# EECS251B : Advanced Digital Circuits and Systems

Lecture 15 – Variability

Borivoje Nikolić, Vladimir Stojanović, Sophia Shao
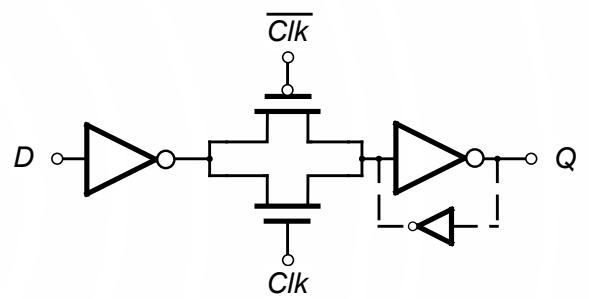
Apple Announces M1Ultra
March 8, 2022, AnandTech

Berkeley
UNIVERSITY OF CALIFORNIA
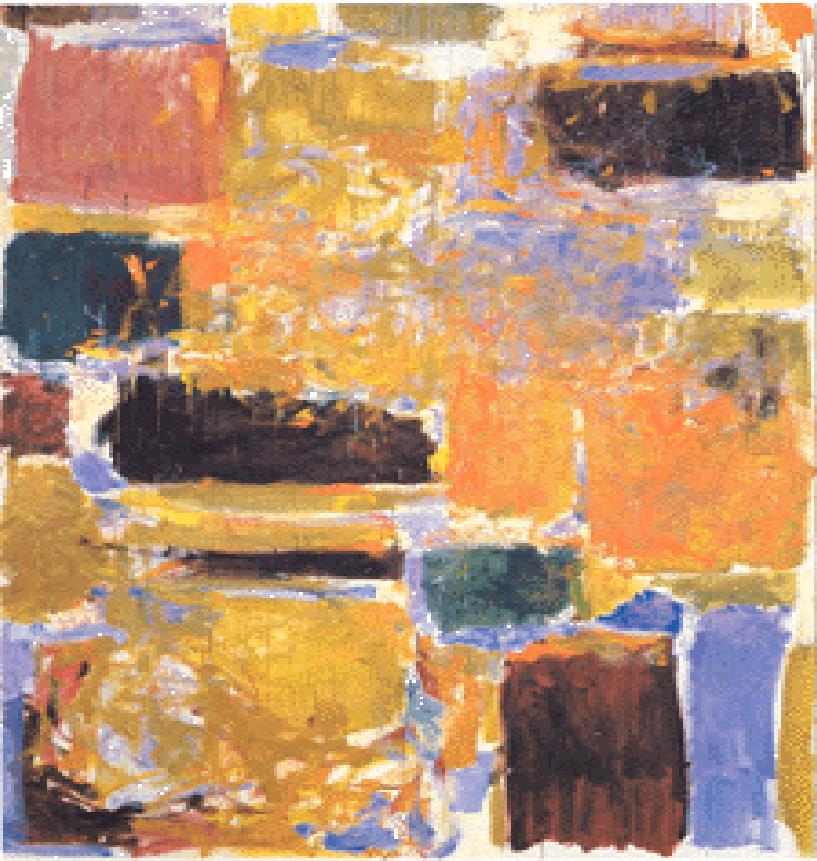
Recap

- Logical effort can be used to analyze latch timing
  - Clk-Q, D-Q delays are ~1FO4 delay (with F=1)

t_{setup}

$$t_{clk-q} = 2.7\ t_u + 2.2\ t_u$$
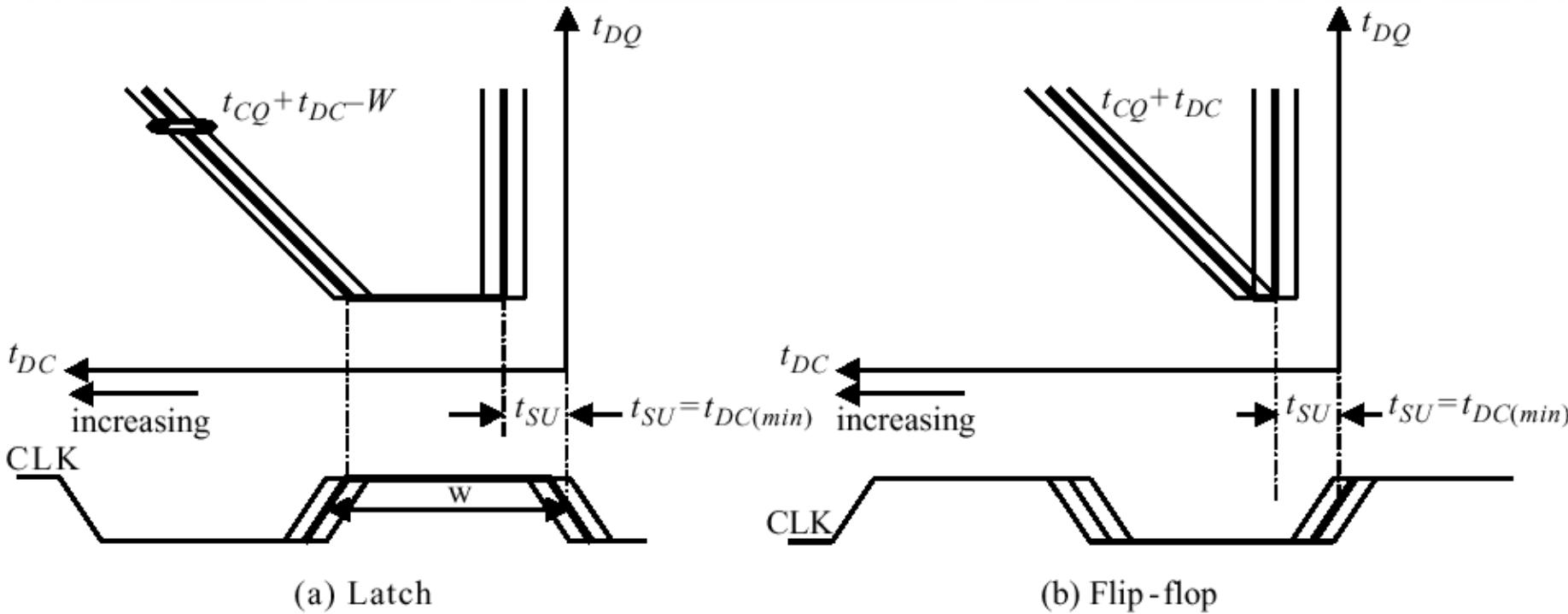
Flip-Flops

Key Point

- Two ways to design a flip-flop

  - Latch pair (large majority)

  - Pulsed latch

Latch vs. Flip-Flop

(a) Latch
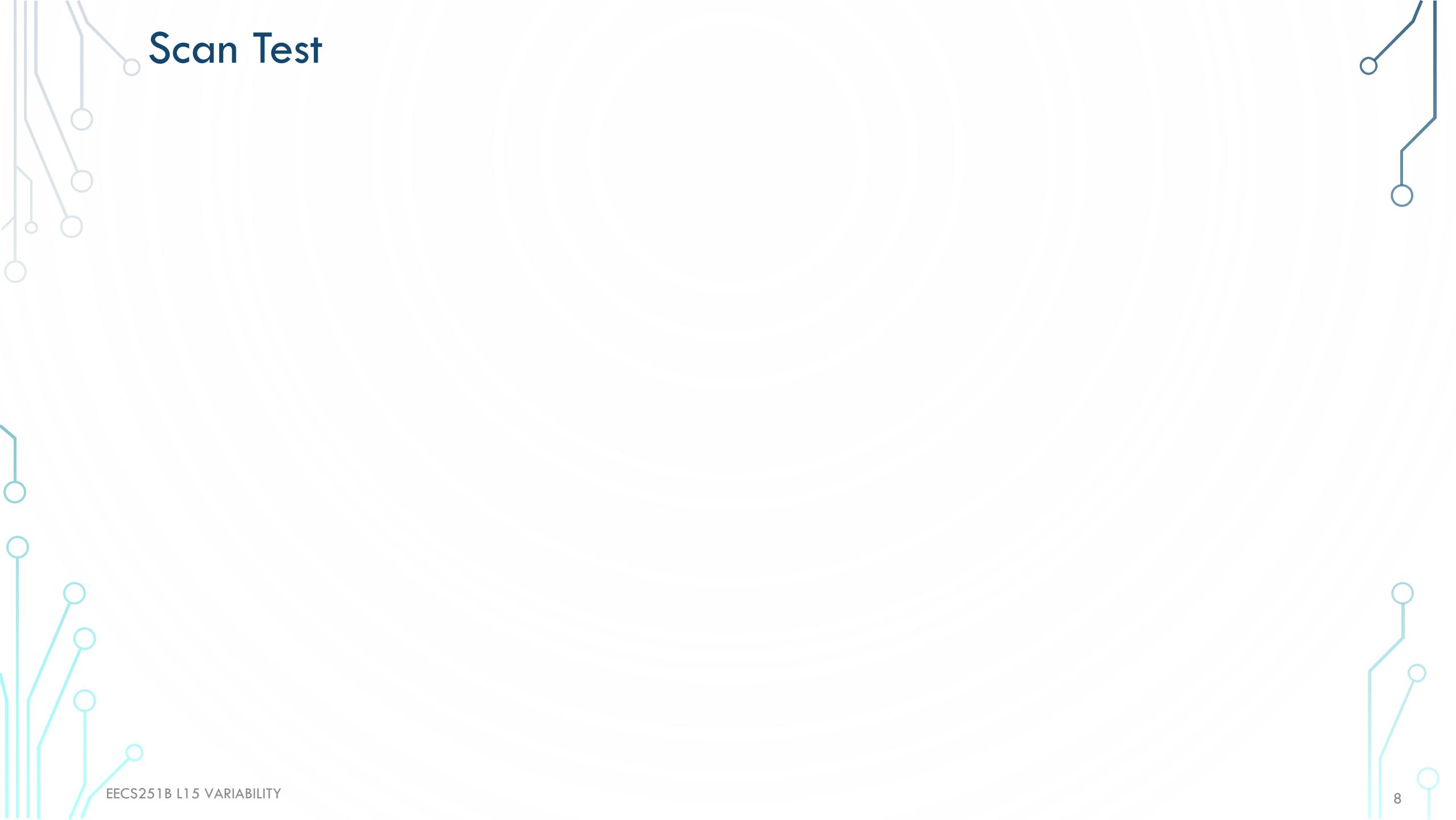
(b) Flip-flop

Courtesy of IEEE Press, New York. © 2000

Flip-Flops
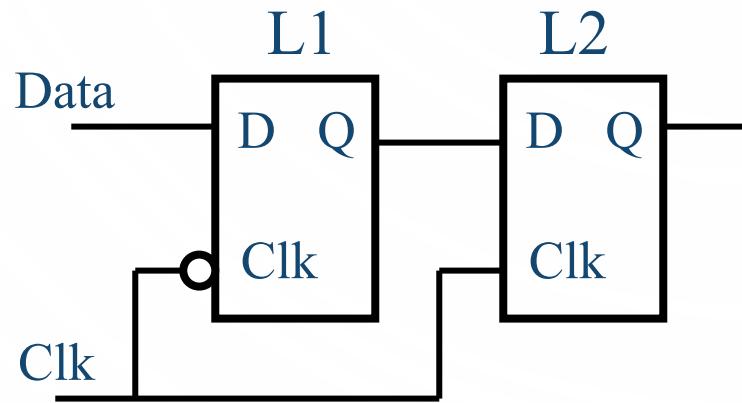
- Performance metrics

- Delay metrics
  - Insertion delay
  - Inherent race immunity
  - 'Softness' (Clock skew absorption)
  - Inclusion of logic
  - Small (+constant) clock load

- Power/Energy Metrics
  - Power/energy

- Design robustness
 - Noise immunity

Scan Test
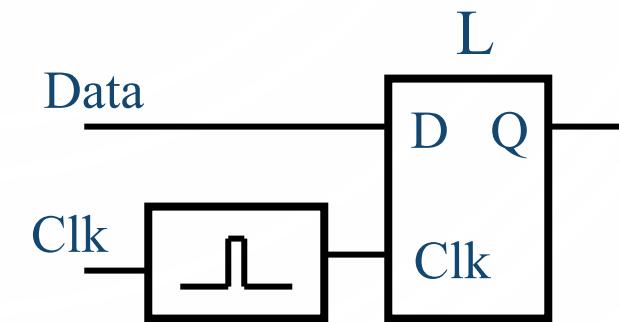
Types of Flip-Flops
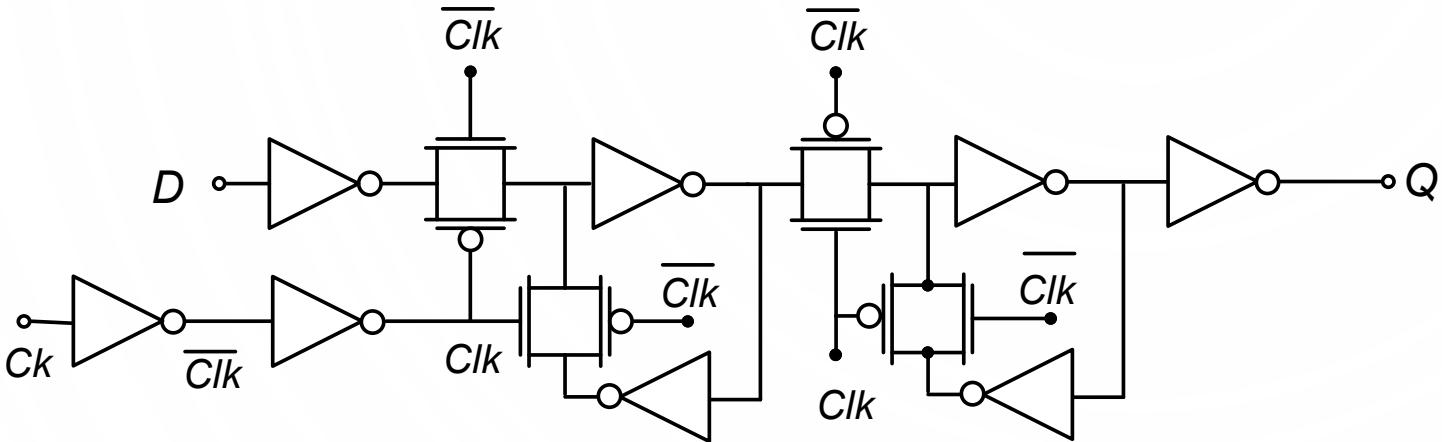
Latch Pair   **Pulse-Triggered Latch**

# Flip-Flop (Latch Pair) Clk-Q, setup, hold

Flip-Flop Timing Characterization

- Combinational logic delay is a function of output load and input slope

- Sequential timing (flip-flop):

