
Timing Design in Digital Systems

Outline

1. Timing design in Synchronous (clocked) Logic
 - Min/Max timing with flip-flops
 - Latch-based design
2. Timing Issues in CMOS circuits
3. Timing verification Flow
4. Techniques to Improve Performance

Course “Mantras”

- One clock, one edge, Flip-flops only
- Design BEFORE coding
- Behavior implies function
- Clearly separate control and datapath

Mantra #1

One clock, one edge; Flip-flops only

- For your design (at least for each module) use one clock source and only one edge of that clock
- Only use edge-triggered flip-flops

Why?

- Moving data between different clock domains requires careful timing design and synthesis “scripting”

If you need multiple clocks in your design

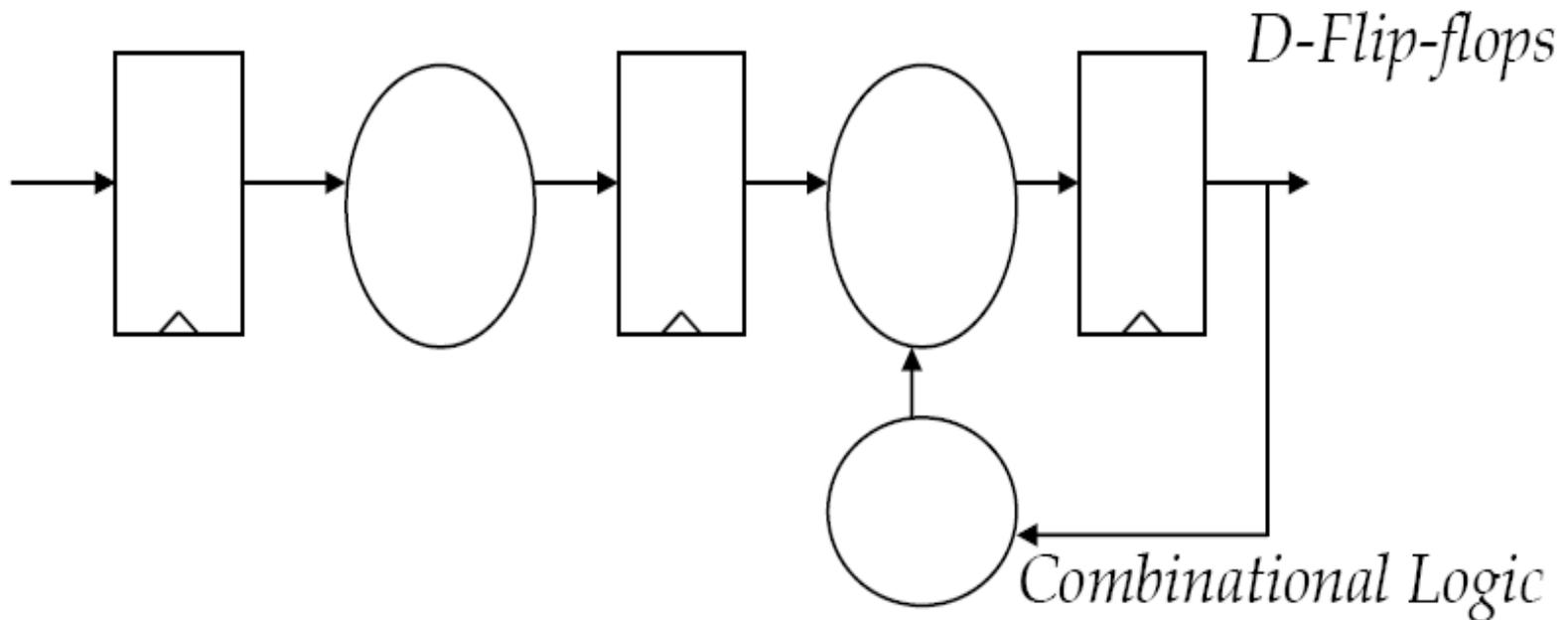
- Make them related by a powers of 2
 - E.g 50, 100 and 200 MHz
- Consider one clock per module
- Consider resynchronizing using flip-flops between clock domains

Caveat

- Tools and designers are getting better at using latches and multi-phase clocks. However, this requires some experience to get correct.

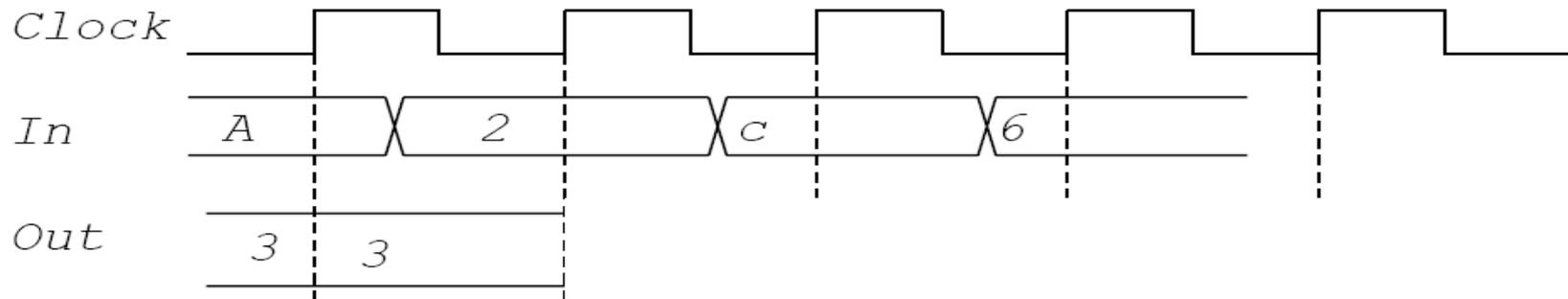
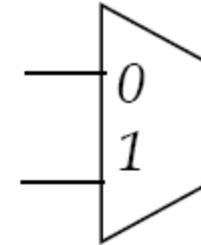
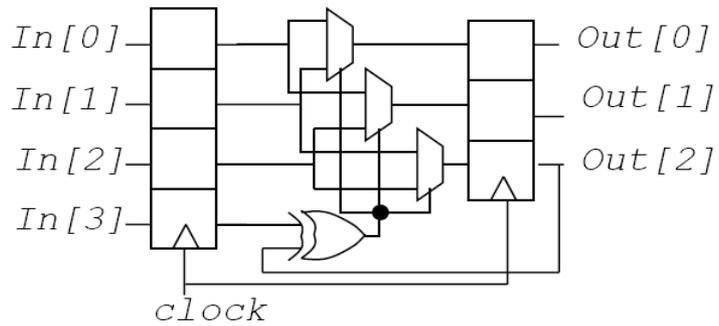
General Approach to Timing Design

