Statistical Static Timing Analysis and Modeling of Parameter Variations

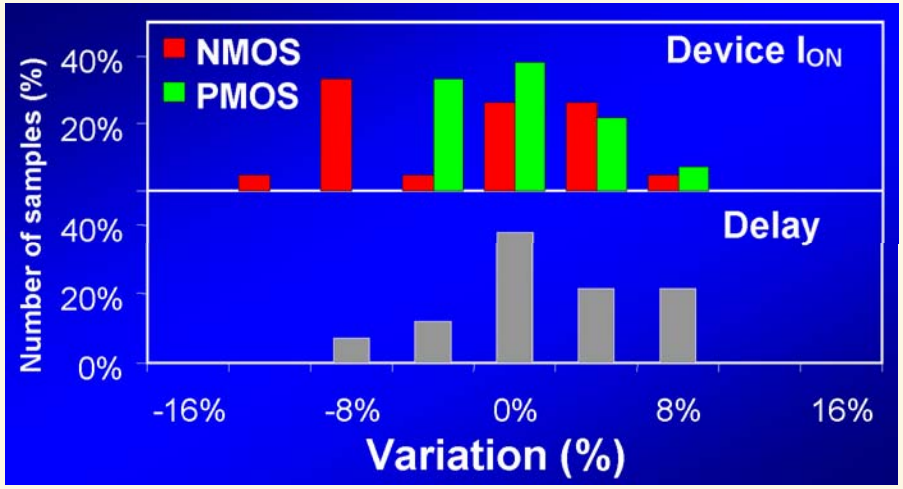
Frank Sill

14th November 2005

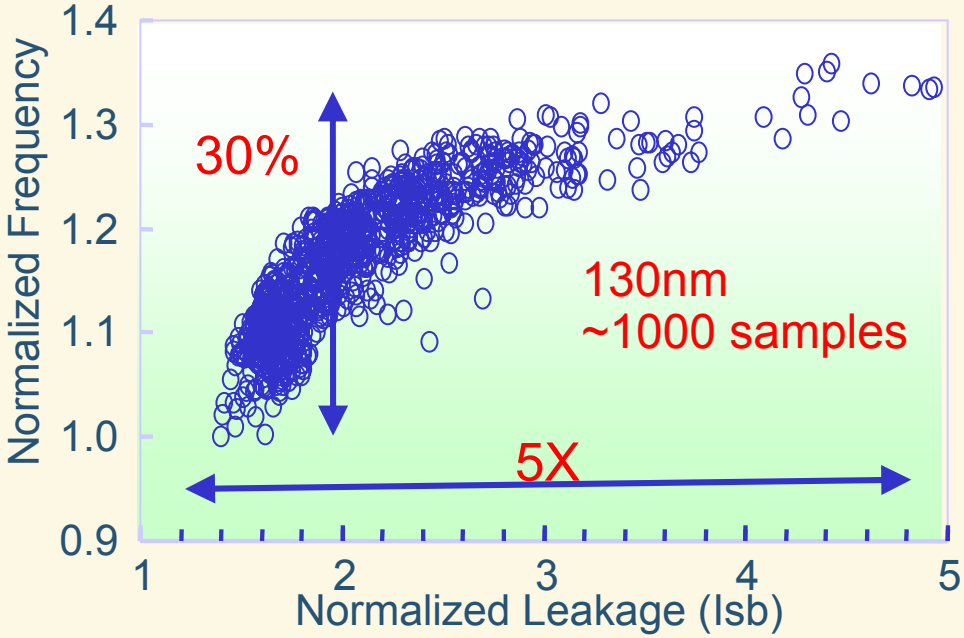
Outline

1) Parameter Variations in Upcoming Technologies
2) Static Timing Analysis (STA)
3) Statistical Static Timing Analysis (SSTA)
4) MVT and SSTA
5) Summary