  - t_{clk-q} is function of output load and clock rise time
  - t_{Su}, t_H are functions of D and Clk rise/fall times
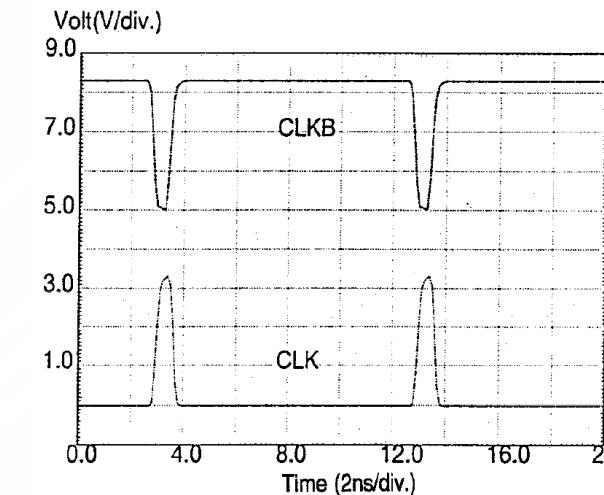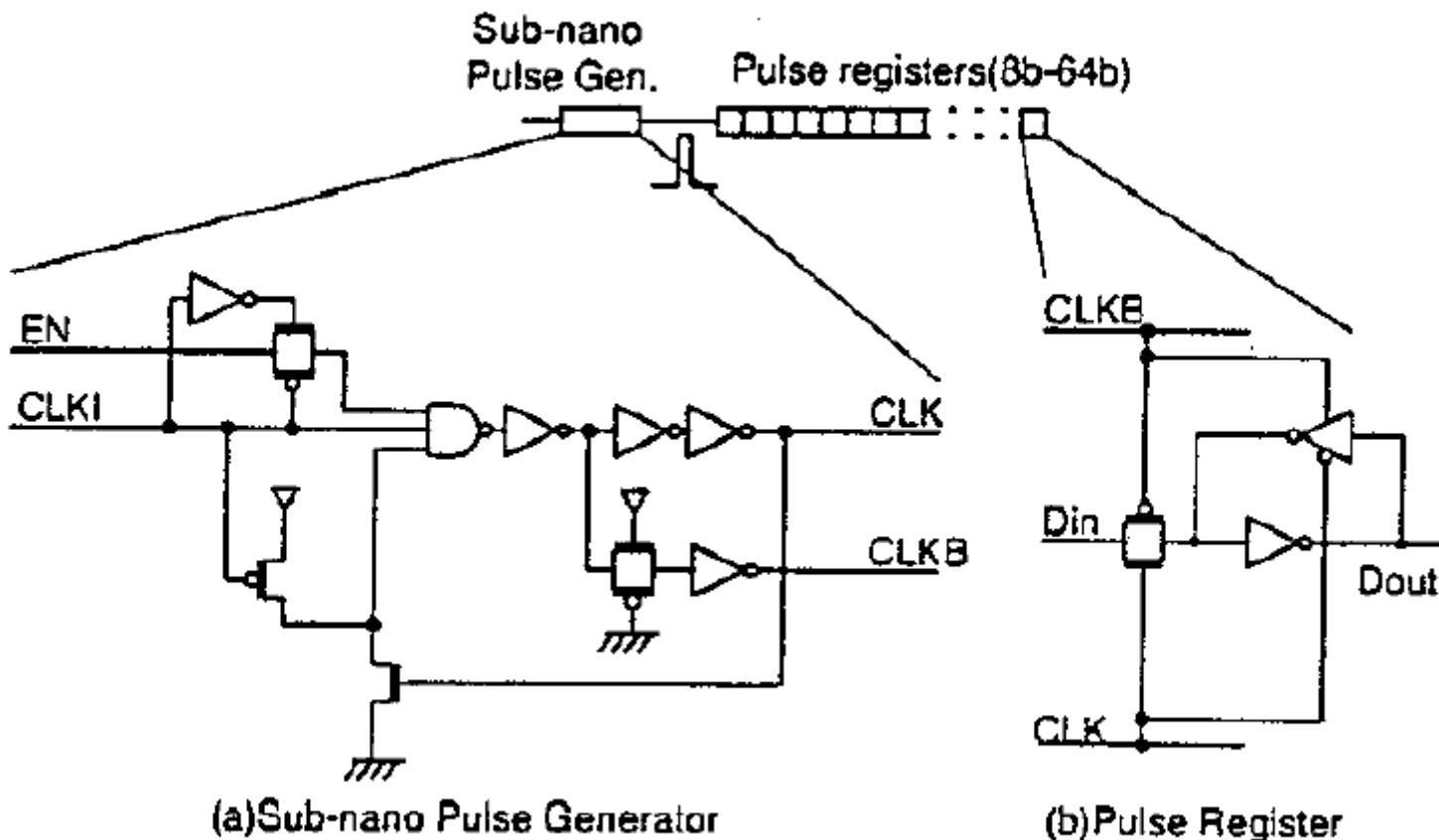
Pulse-Triggered Latches

- First stage is a pulse generator

  - generates a pulse (glitch) on a rising edge of the clock

- Second stage is a latch

  - captures the pulse generated in the first stage

- Pulse generation results in a negative setup time

- Frequently exhibit a soft edge property

- Note: power is always consumed in the pulse generator

  - Often shared by a group (register)
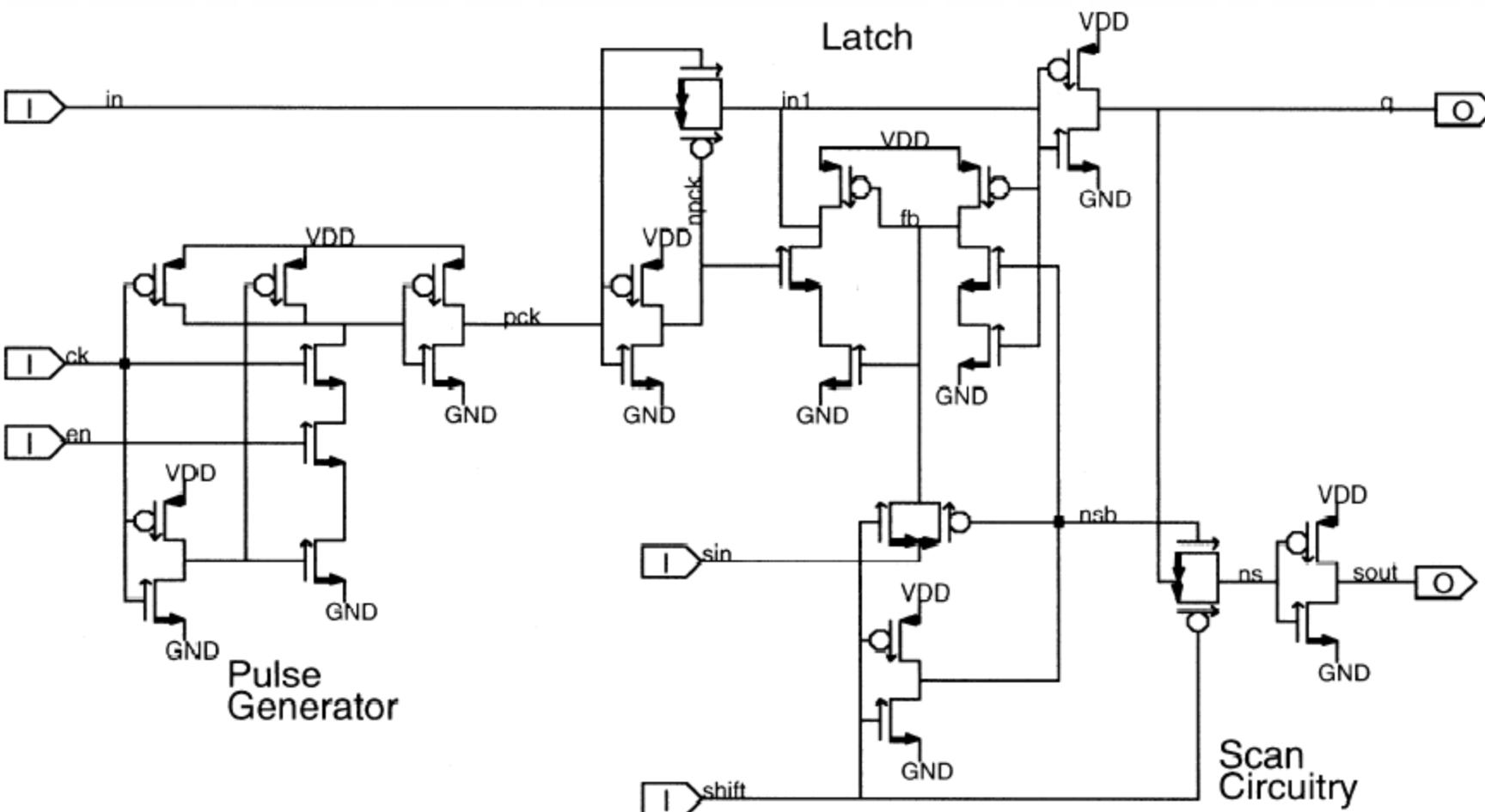
Pulsed Latch

Simple pulsed latch

(a) Sub-nano Pulse Generator

(b) Pulse Register

Kozu, ISSCC'96
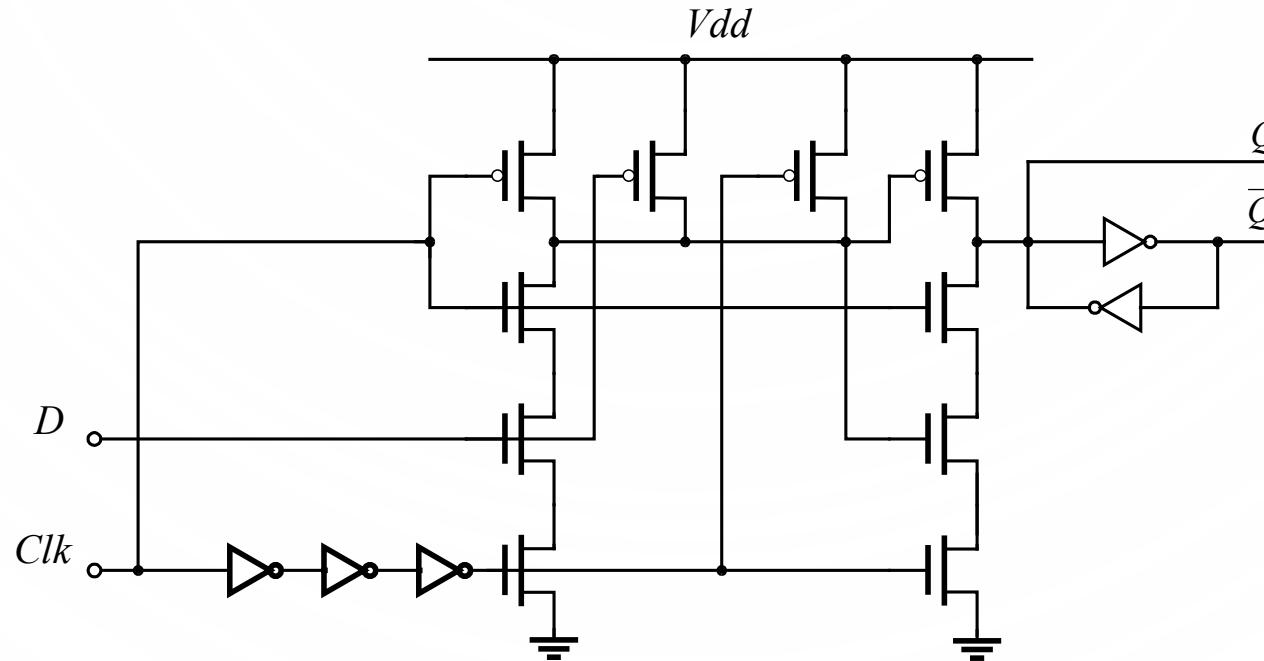
Naffziger, ISSCC'02
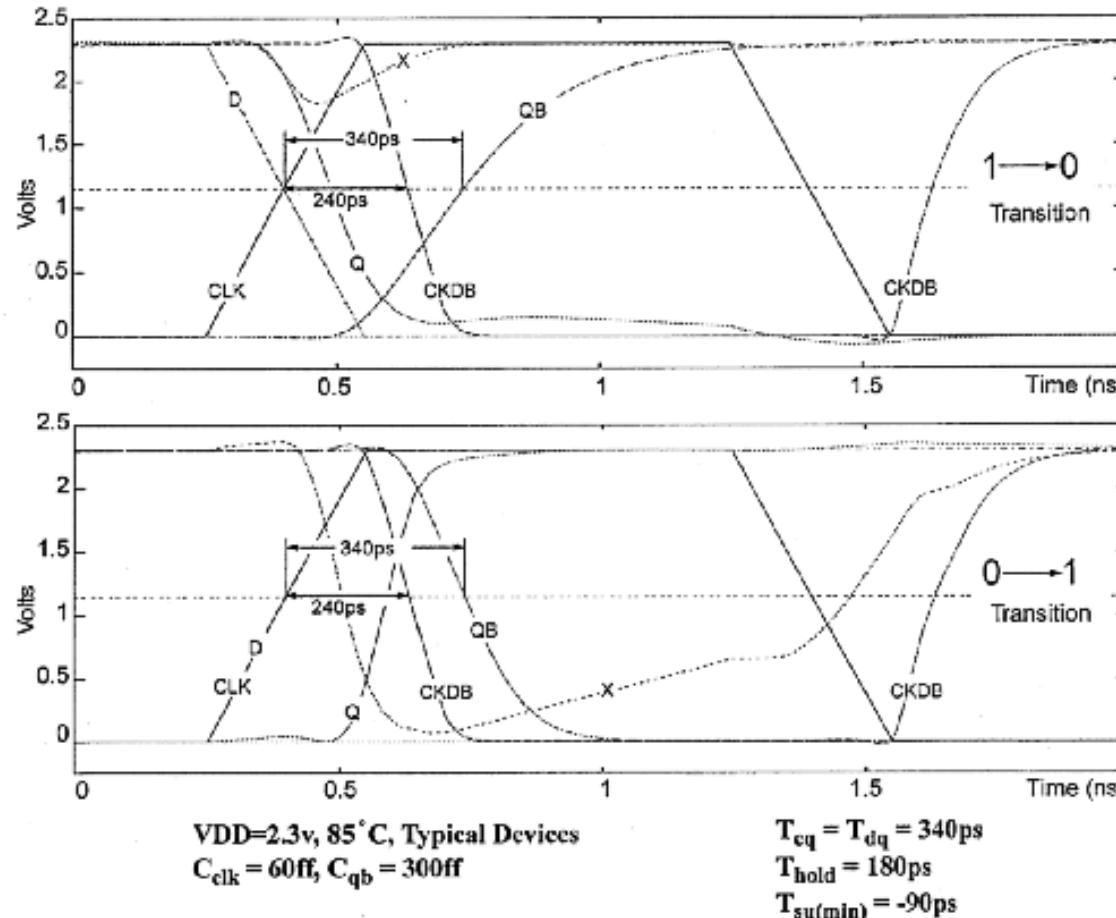
Pulsed Latches

