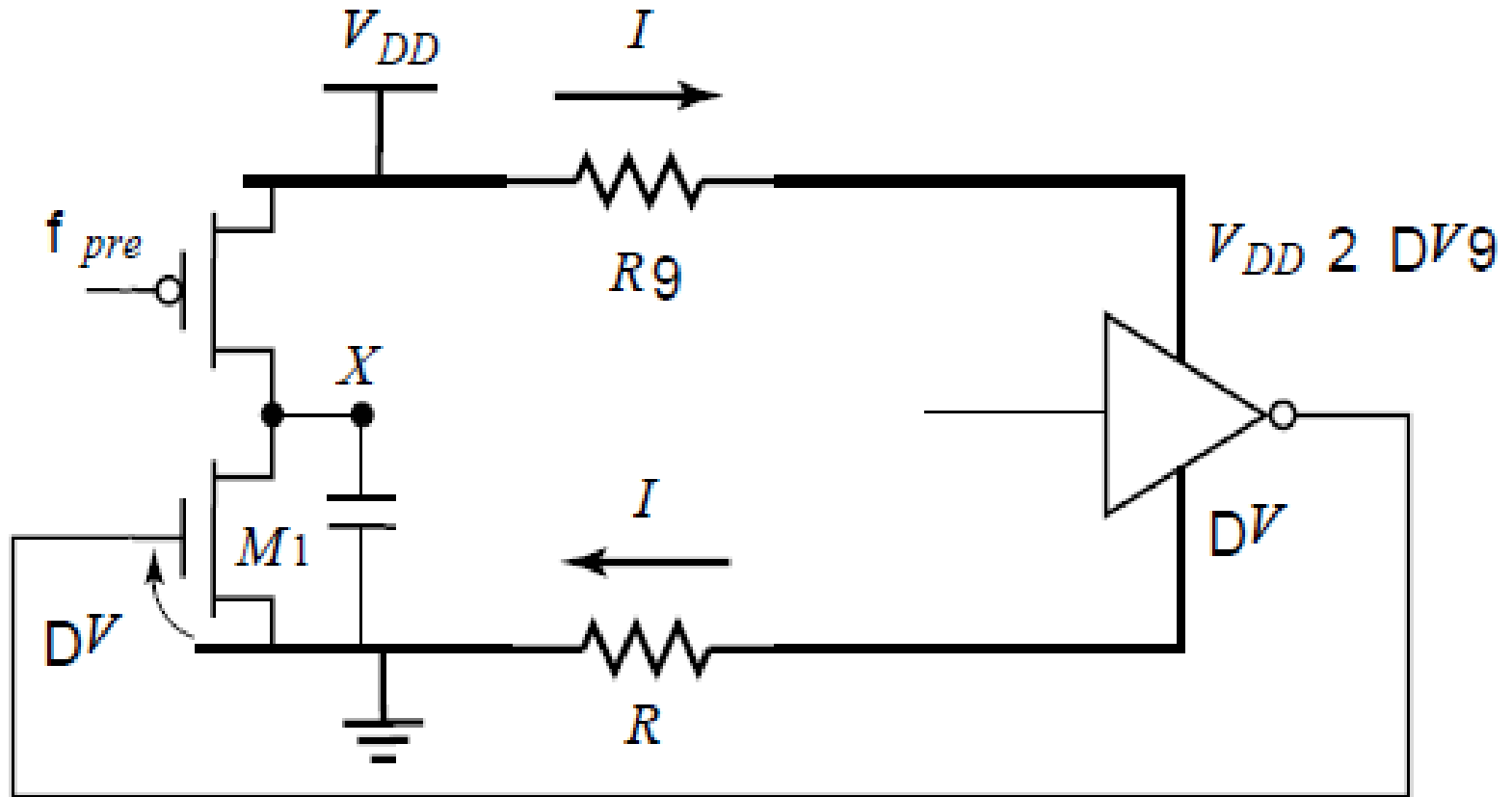
Resistance and the Power Distribution Problem



- Requires fast and accurate peak current prediction
- Heavily influenced by packaging technology

<https://www.apache-da.com/products/redhawk>

RI Introduced Noise



Power Distribution

- Low-level distribution is in Metal 1
- Power has to be 'strapped' in higher layers of metal.
- The spacing is set by IR drop, electromigration, inductive effects
- Always use multiple contacts on straps

Where Does Power Go in CMOS?

- **Dynamic Power Consumption**

Charging and Discharging Capacitors

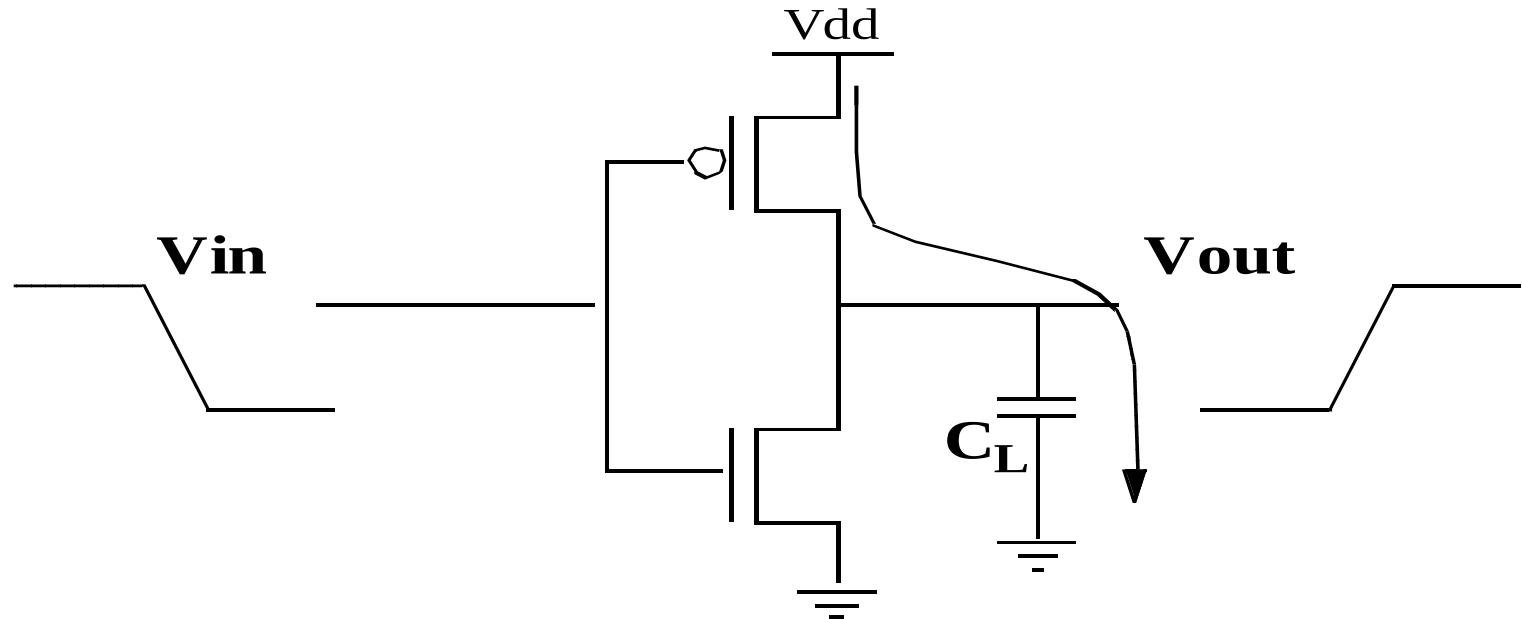
- **Short Circuit Currents**

Short Circuit Path between Supply Rails during Switching

- **Leakage**

Leaking diodes and transistors

Dynamic Power Consumption



$$\text{Energy/transition} = C_L * V_{dd}^2$$

$$\text{Power} = \text{Energy/transition} * f = C_L * V_{dd}^2 * f$$

- Not a function of transistor sizes!
- Need to reduce C_L , V_{dd} , and f to reduce power.

Dynamic Power Consumption - Revisited

Power = Energy/transition * transition rate

$$= C_L * V_{dd}^2 * f_{0 \rightarrow 1}$$

$$= C_L * V_{dd}^2 * P_{0 \rightarrow 1} * f$$

$$= C_{EFF} * V_{dd}^2 * f$$

Power Dissipation is Data Dependent
Function of *Switching Activity*

$$C_{EFF} = \text{Effective Capacitance} = C_L * P_{0 \rightarrow 1}$$

$$P_{0 \rightarrow 1} = P_0 \cdot P_1$$

Factors Affecting Transition Activity

- “Static” component (does not account for timing)
 - ⇒ Type of Logic Function (NOR vs. XOR)
 - ⇒ Type of Logic Style (Static vs. Dynamic)
 - ⇒ Signal Statistics
 - ⇒ Inter-signal Correlations
- “Dynamic” or timing dependent component
 - ⇒ Circuit Topology
 - ⇒ Signal Statistics and Correlations

Type of Logic Function: NOR vs. XOR

Example: Static 2-input NOR Gate

A	B	Out
0	0	1
0	1	0
1	0	0
1	1	0

Assume signal probabilities

$$p_{A=1} = 1/2$$

$$p_{B=1} = 1/2$$

Then transition probability

$$p_{0 \rightarrow 1} = p_{Out=0} \times p_{Out=1}$$

$$= 3/4 \times 1/4 = 3/16$$

If inputs switch every cycle

$$\alpha_{0 \rightarrow 1} = 3/16$$

Example: Static 2-input XOR Gate

A	B	Out
0	0	0
0	1	1
1	0	1
1	1	0

Assume signal probabilities

$$p_{A=1} = 1/2$$

$$p_{B=1} = 1/2$$

Then transition probability

$$p_{0 \rightarrow 1} = p_{Out=0} \times p_{Out=1}$$

$$= 1/2 \times 1/2 = 1/4$$

If inputs switch in every cycle

$$\alpha_{0 \rightarrow 1} = 1/4$$

Power Consumption is Data Dependent

Example: Static 2 Input NOR Gate

A	B	Out
0	0	1
0	1	0
1	0	0
1	1	0

Truth Table of 2 input NOR gate

Assume:

$$P(A=1) = 1/2$$

$$P(B=1) = 1/2$$

Then:

$$P(\text{Out}=1) = 1/4$$

$$P(0 \rightarrow 1) = P(\text{Out}=0) \cdot P(\text{Out}=1)$$

$$= 3/4 \times 1/4 = 3/16$$

$$P_{0 \rightarrow 1} = P_0 \cdot P_1$$

$$C_{\text{EFF}} = 3/16 * C_L$$

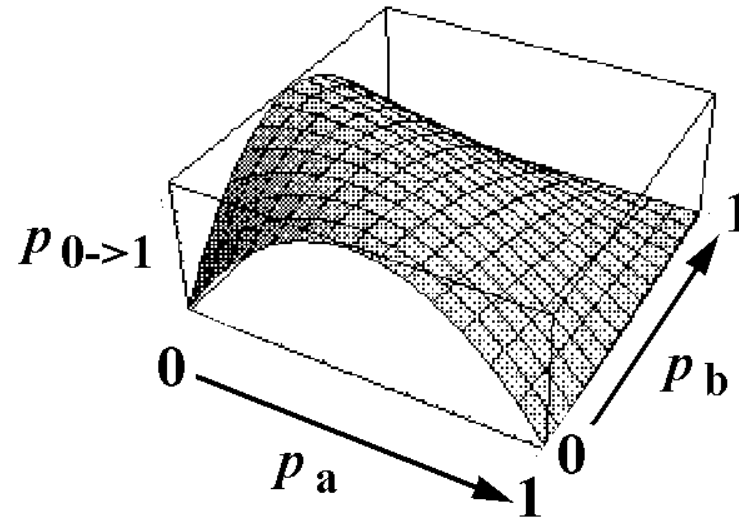
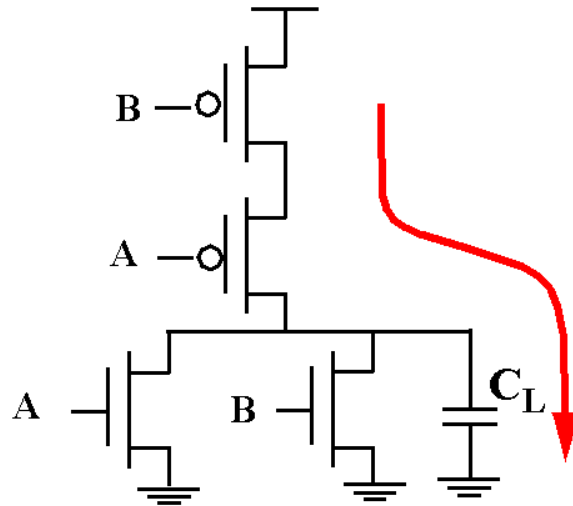
Transition Probabilities for Basic Gates

	$P_{0 \rightarrow 1}$
AND	$(1 - P_A P_B) P_A P_B$
OR	$(1 - P_A)(1 - P_B)(1 - (1 - P_A)(1 - P_B))$
EXOR	$(1 - (P_A + P_B - 2P_A P_B))(P_A + P_B - 2P_A P_B)$

Switching Activity for Static CMOS

$$P_{0 \rightarrow 1} = P_0 \cdot P_1$$

Transition Probability of 2-input NOR Gate

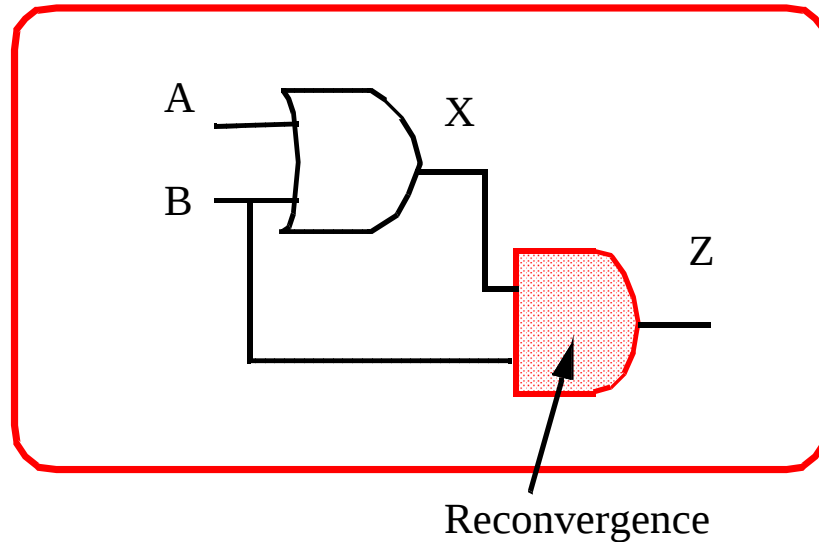


$$p_1 = (1-p_a) (1-p_b)$$

$$p_{0 \rightarrow 1} = p_0 p_1 = (1-(1-p_a) (1-p_b)) (1-p_a) (1-p_b)$$

- $\alpha_{0 \rightarrow 1}$ is a strong function of signal statistics

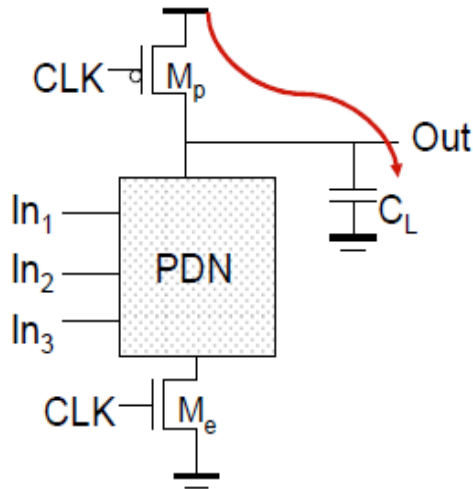
Problem: Reconvergent Fanout



$$P(Z=1) = P(B=1) \cdot P(X=1 \mid B=1)$$

Becomes complex and intractable real fast

How about Dynamic Circuits?