Parameter Variations Increase



from: S. Brokar, VLSI'05

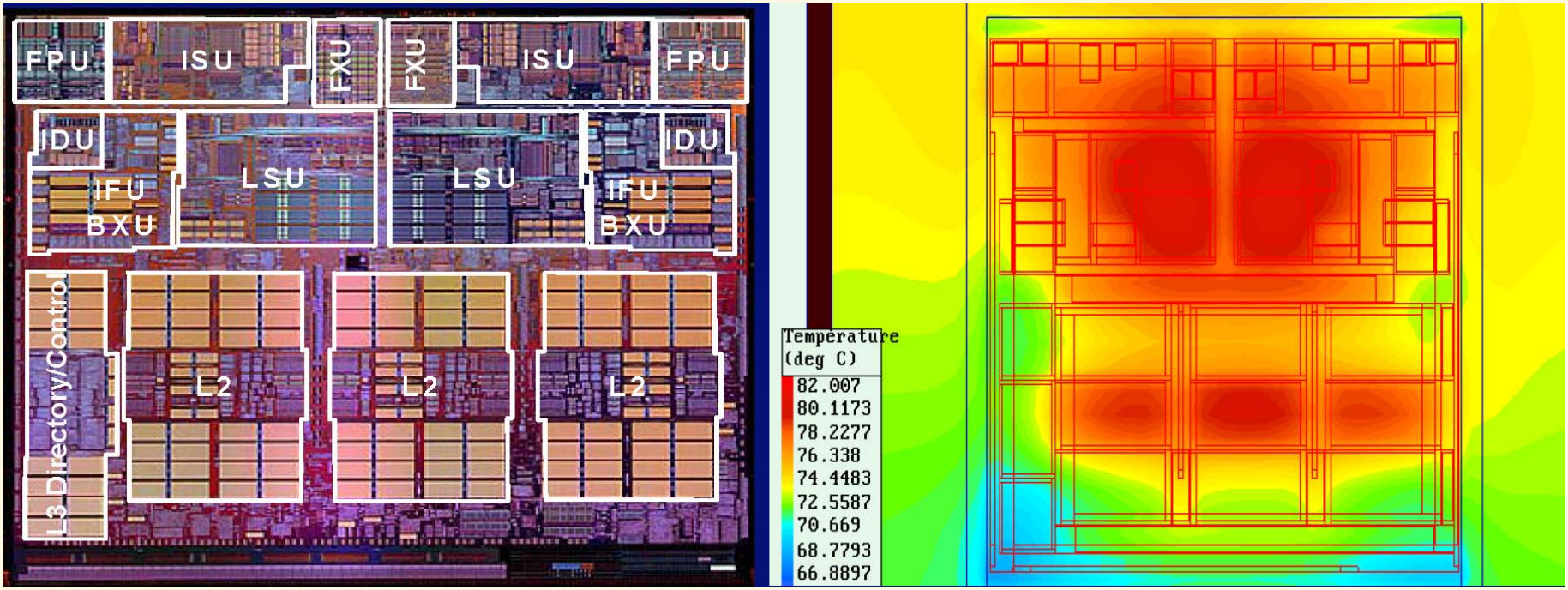


L (nm)	250	180	130	90	65	45
Vt (mV)	450	400	330	300	280	200
σ-Vt (mV)	21	23	27	28	30	32
σ-Vt/Vt	4.7%	5.8%	8.2%	9.3%	10.7%	16%

from: ITRS'03

Sources of Variability (1)

Temperature Variations + Hot spots

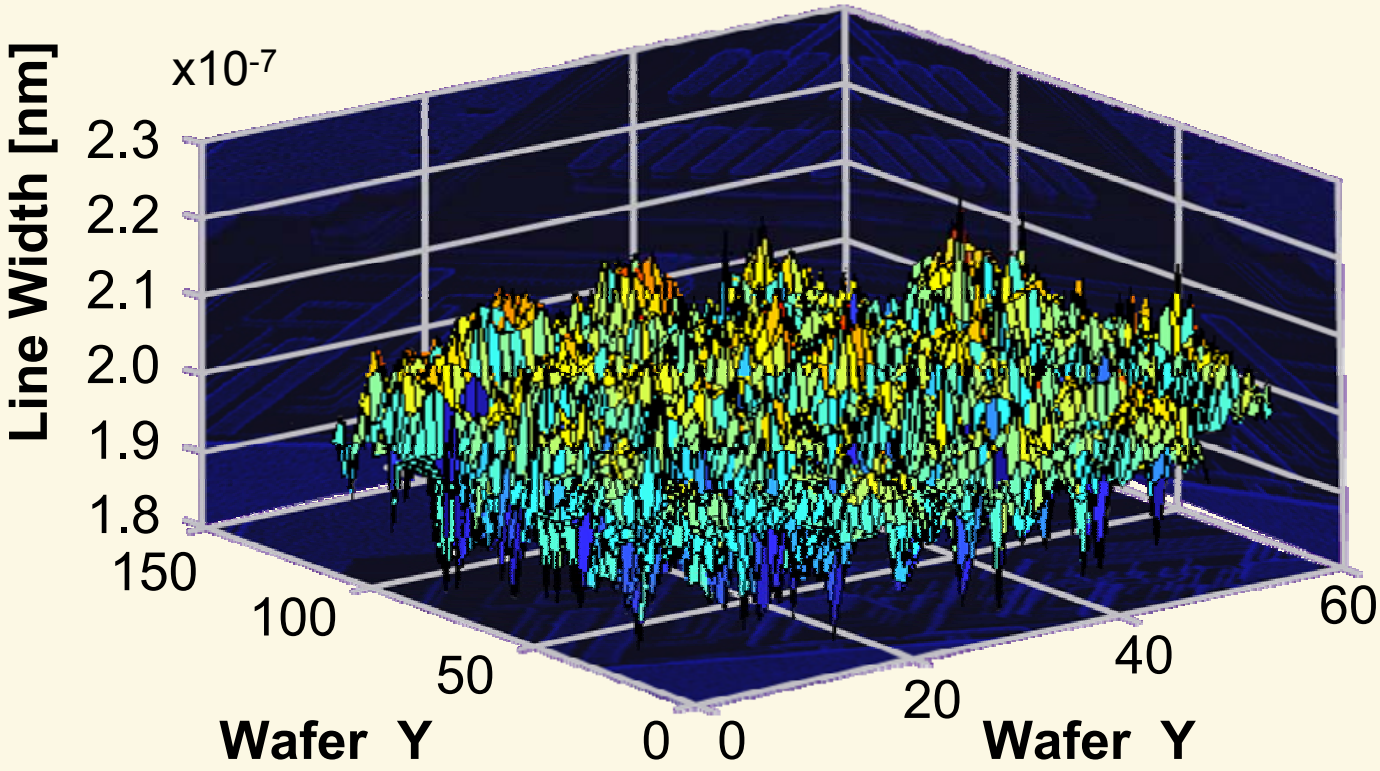


Power4 Server Chip

from: Devgan, ICCAD'03

Sources of Variability (2)

