EE241B : Advanced Digital Circuits

Lecture 9 – Timing

Borivoje Nikolić

February 14, 2020, EETimes: Five Chip Companies Hold 53% of Global Wafer Capacity

An increasing percentage of the world's capacity is getting concentrated in the hands of the largest producers.

Worlwide Wafer Capacity Leaders
(Monthly Installed Capacity in Dec 2019, 200mm-equivalents)

2019 Rank	2018 Rank	Company	Headquarters Region	Dec-2018 Capacity (K w/m)	Dec-2019 Capacity (K w/m)	Yr/Yr Change	Share of Worldwide Total	Inclusion or Exclusion of Capacity Shares from JV Fabs
1	1	Samsung	South Korea	2,934	2,935	0%	15.0%	
2	2	TSMC	Taiwan	2,439	2,505	3%	12.8%	shares of SSMC & VIS
3	3	Micron	North America	1,685	1,841	9%	9.4%	share of IM Flash in '18
4	4	SK Hynix	South Korea	1,630	1,743	7%	8.9%	
5	5	Kioxia/WD	Japan	1,361	1,406	3%	7.2%	

Source: Companies, IC Insights' *Global Wafer Capacity 2020-2024 Report*

Announcements

- Project abstracts due today, by e-mail
  - Teams of 2
  - Title
  - One paragraph
  - 5 relevant references
- Can also combine with CS252 or EE290 projects
- Quiz 1 on Tuesday, Feb 25, in class
- Office hour moved to 11am on Monday

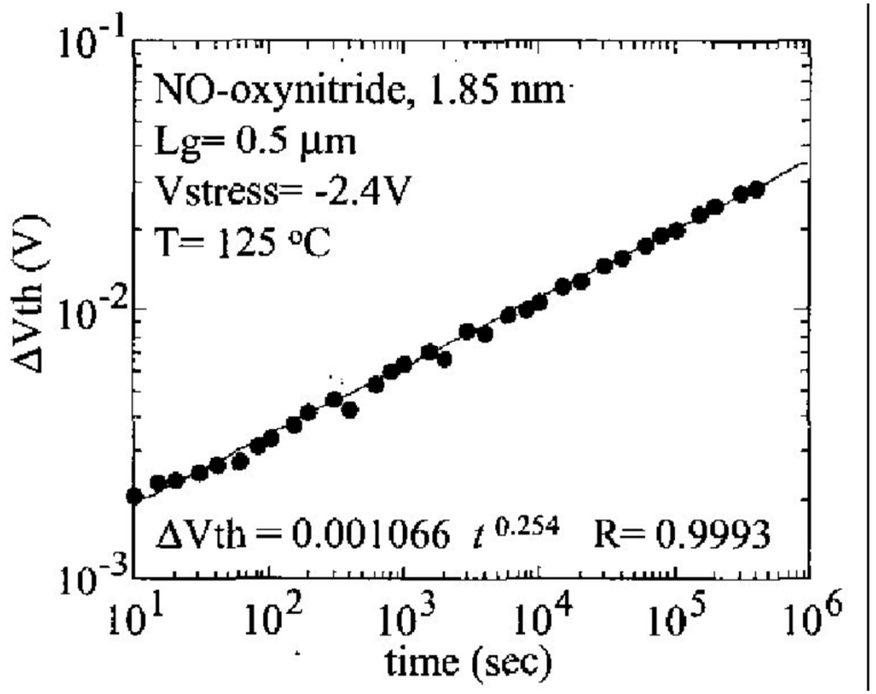
Outline

- ISSCC recap
- Module 2
  - Technology variability
- Module 3
  - Flip-flop timing

2.P Design Variability
Some Random Effects

Negative Bias Temperature Instability

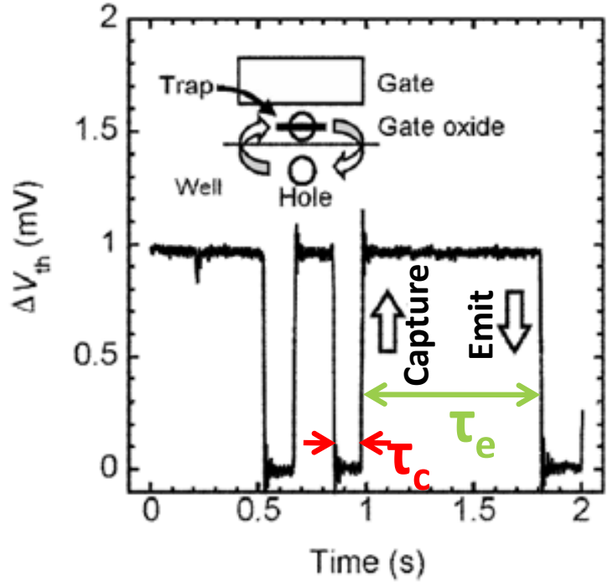
- PFET V_{Th}'s shift in time, at high negative bias and elevated temperatures
- The mechanism is thought to be the breaking of hydrogen-silicon bonds at the Si/SiO2 interface, creating surface traps and injecting positive hydrogen-related species into the oxide.
- Also other charge trapping and hot-carrier defect generation
- Systematic + random shifts



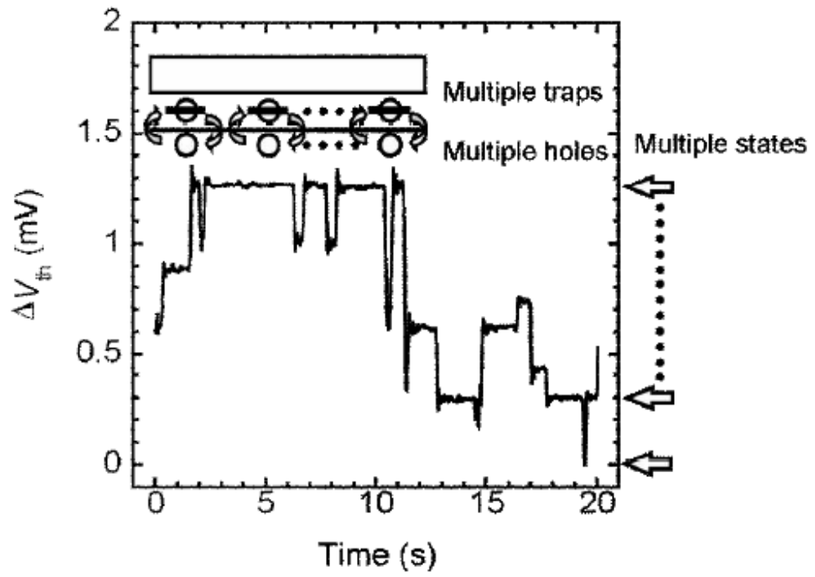
Tsujikawa, IRPS'2003

Random Telegraph Signal (RTS)

Single Hole Trap



Multiple Hole Trap



- Trapping of a carrier in oxide traps modulates V_{th} or I_{ds}
- τ_e and τ_c are random and follow exponential distributions