Dynamic Power Consumption is Data Dependent

Dynamic 2-input NOR Gate

A	B	Out
0	0	1
0	1	0
1	0	0
1	1	0

Assume **signal probabilities**

$$P_{A=1} = 1/2$$

$$P_{B=1} = 1/2$$

Then **transition probability**

$$P_{0 \rightarrow 1} = P_{\text{out}=0} \times P_{\text{out}=1}$$

$$= 3/4 \times 1 = 3/4$$

Switching activity always **higher** in dynamic gates!

$$P_{0 \rightarrow 1} = P_{\text{out}=0}$$

Power only dissipated when previous Out = 0

Power is Only Dissipated when Out=0!

$$C_{\text{EFF}} = P(\text{Out}=0) \cdot C_L$$

4-input NAND Gate

Example: Dynamic 2 Input NOR Gate

A	B	Out
0	0	1
0	1	0
1	0	0
1	1	0

Truth Table of 2 input NOR gate

Assume:

$$P(A=1) = 1/2$$

$$P(B=1) = 1/2$$

Then:

$$P(\text{Out}=0) = 3/4$$

$$C_{\text{EFF}} = 3/4 * C_L$$

Switching Activity Is Always Higher in Dynamic Circuits

Transition Probabilities for Dynamic Gates

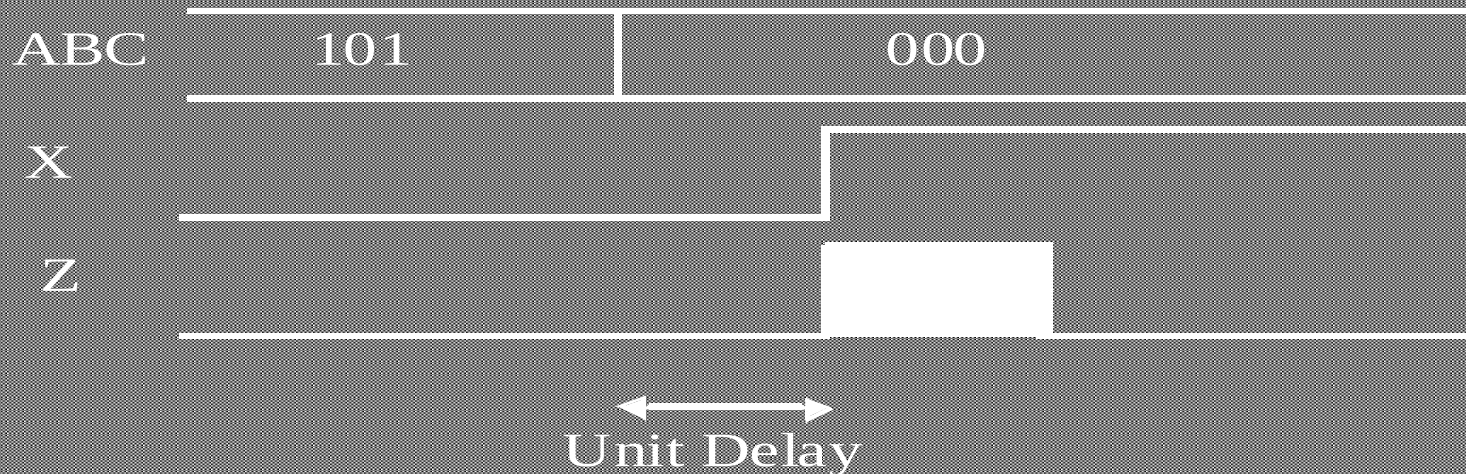
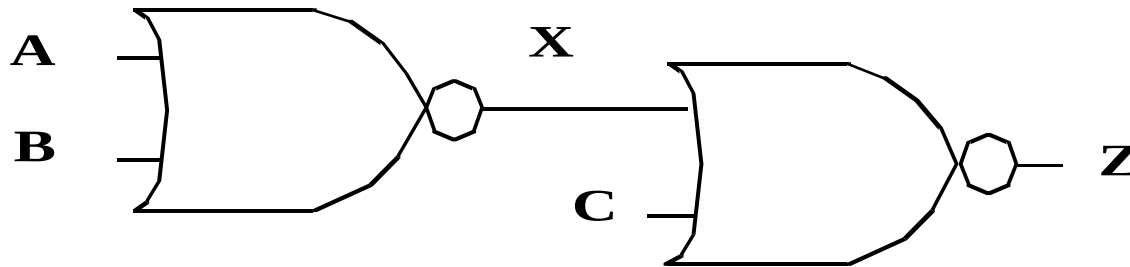
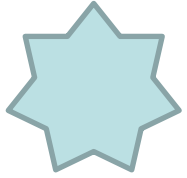
	$P_{0 \rightarrow 1}$
AND	$(1 - P_A P_B)$
OR	$(1 - P_A)(1 - P_B)$
EXOR	$(1 - (P_A + P_B - 2P_A P_B))$

Switching Activity for Precharged Dynamic Gates

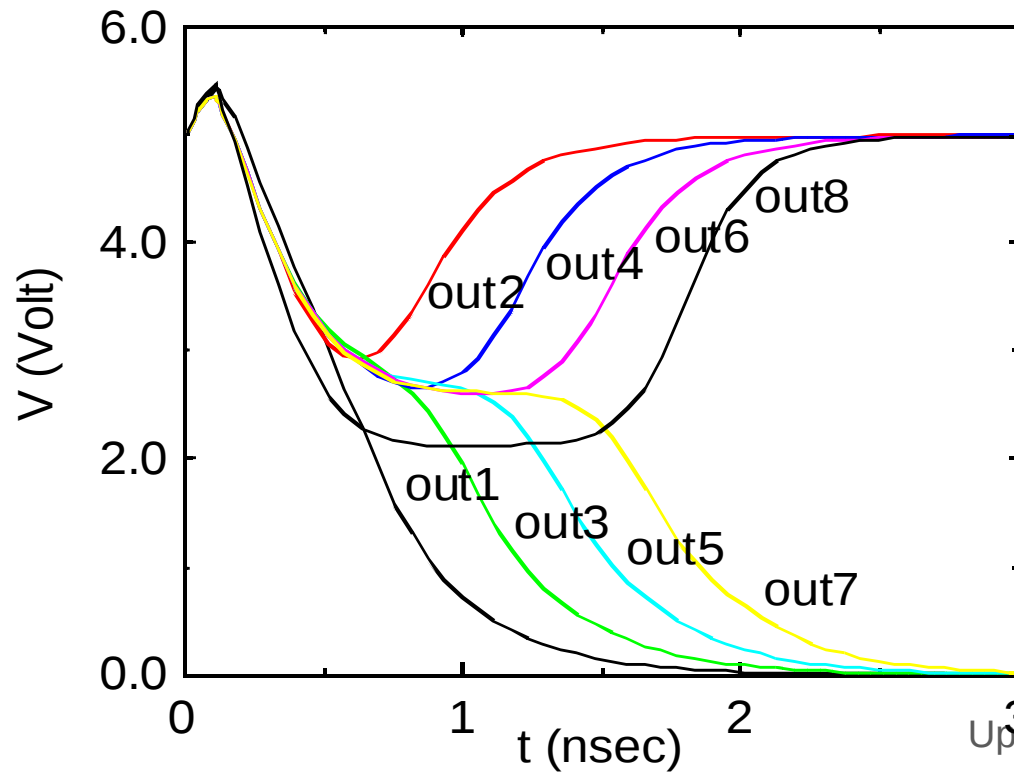
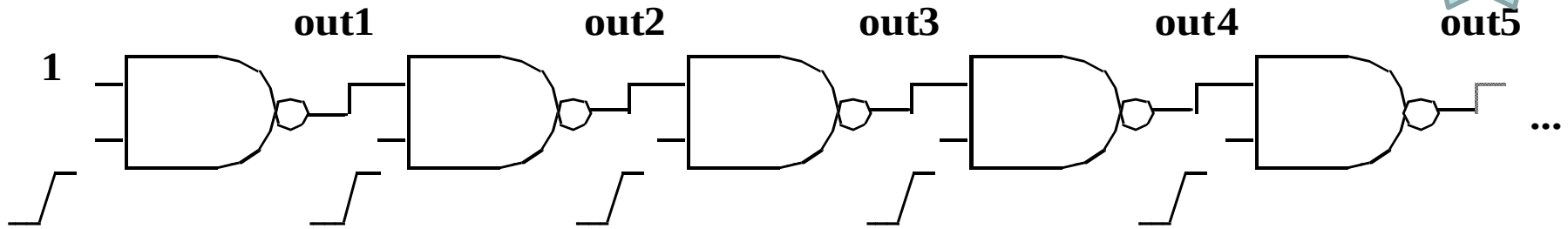
$$P_{0 \rightarrow 1} = P_0$$

Glitching in Static CMOS

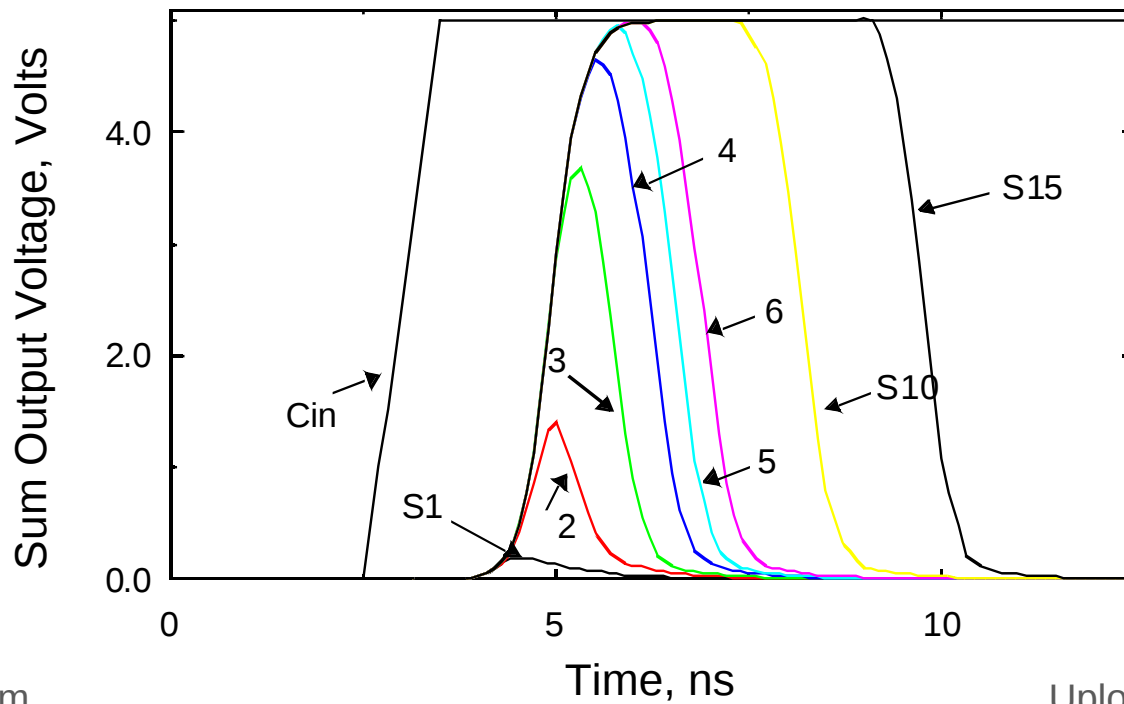
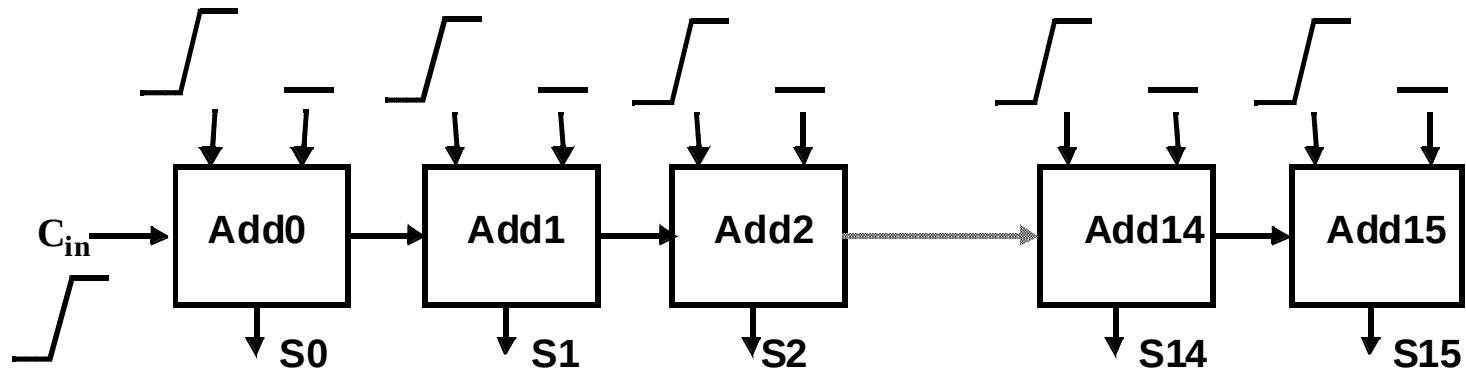
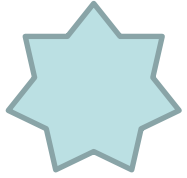
also called: dynamic hazards



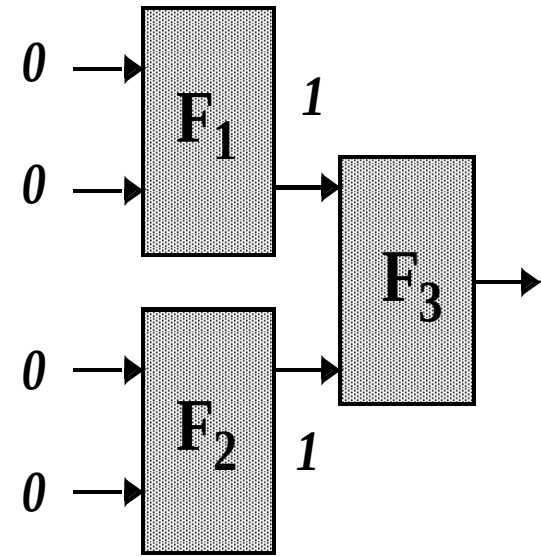
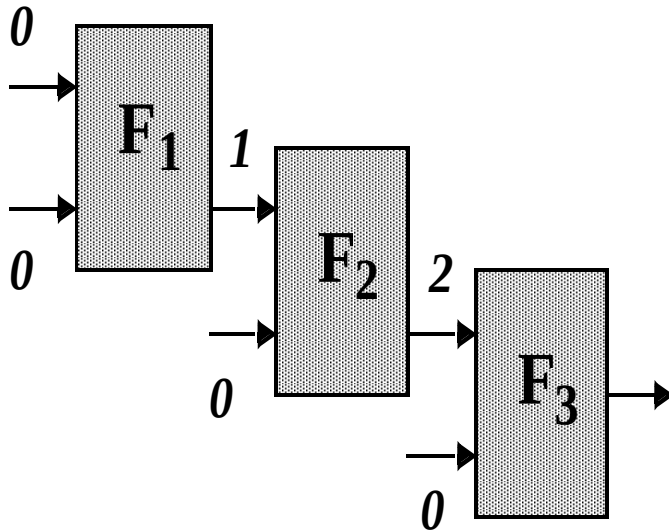
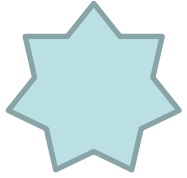
Example 1: Chain of NOR Gates