- In general, all signals start and end in registers every clock period



Clock Level Timing

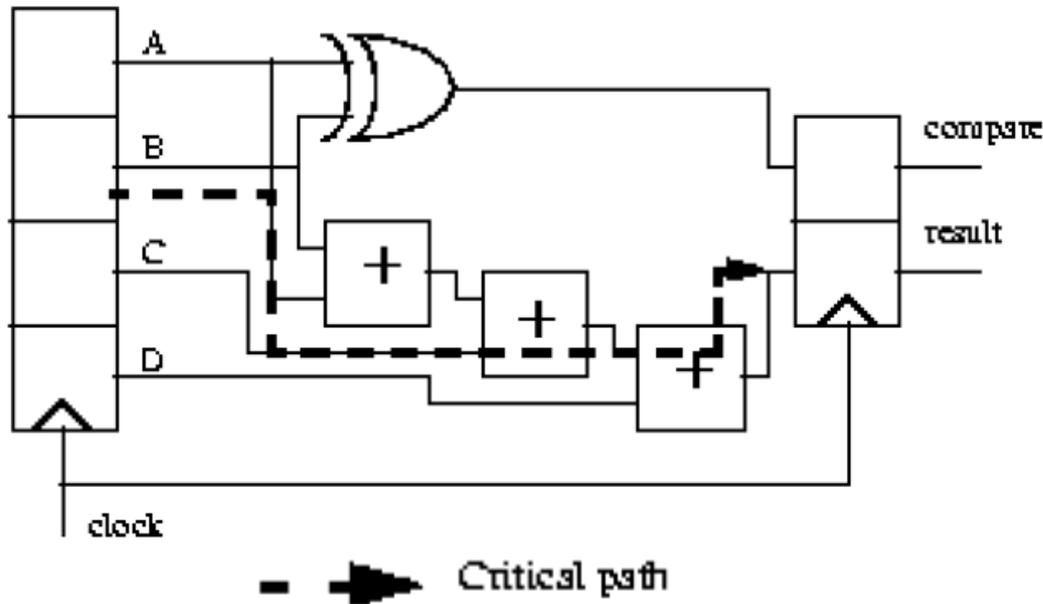
Example (Revision):



Critical Path

Thus, the clock speed is determined by the slowest feasible path between registers in the design

- Often referred to as “the critical path”



Critical path is longer with increased *logic depth* (# gates in series)

Synchronous Clock Distribution

The goal of clock tree is for the clock to arrive at every leaf node at the same time:

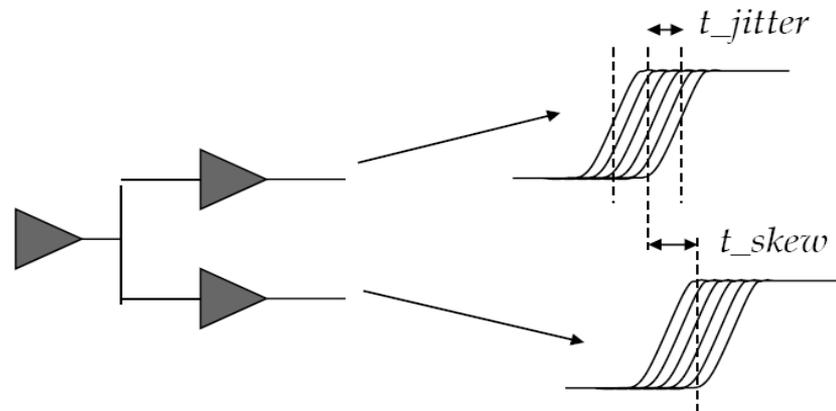
Usually designed after synthesis: Matched buffers; matched capacitance loads

Common design method:

- “H tree”
- Clock tree done after design
- Balance RC delays
- Balance buffer delays

Clock skew and jitter

- Clock skew = systematic clock edge variation between sites
 - Mainly caused by delay variations introduced by manufacturing variations
 - Random variation
- Clock jitter = variation in clock edge timing between clock cycles
 - Mainly caused by noise



Comments on Clock Skew

- ASIC design relies on automatic clock tree synthesis
 - Works to guarantee a global skew target
- Custom clock distribution can be used to add the following features to a clock:
 - Smaller skews
 - Local skews < Global skew
 - Multiple non-overlapping clock phases
 - Deliberate non-random skew at flip-flop/latch level
 - Future automatic clock tree synthesis tools might include features like this

Flip-Flop based design

Edge triggered D-flip-flop

- Q becomes D after clock edge

Set-up time:

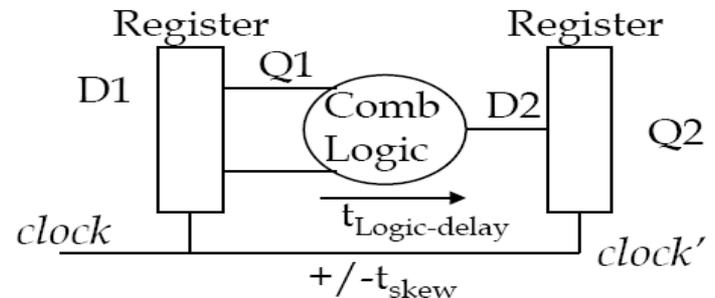
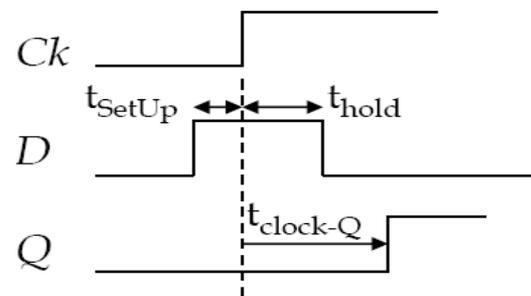
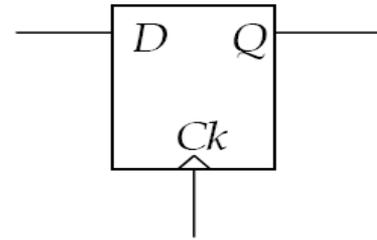
- Data can not change no later than this point before the clock edge.

Hold time:

- Data can not change during this time after the clock edge.