Varying Line Width

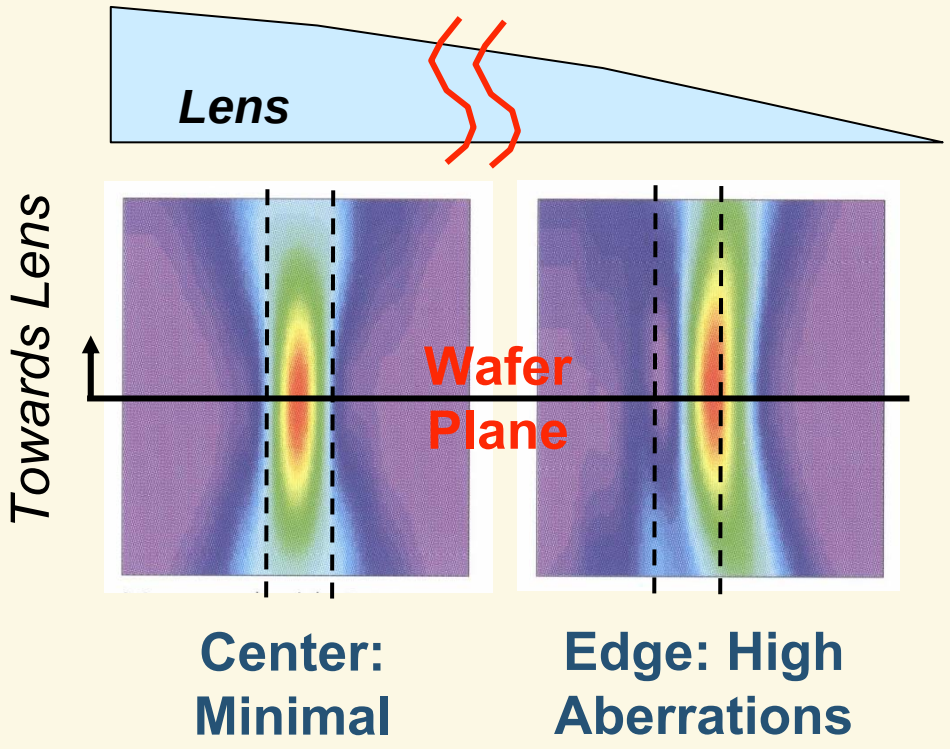


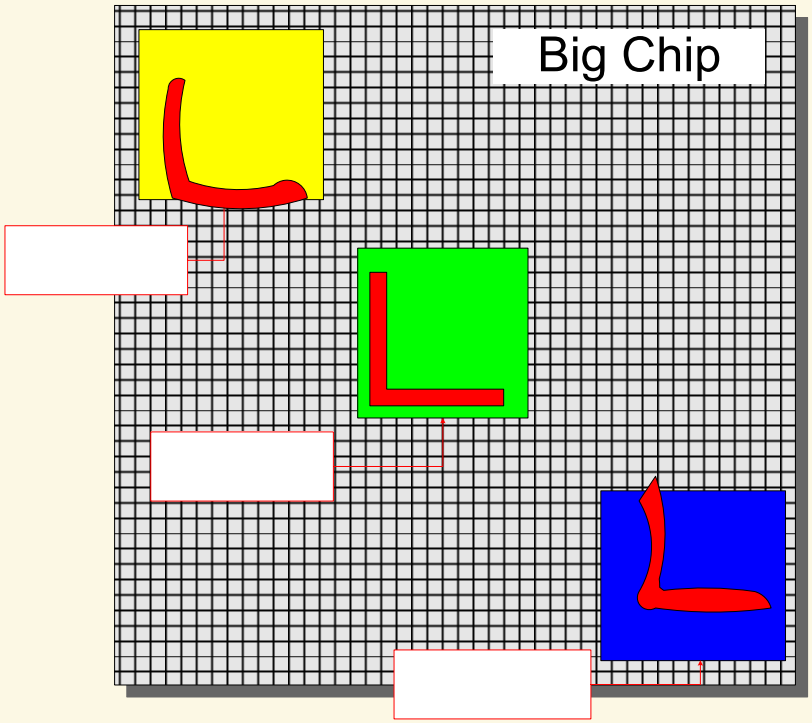
Spanos, UCB

Sources of Variability (3)

Field-dependent aberrations

$$\text{CELL_A}(X_1, Y_1) \neq \text{CELL_A}(X_0, Y_0) \neq \text{CELL_A}(X_2, Y_2)$$