Example 2: Adder Circuit

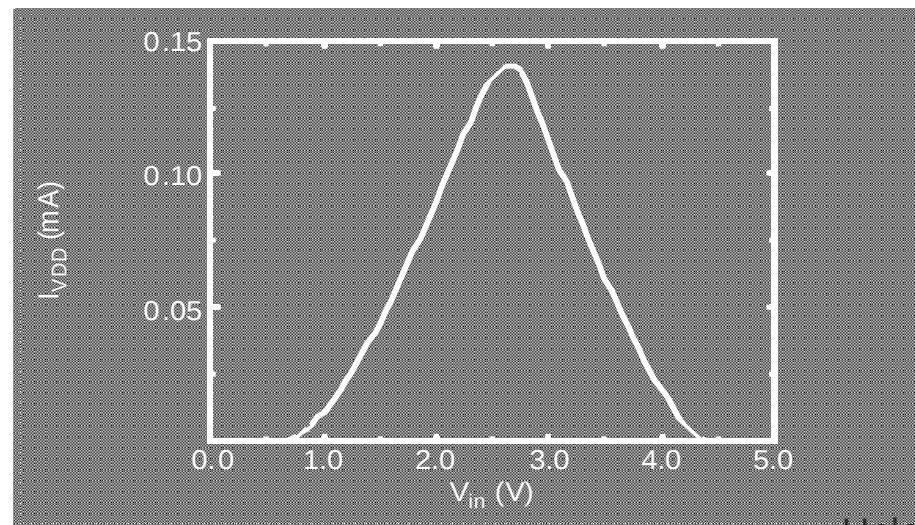
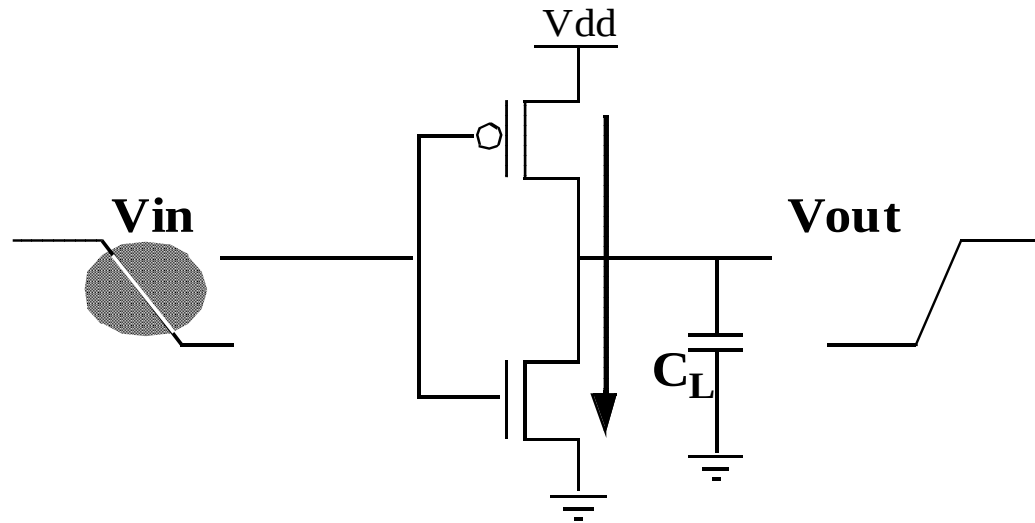
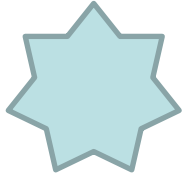


How to Cope with Glitching?

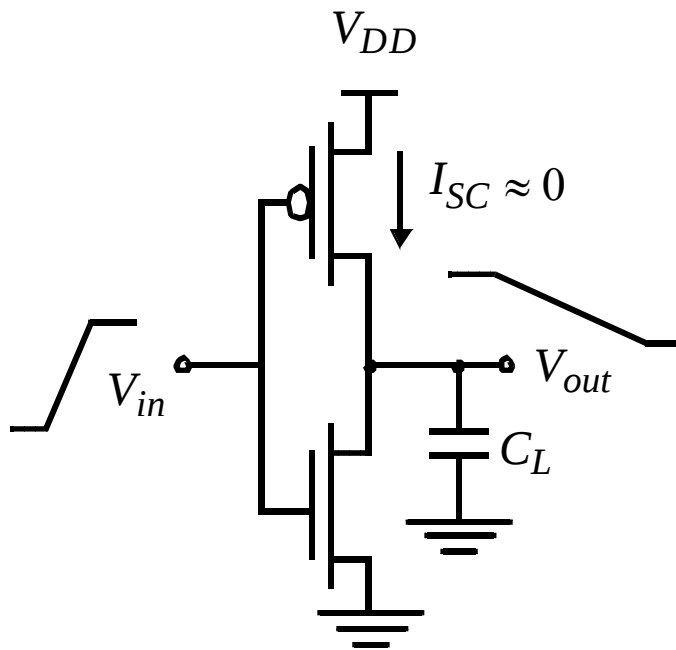


Equalize Lengths of Timing Paths Through Design

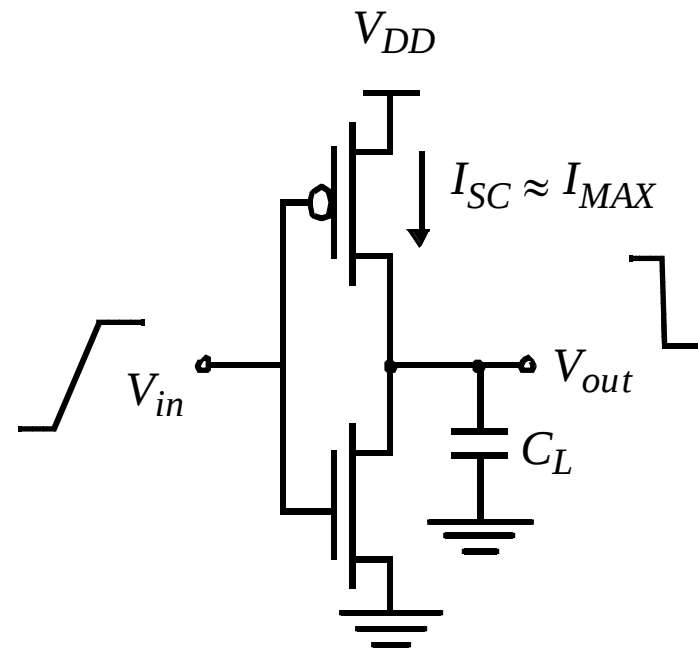
Short Circuit Currents



Impact of rise/fall times on short-circuit currents

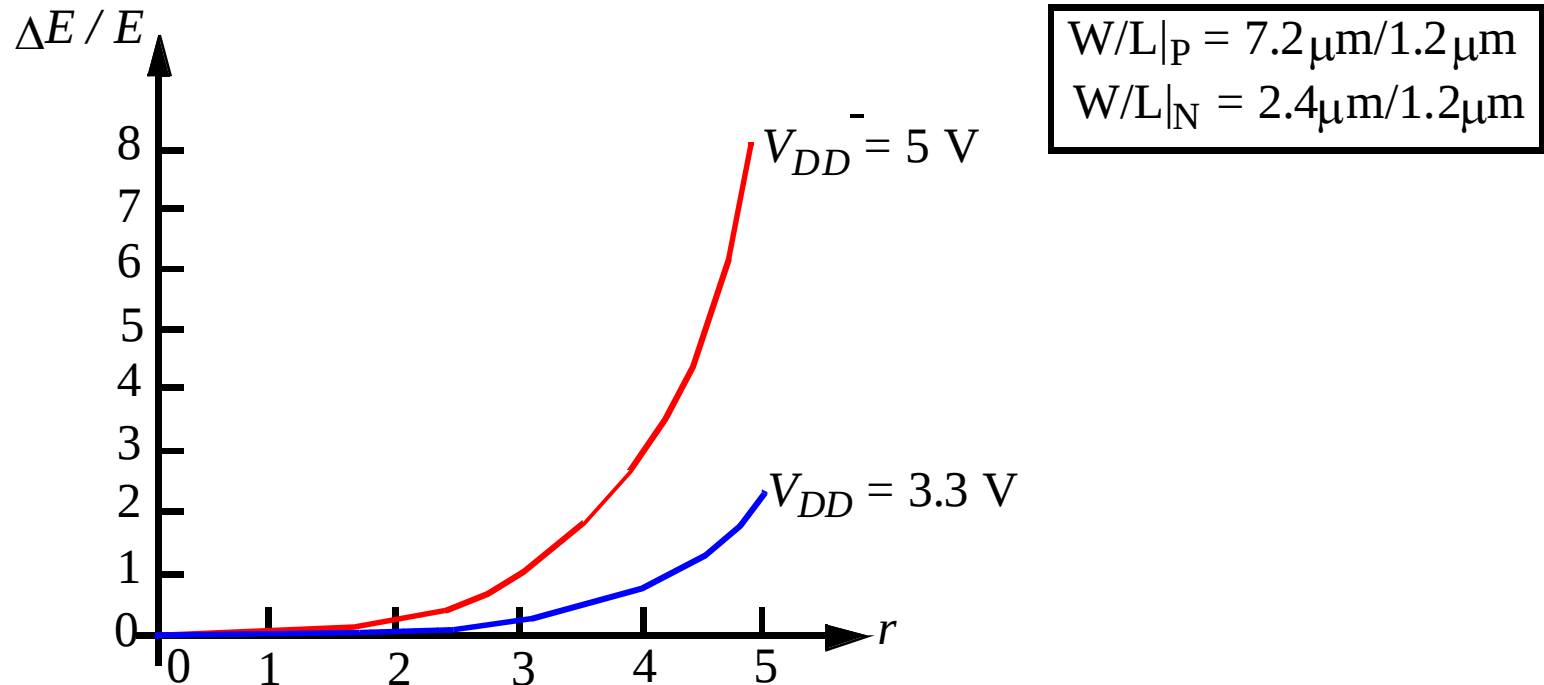


Large capacitive load



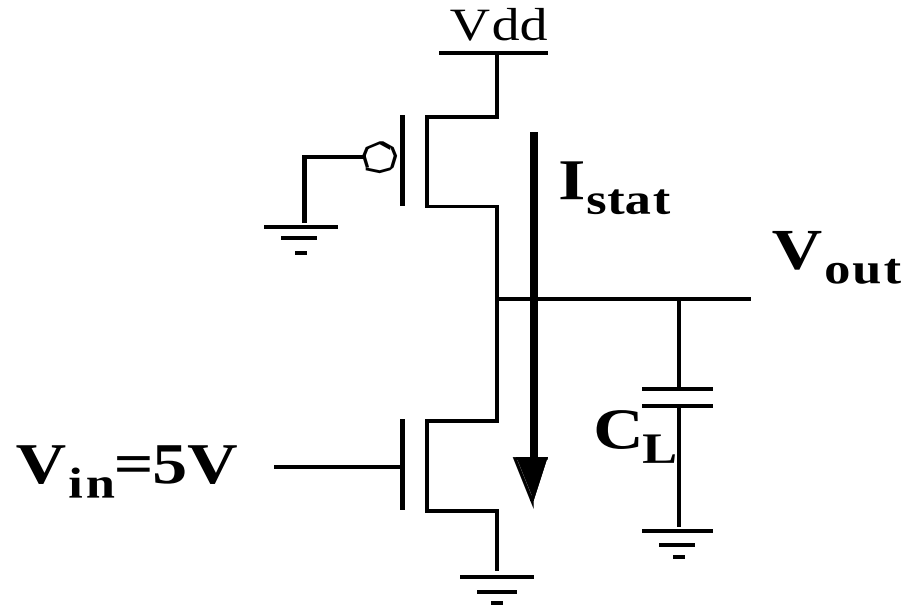
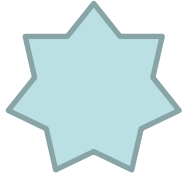
Small capacitive load

Short-circuit energy as a function of slope ratio



The power dissipation due to short circuit currents is minimized by matching the rise/fall times of the input and output signals.

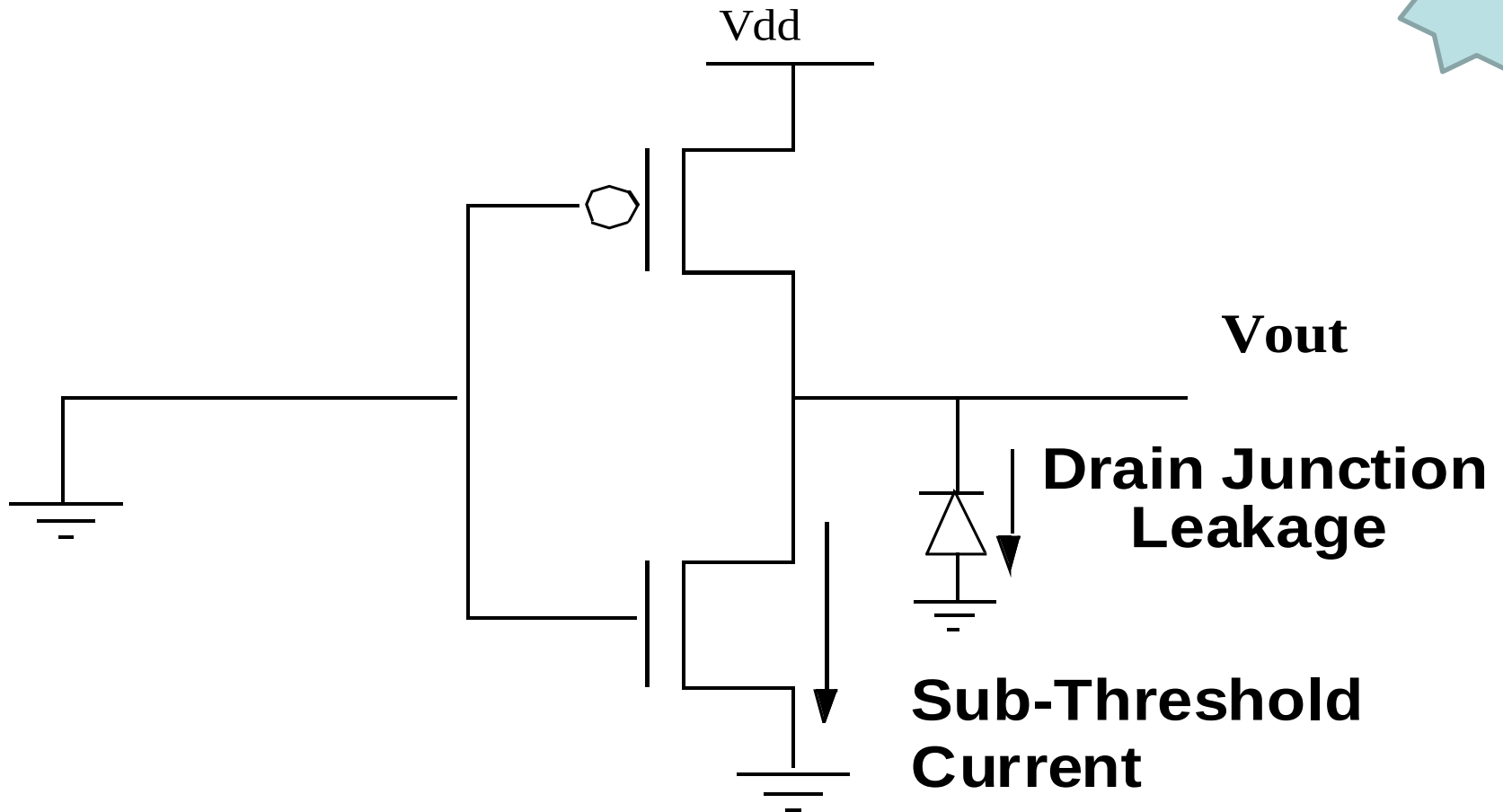
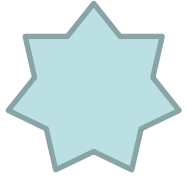
Static Power Consumption