$t_{clock-Q}$

- Delay on output (Q) changing from positive clock edge



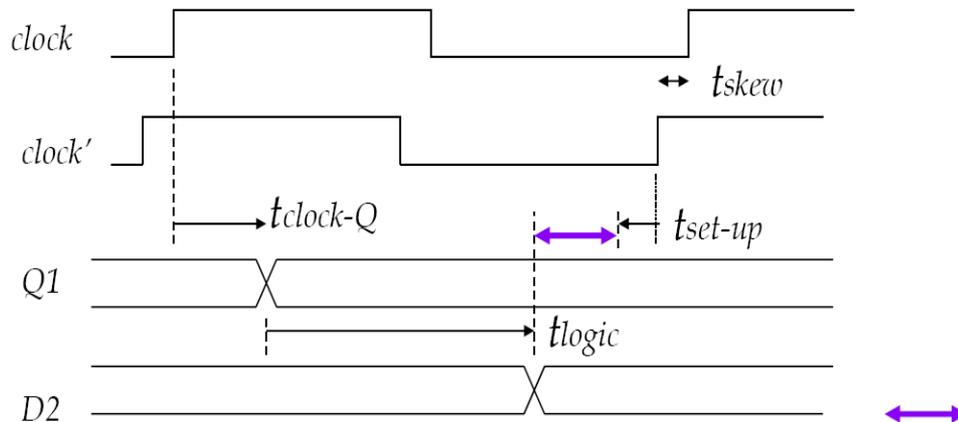
Preventing Set-Up Violations

Set-up violation:

Logic is too slow for the correct logic value to arrive at the inputs to the register on the right before one set-up time before the clock edge

Constraint to prevent this:

$$t_{clock} \geq t_{clock-Q-max} + t_{logic-max} + t_{set-up} + t_{skew}$$



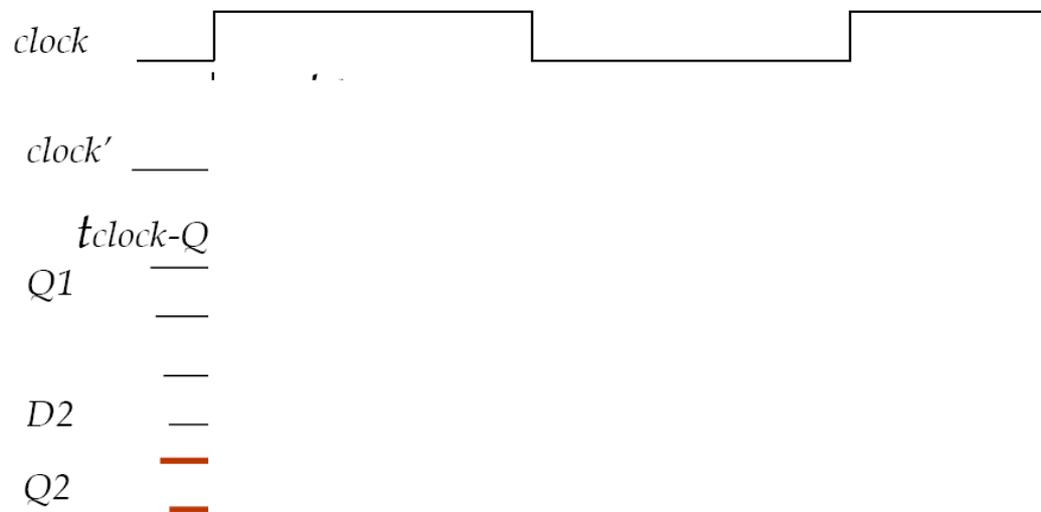
The amount of time required to turn ' $>$ ' into ' $=$ ' is referred to as **timing slack**

Preventing hold Violations

Hold violations occur when race-through is possible

Constraint to prevent hold violations:

$$t_{hold} + t_{skew} \leq t_{clock-Q-min} + t_{logic-min}$$



sometimes have to
insert additional
logic to prevent hold
violations

Latch Based Design

D-latch

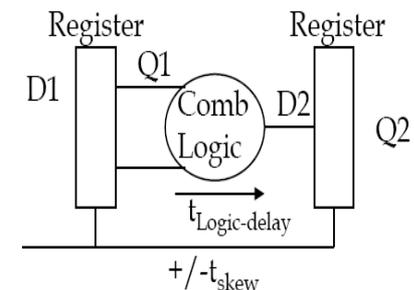
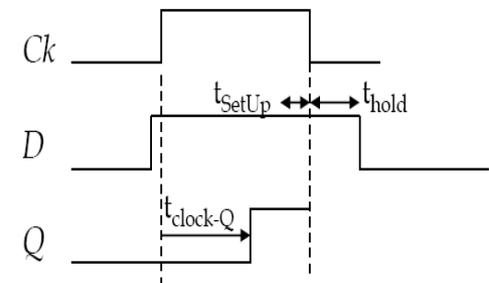
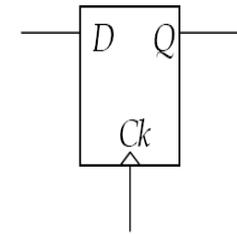
- Q follows D while clock is high (“transparent”)
- Value on D when clock goes low is stored on Q

Set-up and hold times:

- D can not change close to the falling ('latching') clock edge.

$t_{\text{clock-Q}}$

- Delay from clock going high to Q changing



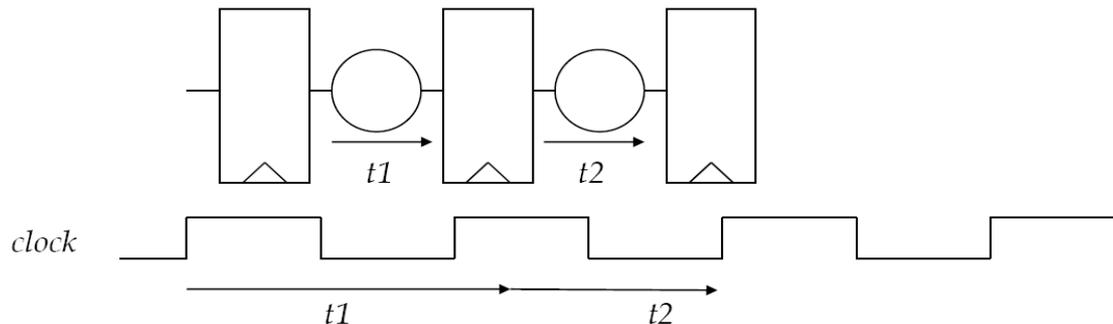
Latch Timing Constraints

- Set-up constraints under nominal design same as for flip-flop

$$t_{clock} \geq t_{clock-Q-max} + t_{logic-max} + t_{set-up} + t_{skew}$$

- “Transparency” of latch can be used to improve flexibility of timing

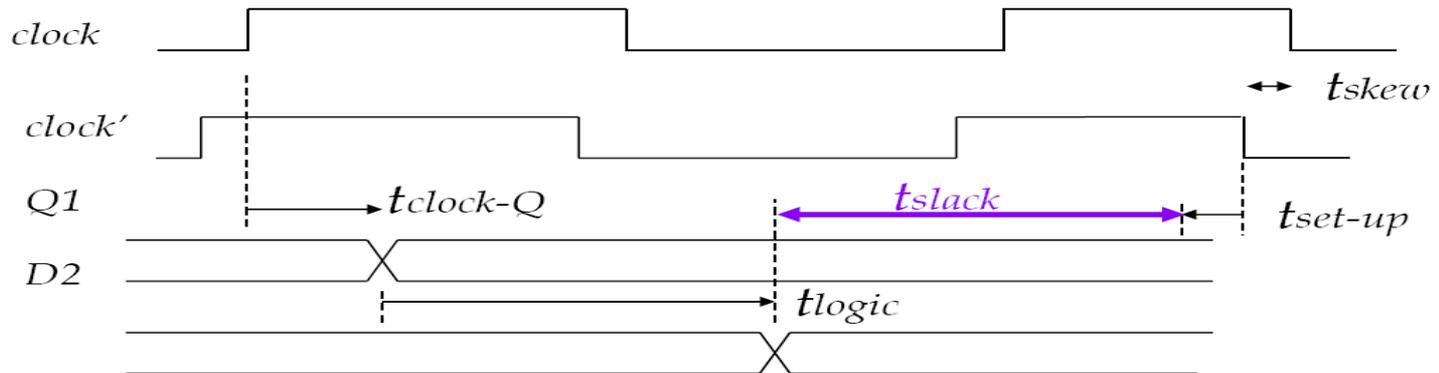
e.g.: If critical path is in logic block 1:



Latch Timing Constraints WITH cycle stealing

To prevent set-up violations:

$$t_{clock} + t_{clock-high-max} \geq t_{clock-Q-max} + t_{logic-max} + t_{set-up} + t_{skew}$$



Notes:

- The percentage of time the clock is high is referred to as the *duty-cycle*
- If part of the following clock-high time is used to allow this logic to be slower, then the logic-block connected to Q2 must be proportionally faster
- Using the clock-high time like this is called *cycle-stealing*

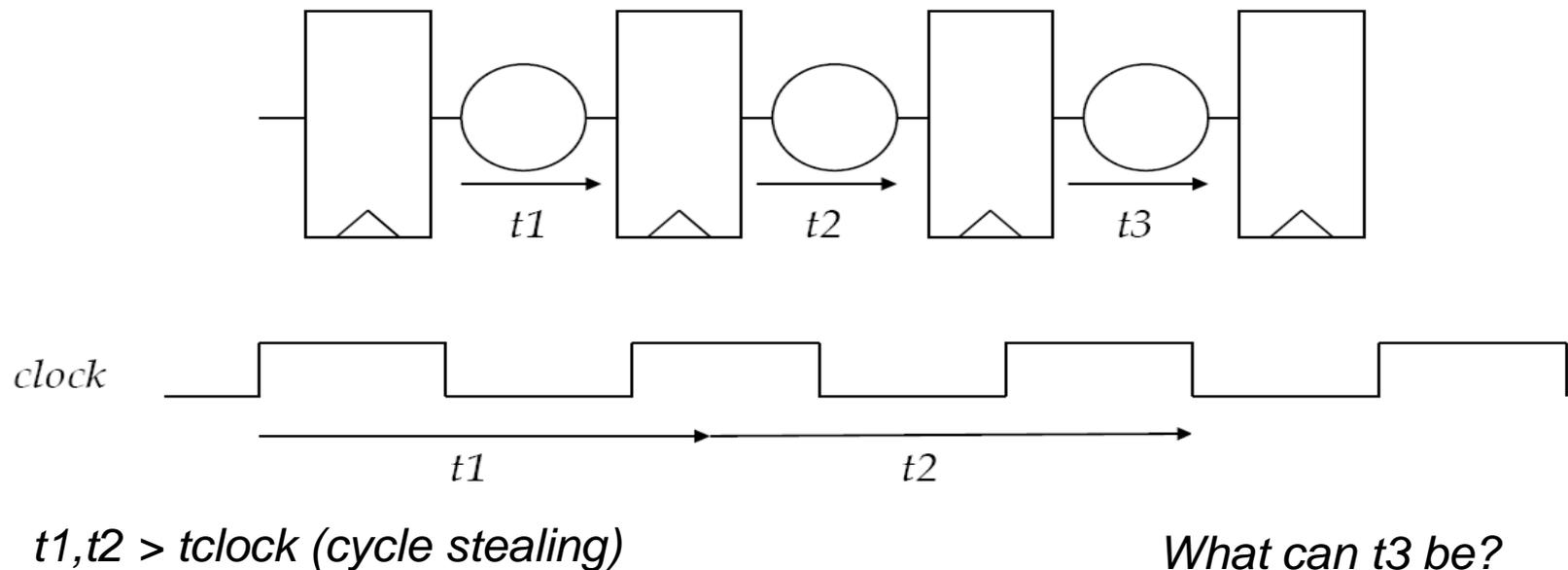
Latch Set Up Violations

Notes:

- The percentage of time the clock is high is referred to as the *duty-cycle*
- If part of the following clock-high time is used to allow this logic to be slower, then the logic-block connected to Q2 must be proportionally faster
- Using the clock-high time like this is called *cycle-stealing*
 - Normally cycle stealing is not enabled

Latches ... Cycle Stealing

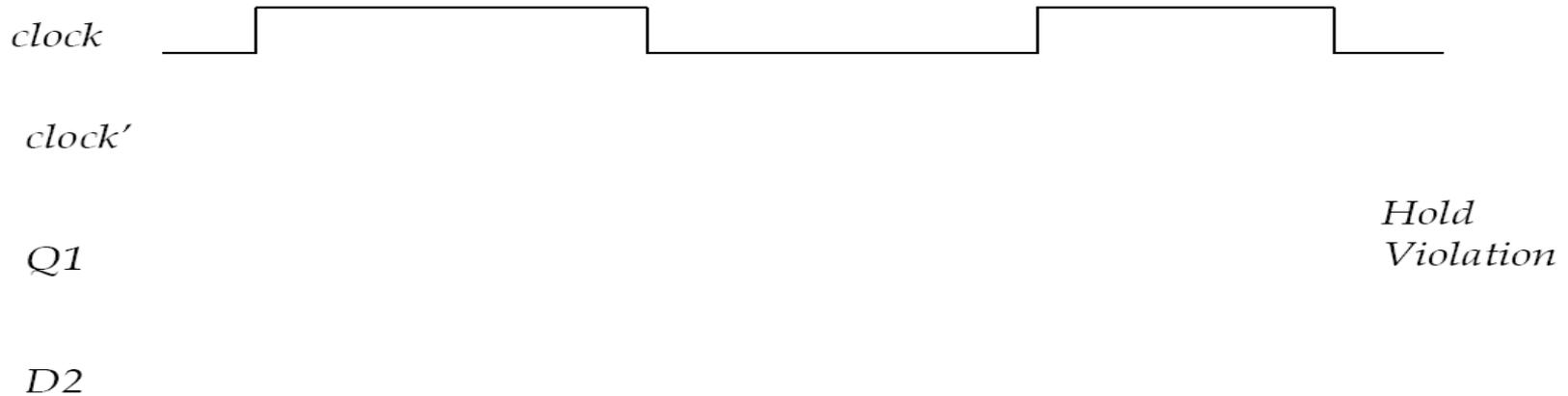
- Can use up to a total of t_clock_high within a pipeline structure, to help in timing closure
 - Example:



...Latch timing constraints

To prevent hold violations:

$$t_{clock-high-max} + t_{hold} + t_{skew} \leq t_{clock-Q-min} + t_{logic-min}$$