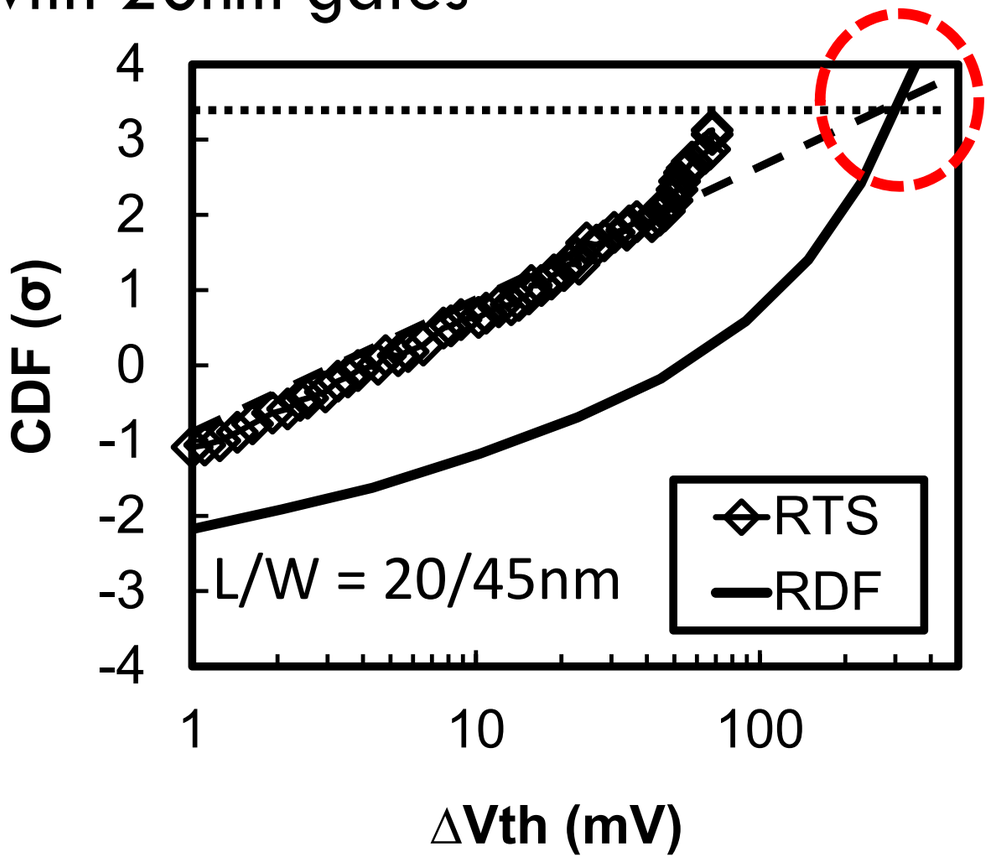
N. Tega et al. IRPS 2008.