Hybrid Latch Flip-Flop, AMD K-6
Partovi, ISSCC'96
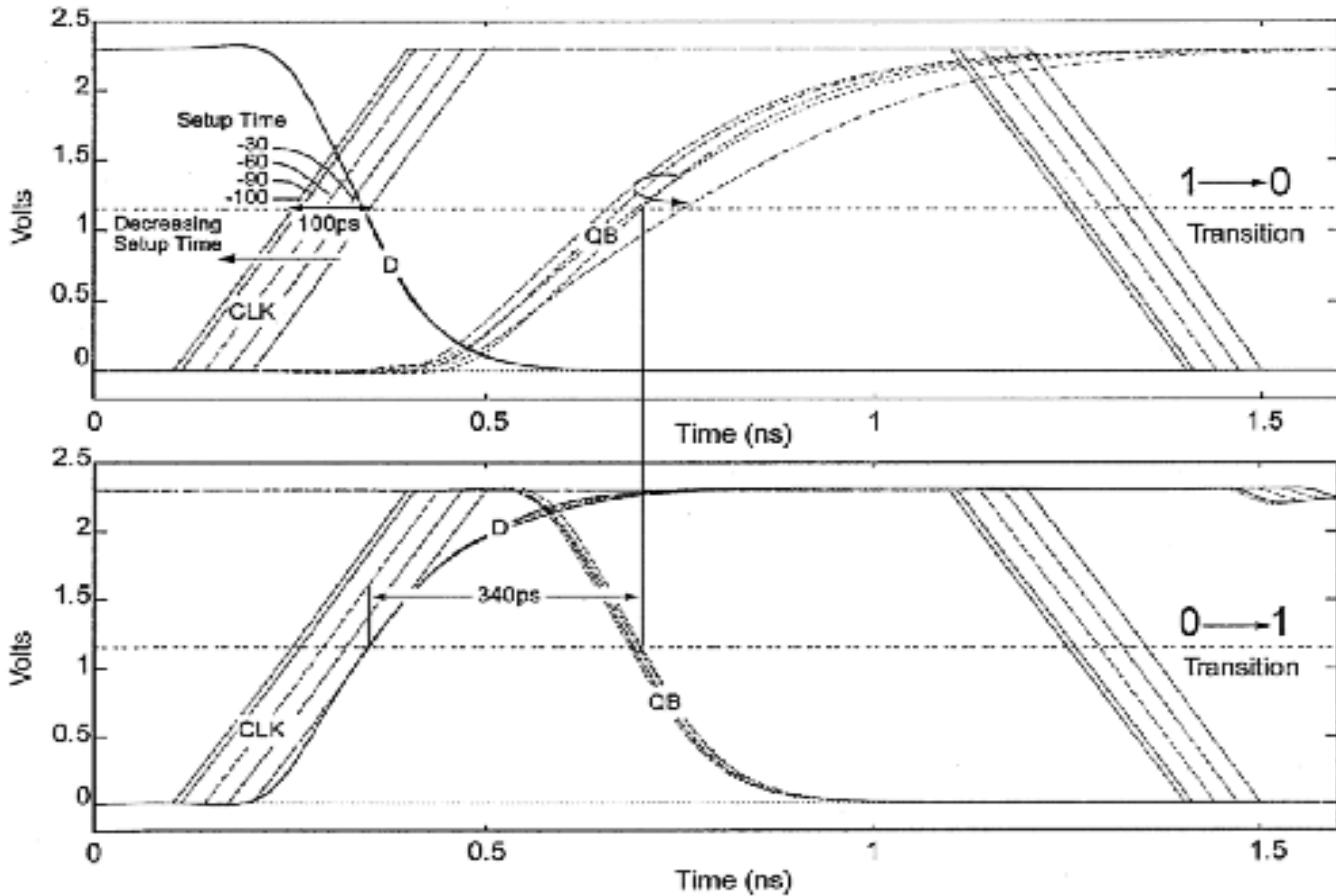
1-0 and 0-1 transitions at the input with 0ps setup time

Hybrid Latch Flip-Flop
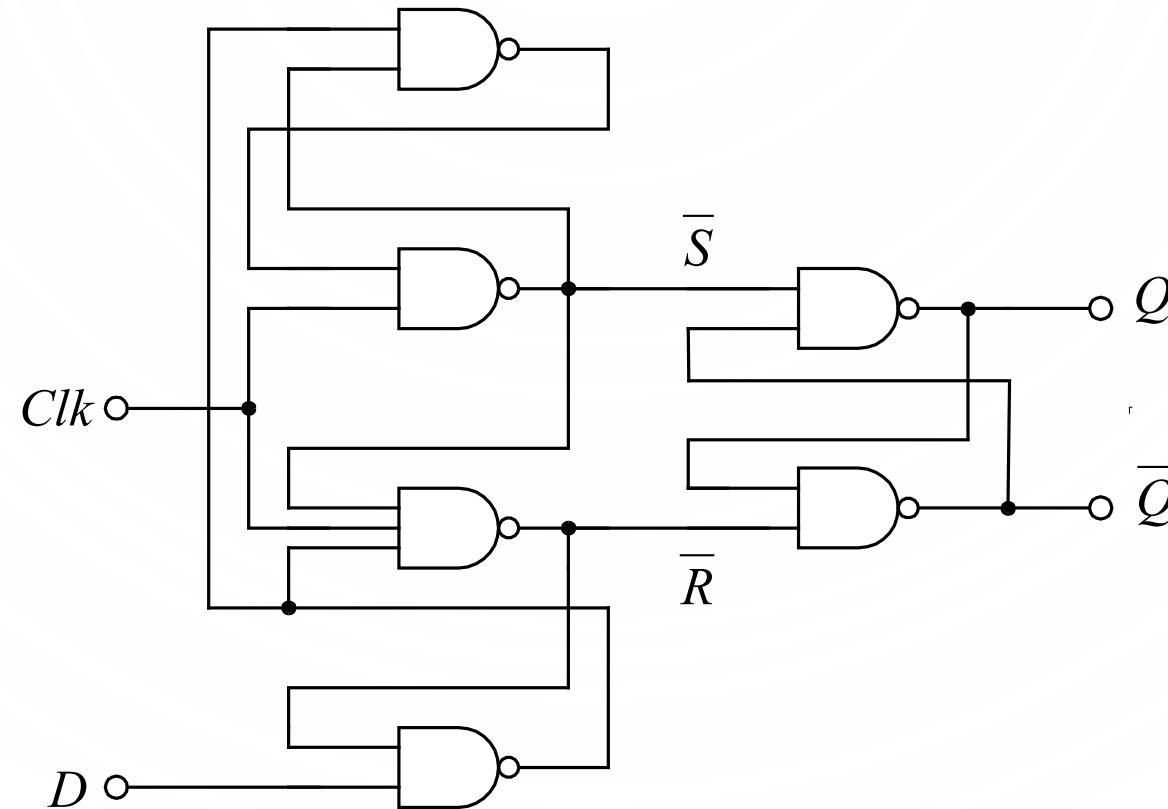
Skew absorption

Partovi *et al*, ISSCC'96

Pulsed Latches

7474, from mid-1960's

Pulsed Latches
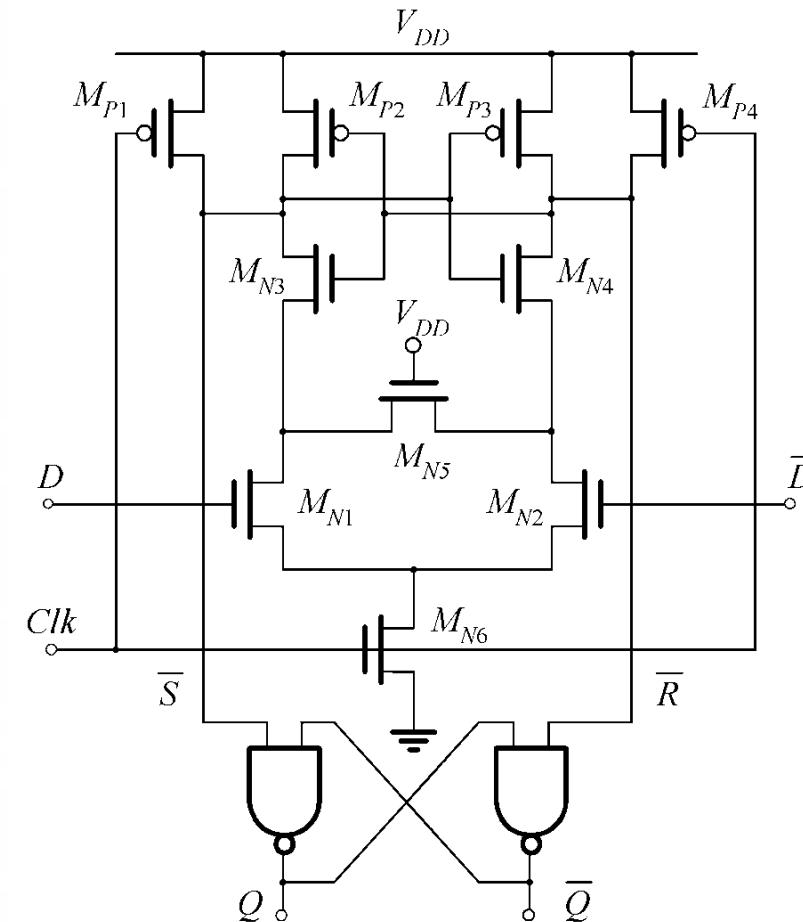
Sense-amplifier-based flip-flop, Matsui 1992.
DEC Alpha 21264, StrongARM 110

First stage is a sense amplifier, precharged to high, when Clk = 0
After rising edge of the clock sense amplifier generates the pulse on
S or R
The pulse is captured in
S-R latch
Cross-coupled NAND has different propagation delays of rising and falling edges
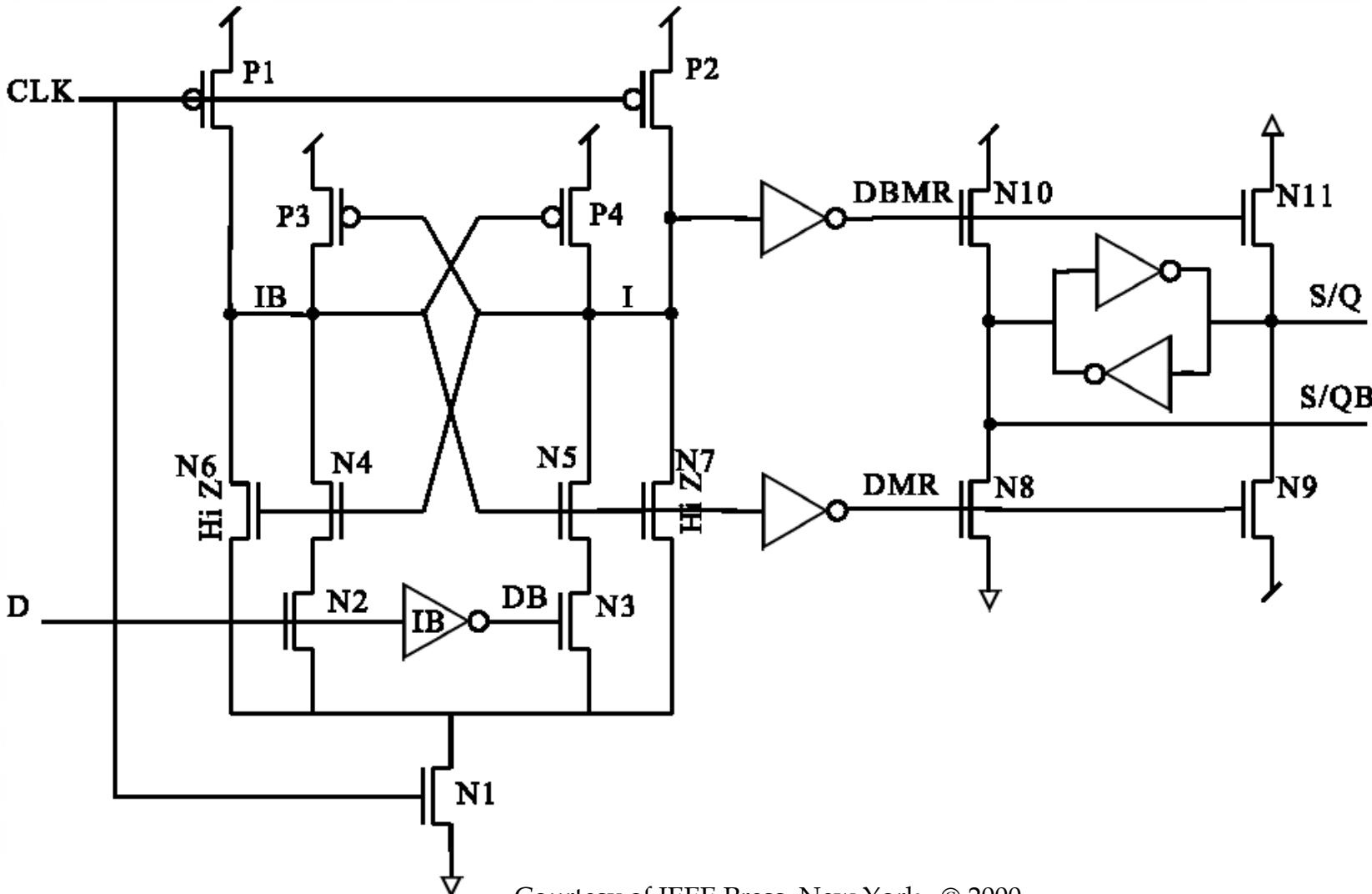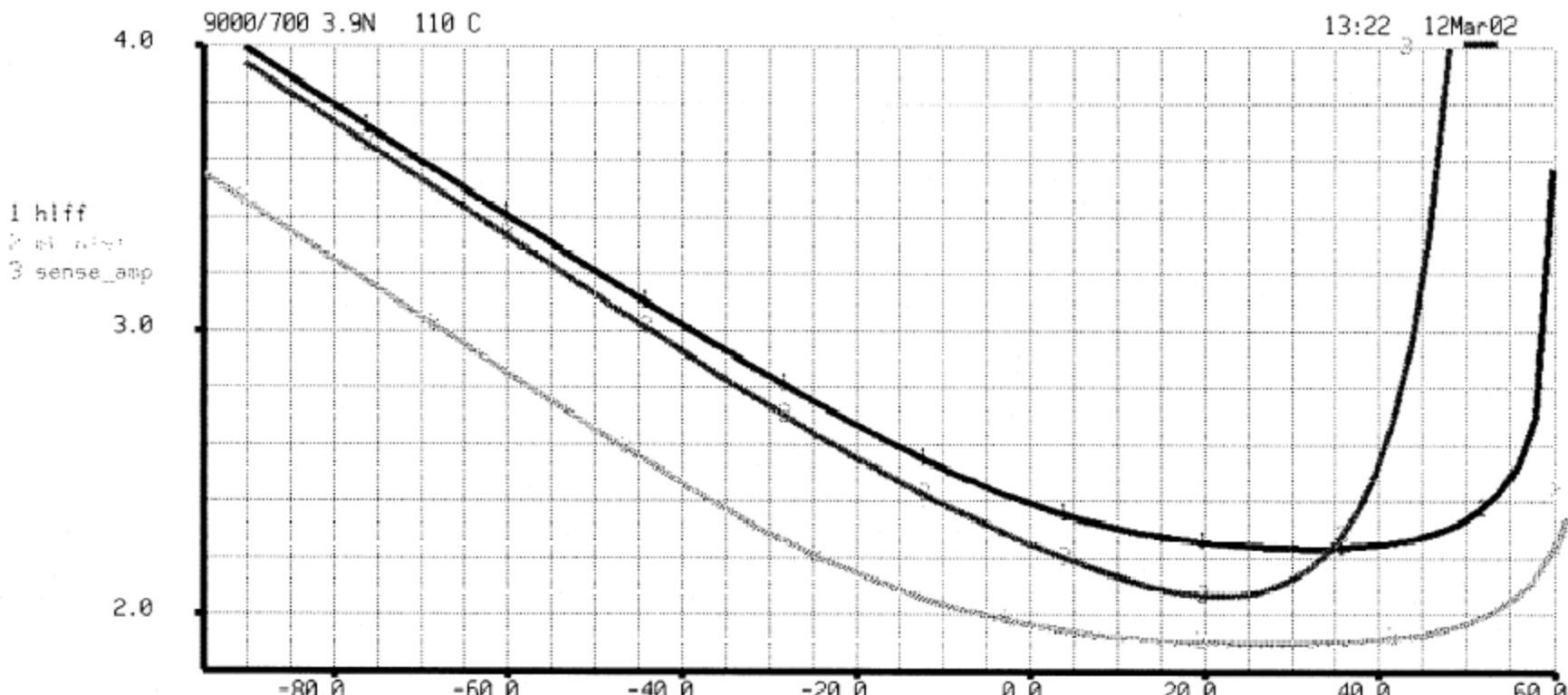
Sense Amplifier-Based Flip-Flop