$$P_{stat} = P_{(I_n=1)} \cdot V_{dd} \cdot I_{stat}$$

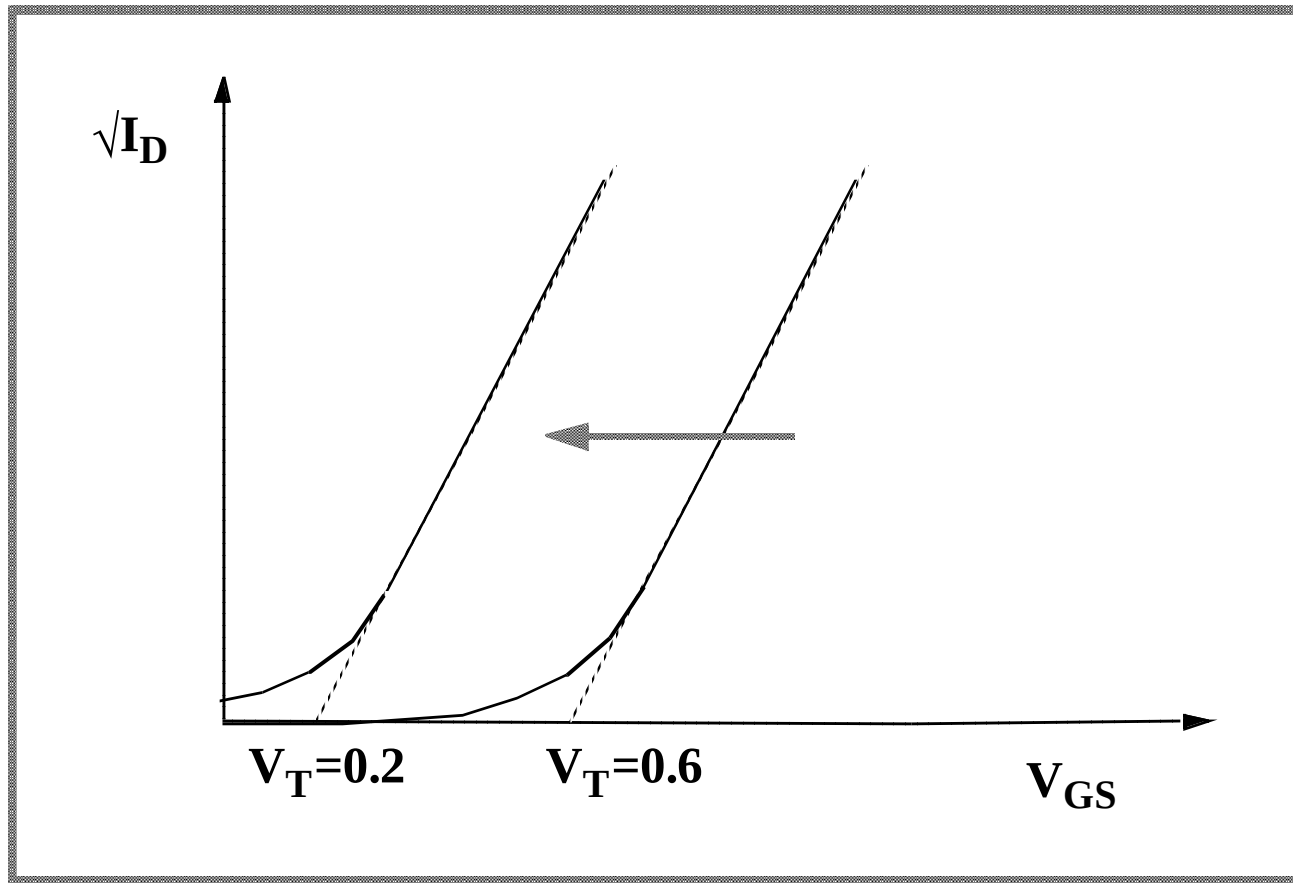
- **Dominates over dynamic consumption**
- **Not a function of switching frequency**

Leakage



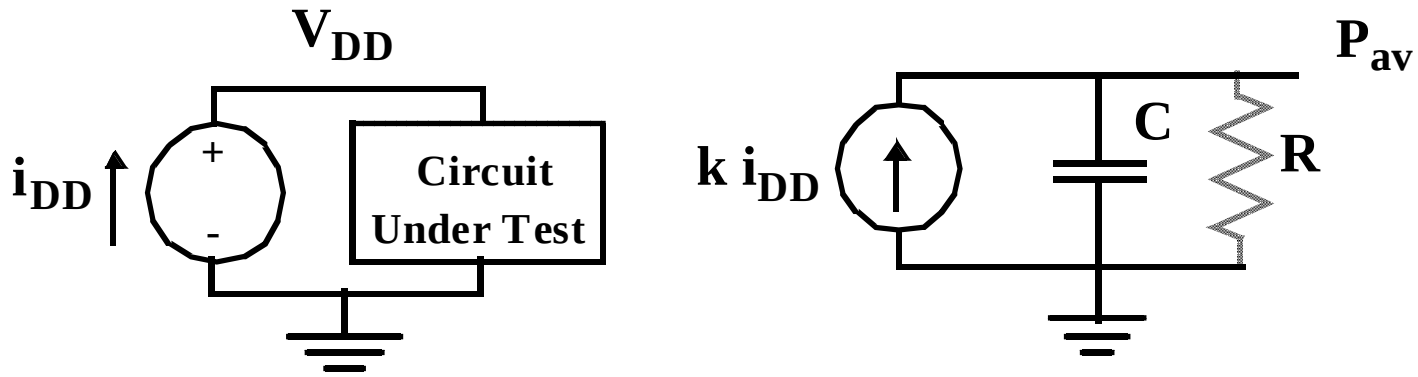
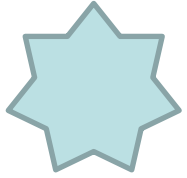
Sub-Threshold Current Dominant Factor 58

Sub-Threshold in MOS



Lower Bound on Threshold to Prevent Leakage

Power Analysis in SPICE



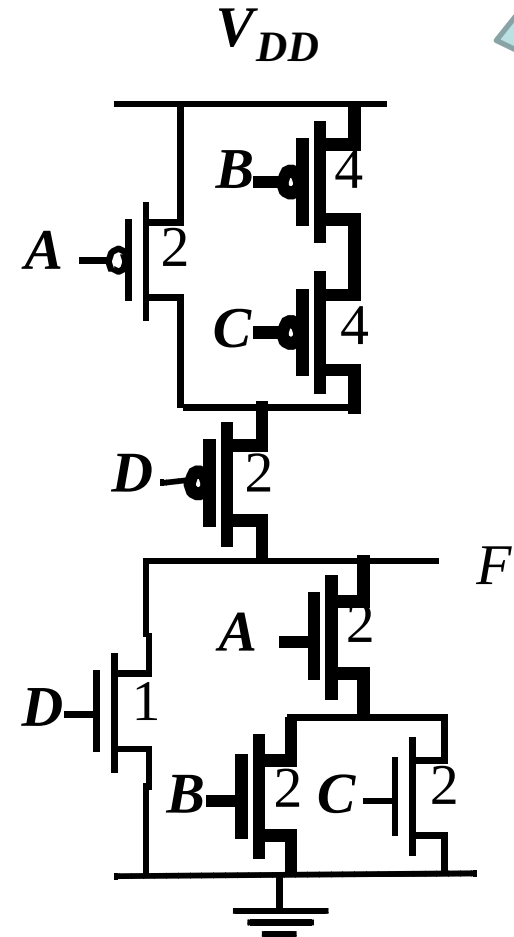
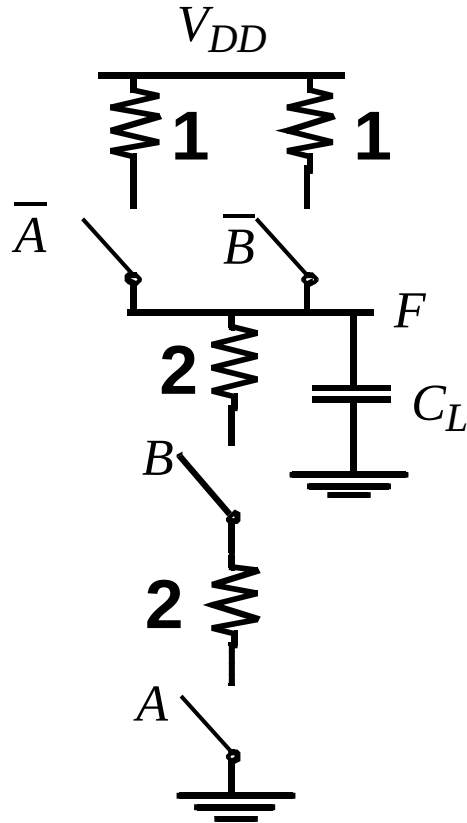
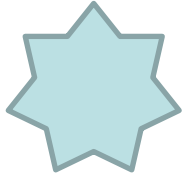
Equivalent Circuit for Measuring Power in SPICE

$$C \frac{dP_{av}}{dt} = k i_{DD}$$

or

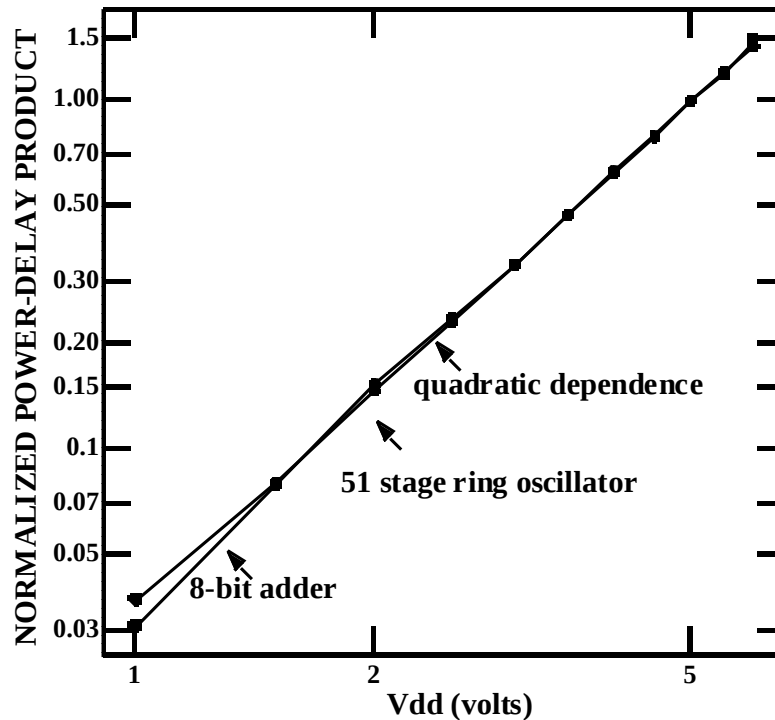
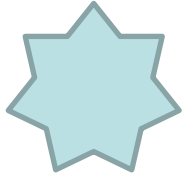
$$P_{av} = \frac{k}{C} \int_0^T i_{DD} dt$$

Design for Worst Case



Here it is assumed that $R_p = R_n$

Reducing V_{dd}



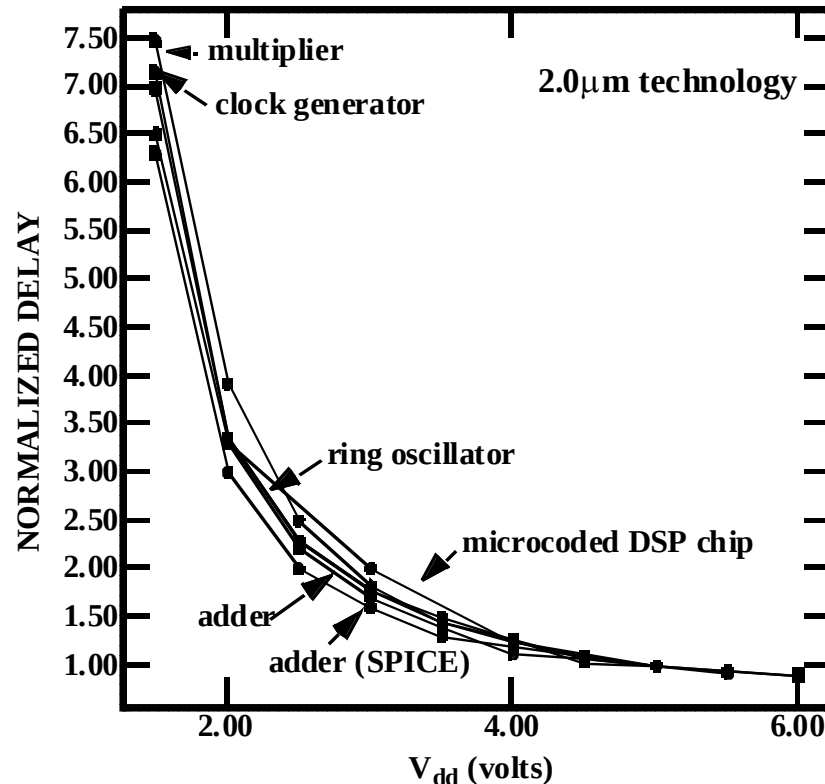
$$P \times t_d = E_t = C_L * V_{dd}^2$$

$$\frac{E_{(V_{dd}=2)}}{E_{(V_{dd}=5)}} = \frac{(C_L) * (2)^2}{(C_L) * (5)^2}$$

$$E_{(V_{dd}=2)} \approx 0.16 E_{(V_{dd}=5)}$$

- Strong function of voltage (V^2 dependence).
- Relatively independent of logic function and style.
- Power Delay Product Improves with lowering V_{DD} .

Lower V_{dd} Increases Delay



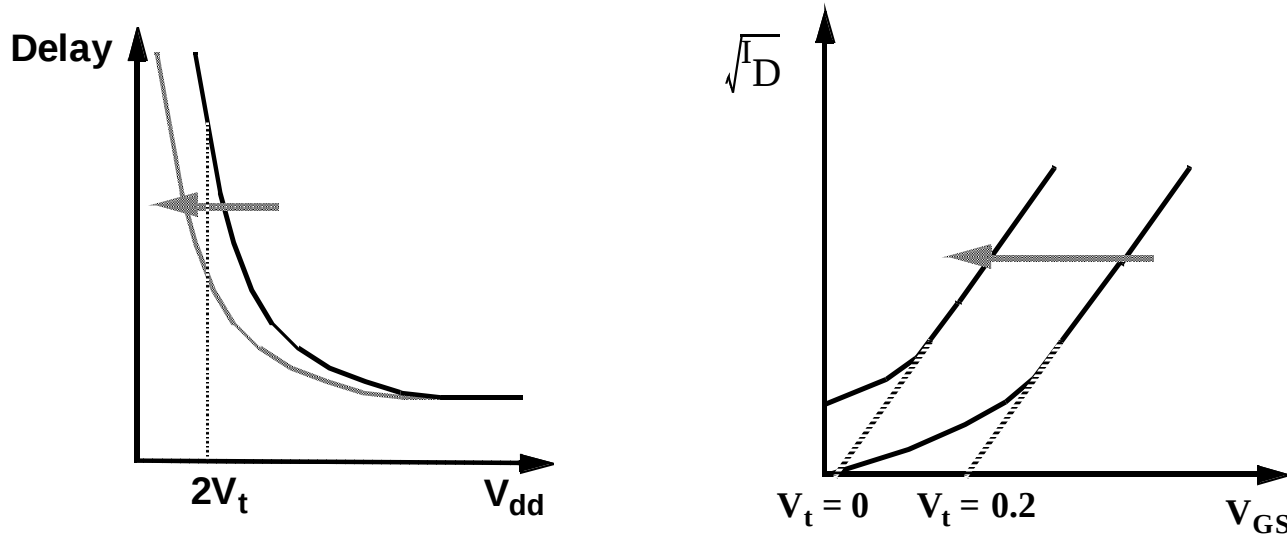
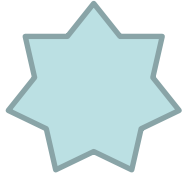
$$T_d = \frac{C_L * V_{dd}}{I}$$

$$I \sim (V_{dd} - V_t)^2$$

$$\frac{T_d(V_{dd}=2)}{T_d(V_{dd}=5)} = \frac{(2) * (5 - 0.7)^2}{(5) * (2 - 0.7)^2} \approx 4$$

- Relatively independent of logic function and style.