RTS and Technology Scaling

- RTS exceeds RDF at 3 sigma with 20nm gates

$$\Delta V_{th,\,RTS} \sim \frac{1}{WL}$$

$$\Delta V_{th,\,RDF} \sim \frac{1}{\sqrt{WL}}$$

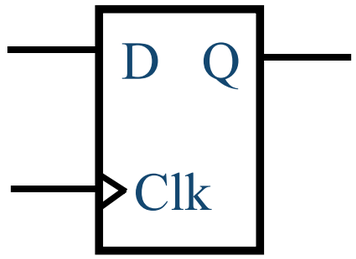


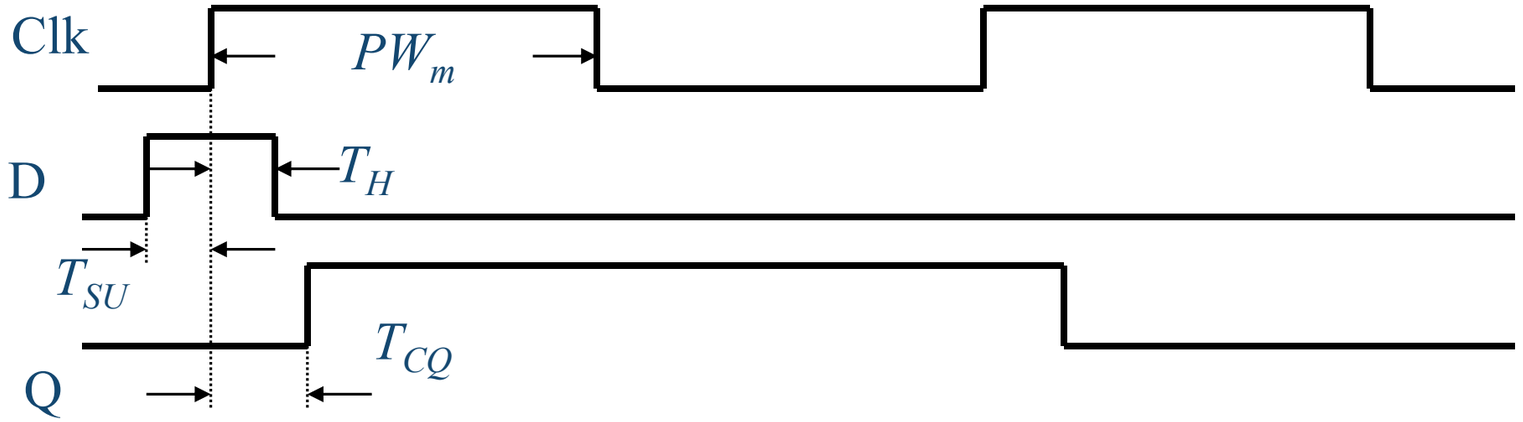
Tega *et. al*, VLSI Tech. 09

3. Design for Performance

3.A Flip-Flop Timing

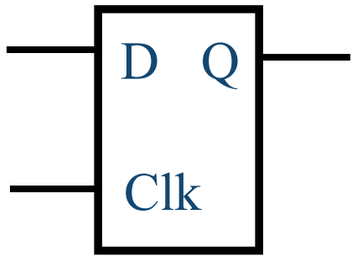
Flip-Flop Parameters



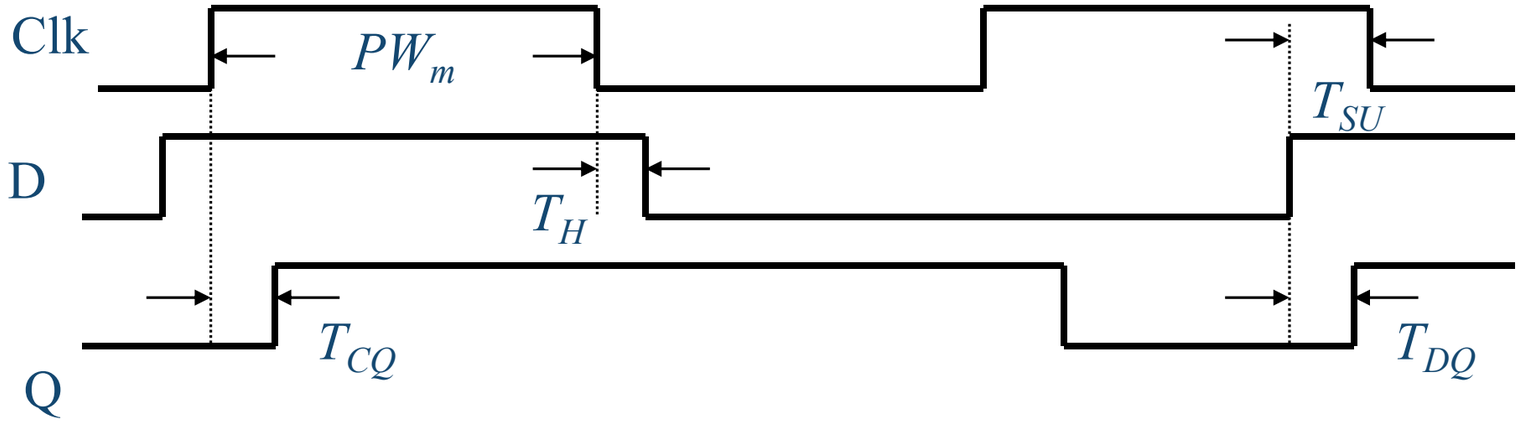


Delays can be different for rising and falling data transitions

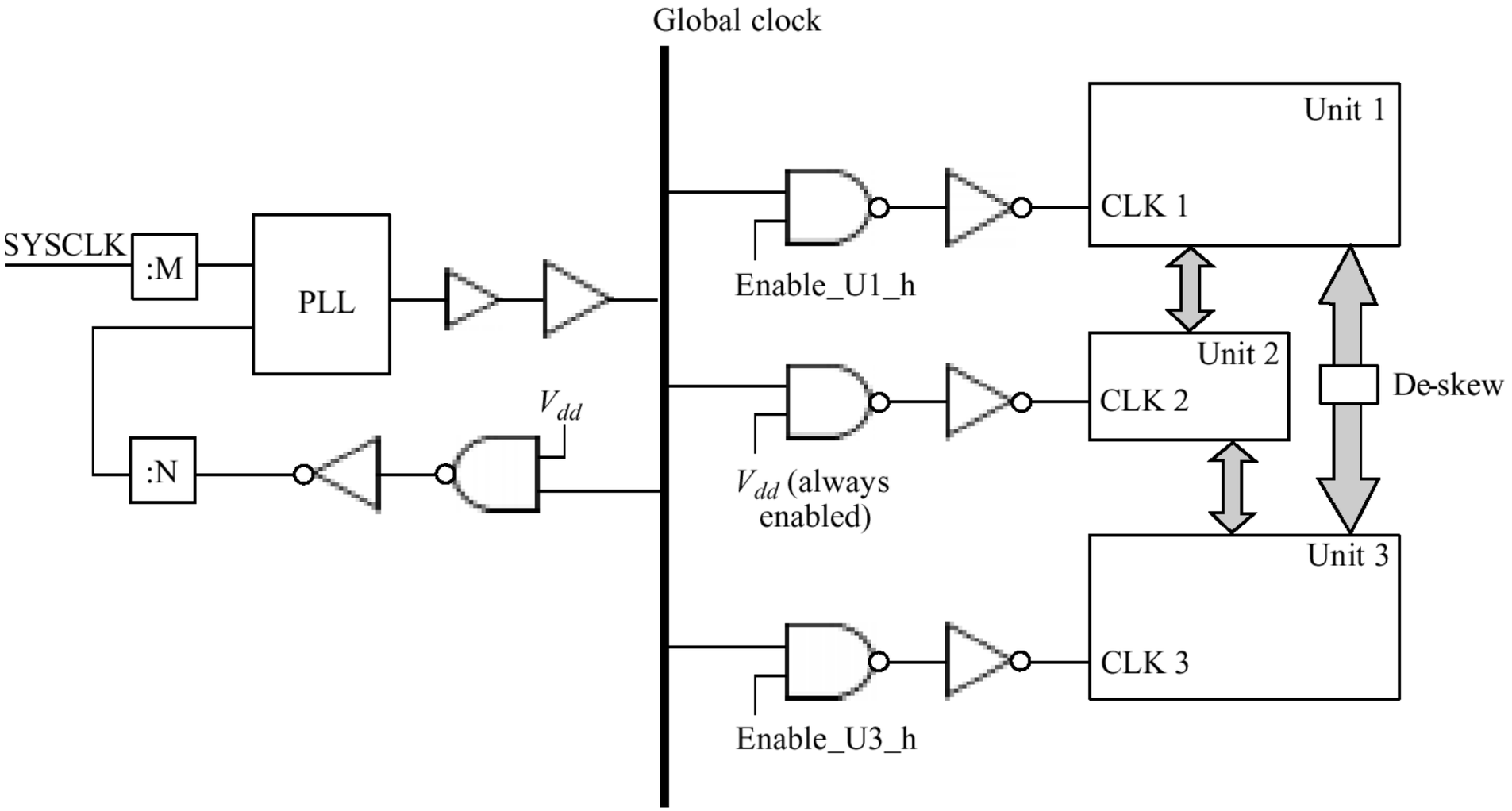
Latch Parameters



Unger and Tan
Trans. on Comp.
10/86



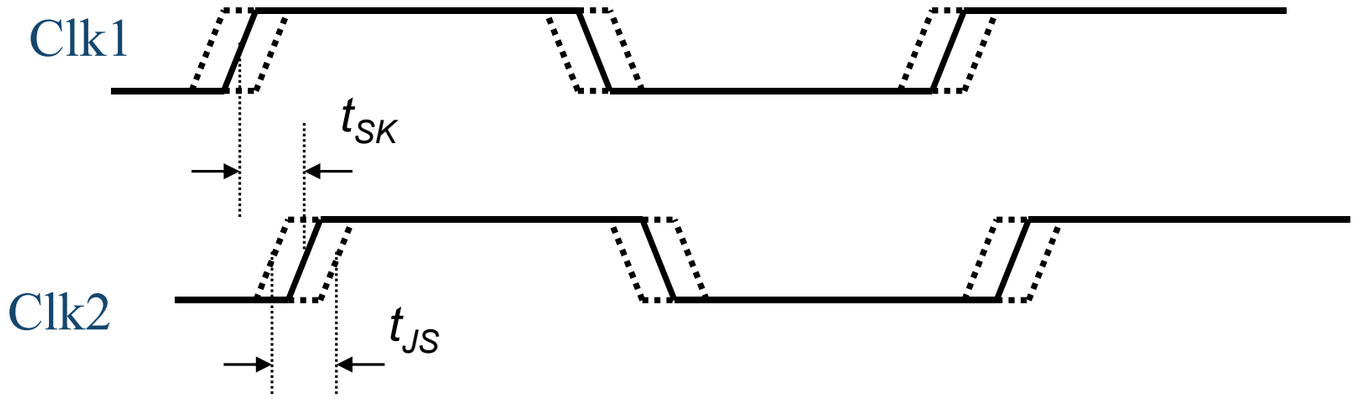
Delays can be different for rising and falling data transitions


Global clock
SYSCLK
:M
PLL
:N
V_{dd}
Enable_U1_h
V_{dd} (always enabled)
Enable_U3_h
Unit 1
CLK 1
Unit 2
CLK 2
Unit 3
CLK 3
De-skew