Courtesy of IEEE Press, New York. © 2000

Sampling Window Comparison

Naffziger, JSSC 11/02

Announcements

- Lab 5 due this week

- Midterm reports due next week
  - 4 pages, conference format

- Assignment 2 posted this week

Design Variability Sources and Impact on Design
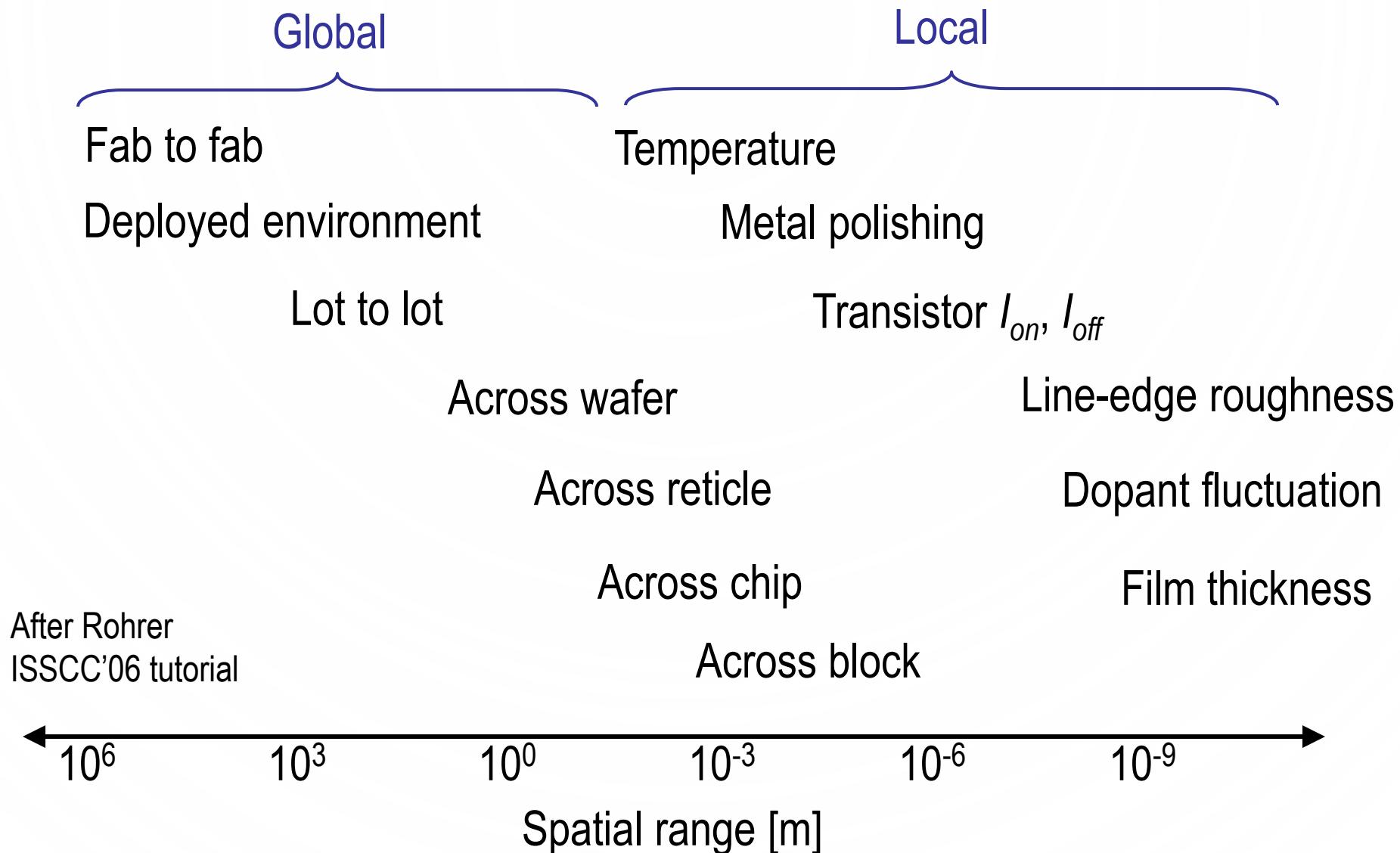
Variability Classification

- Nature of process variability
  - Within-die (WID), Die-to-die (D2D), Wafer-to-wafer (W2W), Lot-to-lot (L2L)
  - Systematic vs. random
  - Correlated vs. non-correlated

- Spatial variability/correlation
  - Device parameters (CD, t_{ox}, ...)
  - Supply voltage, temperature

- Temporal variability/correlation
  - Within-node scaling, Electromigration, Hot-electron effect, NBTI, self-heating, temperature, SOI history effect, supply voltage, crosstalk
    [Bernstein, IBM J. R&D, July/Sept 2006]

- Known vs. unknown
  - Goal of model-to-hw correlation is to reduce the unknowns
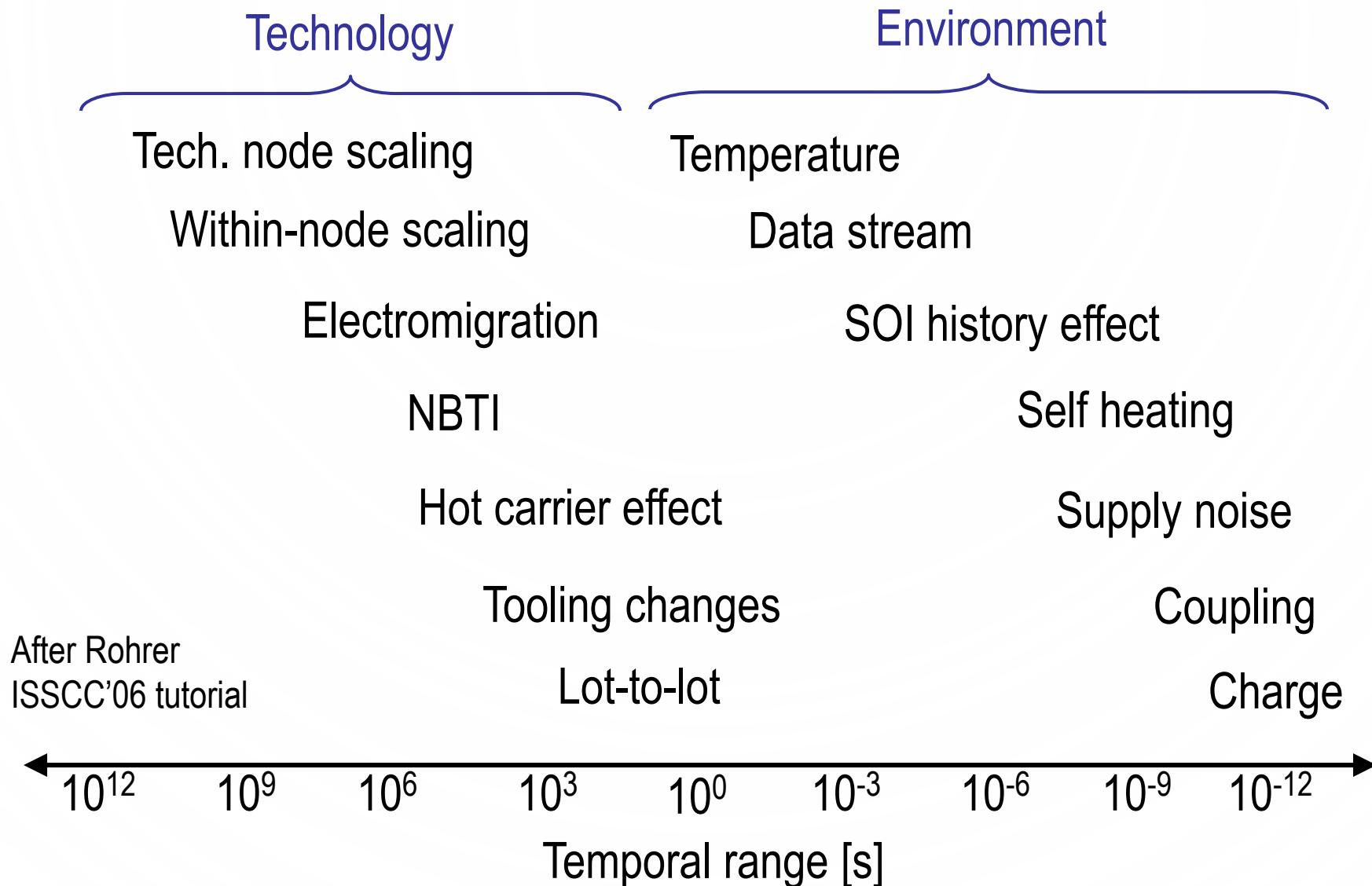
Sources of Variability

- Technology
  - Front-end (Devices)
    - Systematic and random variations in Ion, Ioff, C, …
  - Back-end (Interconnect)
    - Systematic and random variations in R, C

- Environment
  - Supply (IR drop, noise)
  - Temperature

Spatial Variability

Global   Local

Fab to fab   Temperature

Deployed environment   Metal polishing

Lot to lot   Transistor I_{on}, I_{off}

Across wafer   Line-edge roughness

Across reticle   Dopant fluctuation

Across chip   Film thickness

After Rohrer
ISSCC'06 tutorial   Across block

10^6 10^3 10^0 10^{-3} 10^{-6} 10^{-9}

Spatial range [m]

Temporal Variability

Technology

Environment

Tech. node scaling

Temperature

Within-node scaling

Data stream

Electromigration

SOI history effect

NBTI

Self heating

Hot carrier effect

Supply noise

Tooling changes

Coupling

After Rohrer
ISSCC'06 tutorial

Lot-to-lot

Charge

$\longleftrightarrow$

10^{12} 10^9 10^6 10^3 10^0 10^{-3} 10^{-6} 10^{-9} 10^{-12}

Temporal range [s]

Systematic vs. Random Variations
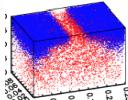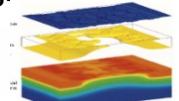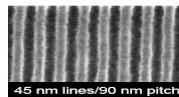
- Systematic
 - A systematic pattern can be traced down to lot-to-lot, wafer-to-wafer, within reticle, within die, from layout to layout,…
  - Within-die: usually spatially correlated

- Random
  - Random mismatch (dopant fluctuations, line edge roughness,…)
  - Things that are systematic, but e.g. change with a very short time constant (for us to do anything about it). Or we don't unedrstand it well enough to model it as systematic. Or we don't know it in advance ("How random is a coin toss?").

- Unknown

Systematic and Random Device Variations
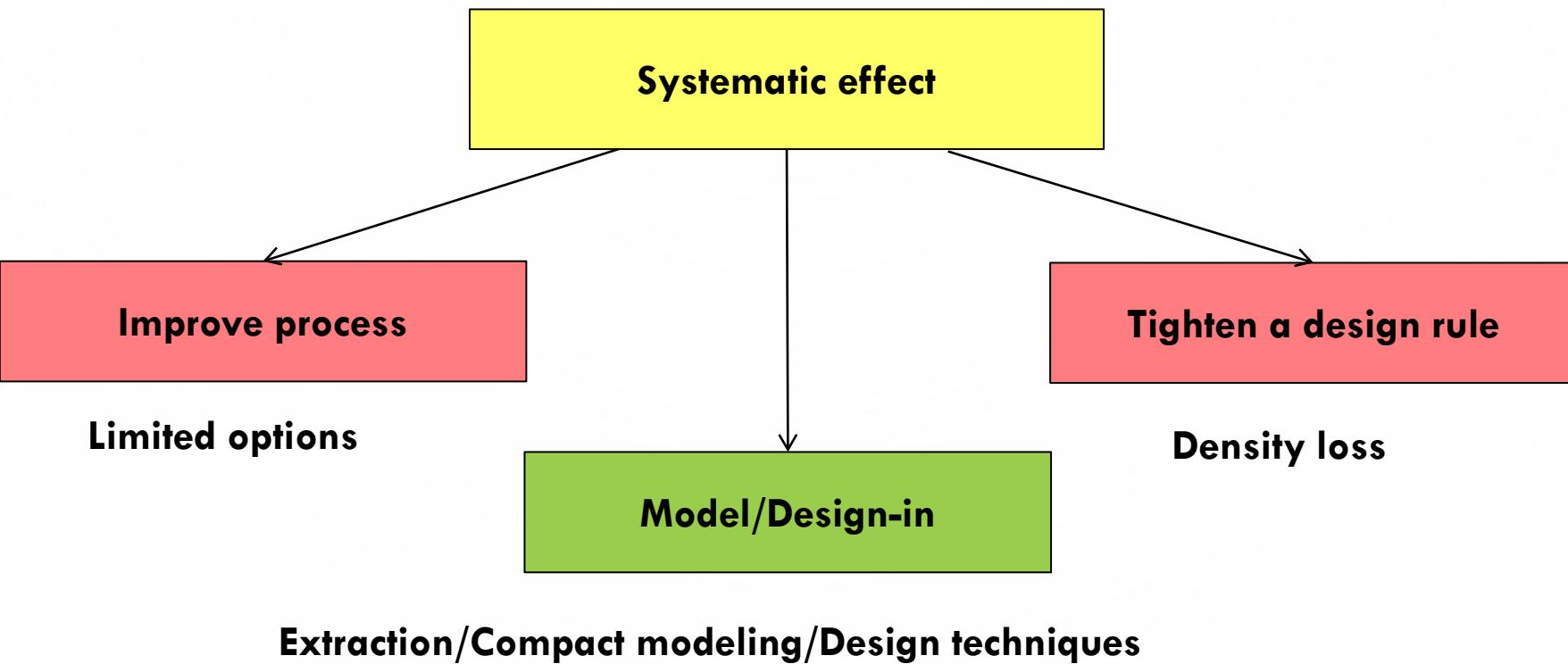
| Parameter	Random	Systematic
Channel Dopant Concentration Nch	Affects σ_{VT} [1]	Non uniformity in the process of dopant implantation, dosage, diffusion
Gate Oxide Thickness Tox	Si/SiO_2 & SiO_2/Poly-Si interface roughness[2]	Non uniformity in the process of oxide growth
Threshold Voltage V_T (non Nch related)	Random anneal temperature and strain effects	Non-uniform annealing temperature[5] (metal coverage over gate) Biaxial strain
Mobility μ	Random strain distributions	Systematic variation of strain in the Si due to STI, S/D area, contacts, gate density, etc
Gate Length L	Line edge roughness (LER)[3] 45 nm lines/90 nm pitch	Lithography and etching: Proximity effects, orientation[4]
Fin geometry/ film thickness variations	Rounding, etc, σ_{VT}, mobility.	Systematic fin thickness Systematic Si film/BOX variations

[1] D. Frank et al, *VLSI Symposium*, Jun. 1999 . [2] A. Asenov et al, *IEEE Trans on Electron Devices,* Jan. 2002.