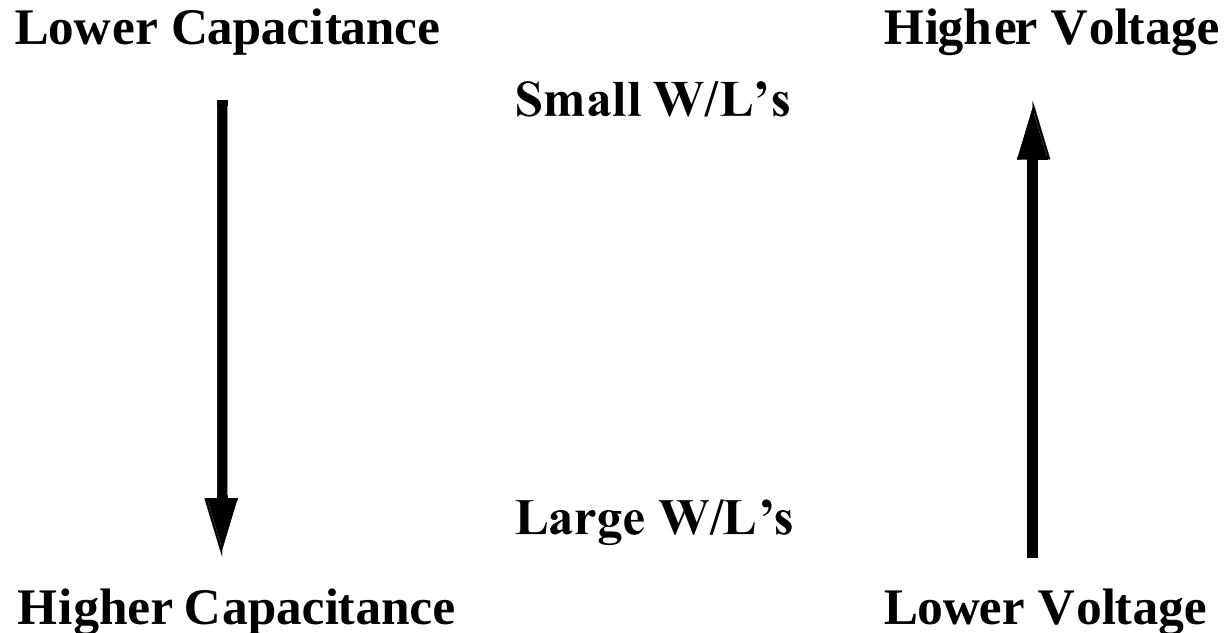
Lowering the Threshold



Reduces the Speed Loss, But Increases Leakage

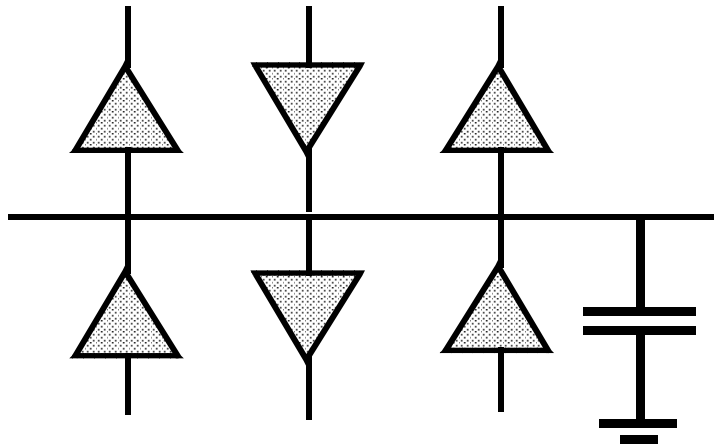
**Interesting Design Approach:
DESIGN FOR $P_{\text{Leakage}} == P_{\text{Dynamic}}$**

Transistor Sizing for Power Minimization

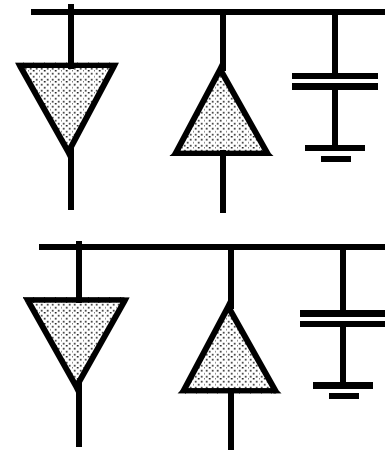


- Larger sized devices are useful only when interconnect dominated.
- Minimum sized devices are usually optimal for low-power.

Reducing Effective Capacitance



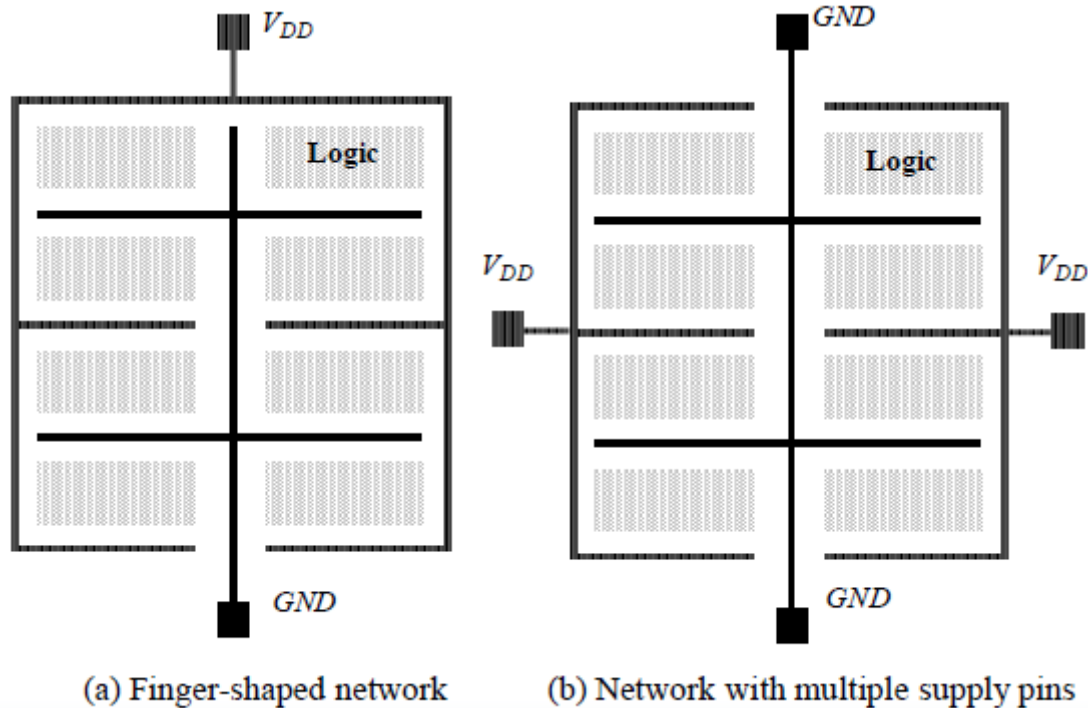
Global bus architecture



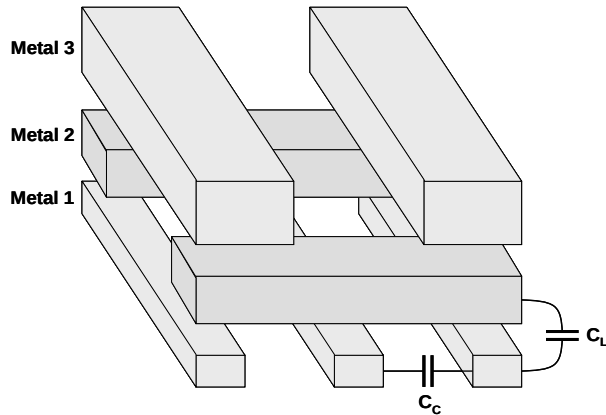
Local bus architecture

Shared Resources incur Switching Overhead

Power and Ground Distribution



Power



$$P = \alpha f C V_{DD}^2$$

Prime choice: Reduce voltage!

- Recent years have seen an acceleration in supply voltage reduction
- Design at very low voltages still open question (0.6 ... 0.9 V by 2010!)
- Reducing thresholds to improve performance increases leakage
- Reduce switching activity
- Reduce physical capacitance

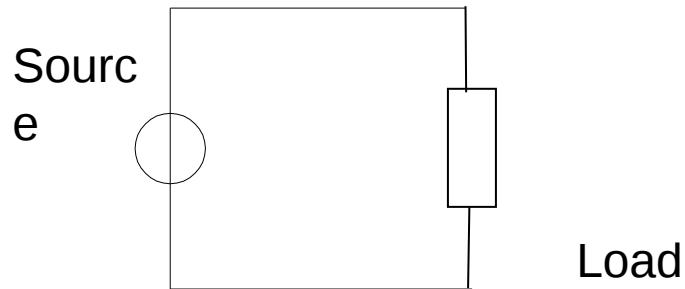
Where

P is the total dynamic power dissipation,
 α the signal transition switching activity,
 f the operating frequency of the bus,
 C the load capacitance of the wire line, and
 V_{DD} the swing voltage. Reducing the bus power consumption is usually achieved by using a low swing voltage and operating frequency as well as reducing capacitance and switching activity

Aspects of power distribution design

What is the simplest active network?

- Active element (voltage/current source) & Load (consumer of energy)



- Loads are different elements and devices, for example:
 - Capacitors
 - Resistors
 - Inductors
 - Transistors

Aspects of power distribution design (cont.)

- High level of integration in digital systems is directly affected on the power distribution system.
- There are several aspects of power distribution design:
 - PCB (board) stackup
 - Bypass capacitor (decap) selection and placement
 - Voltage partitioning
 - Package, socket and connector selection
 - Pin assignments, including selection of the signal-to-power and signal-to-ground ratios
 - Pin placement

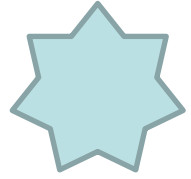
Aspects of power distribution design (cont.)

- The basic goal in power distribution is to minimize inductance & resistance, while optimizing capacitance over a wide frequency range
- Factors that complicate the design:
 - Cost
 - Size limitations
 - Nonideal component behavior (especially for capacitors)
 - Limits on pin counts
 - Limits on signal-to-power & signal-to-ground ratios
 - Limits on the layer count on the PCB

More Extra

Power delivery design methodology

1. Impedance Concepts



What is an impedance ?

The ratio of voltage to current for exponential waveforms is defined as the impedance **Z**.

For a resistance: **$Z_R = R$** in Ohms

For an inductance: **$Z_L = sL$** in Ohms

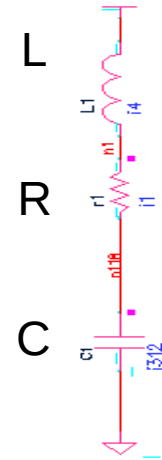
For a capacitance: **$Z_C = 1/sL$** in Ohms

For R, L, C in series: **$Z(s) = R + sL + 1/sL$**

$s = \alpha + j\omega \Rightarrow$ complex frequency

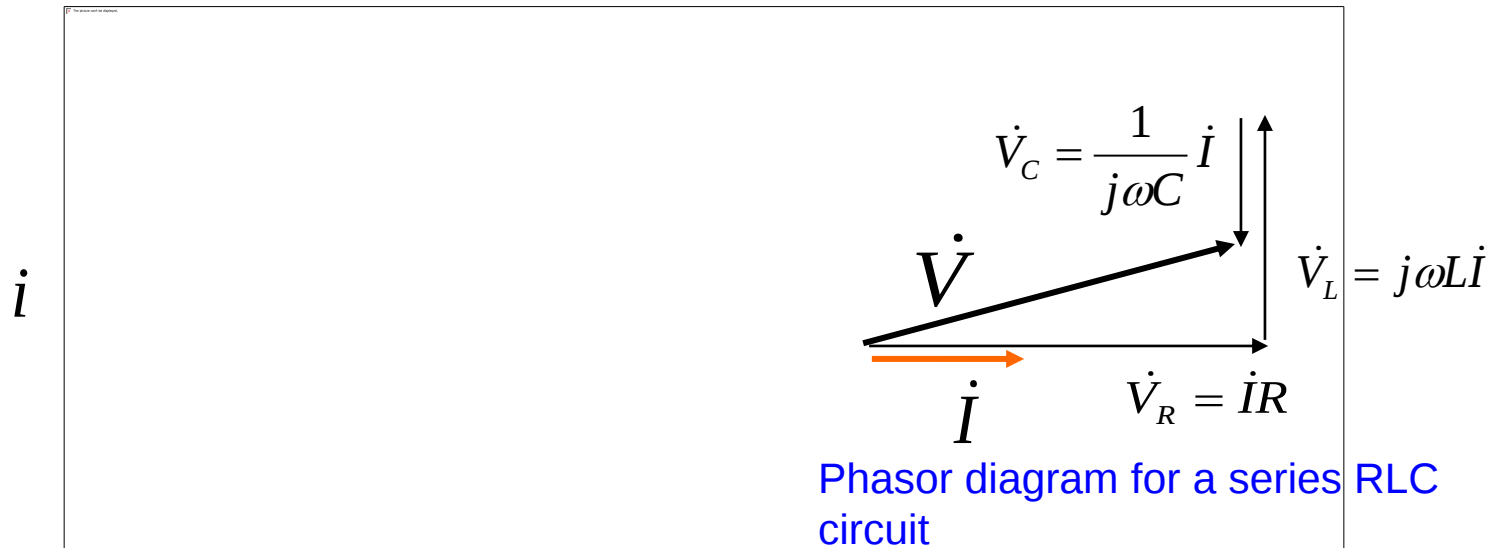
$\alpha = \text{Re}(s) \Rightarrow$ damping factor,

$\omega = \text{Im}(s) \Rightarrow$ angular frequency.