Clock Nonidealities

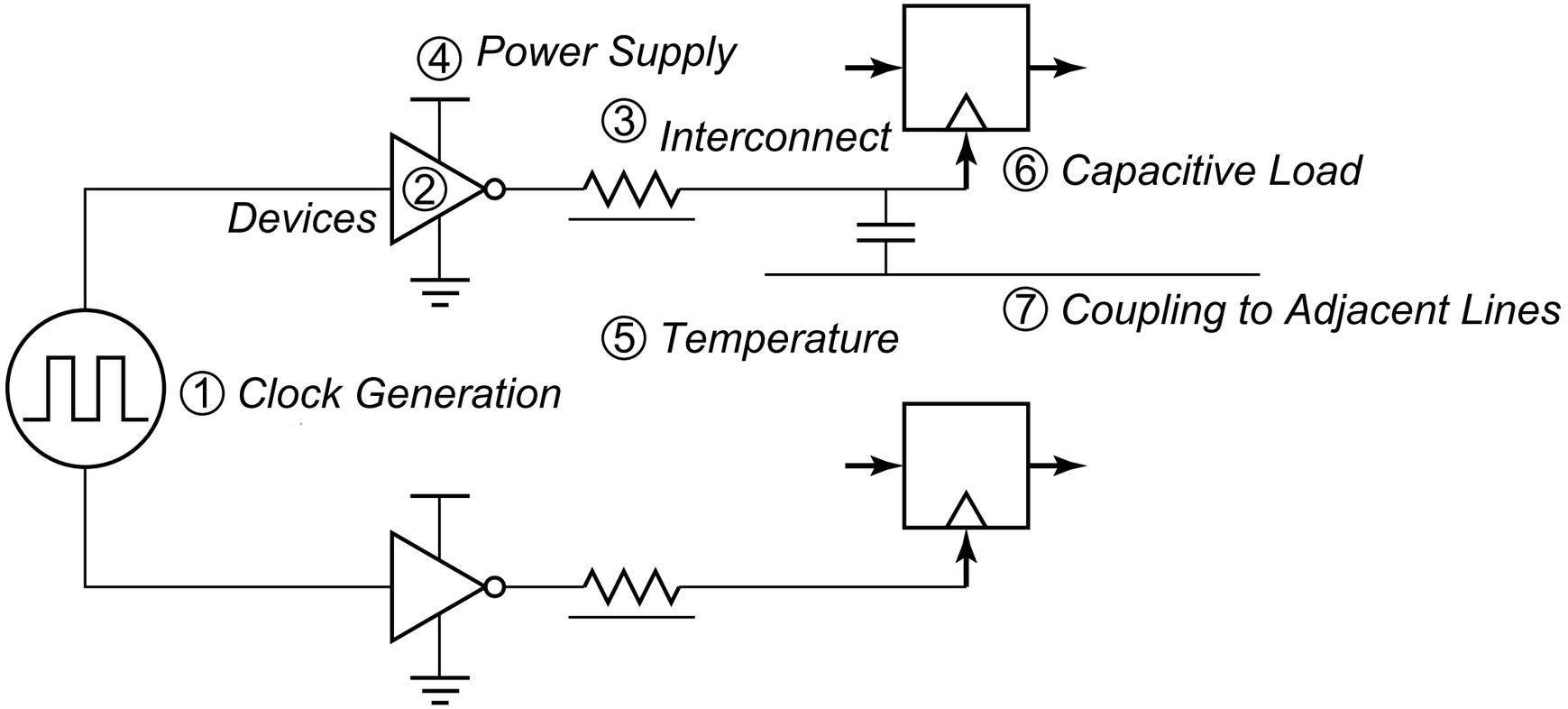
- Clock skew
  - Spatial variation in temporally equivalent clock edges; deterministic + random, t_{SK}
- Clock jitter
  - Temporal variations in consecutive edges of the clock signal; modulation + random noise
  - Cycle-to-cycle (short-term) - t_{JS}
  - Long-term - t_{JL}
- Variation of the pulse width
  - for level-sensitive clocking

Clock Skew and Jitter



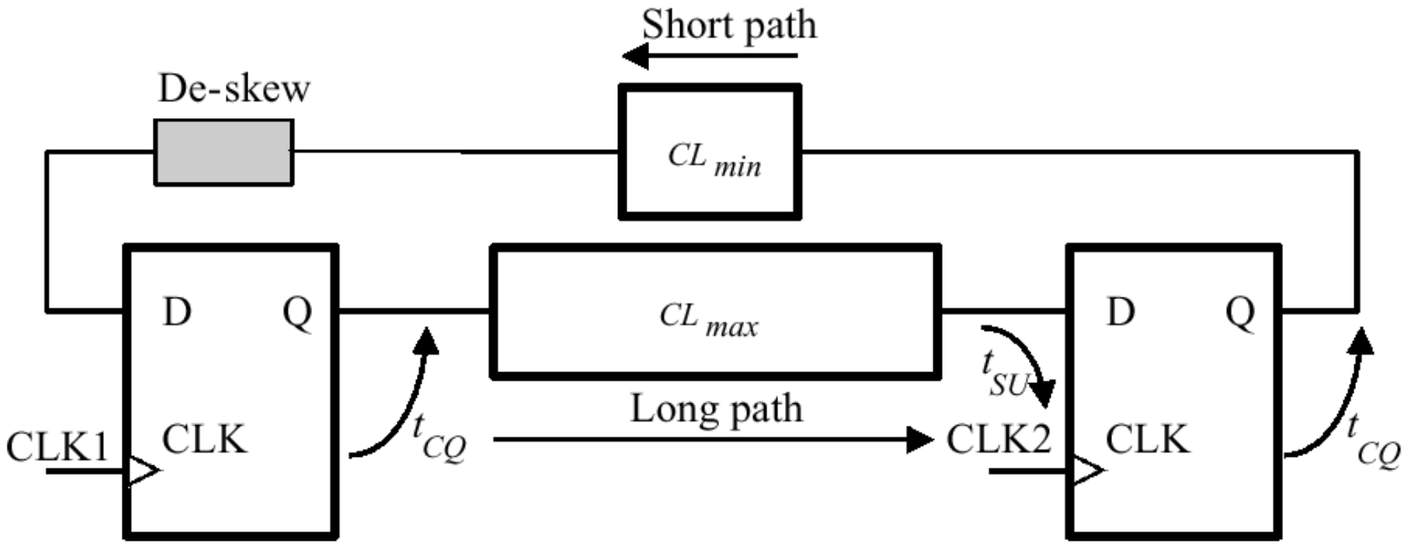
- Both skew and jitter affect the effective cycle time
- Only skew affects the race margin, if jitter is from the source
  - Distribution-induced jitter affects both