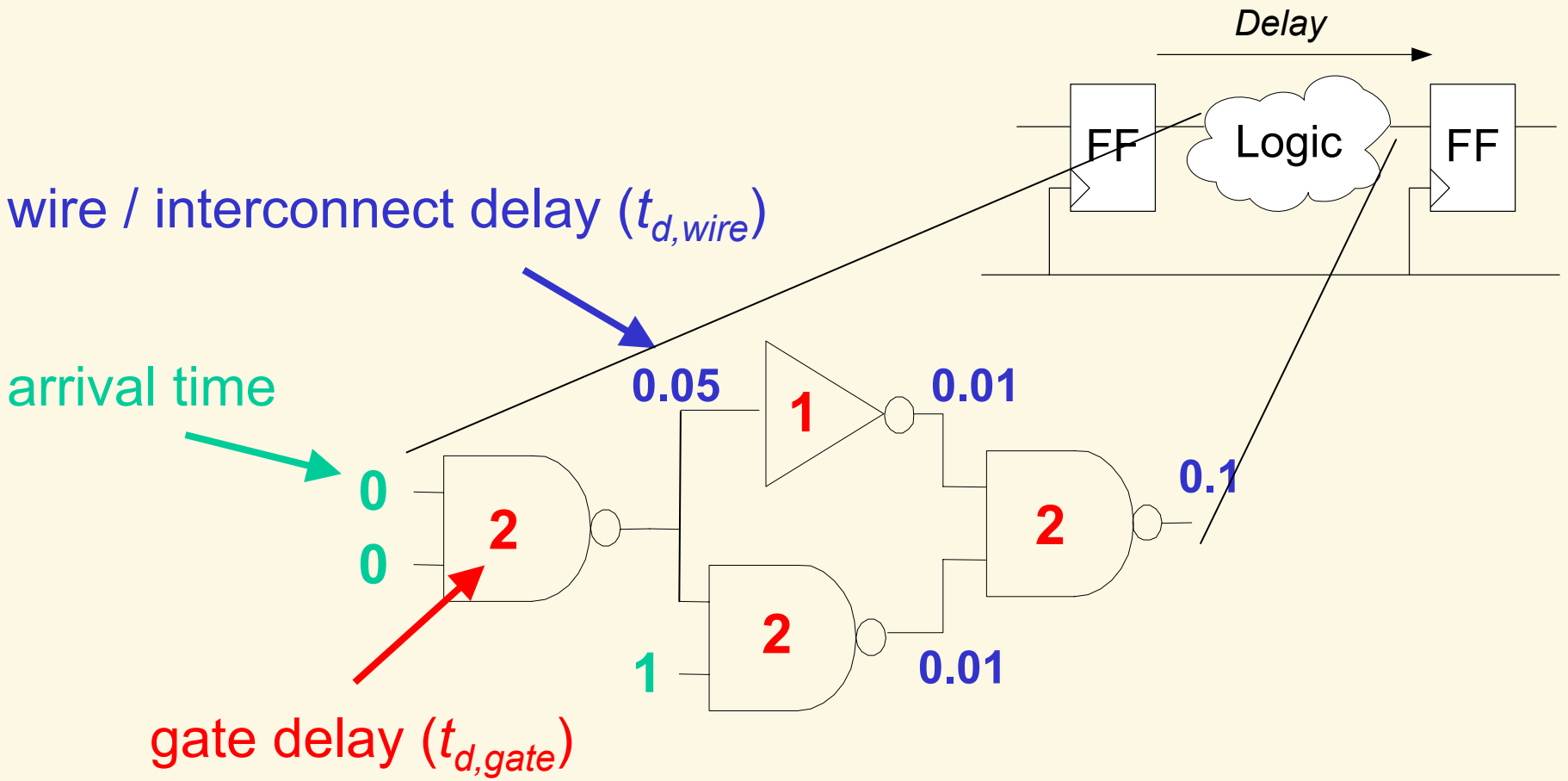


R. Pack, Cadence

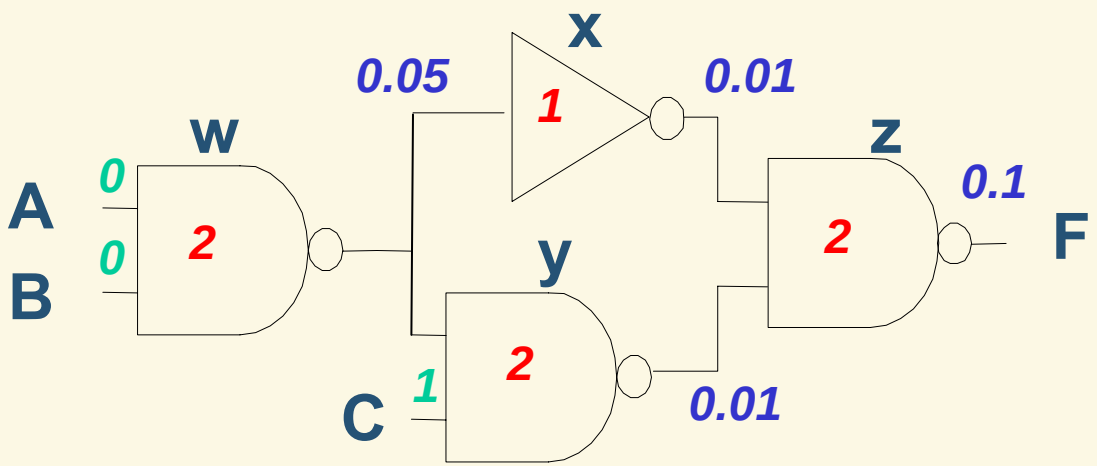
Outline

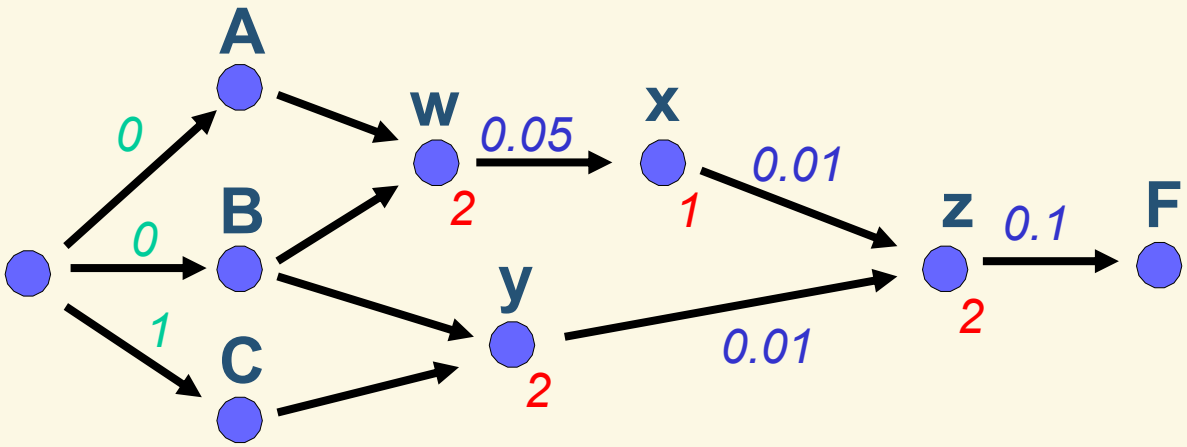
1) Parameter Variations in Upcoming Technologies
2) Static Timing Analysis (STA)
   - Basics
   - Design Corners
3) Statistical Static Timing Analysis (SSTA)
4) MVT and SSTA
5) Summary