[3] P. Oldiges et al, *SISPAD 2000,* Sept. 2000. [4] M. Orshansky et al, *IEEE Trans on CAD*, May 2002. [5] Tuinhout et al, *IEDM*, Dec 1996
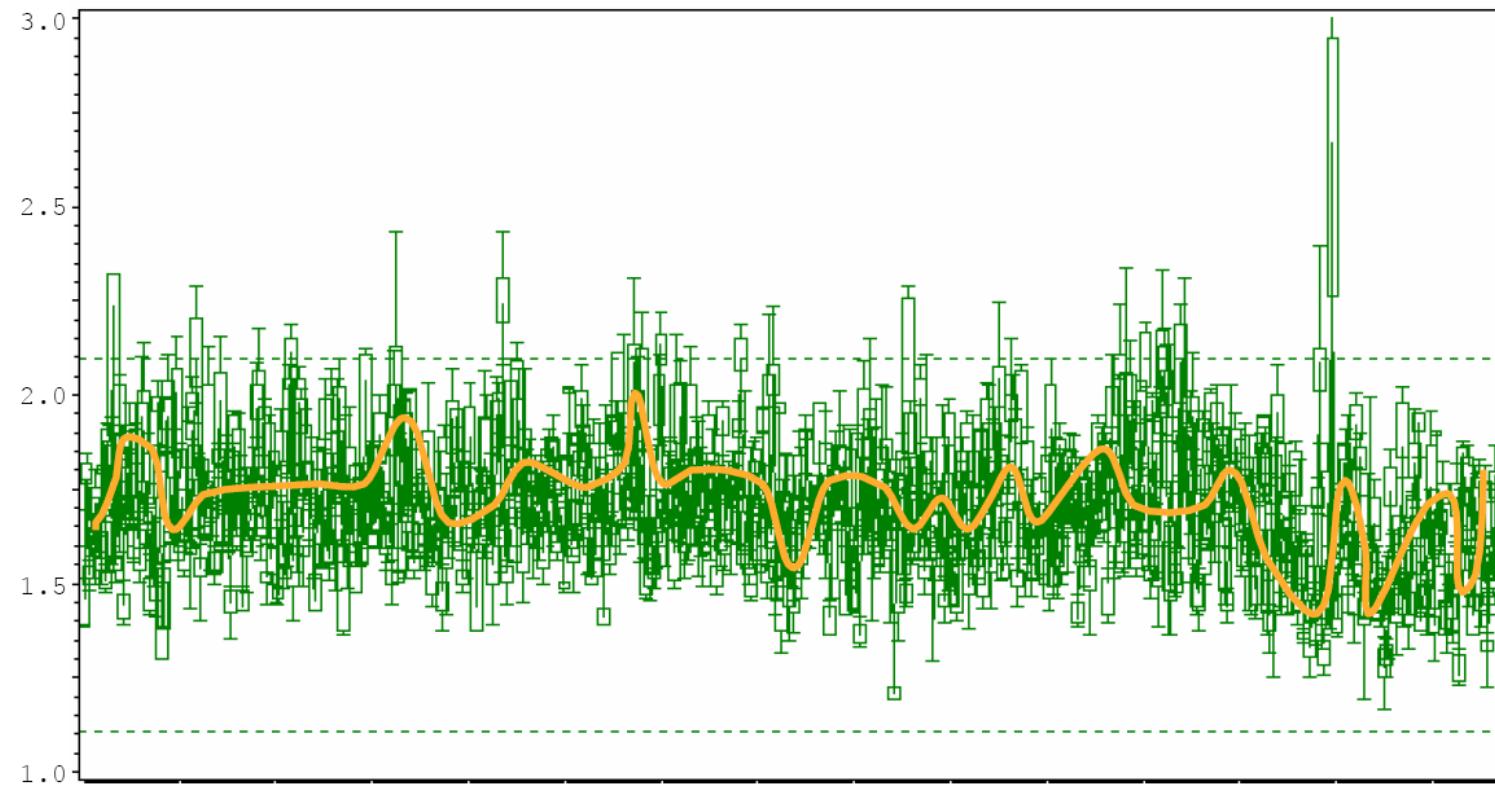
Dealing with Systematic Variations

- Model-to-hardware correlation classifies unknown sources
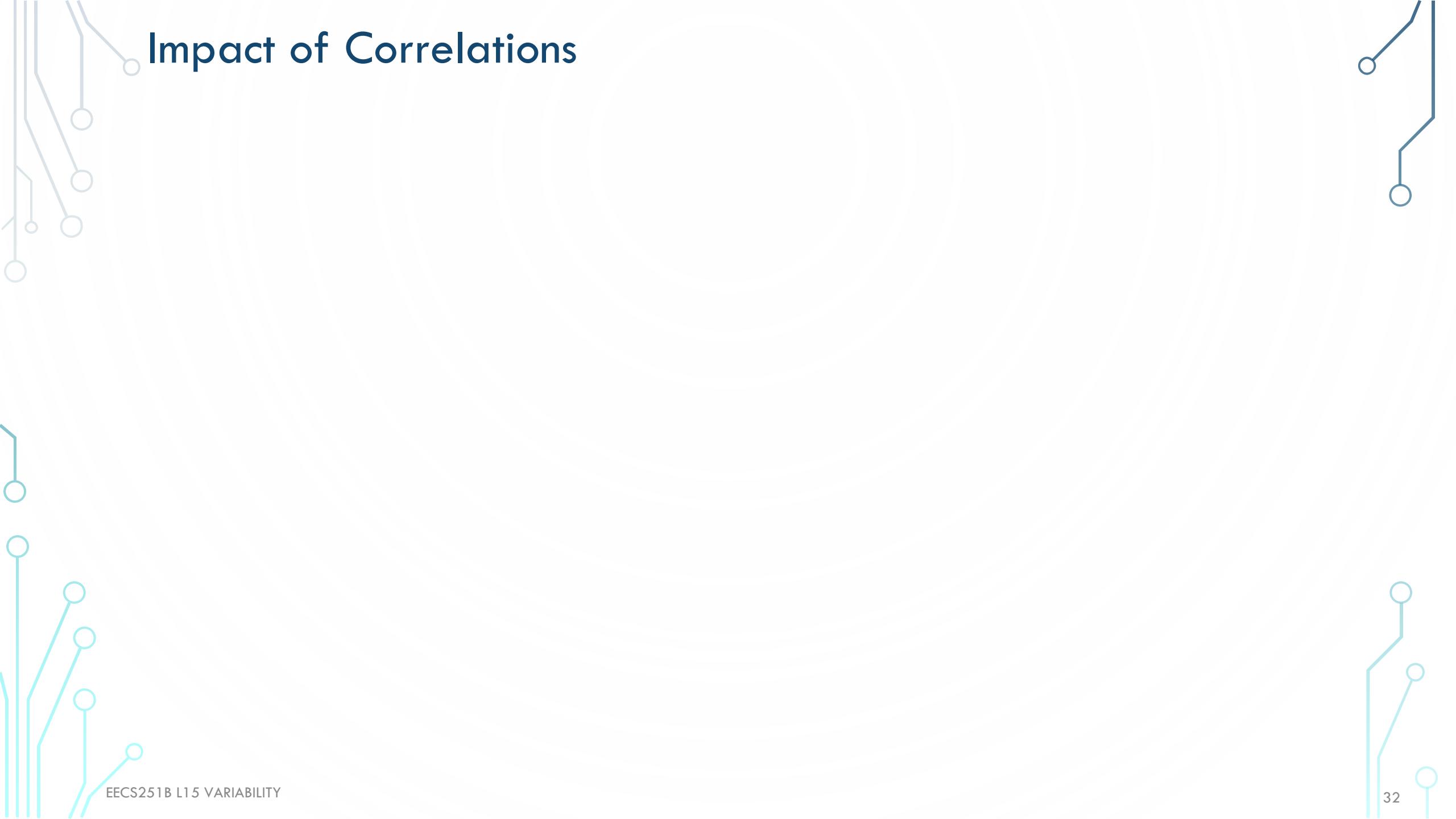
# Systematic (?) Temporal Variability

Metal 3 resistance over 3 months
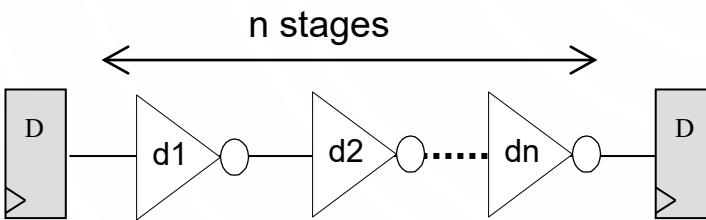
P. Habitz, DAC'06 tutorial

Impact of Correlations
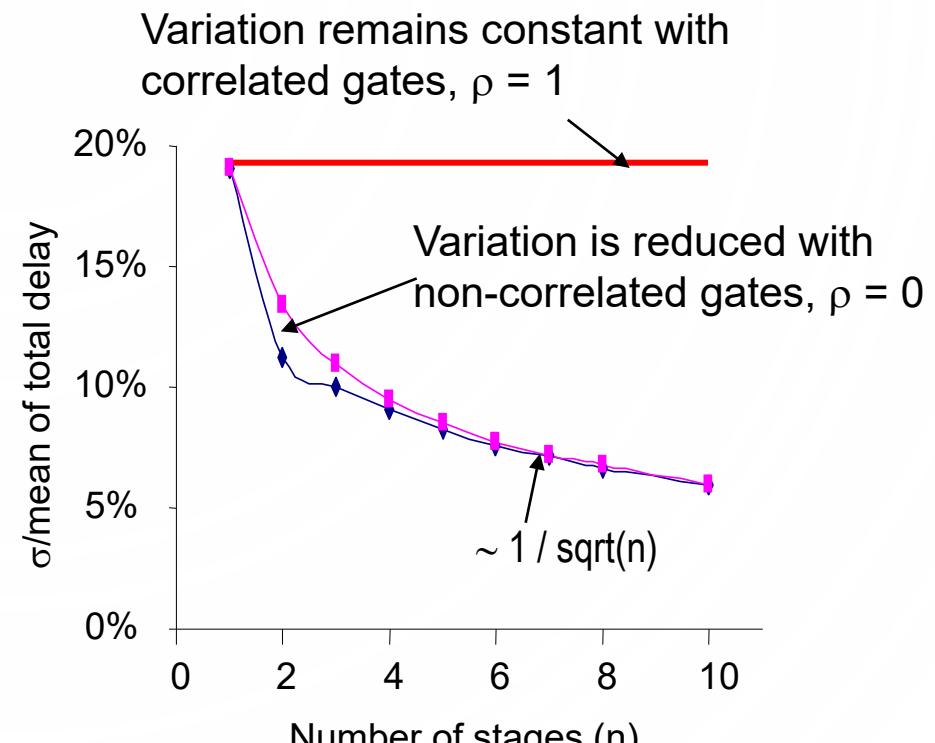
Chip Yield Depends on Inter-Gate Correlation

Variation remains constant with correlated gates, $\rho = 1$

Variation is reduced with non-correlated gates, $\rho = 0$

$\sim 1 / \sqrt{n}$

σ/mean of total delay

20%

15%

10%

5%

0%

Number of stages (n)

n stages

D   d1   d2 ...... dn   D
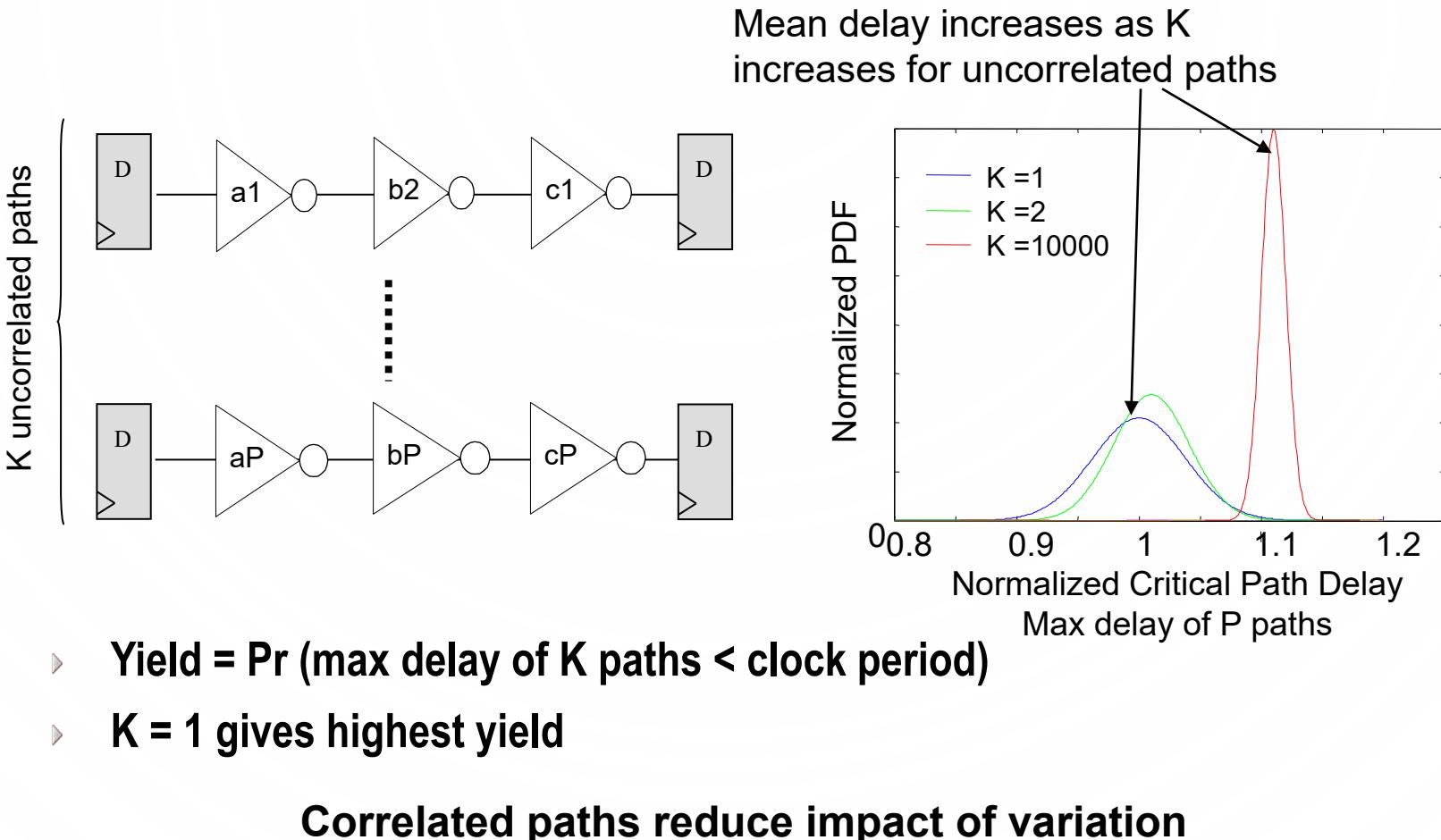
▷ **Yield = Pr (sum of n delays < clock period)**