Clock Uncertainties



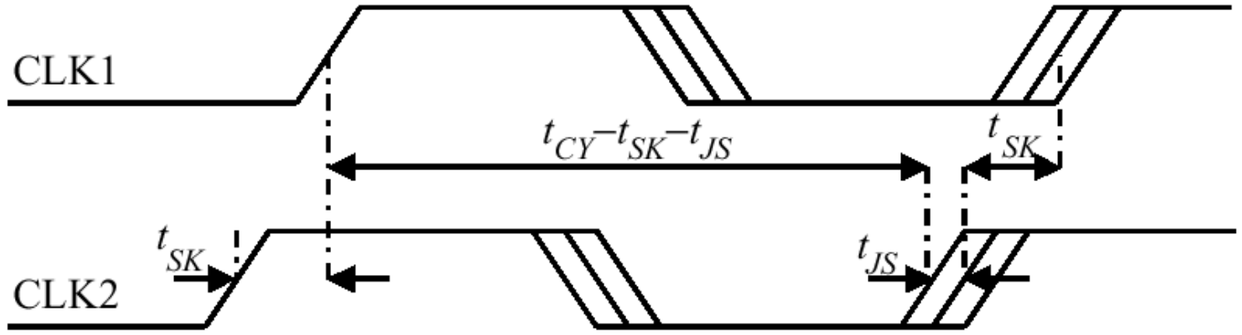
Sources of clock uncertainty

Clock Constraints in Edge-Triggered Systems



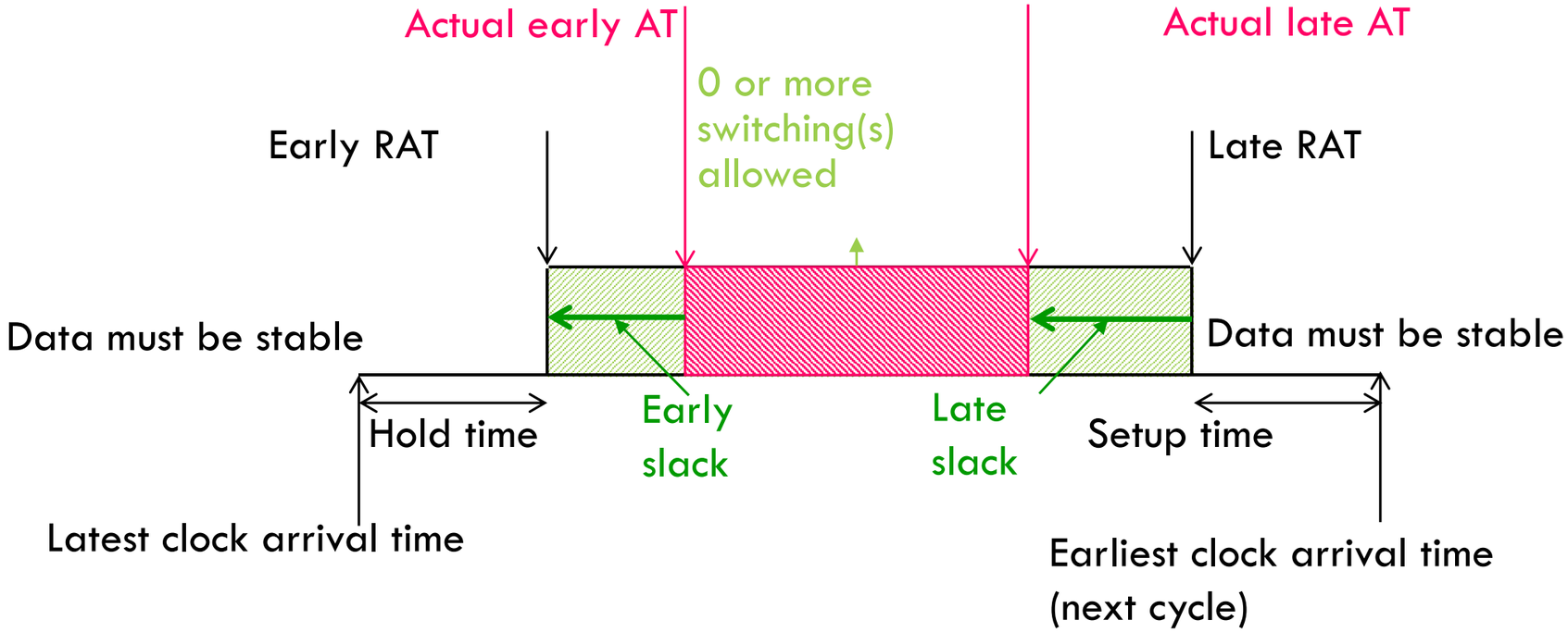
$$t_{CL} \le (t_{CY} - t_{SK} - t_{JS}) - (t_{SU} + t_{CQ})$$

$$t_{CL} \ge t_{SK} + (t_H - t_{CQ})$$



3.B Timing with Uncertainty/Variations

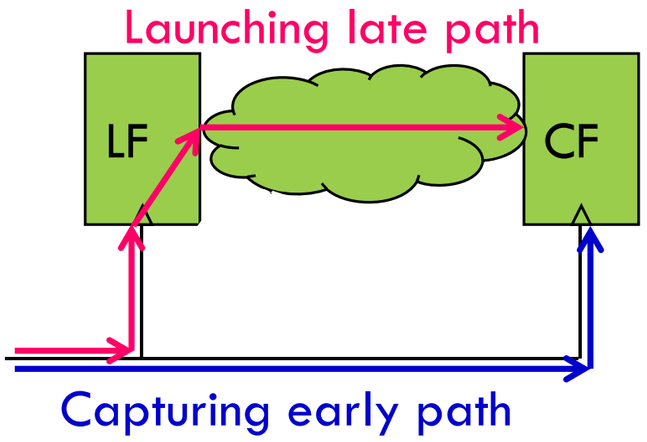
Pictorial View of Setup and Hold Tests



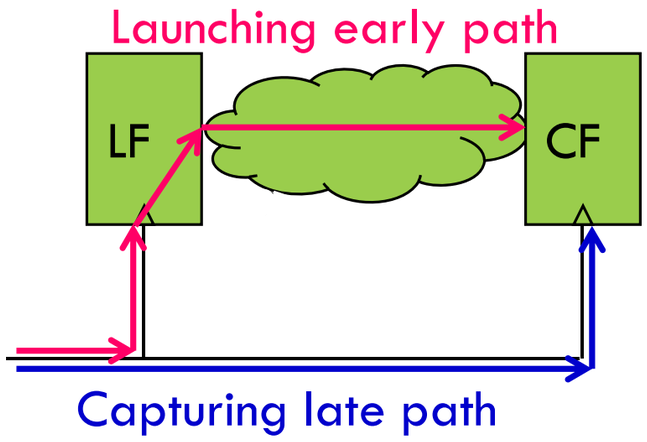
Handling of Across-Chip Variation

- Each gate has a range of delay: [lb, ub]
  - The lower bound is used for early timing
  - The upper bound is used for late timing
- This is called an early/late split
- Static timing obtains bounds on timing slacks
  - Timing is performed as one forward pass and one backward pass