Static Timing Analysis (1)



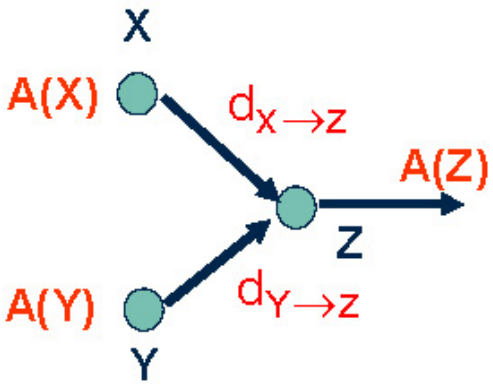
Static Timing Analysis (2)





Static Timing Analysis (3)

- **Arrival time A(v) for a node v is time when signal arrives at node v**

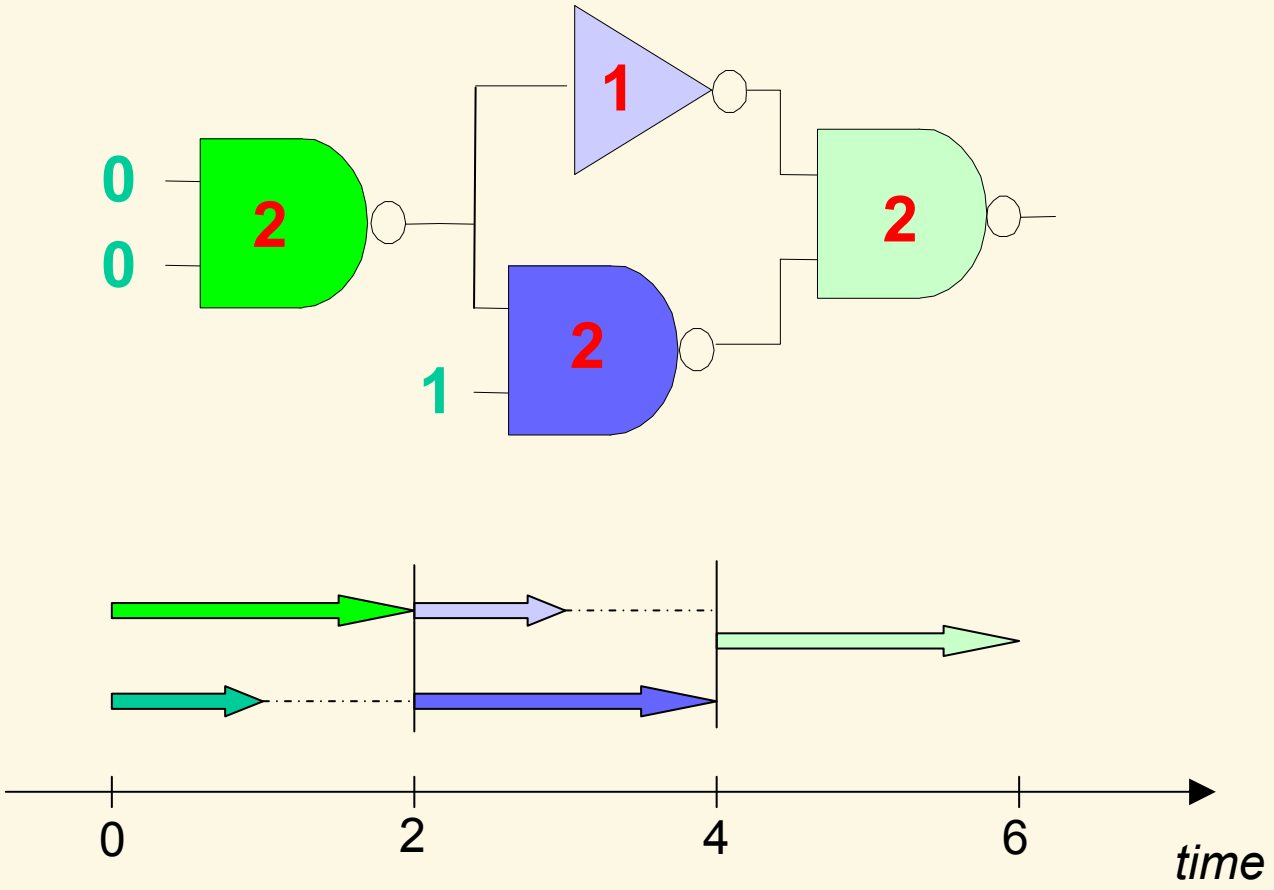


$$A(\upsilon) = \max_{u \in FI(\upsilon)} (A(u) + d_{u \to \upsilon})$$

where $d_{\upsilon \to u}$ is delay from υ to u, $FI(u) = \{X, Y\}$, **and** $\upsilon = \{Z\}$.

[from K. Keutzer]

Static Timing Analysis (4)



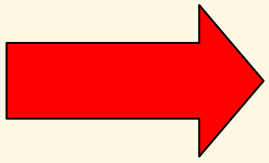
How affect Variations the Timing?

Gate delay:

$t_{d,gate}$ => function of: *temperature*, T_{ox}, W , L , V_{th}, *doping*, V_{DD}, ...