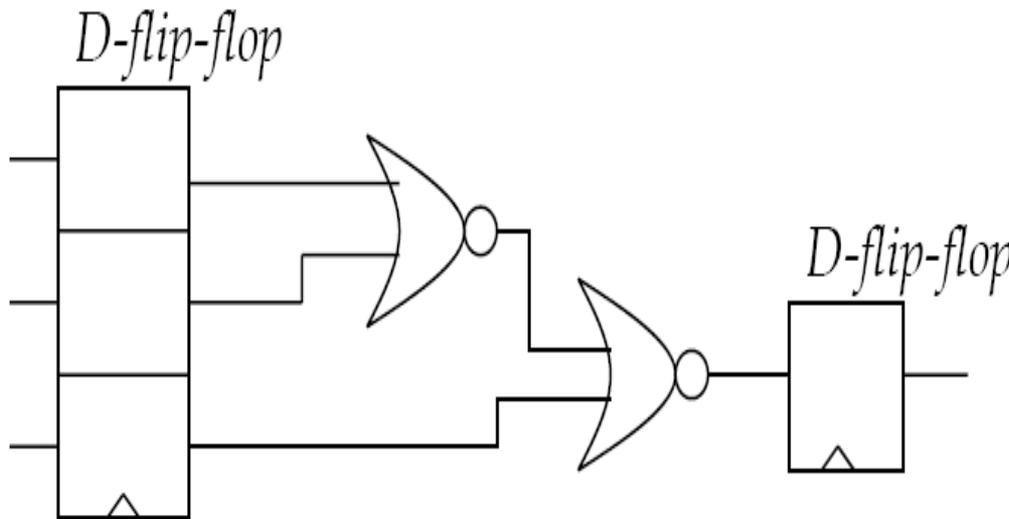
Note:

- Hold violations are harder to prevent in latch-based designs

Revision – So far

- What is a set-up violation?
- How is a set-up violation fixed?
- What is a hold-violation?
- How is a hold-violation fixed?
- Why are edge-triggered flip-flops preferred over latches?

Example:



min : typ : max

$T_{\text{clock-Q}} = 3 : 4 : 5$

$T_{\text{NOR}} = 1 : 2 : 3$

$T_{\text{su_max}} = 1$

$T_{\text{hold_max}} = 2$

$T_{\text{skew}} = 1 \text{ ns}$

If this is the critical path, what is the fastest clock frequency?

Is there potential for a hold violation?

CMOS Drive Strength

Revision: CMOS transistors operating in the linear region:

$$I_{ds} = \beta((V_{gs} - V_t)V_{ds} - V_{ds}^2 / 2)$$

where $\beta = (\mu\epsilon/t_{ox})(W / L)$

where W is the transistor width, and L is the channel length

i.e. To a first approximation, $I_{ds} \approx V_{ds} / R_{on}$ $R_{on} \approx 1 / \beta(V_{GS} - V_T)$

Thus, delay in CMOS circuits depends largely on W/L of the drive transistor and the capacitance of the load it is driving. That capacitance consists of:

- Input gates of cells being driven, and
- Capacitance of wiring

$\tau = R_{on} C_{load}$

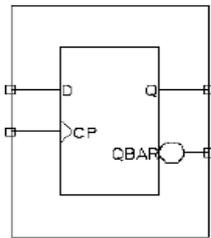
Cell Library Example

AμE Advanced Microelectronics
A Division of ITD

ms080cmosxCells
CMOSX 0.8 Micron
Standard Cell Library

D flip-flop

DDF



Logic Equation:
 $Q = [(D \& \text{CP}[\text{rise}])] | (Q' \& \text{CP}[\text{rise}])$
 $QBAR = !Q$

Truth Table

CP	Q	QBAR
01	1	!D
10	Q	QBAR

Size in microns (W x H):
63.4 x 45.6

Pin Capacitance (fF)

pin	best	typical	worst
CP	14.7	21.6	36.9
D	14.4	13.5	18.4
Q	5.01	10.6	11.7
QBAR	11.4	12.7	22.3

Delay Information

Data	Timing	best, 5.5V, -55C, load 0.35pF (t _{0.50} ns, t _{0.63} ns)			typical, 5V, 25C, load 0.35pF (t _{0.50} ns, t _{0.63} ns)			worst, 4.5V, 125C, load 0.35pF (t _{0.50} ns, t _{0.63} ns)			
		0.25 * load	1 * load	4 * load	0.25 * load	1 * load	4 * load	0.25 * load	1 * load	4 * load	
cp->q	01->01	PD	1.240	0.313	0.647	0.553	0.679	1.14	1.51	1.76	2.94
	TR	0.0997	0.236	0.811	0.144	0.354	1.24	0.381	0.668	2.32	
cp->q	01->01	PD	0.211	0.256	0.545	0.439	0.524	1.25	1.24	1.76	3.34
	TR	0.0174	0.228	0.817	0.148	0.577	1.89	0.456	1.23	4.86	
cp->q	01->01	PD	1.222	0.251	0.440	0.506	0.513	1.03	1.44	1.72	2.80
	TR	0.0669	0.154	0.505	0.140	0.347	1.24	0.364	0.901	3.24	
cp->q	01->10	PD	1.209	0.316	0.797	0.429	0.631	1.37	1.16	1.59	2.91
	TR	1.107	0.310	1.39	0.210	0.537	1.92	0.441	1.06	3.55	

Special Timing Information

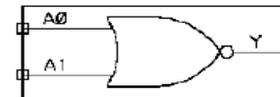
	best, 5.5V, -55C	typical, 5V, 25C	worst, 4.5V, 125C
Setup time on D	0.3	0.6	1.4
Hold time on D	0.1	0.05	0.01
Minimum pulse width low on CP	0.2	0.1	0.9
Minimum pulse width high on CP	1.08	0.2	0.6
Minimum period on CP	0.1	0.4	2.2
Maximum fall time on CP	4	38	3.8e+02