Setup test



Hold test

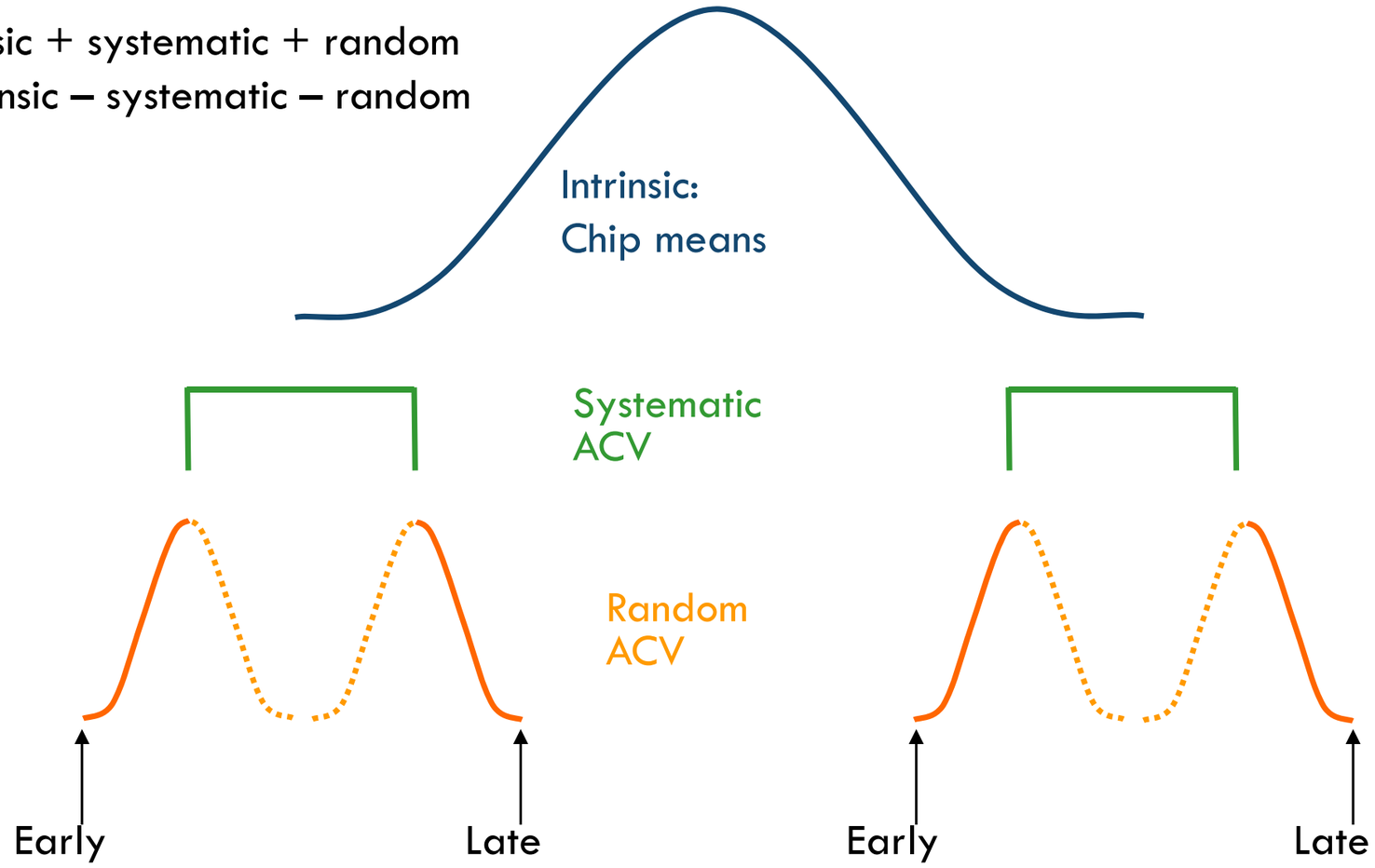


How is the Early/Late Split Computed?

- The best way is to take known effects into account during characterization of library cells
  - History effect, simultaneous switching, pre-charging of internal nodes, etc.
  - This drives separate characterization for early and late; this is the most accurate method
- Failing that, the most common method is derating factors
  - Example: Late delay = library delay * 1.05
    Early delay = library delay * 0.95
- The IBM way of achieving derating is LCD factors (Linear Combination of Delay) (FC=fast chip, SC=slow chip, see next page)
  - Late delay = α_L * FC_delay + β_L * NOM_delay + γ_L * SC_delay
    Early delay = α_E * FC_delay + β_E * NOM_delay + γ_E * SC_delay
  - Across-chip variation is therefore assumed to be a fixed proportion of chip-to-chip variation for each cell type

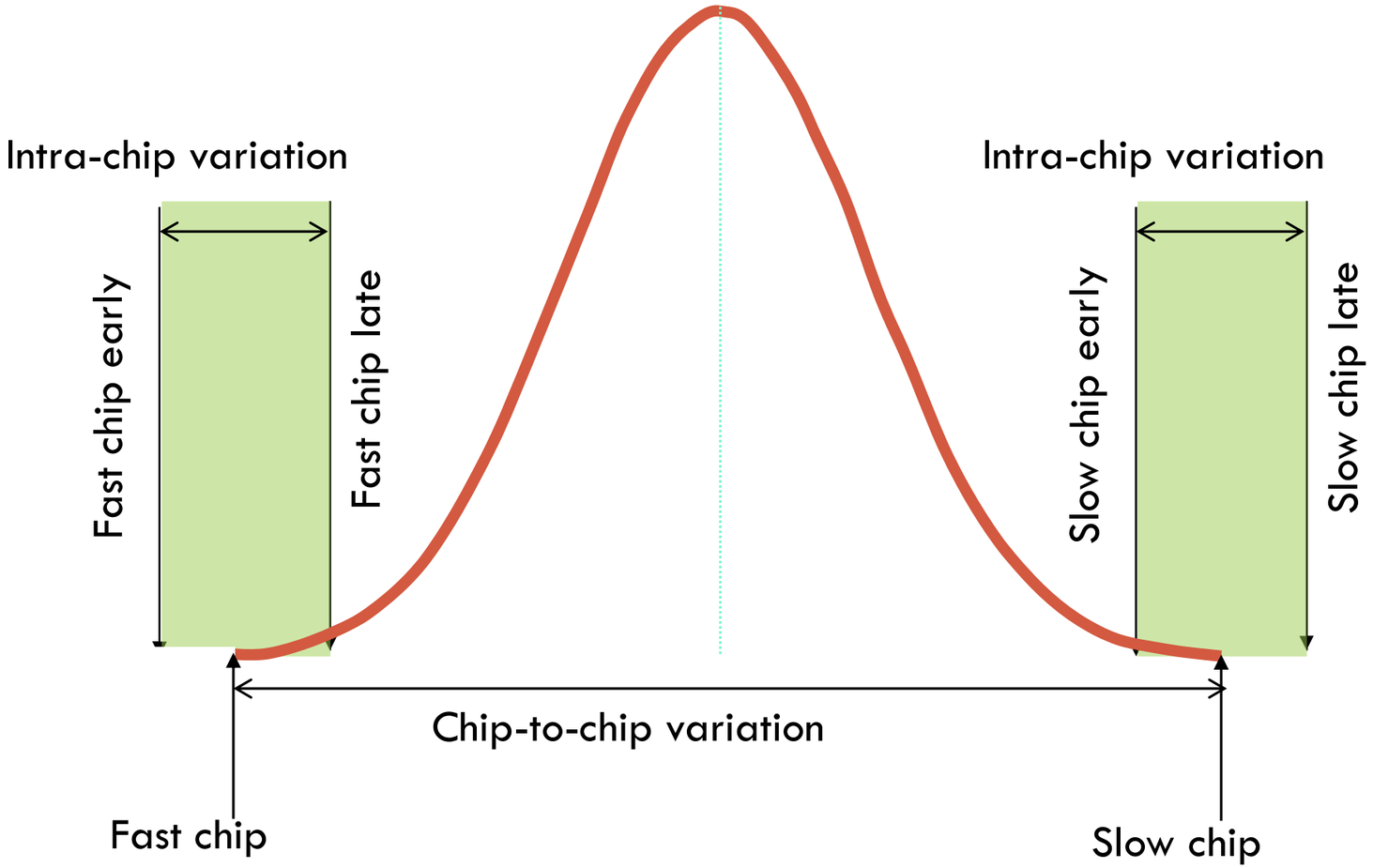
IBM Delay Modeling*

At a given corner
late delay = intrinsic + systematic + random
early delay = intrinsic – systematic – random