Wire/Interconnect delay:

$t_{d,wire}$ => function of: *temperature*, $C_{para.}$, W, …

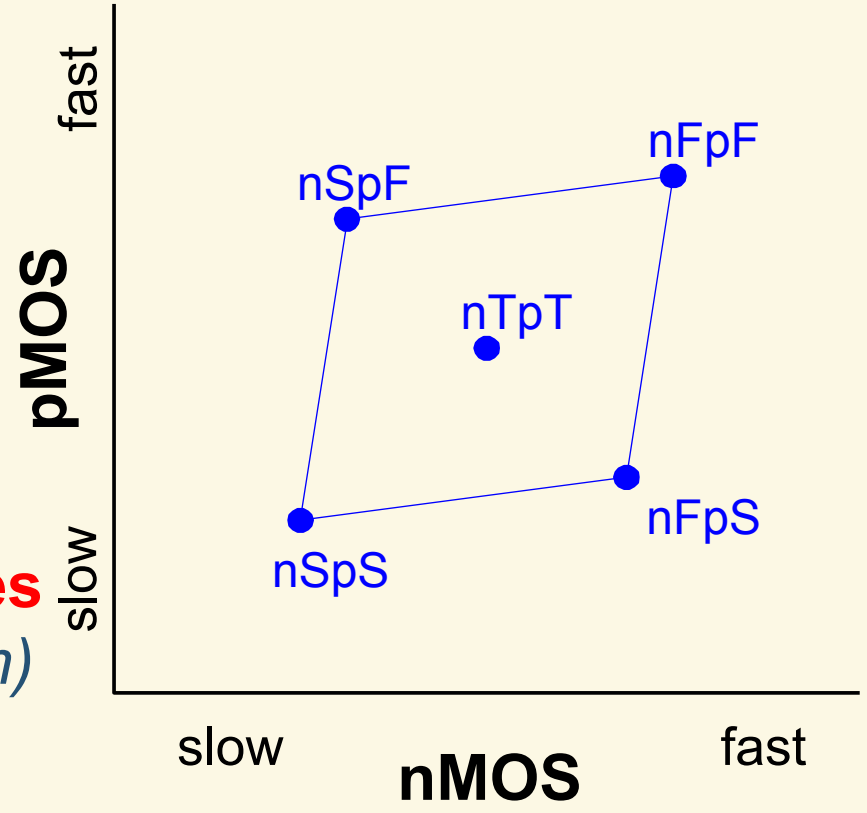


Each parameter can vary!

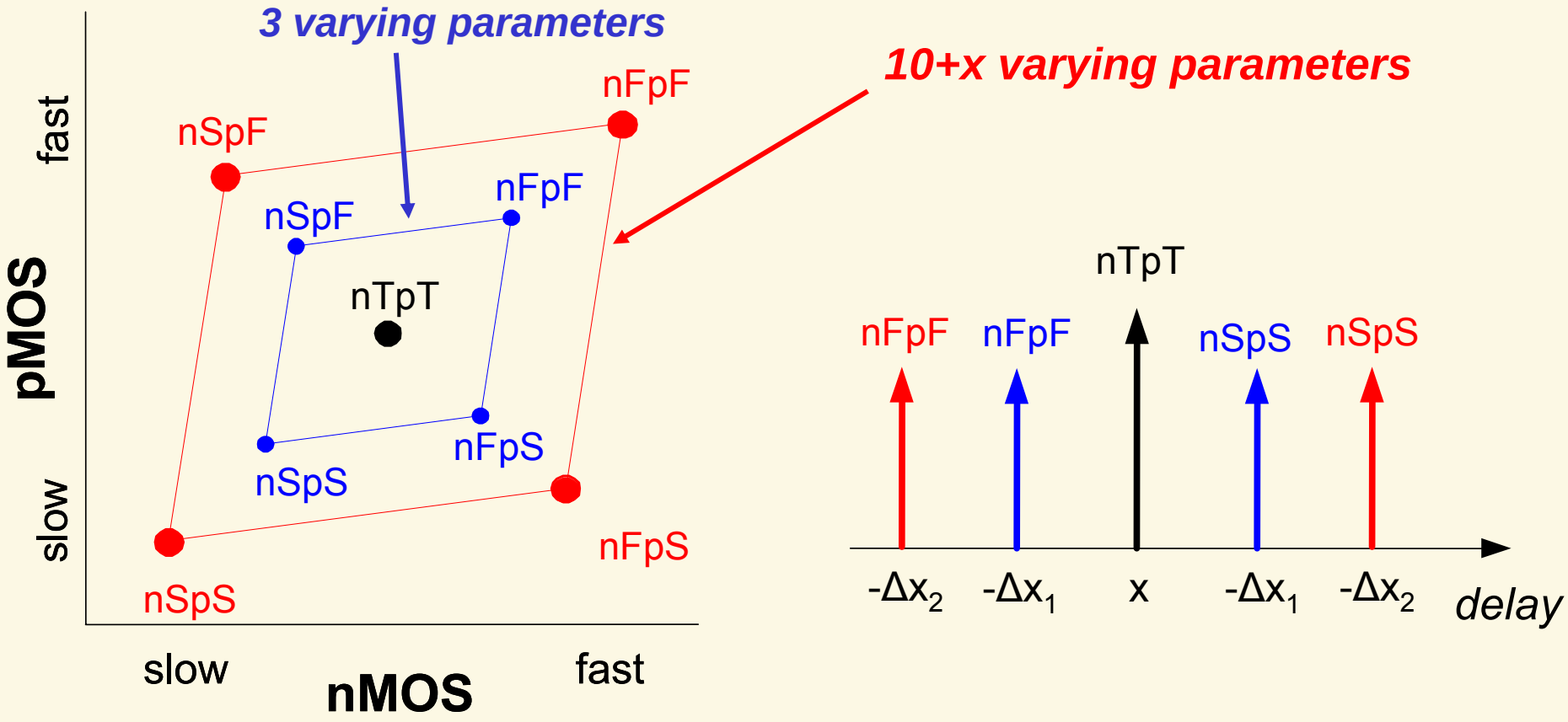
Until now: Design Corners

- *Design Corners = Processing + Environmental Variations*
- Box surrounding guaranteed circuit performance
- STA for all cases (usually 5)
- Typically 3 varying Parameters (*Temp.*, V_{DD}, *Process*)