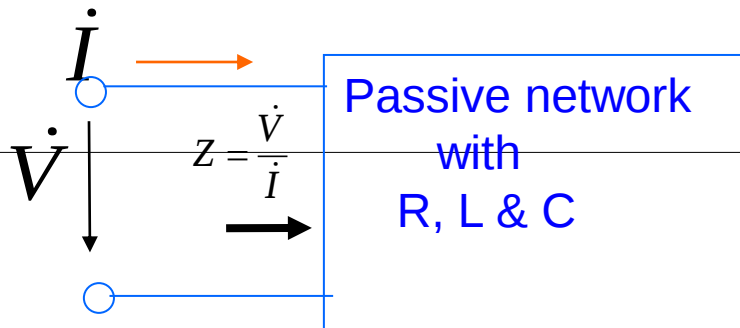


Power delivery design methodology

2. Impedance Concepts



Impedance Z is the ratio of the terminal voltage phasor to the terminal current phasor.



$$v = v_L + v_R + v_C$$

$$v = V_m \sin \omega t$$

$$i = I_m \sin(\omega t - \varphi)$$

$$I_m = \frac{V_m}{Z} = \frac{V_m}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}}$$

$$x_L = \omega L$$

$$x_C = \frac{1}{\omega C}$$

$$Z = R + j(x_L - x_C)$$

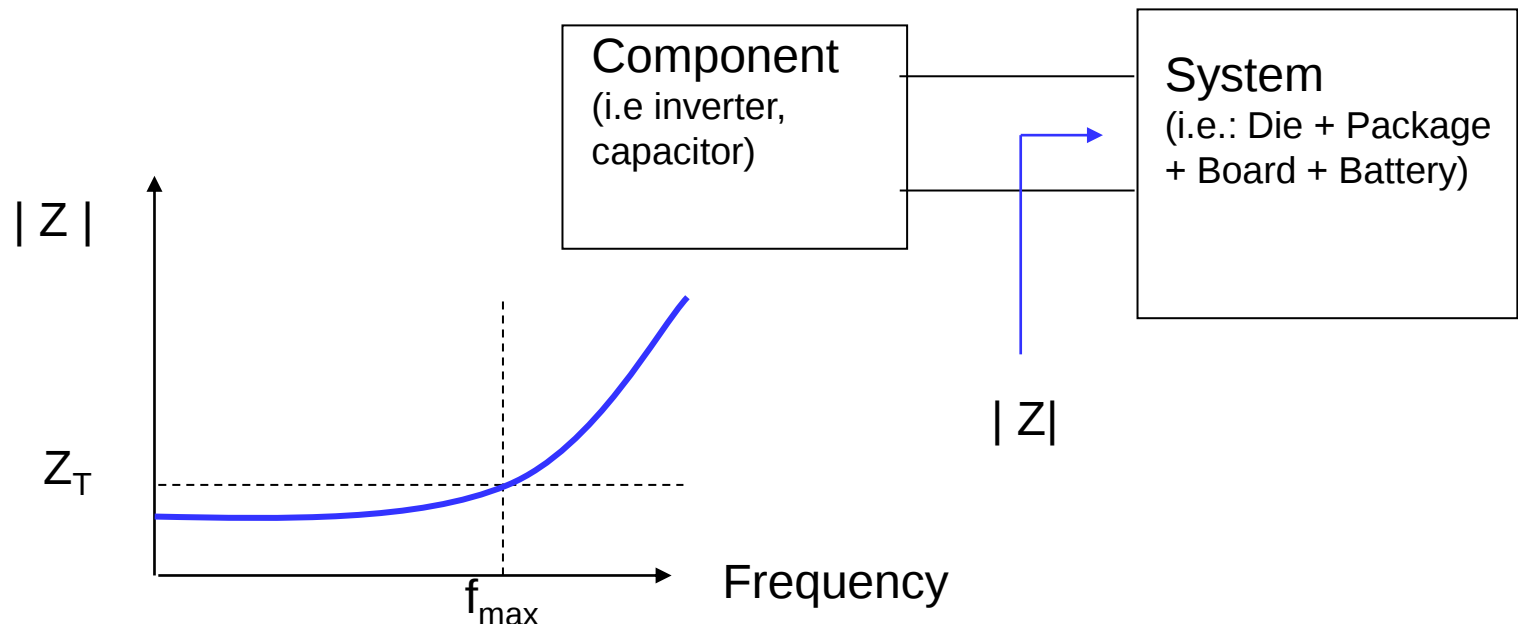
Power delivery design methodology (cont.)

2. Target Impedance

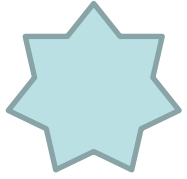


The power distribution system must work as low-impedance voltage source over a bandwidth from DC to several harmonics of the clock frequency (to minimize the generation of noise, radiation of electromagnetic energy and EMI).

Target impedance Z_T is the maximum allowed impedance for the system to meet a specified noise level.



Power delivery design methodology (cont.)



3. Techniques for lowering $|Z|$:

- a) packaging,
- b) PCB stackup,
- c) bypass capacitor

a) Packaging (include integrated circuit packages, sockets & edge connectors)

$|Z|$ is lowered:

- by choosing a packaging component that offers a shorter connection for lower partial self inductance
- by assigning more pins to power and ground connections (lower signal/ground & signal/power).

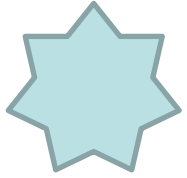
Package styles:

BGA – ball-grid array (C4 bump)

QFP – quad flat pack

BGAs have more pins than GFPs

Power delivery design methodology (cont.)

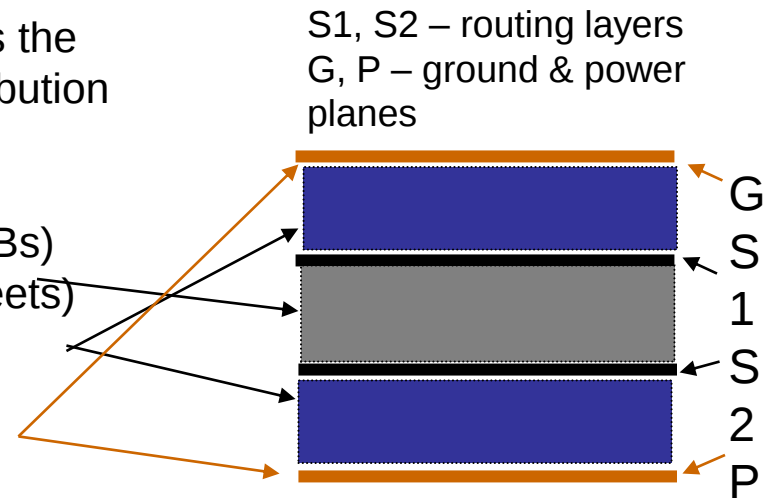


b) PCB stackup

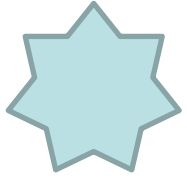
The PCB stackup strongly affects the impedance $|Z|$ of the power distribution system

Multilayer PCB's elements are:

- cores (two-sided PCBs)
- pregreg (dielectric sheets)
- copper sheets



Power delivery design methodology (cont.)



c) Bypass capacitance (decap)

High-performance design require **bypass capacitance** for four reasons:

- 1) supplying current bursts for fast switching circuits (near the component, large value) ;
 - 2) providing an AC connection between power and ground planes for return currents (near the power pins of high-speed component, small value);
 - 3) controlling EMI (distributed, small value);
 - 4) lowering the impedance $|Z|$ of the power distribution system (system design).
- For all of this uses, capacitors connect the power & the ground planes. The difference is the size and quality of capacitors and their locations.

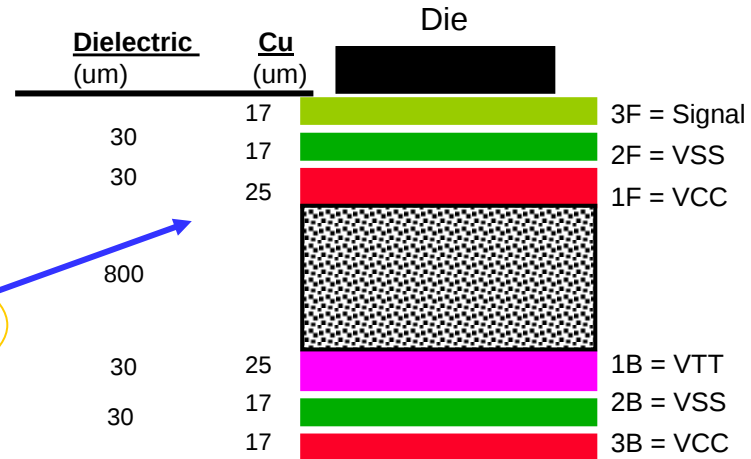
Power delivery design methodology (cont.)

Example of Package template

□ Package

- ⇒ FCPGA 35 x 35mm
- ⇒ 2-2-2 Stack up

**Package Assumptions
(slide of ATD group)**



- ⇒ POR X62 FCPGA3 Design Rules
- ⇒ 17um Build up Layer thickness (stretch target)
- ⇒ **Capacitors**

- 0805 IDC+ Under the die $L = 57\text{pH}$; $R = 6.5\text{mOhm}$

- Outside the die

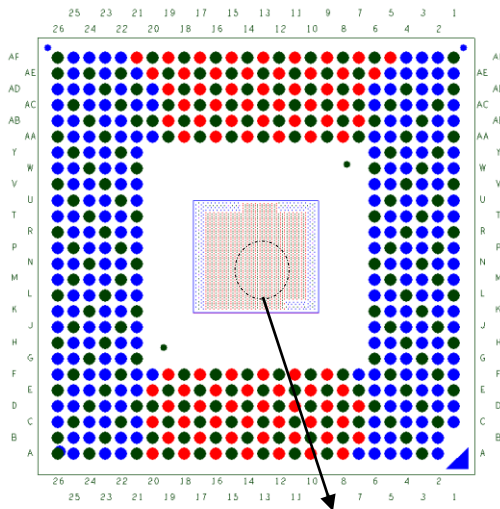
0805 IDC+ $L = 2.5 * 57\text{pH}$; $R = 6.5\text{mOhm}$

1206 $L = 220\text{pH}$; $R = 5\text{mOhm}$ (*L is too optimistic; Ansoft modeling yields 600pH - 800pH*)

Did not account for VTT or PLL capacitor requirements.

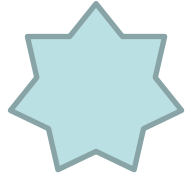
Power delivery design methodology (cont.)

□ C4 bump pattern:



296u: provide an 86um trace width bussing on surface layer

270u: provide a 76um trace width bussing on 2F layer



Surface

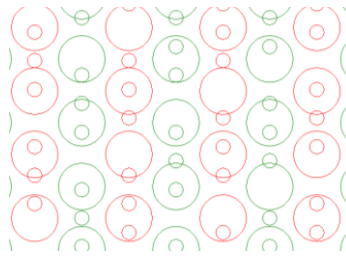
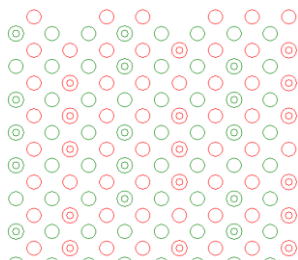
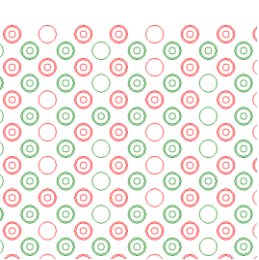
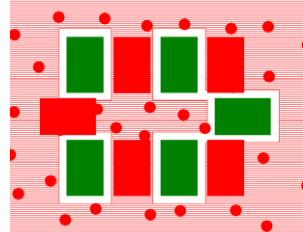
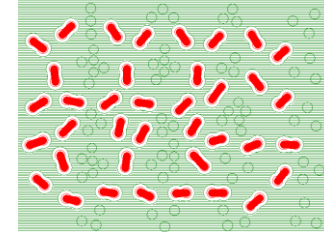
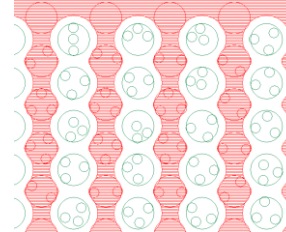
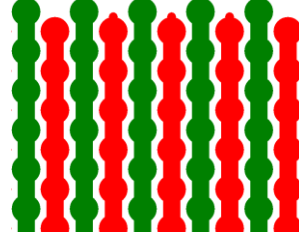
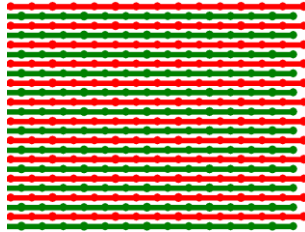
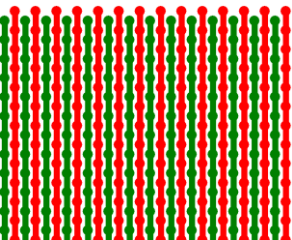
2F

1FC

1BC

2B

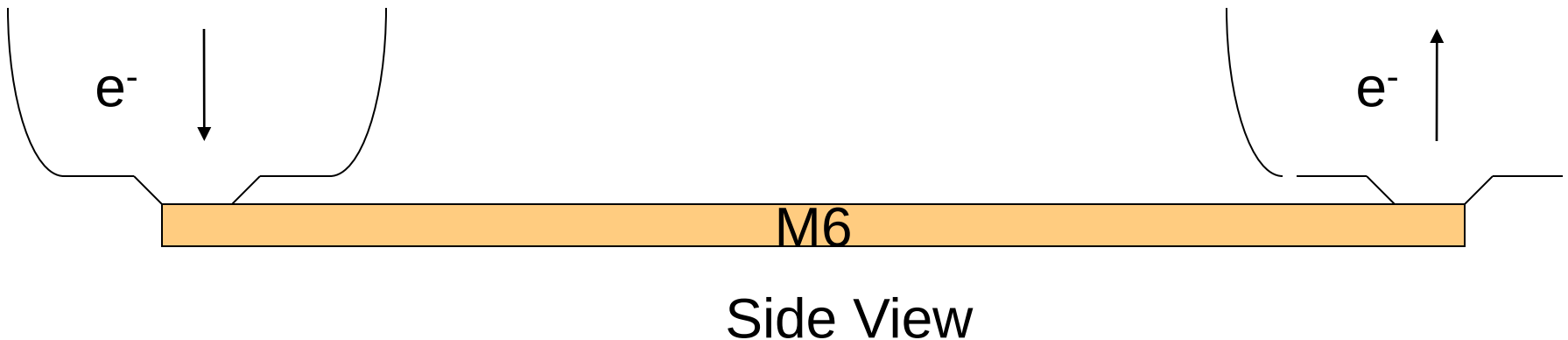
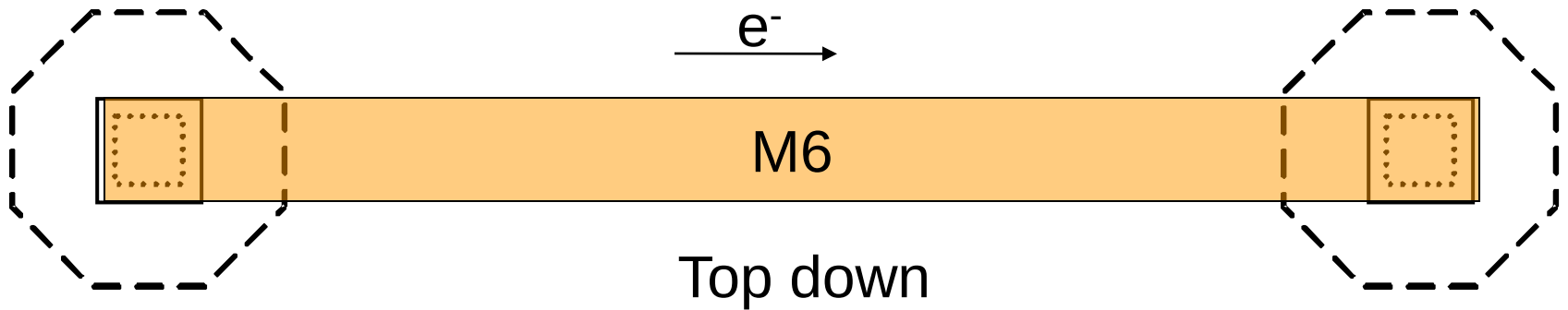
Base