*P. S. Zuchowski, ICCAD'04

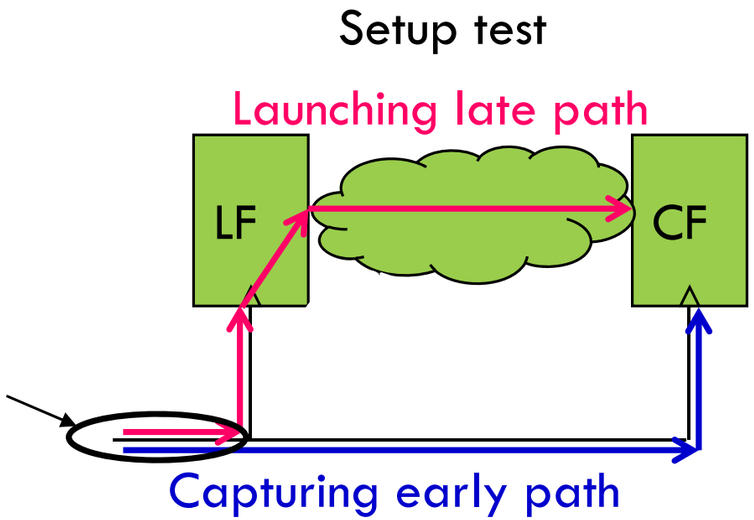
Traditional Timing Corners



The Problem with an Early/Late Split

- The early/late split is very useful
  - Allows bounds during delay modeling
  - Any unknown or hard-to-model effect can be swept under the rug of an early/late split
- But, it has problems
  - Additional pessimism (which may be desirable)
  - Unnecessary pessimism (which is never desirable)

This physically common portion can't be both fast and slow at the same time

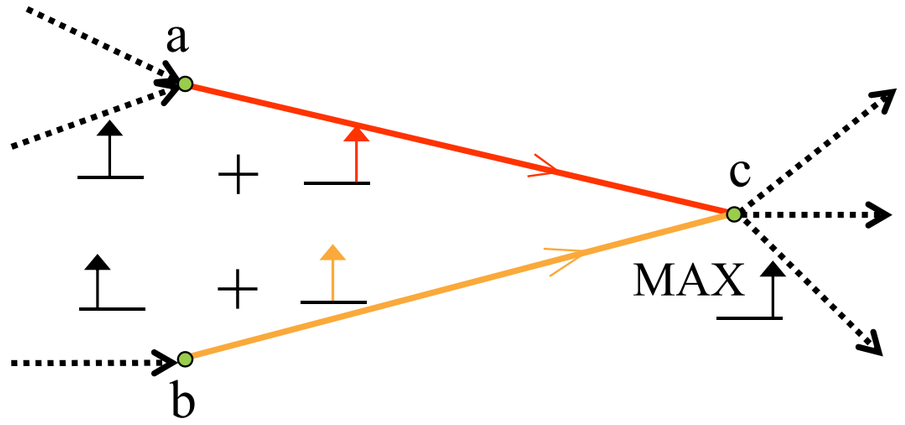


How to Have Less Pessimism?

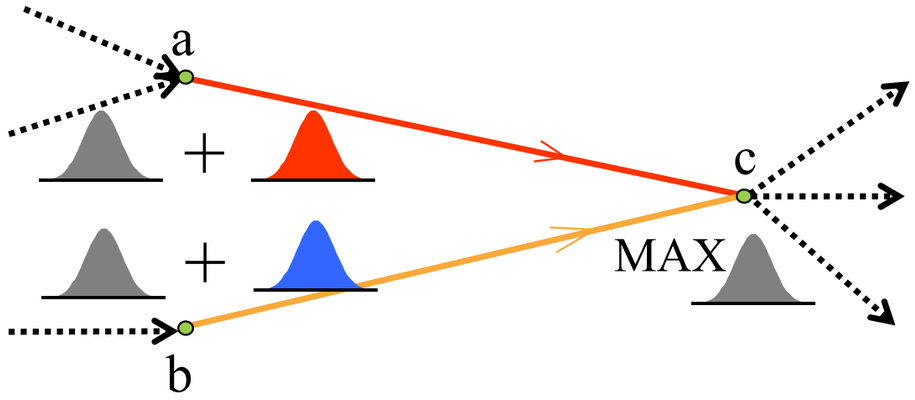
- Common path pessimism removal
- Account for correlations
- Credit for statistical averaging of random