too less for *DSM*-Technologies
(*D*eep *S*ub-*M*icron, L_{eff} < *100nm)*

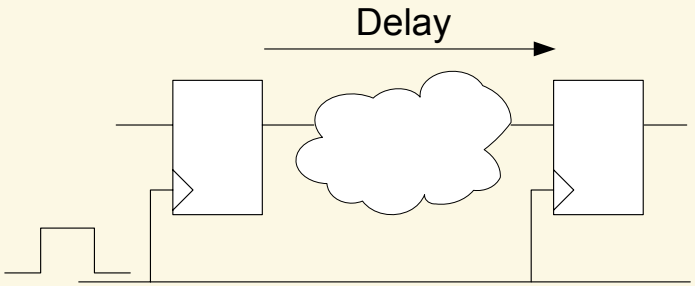


Design Corners in DSM (1)



Design Corners in DSM (2)

Problem: predicted delay very high => maximum allowed frequency is very low



$$frequency_{CLK} = \frac{1}{max(\,Delay\,)}$$

Solution: Probabilistic determination of timing behavior

Outline

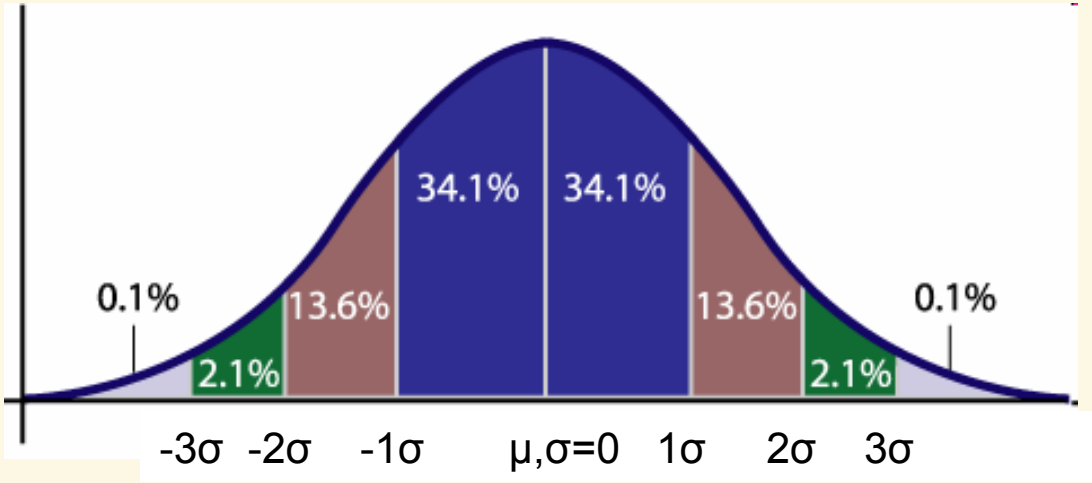
1) Parameter Variations in Upcoming Technologies
2) Static Timing Analysis (STA)
3) Statistical Static Timing Analysis (SSTA)
   - Basics
   - SSTA at Multi-Input Gates
   - Delay Extraction
4) MVT and SSTA
5) Summary