▷ **ρ = 0 gives highest yield through averaging**

Non-correlated gates in a path reduce impact of variation

Bowman et al, *JSSC*, Feb 2002 .
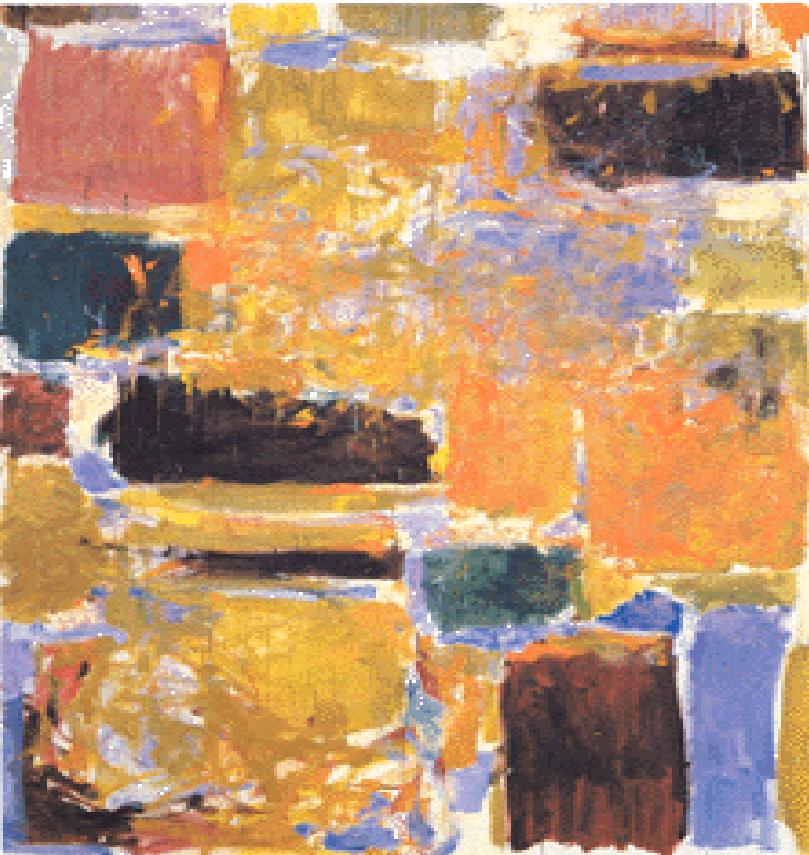
# Chip Yield Depends on Inter-Path Correlation

Mean delay increases as K increases for uncorrelated paths

K uncorrelated paths

D — a1 — b2 — c1 — D

⋮

D — aP — bP — cP — D

Normalized PDF

K =1
K =2
K =10000

0.8 0.9 1 1.1 1.2

Normalized Critical Path Delay
Max delay of P paths

▹ **Yield = Pr (max delay of K paths < clock period)**

▹ **K = 1 gives highest yield**

Correlated paths reduce impact of variation

Bowman et al, *JSSC*, Feb 2002 .
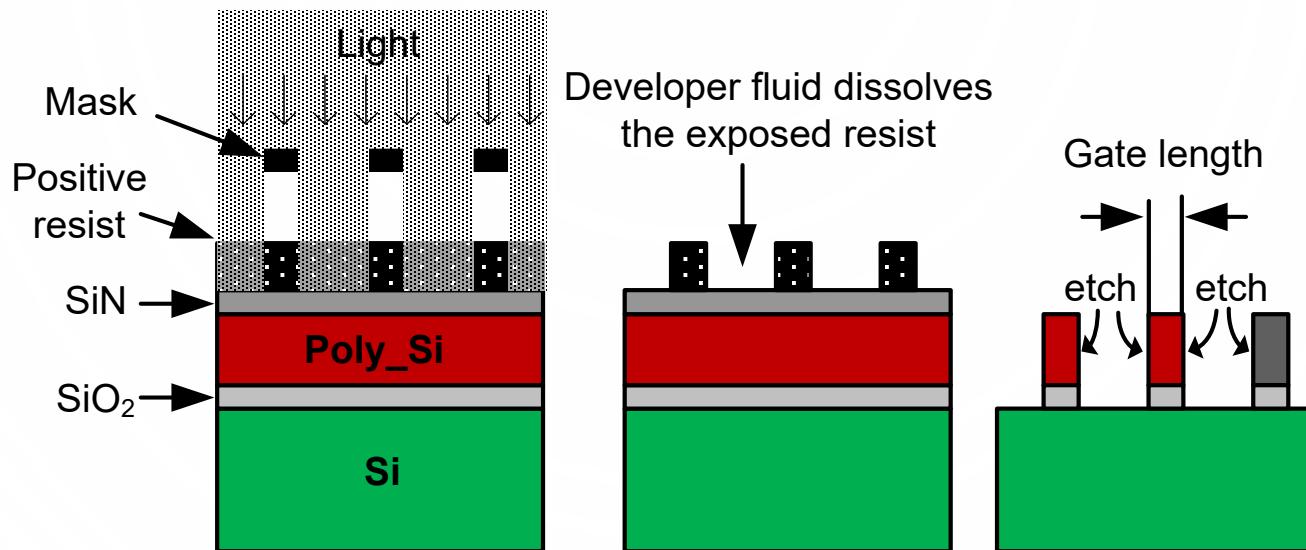
# Process Corners

Design Variability
Some Systematic Effects

Layout: Poly Proximity Effects

- Gate CD is a function of its neighborhood

Light

Mask

Positive resist

SiN

Poly_Si

SiO$_2$

Si

Developer fluid dissolves the exposed resist

Gate length

etch   etch

Gate length depends on

> **Light intensity profile falling on the resist**

> **Resist: application of developer fluid[1], post exposure bake (PEB) temperature[2]**

> **Dry etching: microscopic loading effects[3]**
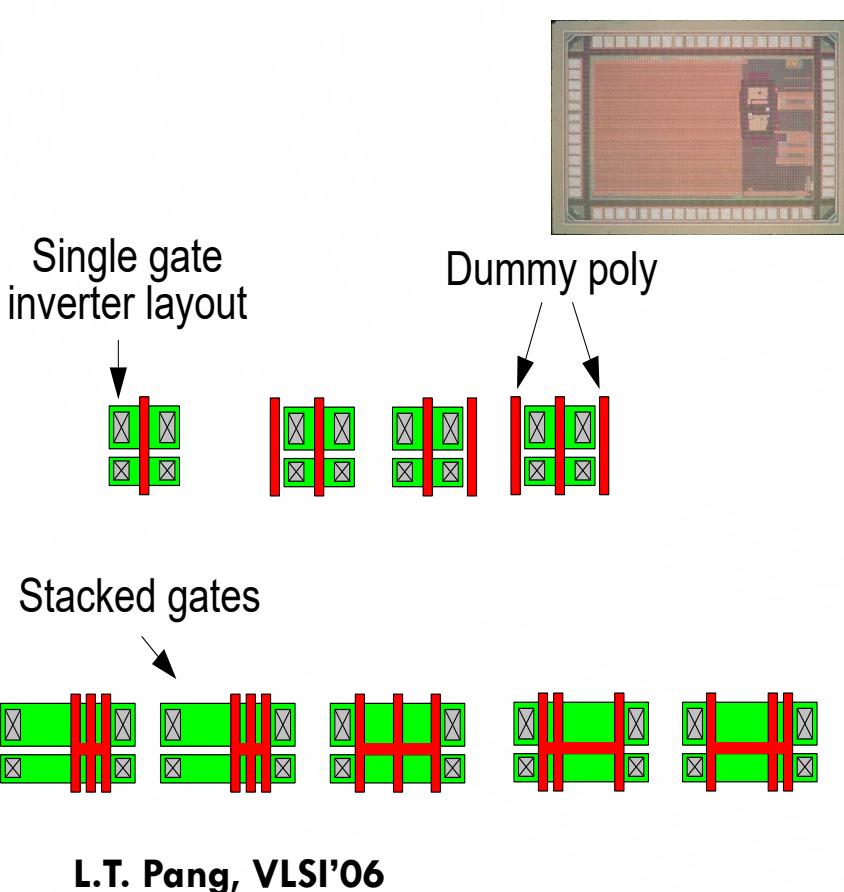
[1] J.Cain, M.S. Thesis, UC Berkeley
[2] D. Steele et al, *SPIE,* vol.4689, July 2002.
[3] J. D. Plummer, M.D. Deal, P.B. Griffin, Silicon VLSI Technology, Prentice-Hall, 2000.
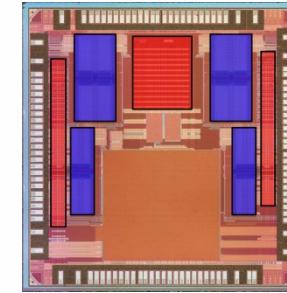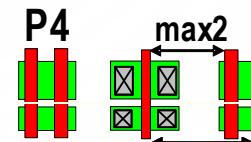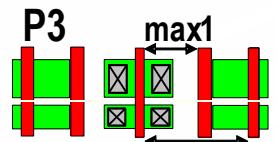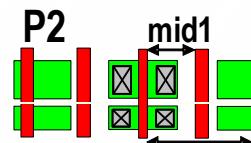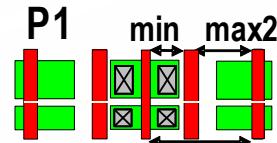
Layout: Proximity Test Structures

- **90nm experiments**

Single gate inverter layout

Dummy poly

Stacked gates

L.T. Pang, VLSI'06
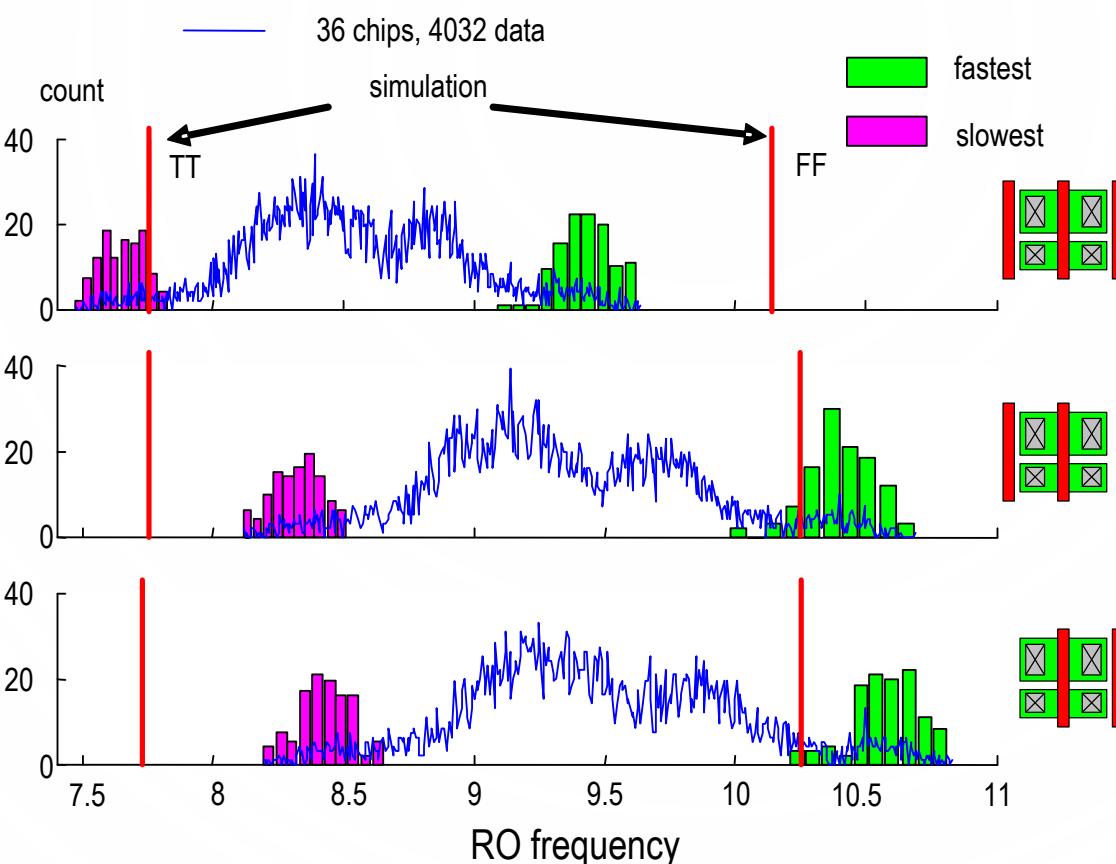
▷ **45nm experiments**

No single gates allowed

P1   min   max2

P2   mid1

P3   max1

P4   max2

L.T. Pang, CICC'08
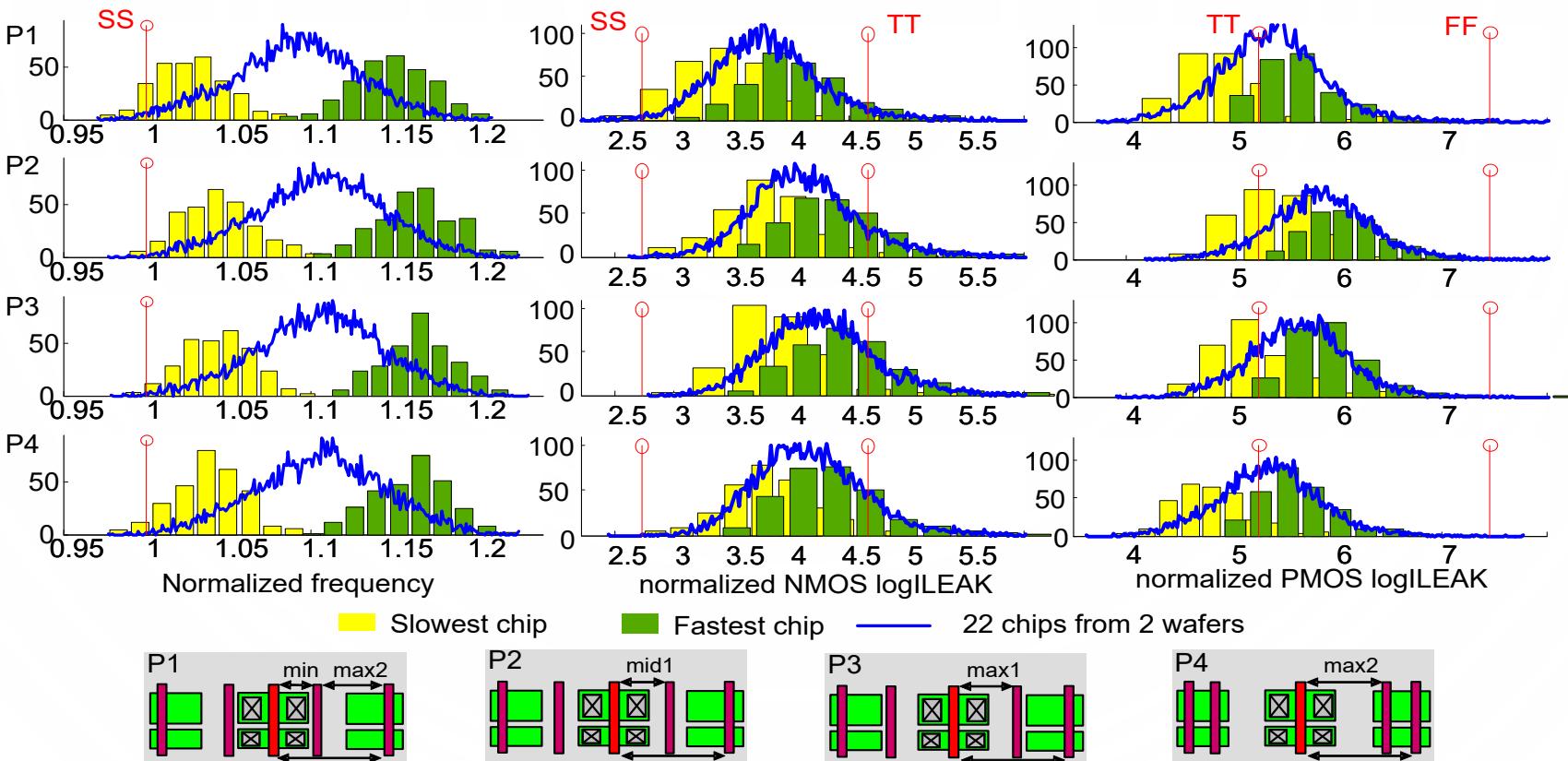
▷ **Ring oscillators and individual transistor leakage currents**