Statistical Timing

- Deterministic



- Statistical



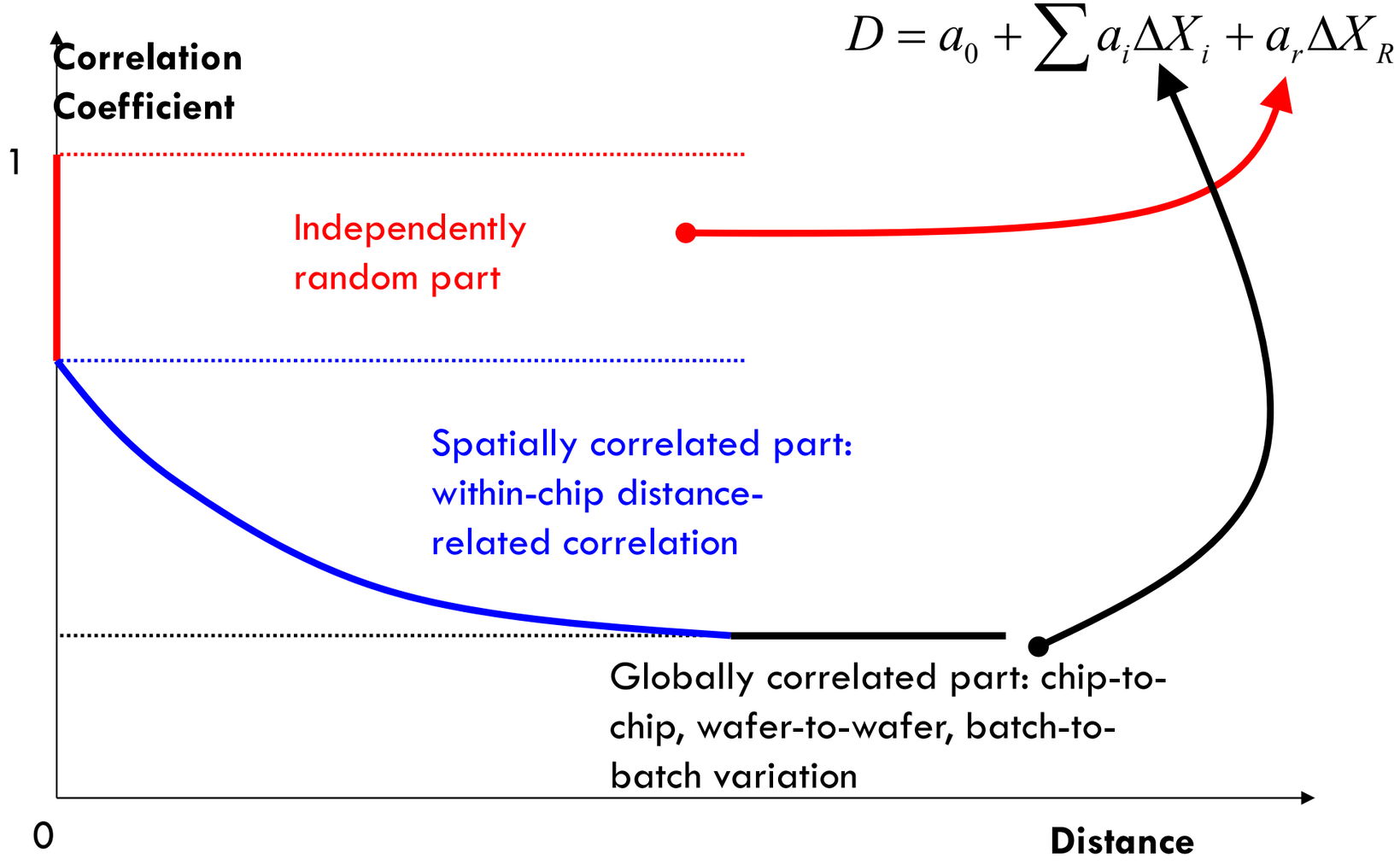
Statistical Max Operation

$$
\begin{aligned}
A &= a_0 + \sum_{i=1}^{n} a_i \Delta X_i + a_{n+1} \Delta R_a \\
B &= b_0 + \sum_{i=1}^{n} b_i \Delta X_i + b_{n+1} \Delta R_b \\
\sigma_A &= \sqrt{\sum_{i=1}^{n+1} a_i^2} \\
\sigma_B &= \sqrt{\sum_{i=1}^{n+1} b_i^2} \\
\rho &= \frac{\sum_{i=1}^{n} a_i b_i}{\sigma_A \sigma_B} \\
\theta &\equiv (\sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B)^{1/2} \\
t &= \Phi\left[\frac{a_0 - b_0}{\theta}\right] \\
E[max(A,B)] &= a_0 t + b_0(1-t) + \theta\phi\left[\frac{a_0 - b_0}{\theta}\right] \\
E[max(A,B)]^2 &= (\sigma_A^2 + a_0^2)t + (\sigma_B^2 + b_0^2)(1-t) + (a_0 + b_0)\theta\phi\left[\frac{a_0 - b_0}{\theta}\right]
\end{aligned}
$$

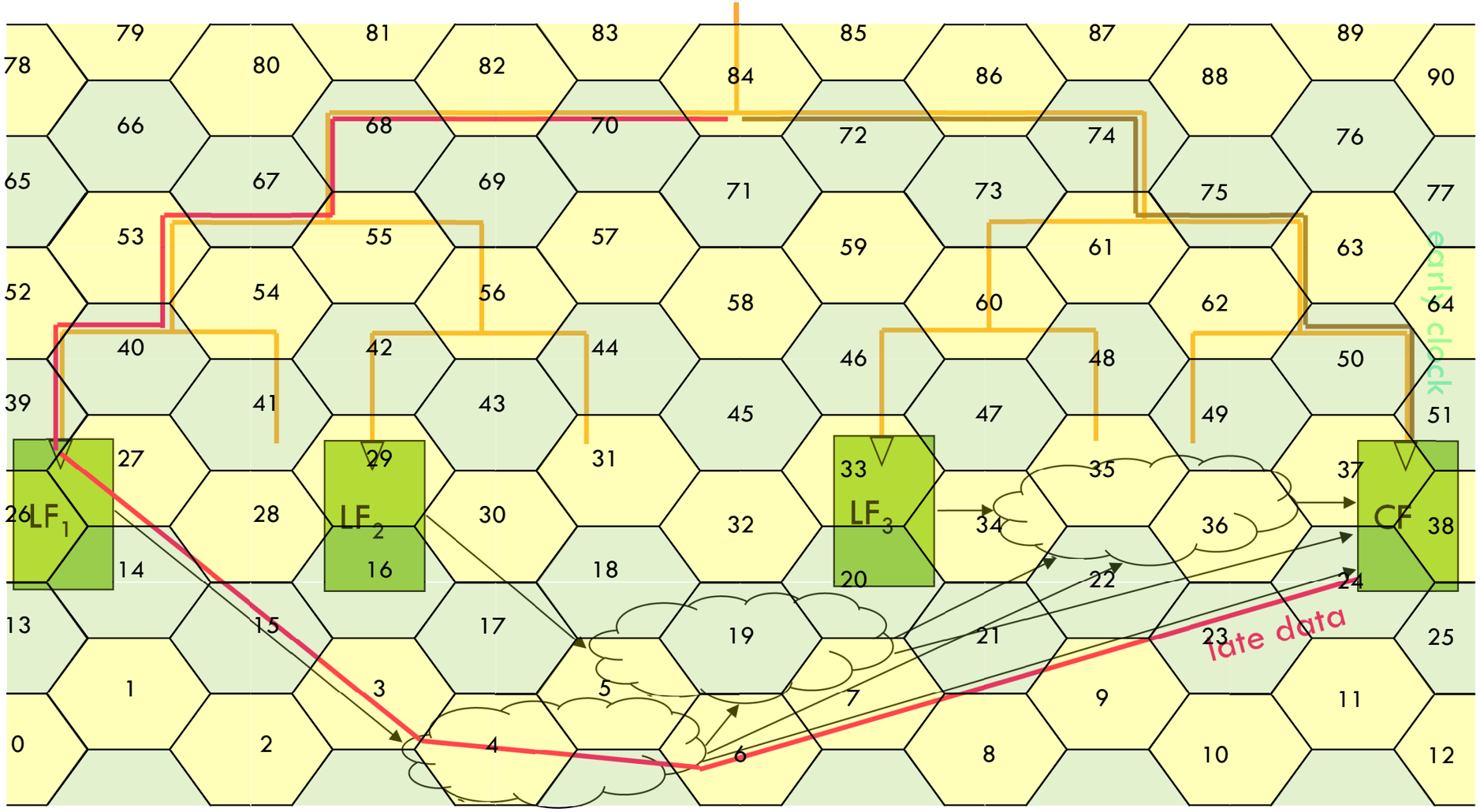
*C. E. Clark, “The greatest of a finite set of random variables,” OR Journal, March-April 1961, pp. 145—162

**M. Cain, “The moment-generating function of the minimum of bivariate normal random variables,” American Statistician, May ’94, 48(2)

Unified View of Correlations



Spatial Correlation vs. Early/Late Split



Dependence on common virtual variables cancels out at the timing test

Next Lecture

- Latch-based timing
- Flip-flops