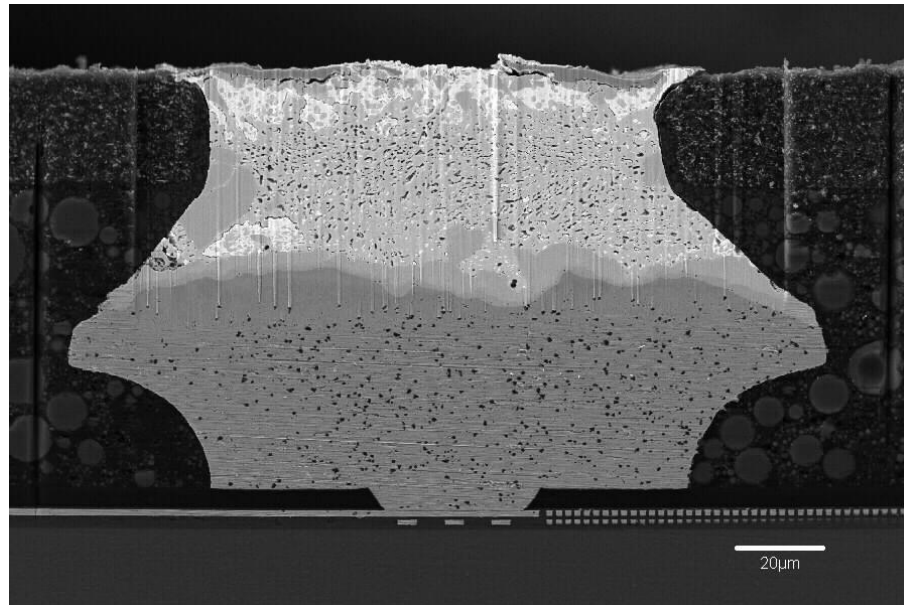
- 75% via connection to die bumps from 2F.
- 25% via connection to die bumps from 1FC or PTH.
- ~ 3 via connection to 2 PTHs

Example: P1262 DT Power Delivery WW17.4
Power Bussing In Netlist, slide of ATD group

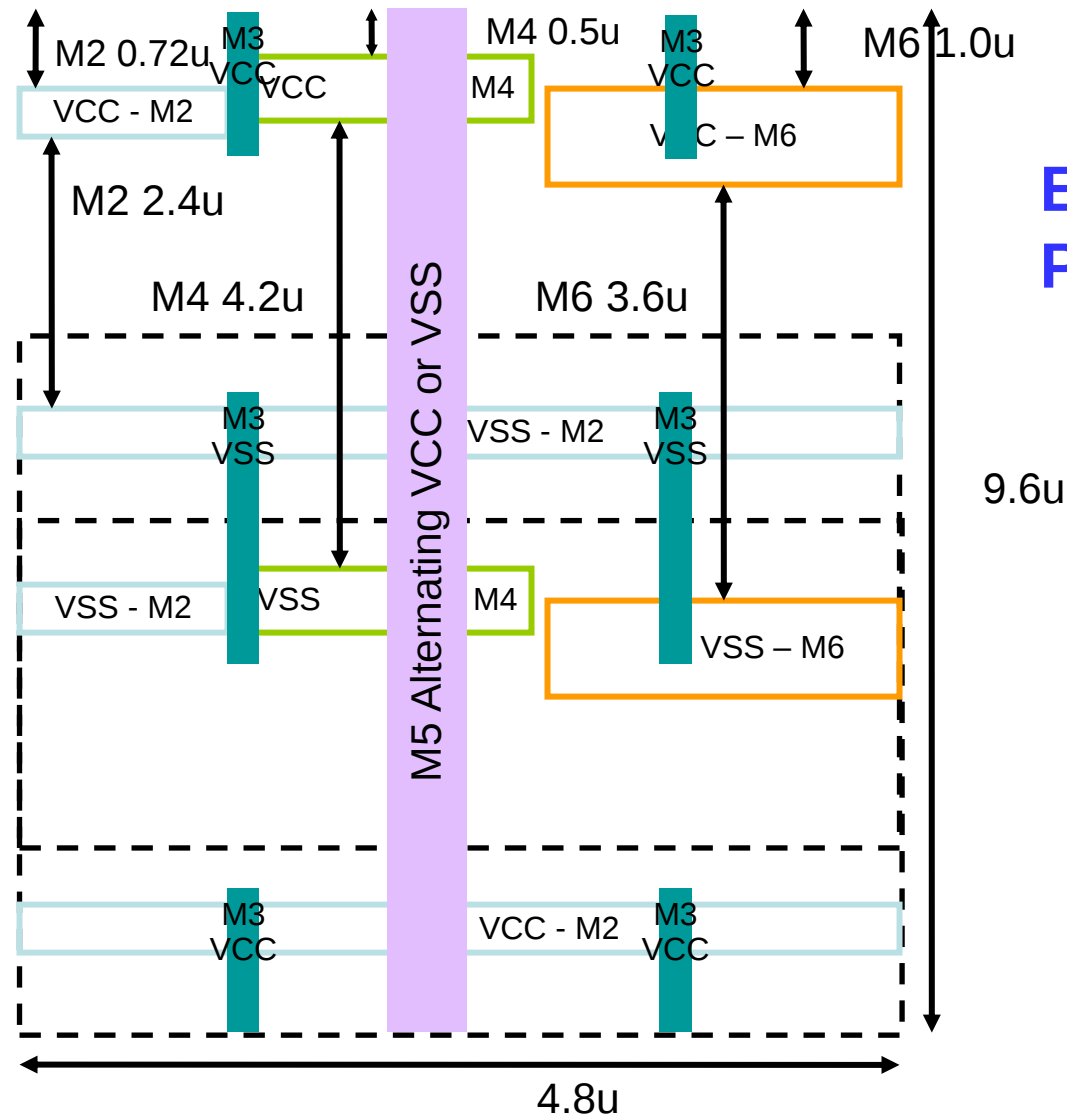
C4 Bump & M6 of Die



C4 Bump Cross Section

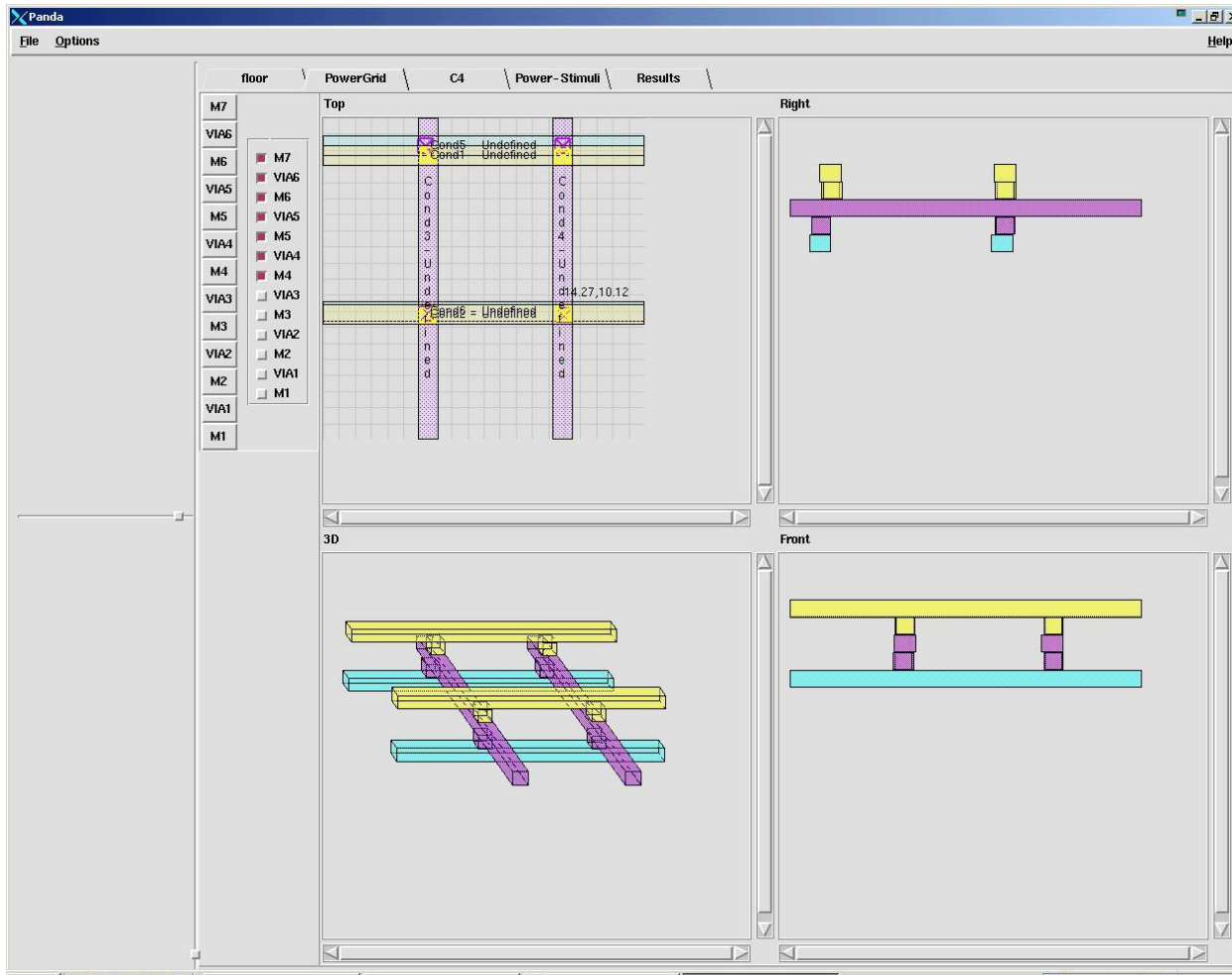


P1262 M2-M6 Power-B. Martell

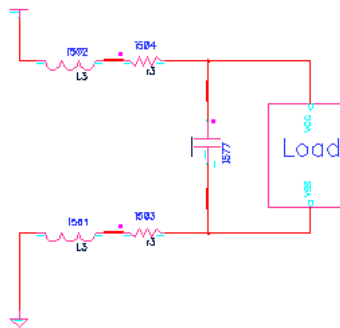
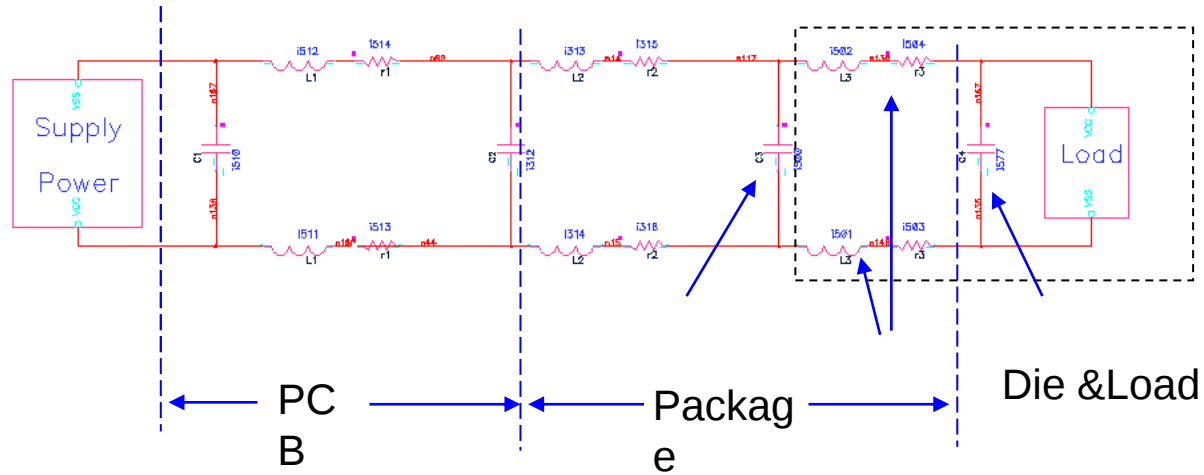


**Example of
PG template**

3D PG: Template for M6, M5 & M4 Panda, GUI



Power Delivery 3-levels Model



Power delivery package to die and die model

Component Behavior

Conductors for power supply & for impulse signals:

$$Z = R + j\omega L \quad \Rightarrow R, RL, RLM$$

Pins $\Rightarrow R$

Capacitors:

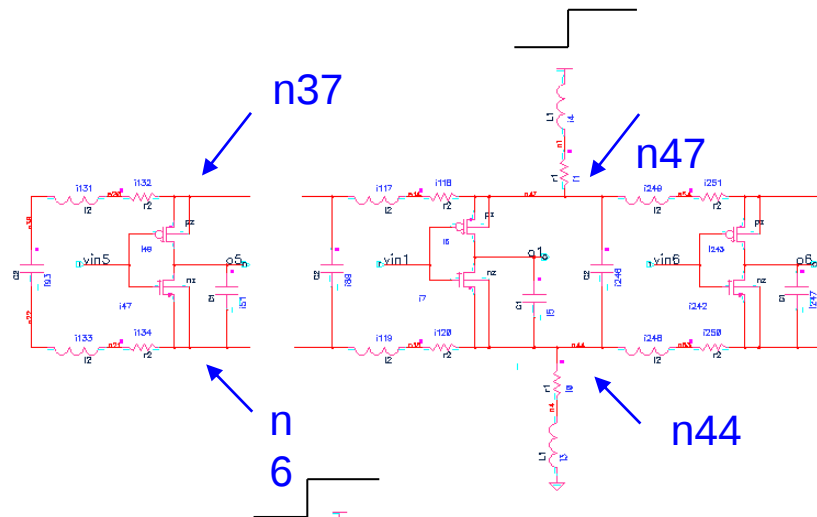
$\Rightarrow$ open circuit, RC, RL

Transistors: nonlinear & linear models

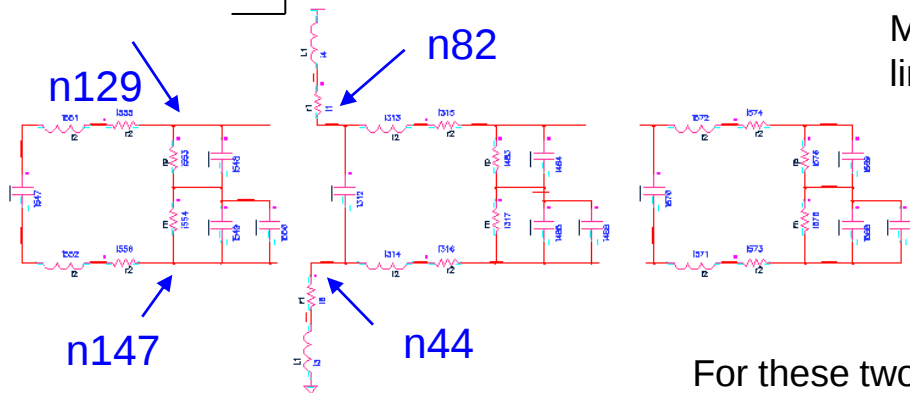
$$Z = R + j\left(\omega L - \frac{1}{\omega C}\right)$$

Component behavior

Linear RLC Model and Comparison with Nonlinear Model



Multi-cascades (9 cascades)
nonlinear model



Multi-cascades (9 cascades)
linear RLC model

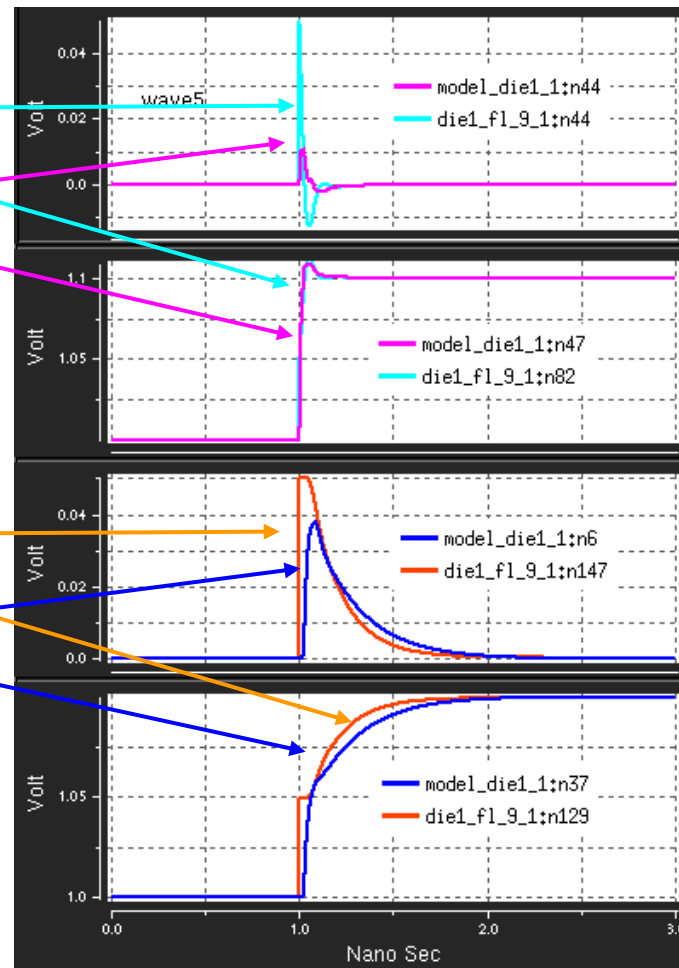
For these two models response voltage waves to VCC-input = 1(t)V step-function were found in vcc & vss on die nodes .