Results: Single Gates in 90nm
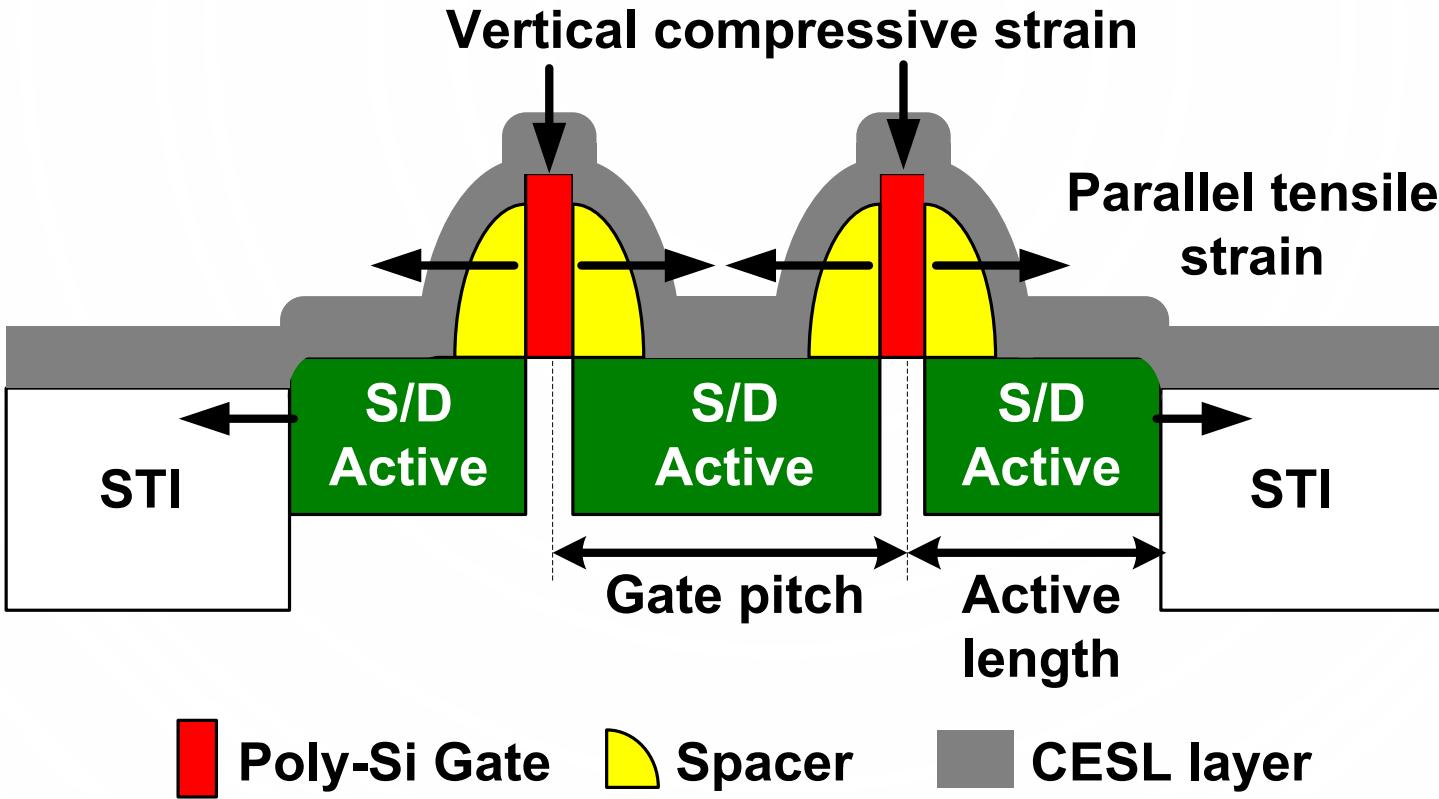
- **Max ∆F between layouts > 10%**
- **Within-die $3\sigma/\mu \sim 3.5\%$, weak dependency on density**

- Slowest chip
- Fastest chip
- 22 chips from 2 wafers
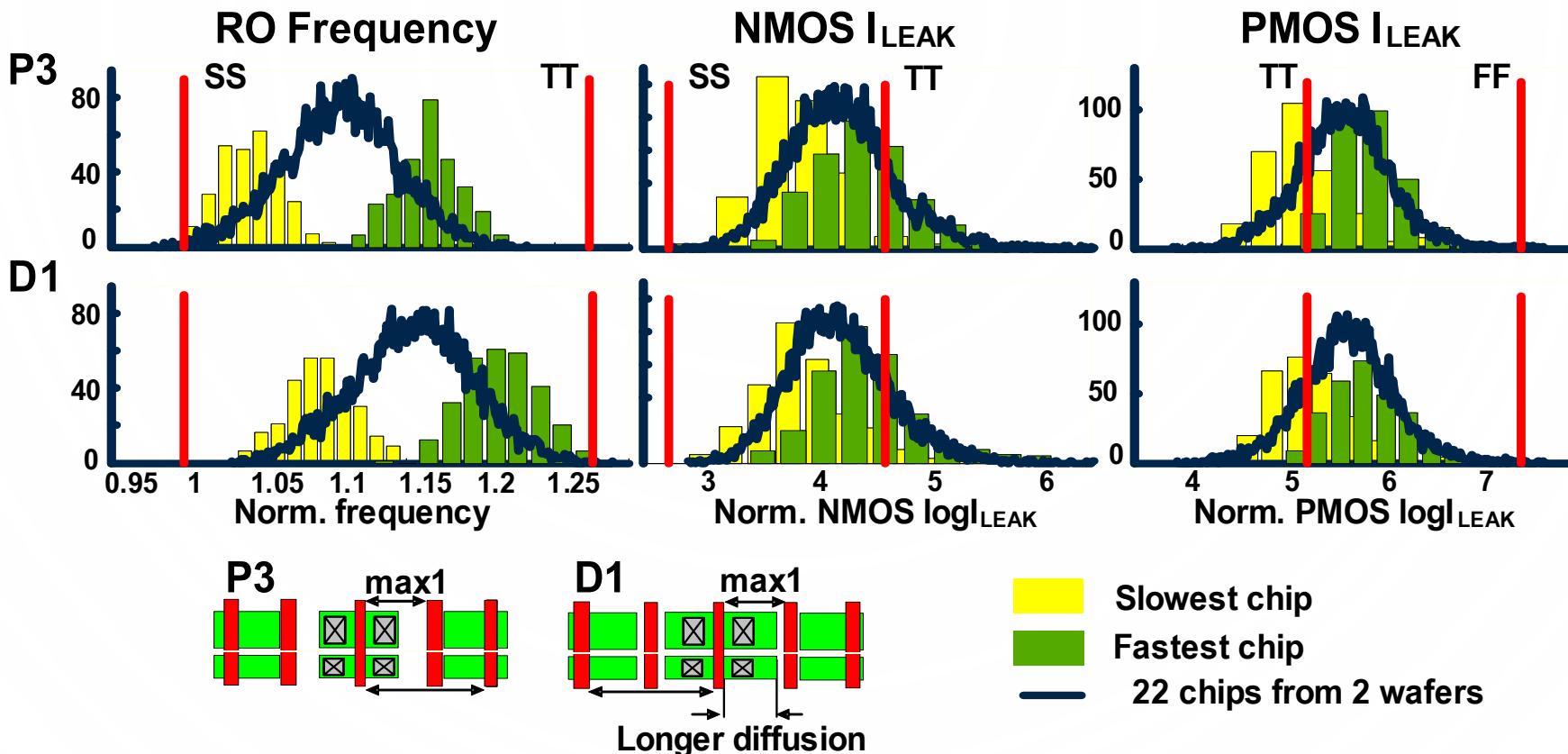
- **Weak effect on performance. ΔF ~ 2%**

- **Small shifts in NMOS leakage and bigger shifts in PMOS leakage**

Impact of Stress

Vertical compressive strain

Parallel tensile strain

S/D Active — S/D Active — S/D Active

STI — STI

Gate pitch — Active length

Poly-Si Gate | Spacer | CESL layer
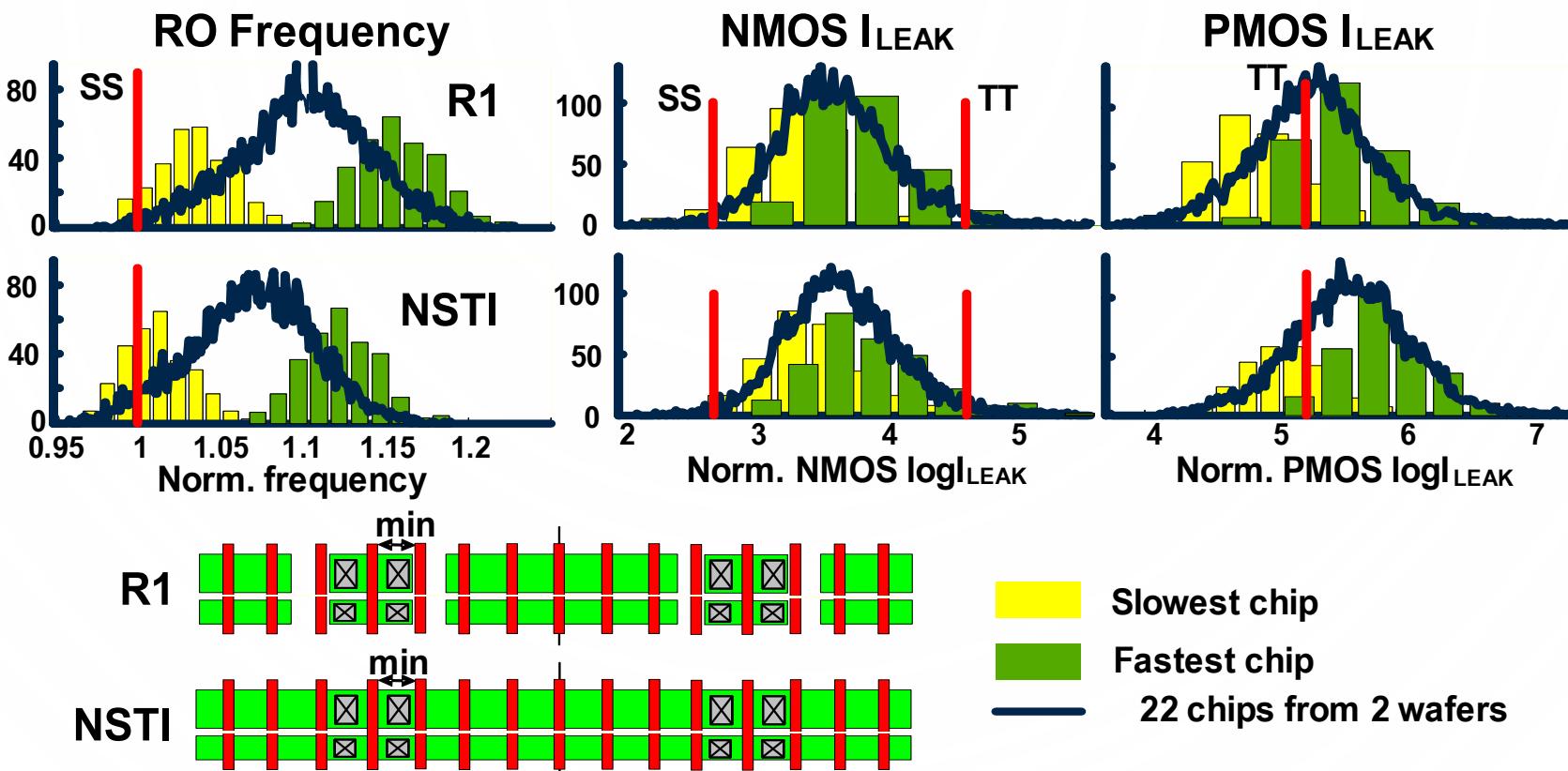
- 45nm STM process: Wafer rotated <100> - higher PMOS mobility
- NMOS strained through capping layer
- Subatmospheric STI – weak tensile stress

Impact of Longer Diffusion in 45nm

P3

RO Frequency

SS   TT

NMOS I_{LEAK}

SS   TT

PMOS I_{LEAK}

TT   FF

D1

Norm. frequency

Norm. NMOS $\log I_{LEAK}$

Norm. PMOS $\log I_{LEAK}$

P3   max1   **D1**   max1

Longer diffusion

- Slowest chip
- Fastest chip
- 22 chips from 2 wafers
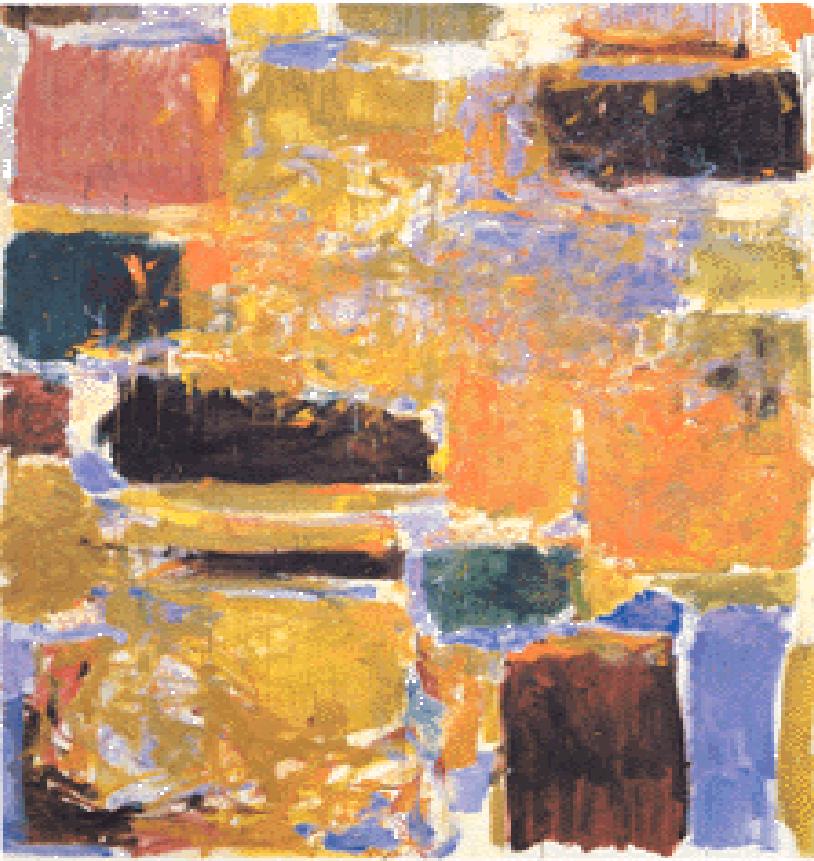
- **Strongest effect measured in 45nm, ΔF ~ 5%**
- **No significant shift in I_{LEAK}**

Impact of Shallow Trench Isolation (STI)

RO Frequency — R1, NSTI
NMOS I_{LEAK} — SS, TT
PMOS I_{LEAK} — TT

Norm. frequency (0.95, 1, 1.05, 1.1, 1.15, 1.2)
Norm. NMOS $\log I_{LEAK}$ (2, 3, 4, 5)
Norm. PMOS $\log I_{LEAK}$ (4, 5, 6, 7)