AμE Advanced Microelectronics
A Division of ITD

ms080cmosxCells
CMOSX 0.8 Micron
Standard Cell Library

2 Input NOR 1x drive

NOR2



Size in microns (W x H):
12.2 x 45.0

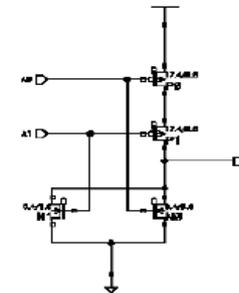
Logic Equation:
 $Y = !(A0 | A1)$

Pin Capacitance (fF)

pin	best	typical	worst
A0	15.5	24.2	37.1
A1	15.9	21.3	34.6
Y	5.81	10.1	13.6

Truth Table

A0	A1	Y
0	0	1
0	1	0
1	0	0
1	1	0

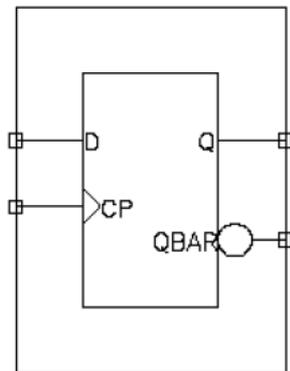


Delay Information

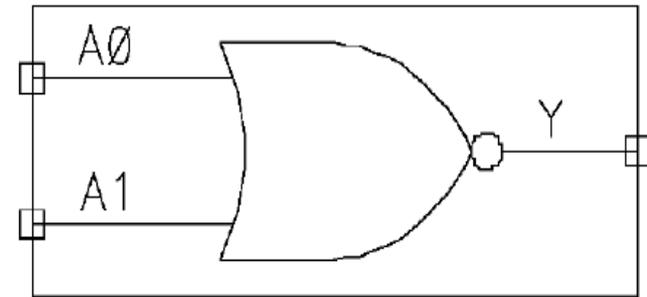
Data	Timing	best, 5.5V, -55C, load 0.35pF (t _{0.50} ns, t _{0.63} ns)			typical, 5V, 25C, load 0.35pF (t _{0.50} ns, t _{0.63} ns)			worst, 4.5V, 125C, load 0.35pF (t _{0.50} ns, t _{0.63} ns)			
		0.25 * load	1 * load	4 * load	0.25 * load	1 * load	4 * load	0.25 * load	1 * load	4 * load	
a1->y	01->01	PD	1.103	0.281	0.827	0.271	0.620	1.84	0.549	1.48	5.07
	TR	0.0975	0.243 * CL	0.188 + 1.21 * CL	0.913 + 3.27 * CL						
a1->y	01->10	PD	1.141	0.357	1.08	0.257	0.615	1.73	0.502	1.25	3.40
	TR	0.0885	0.638 * CL	0.186 + 1.19 * CL	0.420 + 2.17 * CL						
a0->y	01->01	PD	1.297	0.551	1.70	0.491	0.948	2.93	0.929	1.97	5.61
	TR	0.0842	0.225	0.767	0.212	0.549	1.79	0.636	1.50	4.85	
a0->y	01->10	PD	1.286	0.485	1.77	0.451	1.17	3.04	1.10	2.81	10.1
	TR	0.114 + 0.657 * CL	0.230 + 1.01 * CL	0.515 + 2.12 * CL							

Load on flip-flop

- = output capacitance of flip-flop + input capacitance of NOR gate



Pin Capacitance (fF)			
pin	best	typical	worst
CP	14.7	21.6	30.9
D	10.4	13.5	18.4
Q	5.04	10.6	11.7
QBAR	11.4	12.7	22.3



Pin Capacitance (fF)			
pin	best	typical	worst
A0	16.5	24.2	37.1
A1	15.9	23.3	34.6
Y	5.54	10.1	13.6

Timing Tables

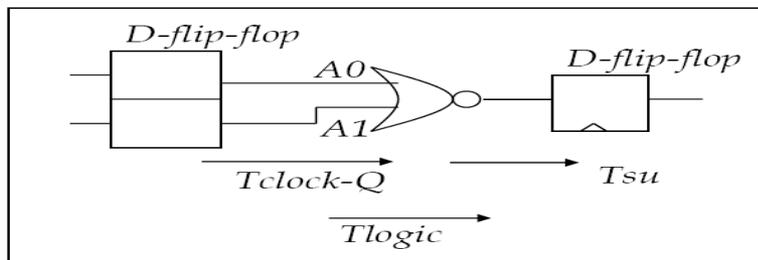
- Predict delay
 - Note delay different for rising and falling edges

Delay Information

Path	Timing		best, 5.5V, -55C, load:0.35pF tr:0.584ns, tf:0.638ns			typical, 5V, 25C, load:0.35pF tr:1.15ns, tf:1.12ns			worst, 4.5V, 125C, load:0.35pF tr:2.78ns, tf:2.27ns		
			0.25 * load	1 * load	4 * load	0.25 * load	1 * load	4 * load	0.25 * load	1 * load	4 * load
cp->qbar	01->10	PD	0.249	0.315	0.647	0.553	0.679	1.18	1.52	1.76	2.61
			0.213 + 0.308 * CL			0.517 + 0.478 * CL			1.46 + 0.831 * CL		
		TR	0.0697	0.206	0.811	0.144	0.354	1.29	0.281	0.668	2.32
cp->q	01->01	PD	0.211	0.266	0.545	0.459	0.624	1.25	1.24	1.70	3.34
			0.180 + 0.258 * CL			0.416 + 0.609 * CL			1.12 + 1.59 * CL		
		TR	0.0874	0.229	0.847	0.198	0.517	1.89	0.456	1.33	4.86
cp->qbar	01->01	PD	0.222	0.258	0.440	0.506	0.613	1.03	1.44	1.72	2.80
			0.202 + 0.169 * CL			0.475 + 0.403 * CL			1.35 + 1.03 * CL		
		TR	0.0669	0.154	0.565	0.140	0.347	1.26	0.364	0.901	3.24
cp->q	01->10	PD	0.209	0.316	0.797	0.429	0.631	1.37	1.16	1.59	2.91
			0.162 + 0.451 * CL			0.378 + 0.714 * CL			1.09 + 1.32 * CL		
		TR	0.107	0.310	1.19	0.210	0.537	1.92	0.441	1.06	3.55

Data Sheet Example

- Using timing approximations in the datasheet, what is the maximum clock frequency for this circuit (ignore wire load, T_{skew}):



$T_{clock-Q}$:

cp->q	01->01	PD	1.24	1.70	3.34
		TR	1.12 + 1.59 * CL		
			0.456	1.33	4.86
cp->q	01->10	PD	1.16	1.59	2.91
		TR	1.09 + 1.32 * CL		
			0.441	1.06	3.55

CL :

Pin Capacitance (fF)

pin	best	typical	worst
A0	16.5	24.2	37.1
A1	15.9	23.3	34.6
Y	5.54	10.1	13.6

DFF

Pin Capacitance (fF)

pin	best	typical	worst
CP	14.7	21.6	30.9
D	10.4	13.5	18.4
Q	5.04	10.6	11.7
QBAR	11.4	12.7	22.3

Q -> A0:

$$CL_{max} = 0.0371 + 0.0117 \text{ pF} = 0.0488 \text{ pF}$$

$$T_{cp-Q}_{max} = \max(1.12 + 1.59 * CL, 1.09 + 1.32 * CL)$$

$$= \max(1.12 + 1.59 * 0.0488, 1.09 + 1.32 * 0.0488)$$

$$= \max(1.2, 1.15) = 1.2 \text{ ns}$$

$$Q \rightarrow A1: T_{cp-Q}_{max} = 1.2 \text{ ns}$$

...Example

Tlogic:

			worst, 4.5V, 125C, load:0.35pF tr:2.78ns, tf:2.27ns		
			0.25 * load	1 * load	4 * load
a1->y	10->01	PD	0.749	1.68	5.07
		TR	0.503 + 3.27 * CL		
a1->y	01->10	PD	0.502	1.25	3.40
		TR	0.420 + 2.17 * CL		
a0->y	10->01	PD	0.636	1.50	4.85
		TR	0.371 + 3.21 * CL		
a0->y	01->10	PD	0.609	1.31	3.43
		TR	0.505 + 2.12 * CL		