Component Behavior (cont.)

Linear and Nonlinear Power Grid Models Results

Liner
Nonliner

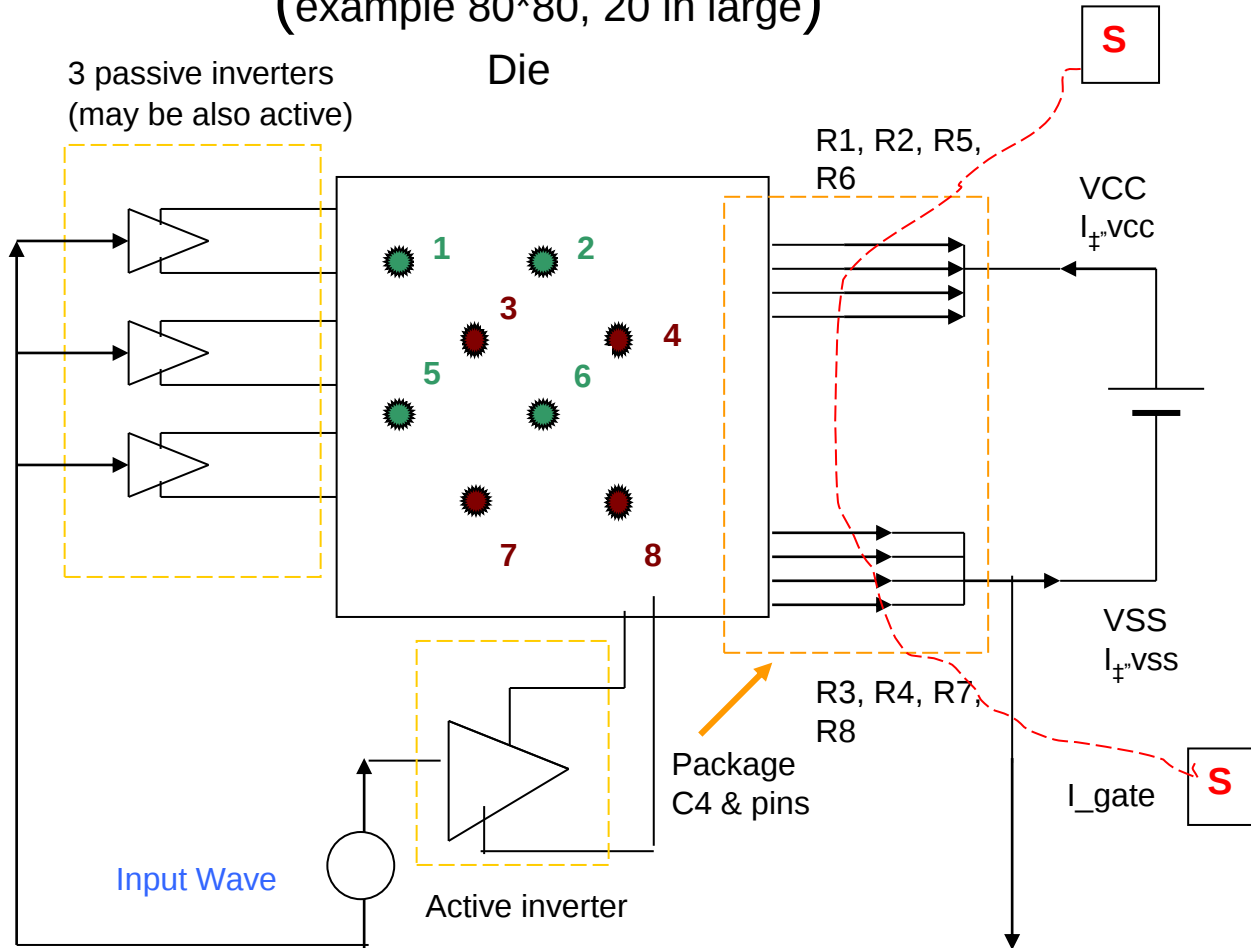
Liner
Nonliner



model_die1 is nonlinear model

die1_fl_9 is linear model

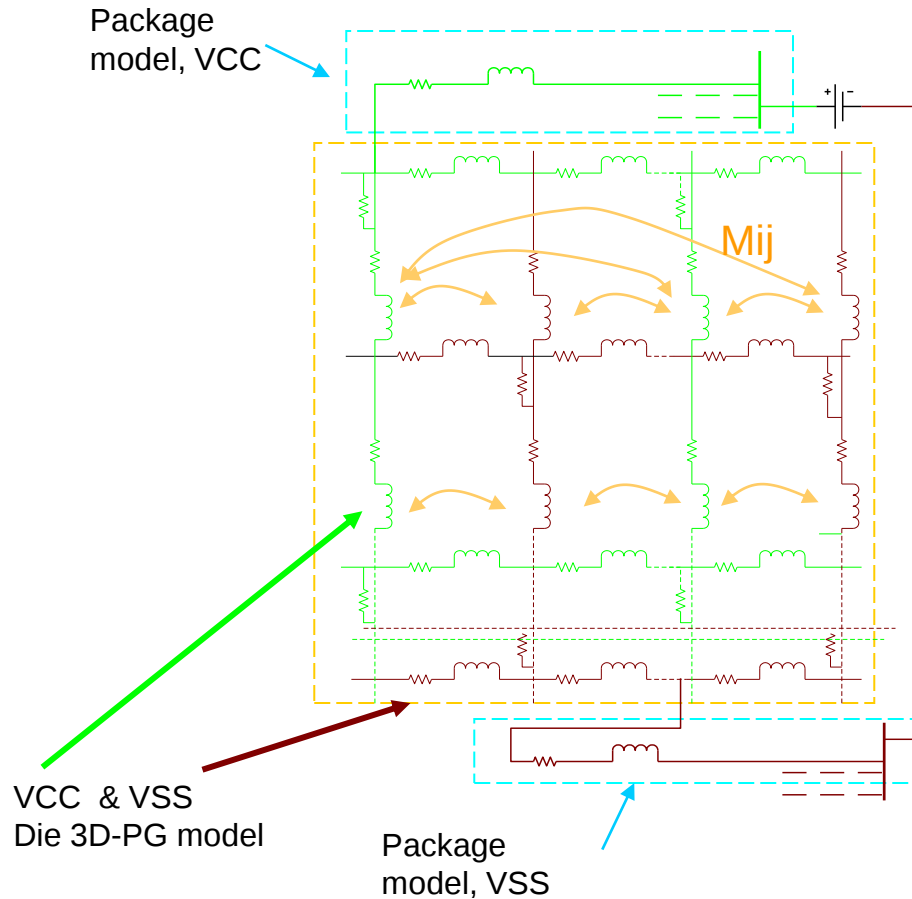
(example 80*80, 20 in large)



$$\sum I_{\text{sectionS}} = 0$$

$$\sum I_{\text{sectionS}} = \sum I_{\text{c4pins}} - \sum I_{\text{gates}}$$

3D Die +Package Distributed Model



2D LM-matrix:

	1	2	3	...	n
1	L_1	M_{12}	M_{13}	...	M_{1n}
2	M_{12}	L_2	M_{23}	...	M_{2n}
3	M_{13}	M_{23}	L_3	...	M_{3n}
.....					
n	M_{1n}	M_{2n}	M_{3n}	...	L_n

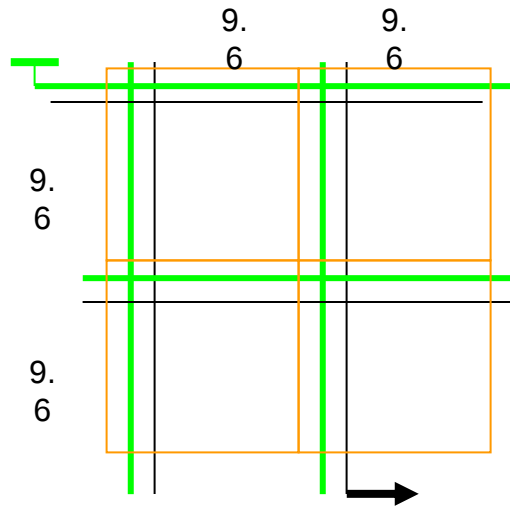
Number M_{ij} :

$$\frac{1}{2} n_H * (n_H - 1) * s_H + \frac{1}{2} n_V * (n_V - 1) * s_V$$

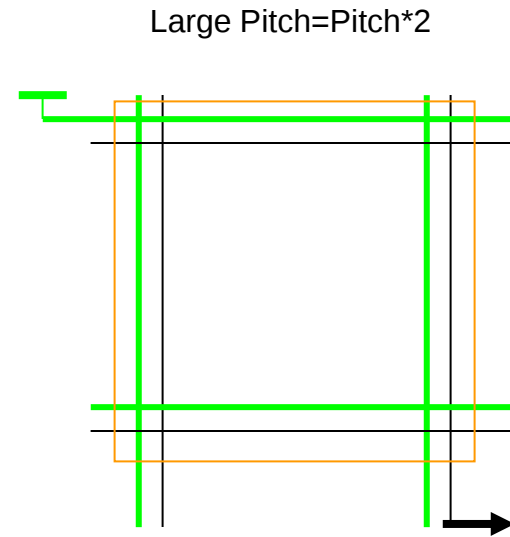
$n_H, n_V \Rightarrow \# \text{conductors}$

$s_H, s_V \Rightarrow \# \text{sections}$

Use large templates: merge several ones into large macros



Four basic templates

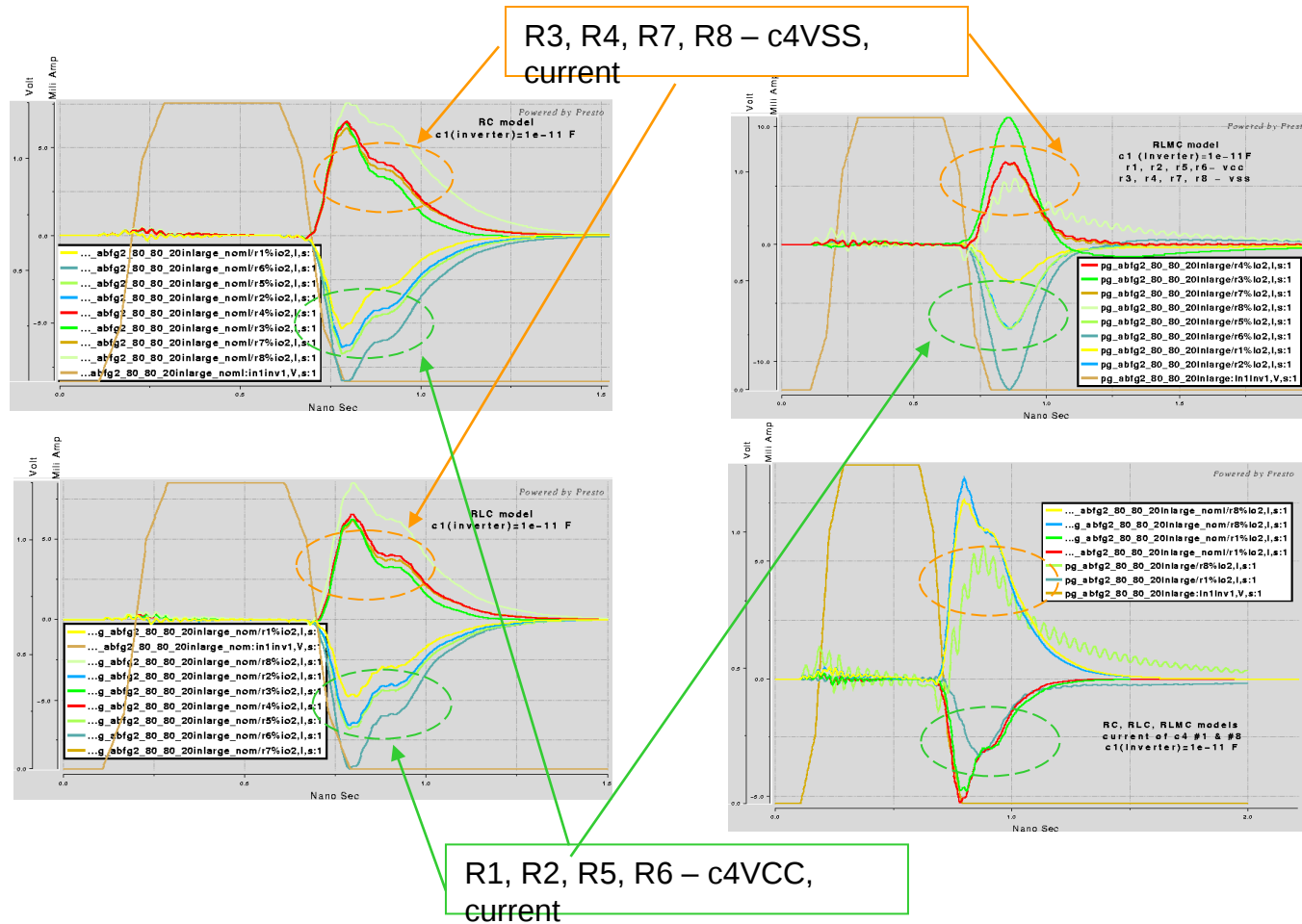


Large Pitch = Pitch * 2

One Large template

Modeling techniques

(RC, RLC, RLNC models. Example 80*80, 20 in large, 3 passive & 1 active inverter, c4 bumps current)



Summary

We have covered the following topics:

Aspects of power distribution design & factors that complicate the design

Power delivery design methodology:

- Impedance Z concept

- Techniques for lowering $|Z|$: packaging, PCB stackup, bypass capacitor

Components behavior: conductors, capacitors, transistors

Modeling techniques:

- Package+die 3D model

- Use large template for die

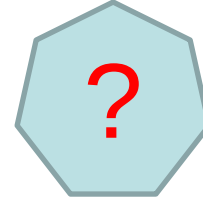
- Results of simulations (current & voltage responses, input impedance)

SRAM PROJECT:

<https://www.youtube.com/watch?v=68Dn1x6cZ4g>

<https://www.youtube.com/watch?v=SHJPFNI5Mzo>

<https://www.youtube.com/watch?v=KrqyvpU9Cu0>



Multiplier -A Comparative Study On Low Power Multiplier Using Microwind Tool

- https://www.youtube.com/watch?v=4-l_PGPog9o
- <https://www.youtube.com/watch?v=MCFG7XD16Ek>
- <http://www.ijsret.org/pdf/EATHD-15026.pdf>
- <https://www.youtube.com/watch?v=rqwkrUcNyH4>
- https://www.youtube.com/watch?v=4-l_PGPog9o
- <https://www.youtube.com/watch?v=WxSR2Yhnqk4&t=30s>
- https://www.acsu.buffalo.edu/~phaniram/bootstrap-prestructure22_files/images/paper_1.pdf