R1 — min
NSTI — min

- Slowest chip
- Fastest chip
- 22 chips from 2 wafers
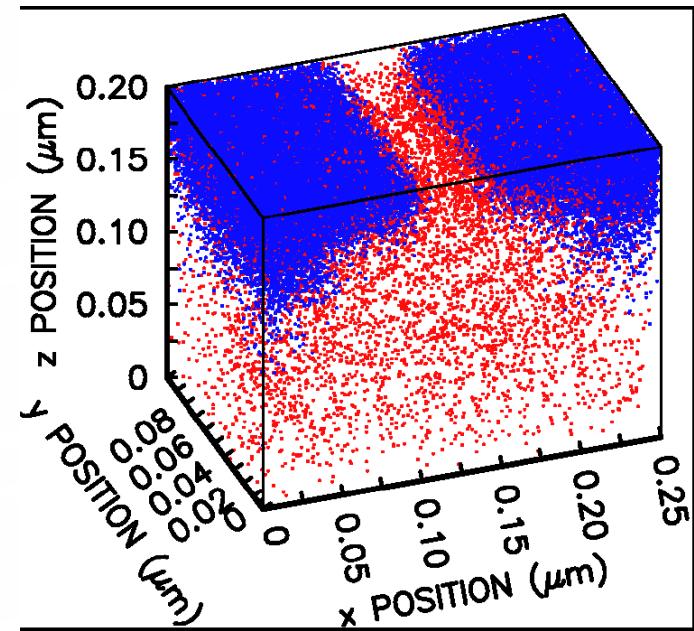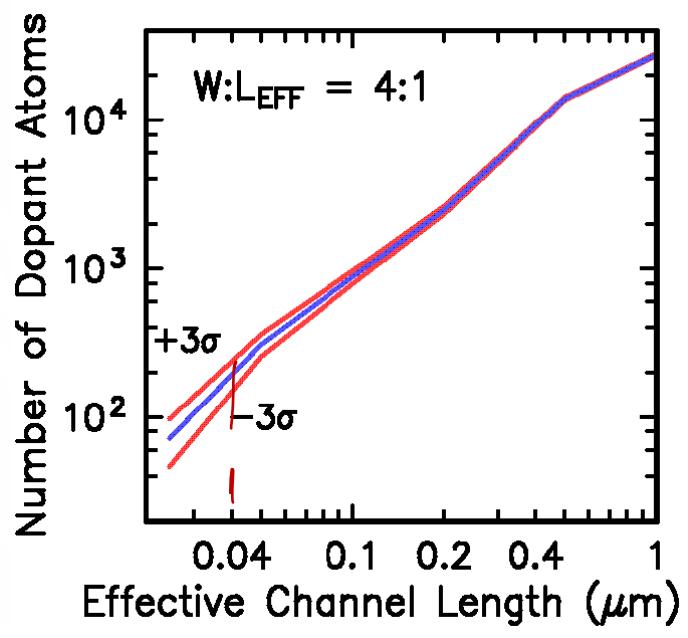
- $\Delta F \sim 3\%$, small changes in I_{LEAK}
- Due to STI-induced stress

Design Variability
Some Random Effects
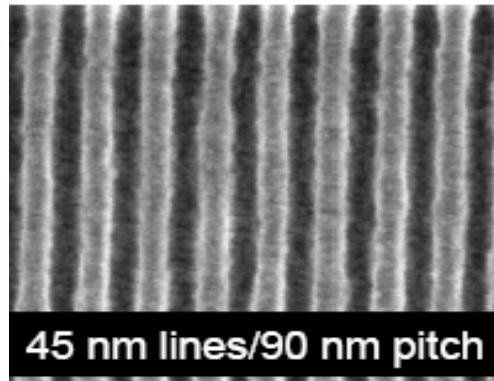
Random Dopant Fluctuations

- Number of dopants is finite


Frank, IBM J R&D 2002
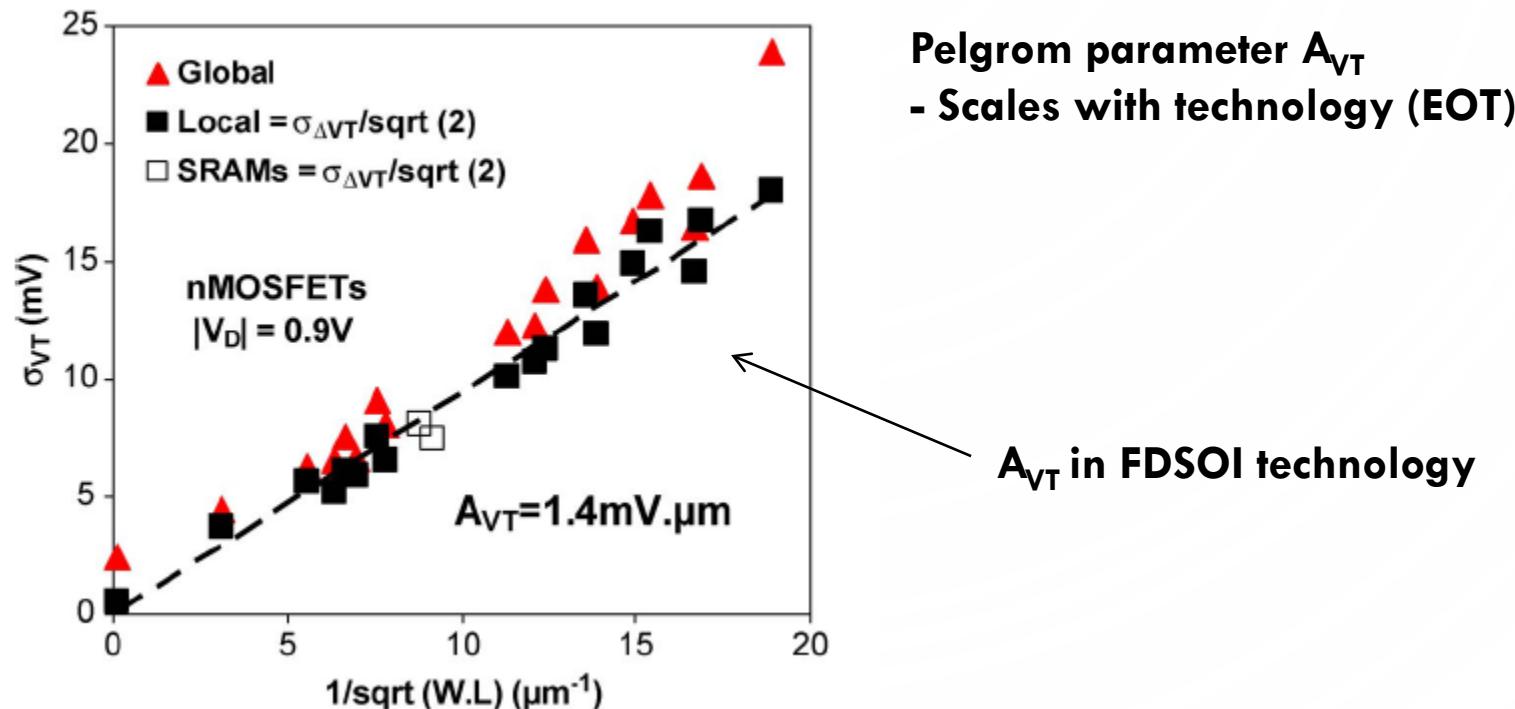
Processing: Line-Edge Roughness

45 nm lines/90 nm pitch

- Sources of line-edge roughness:
- Fluctuations in the total dose due to quantization
- Resist composition
- Absorption positions

Effect:
- Variation (random) in leakage and power

Transistor Matching

- V_{Th} matching of geometrically identical transistors varies with size $\sim\sqrt{WL}$ and distance


Pelgrom parameter A_{VT}
- Scales with technology (EOT)
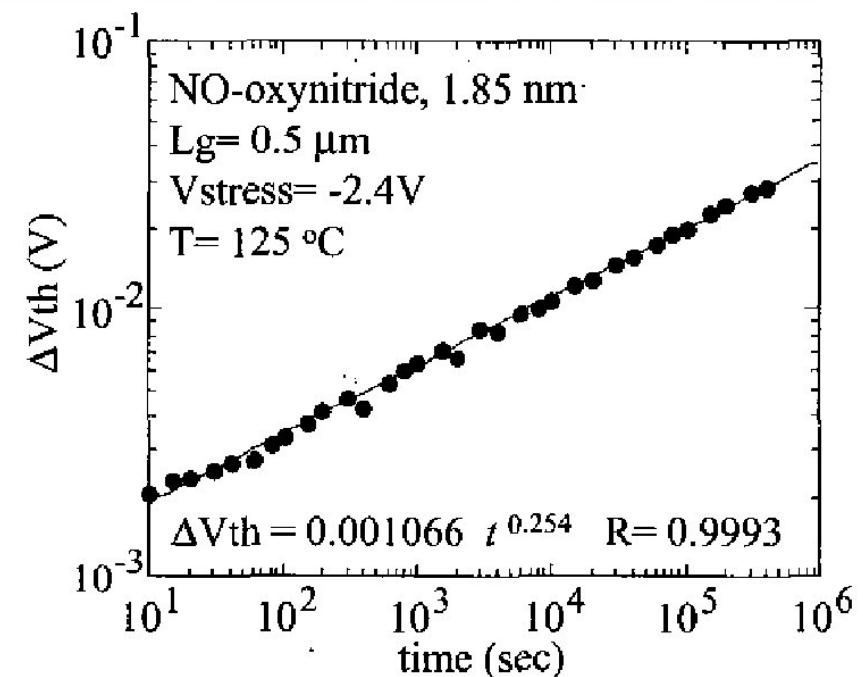
A_{VT} in FDSOI technology

J. Mazurier, Trans E.D., 2011.

Negative Bias Temperature Instability
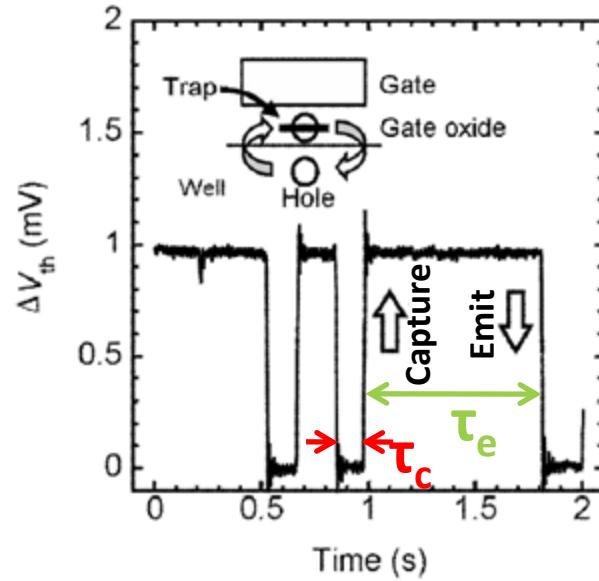
- PFET V_{Th}'s shift in time, at high negative bias and elevated temperatures

- 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

NO-oxynitride, 1.85 nm
Lg= 0.5 μm
Vstress= -2.4V
T= 125 °C

$\Delta Vth = 0.001066 \ t^{0.254}$ R= 0.9993
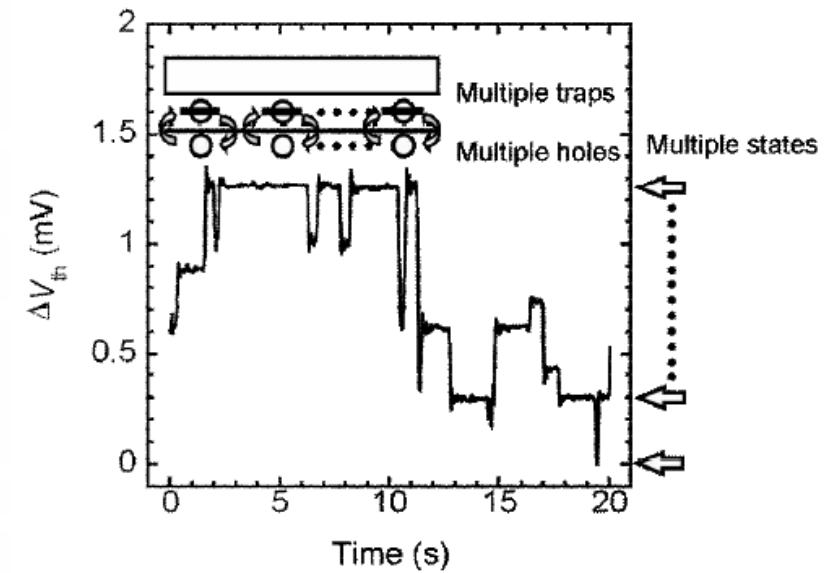
Random Telegraph Signal (RTS)

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
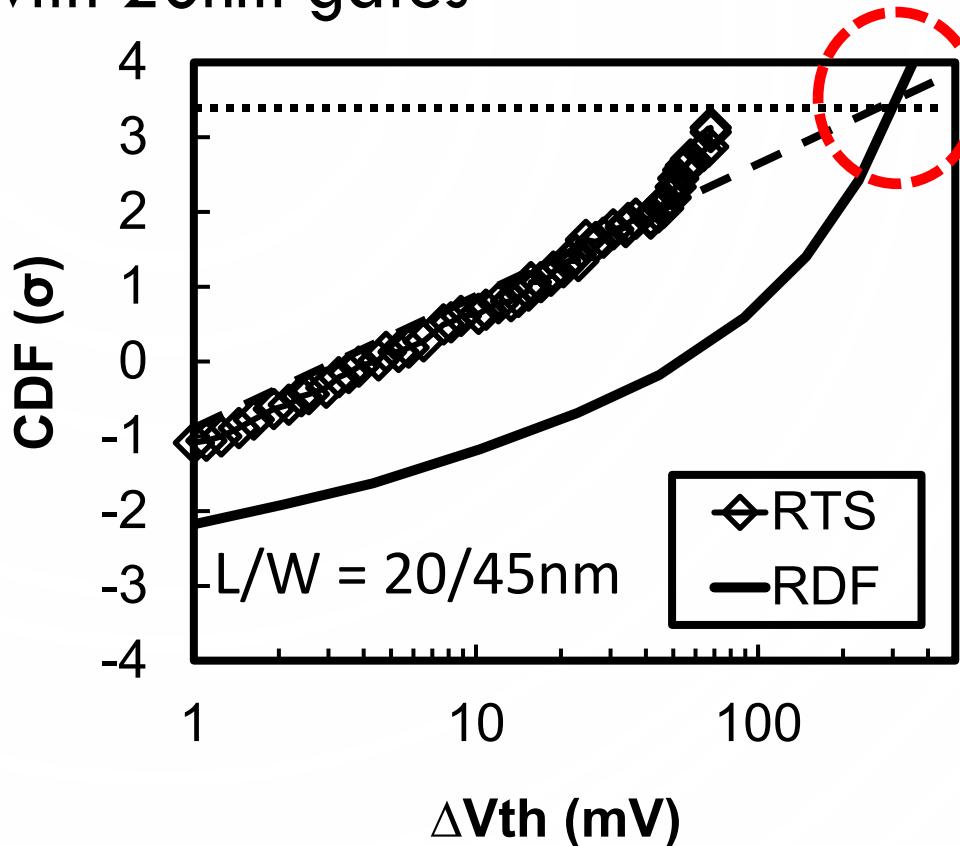
N. Tega et al, IRPS 2008.

RTS and Technology Scaling

- RTS exceeds RDF at 3 sigma with 20nm gates

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

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

L/W = 20/45nm

RTS

RDF

ΔVth (mV)

CDF (σ)

Tega *et. al*, VLSI Tech. 09

Summary

- Flip-flops:
  - Latch pairs
  - Pulse triggered

- Variability
  - Systematic
  - Random

Next Lecture

- Memories