CL:

Pin Capacitance (fF) NOR2

pin	best	typical	worst
A0	16.5	24.2	37.1
A1	15.9	23.3	34.6
Y	5.54	10.1	13.6

DFE

Pin Capacitance (fF)

pin	best	typical	worst
CP	14.7	21.6	30.9
D	10.4	13.5	18.4
Q	5.04	10.6	11.7
QBAR	11.4	12.7	22.3

NOR2

$$CL = 0.0136 + 0.0184 = 0.032 \text{ pF}$$

$$\text{From A0 : } T_{logic} = \max(0.503 + 3.27 * 0.032, 0.420 + 2.17 * 0.032) \\ = 0.61 \text{ ns}$$

$$\text{From A1 : } T_{logic} = \max(0.505 + 2.12 * 0.032, 0.371 + 3.21 * 0.032) \\ = 0.57 \text{ ns}$$

...Example

T_{su}

Special Timing Information

	best, 5.5V, -55C	typical, 5V, 25C	worst, 4.5V, 125C
Setup-time on D	0.3	0.6	1.4
Hold-time on D	0.1	0.05	0.01
Minimum-pulse-width-low on CP	0.2	0.3	0.9
Minimum-pulse-width-high on CP	0.08	0.2	0.6
Minimum-period on CP	0.4	0.8	2.2
Maximum-fall-time on CP	4	39	3.8e+02



$$\begin{aligned} T_{\text{clock}} &> T_{\text{cp-Q_max}} + T_{\text{logic_max}} + T_{\text{su_max}} \\ &= 1.2 + 0.61 + 1.4 \\ &= 3.21 \text{ ns} \end{aligned}$$

$$F_{\text{clock}} < 311 \text{ MHz}$$

Estimating and Improving Performance

- With a focus on timing:
- Topics:
 - Metrics : FO-4
 - Typical timing budgets
 - Pipelining and Parallelism
 - Logic style

Delay Metric

- Usual Metric for delay:
 - Fanout of 4 inverter delay: FO4
- Estimating FO4:
 - Typical $\sim 360 \times L_{\text{eff}}$ (ps)
 - Worst Case $\sim 600 \times L_{\text{eff}}$ (ps)
 - L_{eff} = Effective gate length in $\mu\text{m} \sim 0.7 \times L_{\text{drawn}}$
 - E.g. In a 0.18 μm process, $L_{\text{eff}} = 0.126 \mu\text{m}$ and FO-4 ≤ 75 ps

- Exemplar delays:

Inverter = FO-4

1-bit adder = 10 FO4

Flip-flop $t_{\text{cp-Q}}$ = 4 FO

Clock skew = 4 FO4

Clock jitter = 2 FO4

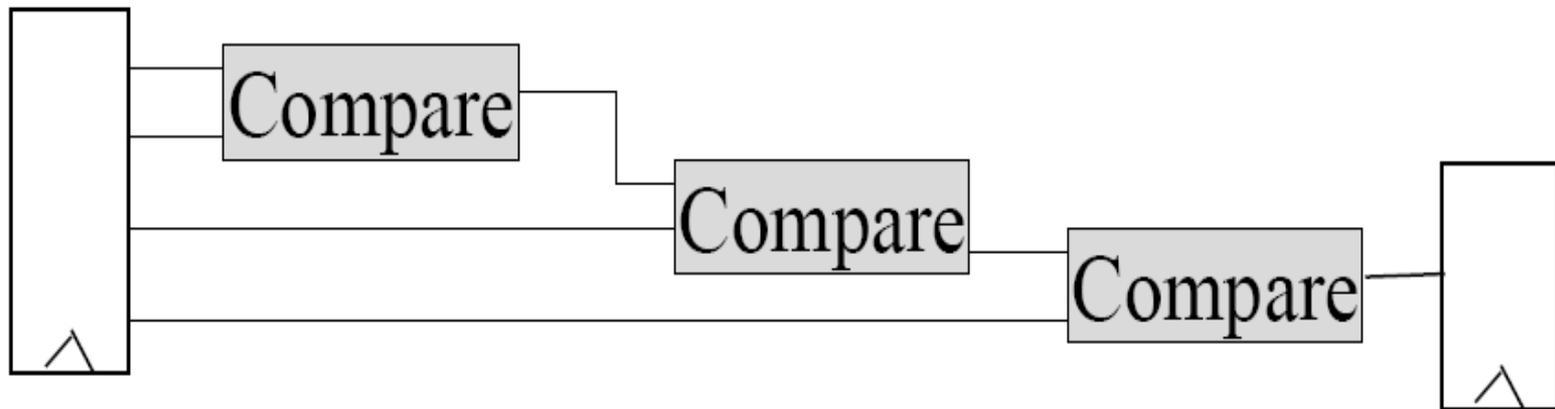
2-input NAND gate = 2FO4

2-input Multiplexer = 4 FO4

4 Flip-flop $t_{\text{su}} / t_{\text{h}}$ = 2 FO4

Examples of Improving Timing Performance

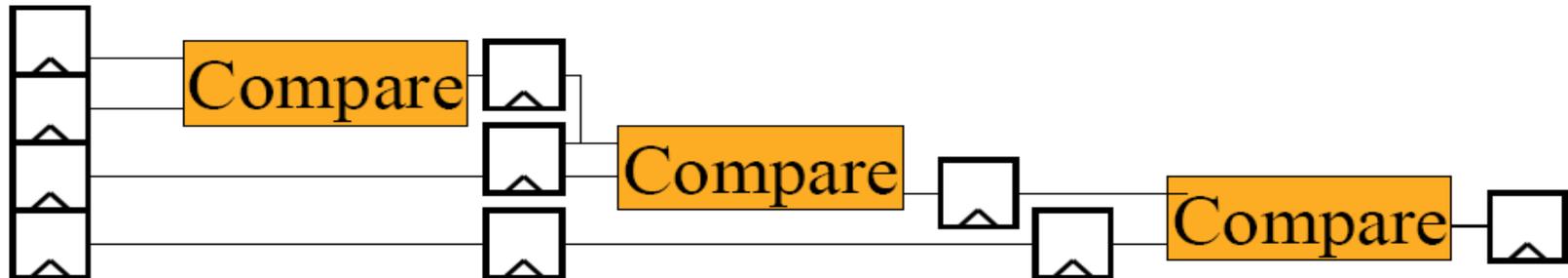
- Example 1 : Benefits of Pipelining and Parallelism
- Example:



If $t_{\text{comparator}} = 20 \text{ FO4}$, what is the clock period?
(Use values on previous page)

Pipelining

- Replace with:

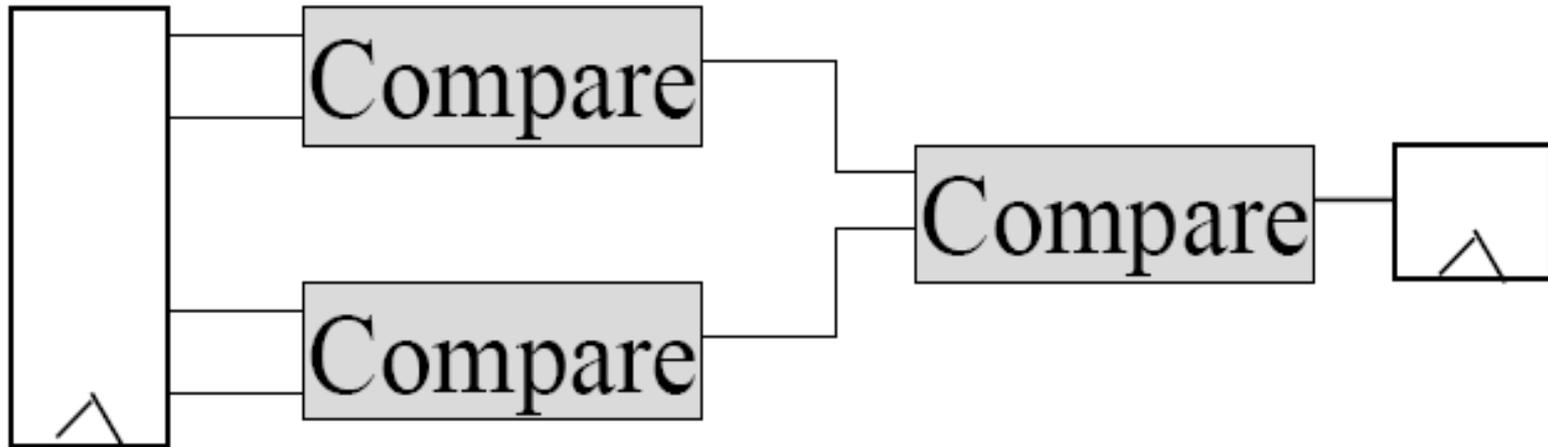


$$T_{cp} = t_{ck-Q} + t_{logic} + t_{su} + t_{skew} + t_{jitter}$$
$$= 4 + 20 + 2 + 4 + 2 = 32 \text{ FO4}$$

- What is the delay improvement?
- What is the drawback?

Logic Level Parallelism

- Replace with:



- Clock Period = 52 FO-4
- No increase in area

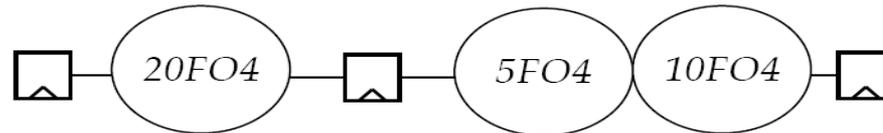
Retiming

- Impact of critical paths can often be reduced by retiming or rebalancing a design:
- Example:
 - Before:



$$T_{cp} = 4 + 20 + 5 + 2 + 4 + 2 = 37 FO4$$

- After:



$$T_{cp} = 4 + 20 + 2 + 4 + 2 = 32 FO4$$

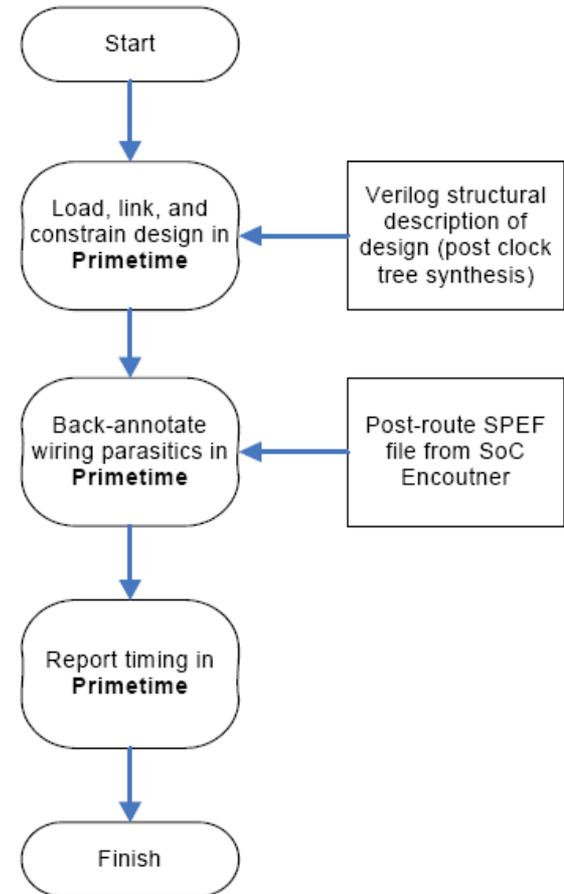
Note: Clock level logic sequence has been changed

Timing in CAD Flow

- During synthesis
 - Tool calculates path delays under worst-case delay conditions
 - Determines critical path
 - Moves logic to a faster path if setup violation predicted
- After synthesis:
 - Perform Static Timing Analysis
 - Determine no setup violations exist under worst case conditions
 - Determine no hold violations exist under best case conditions
- After place and route:
 - Perform Timing Analysis
 - i.e. Run timing verification tools on netlist with actual delays
 - Back-annotate actual delays to netlist from later tools

Initial Delay Estimation Flow

- Primetime: Gate-level static timing analysis tool
 - Report timing for critical path
- Back-annotate wire parasitics for more accuracy
 - SPEF file from place and route tool



Sidebar

Not Examinable!

- What is asynchronous design?
- Can we use deliberate local clock skew in a design?
- Are you sure flip-flops are better, cycle stealing sounds useful?

Message: The CAD tools' capabilities constrain your design flexibility.

Summary

1. What determines the maximum clock frequency?
2. What is a hold violation?
3. Why do we prefer flip-flop designs over using latches?
4. What tool is used to check timing at design closure?

Remember

Methodology for purposes of ASIC Design class

- If at all possible one-edge of one clock
- If you need multiple clocks, they must have a common root and be related by factors of 2
 - E.g. Root clock : Tclock = 5 ns
 - This is BEST: Tclock only!!
 - This is OK:
 - Tclock, Tclock10, Tclock20
 - Tclock10 = 10 ns (Tclock*2)
 - Tclock20 = 20 ns (Tclock*4)
 - This is NOT OK
 - Tclock, Tclock15, Tclock17
 - Tclock15 = 15 ns (Tclock*3)
 - Tclock17 = 17 ns (??)