Gaussian Distribution

Most variations can be described as *Gaussian Distribution*

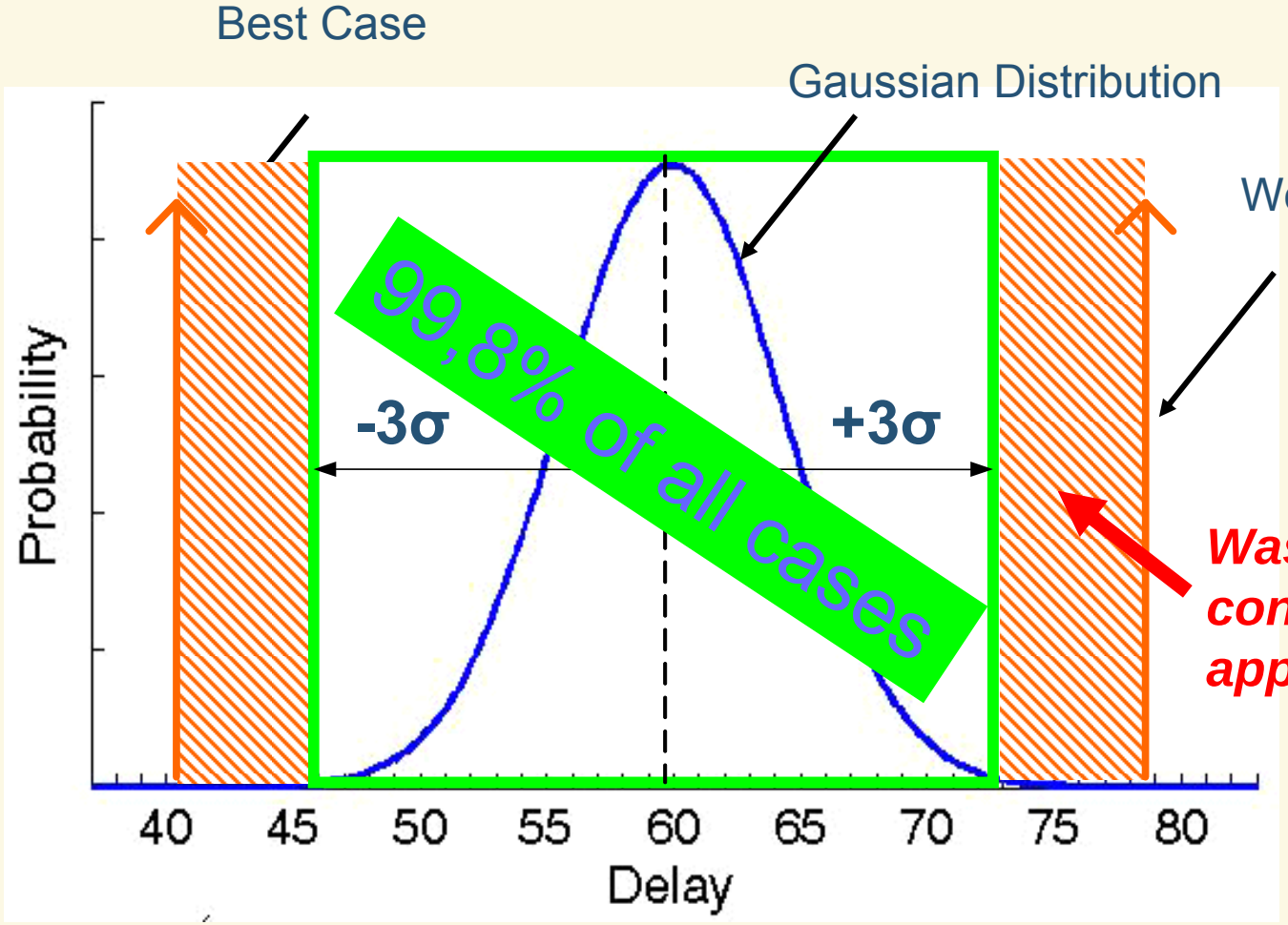
$$y(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

μ - mean value, σ – standard deviation



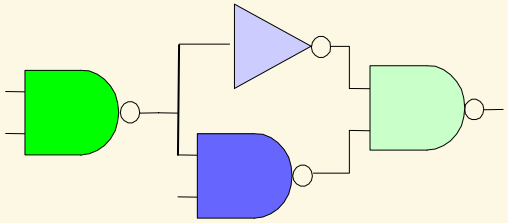
Copyright by wikipedia.com

Deterministic vs. Statistical STA

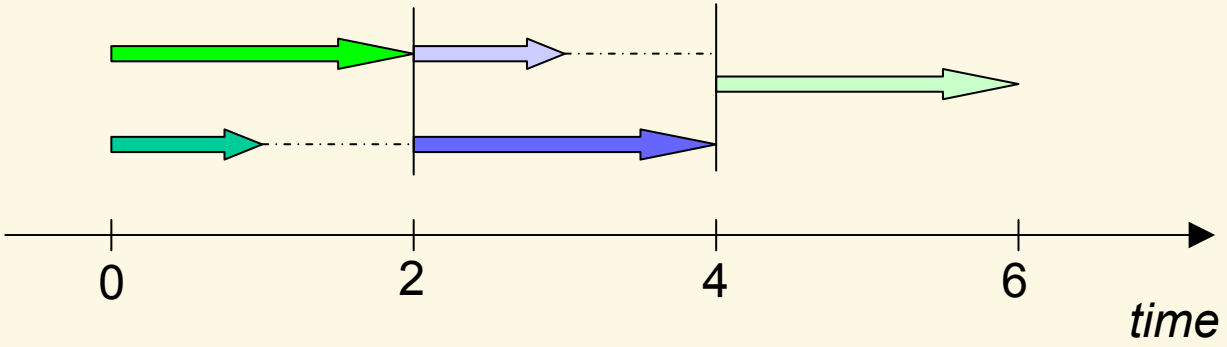


#10k Monte-Carlo simulations of NAND2 (Variation of NDEP, L_{eff}, W_{eff}, *Temp*, V_{th})

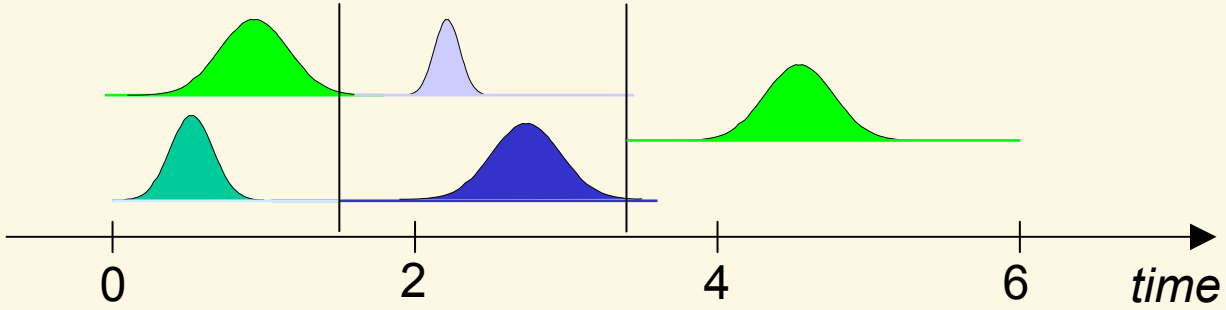
Statistical Static Timing Analysis



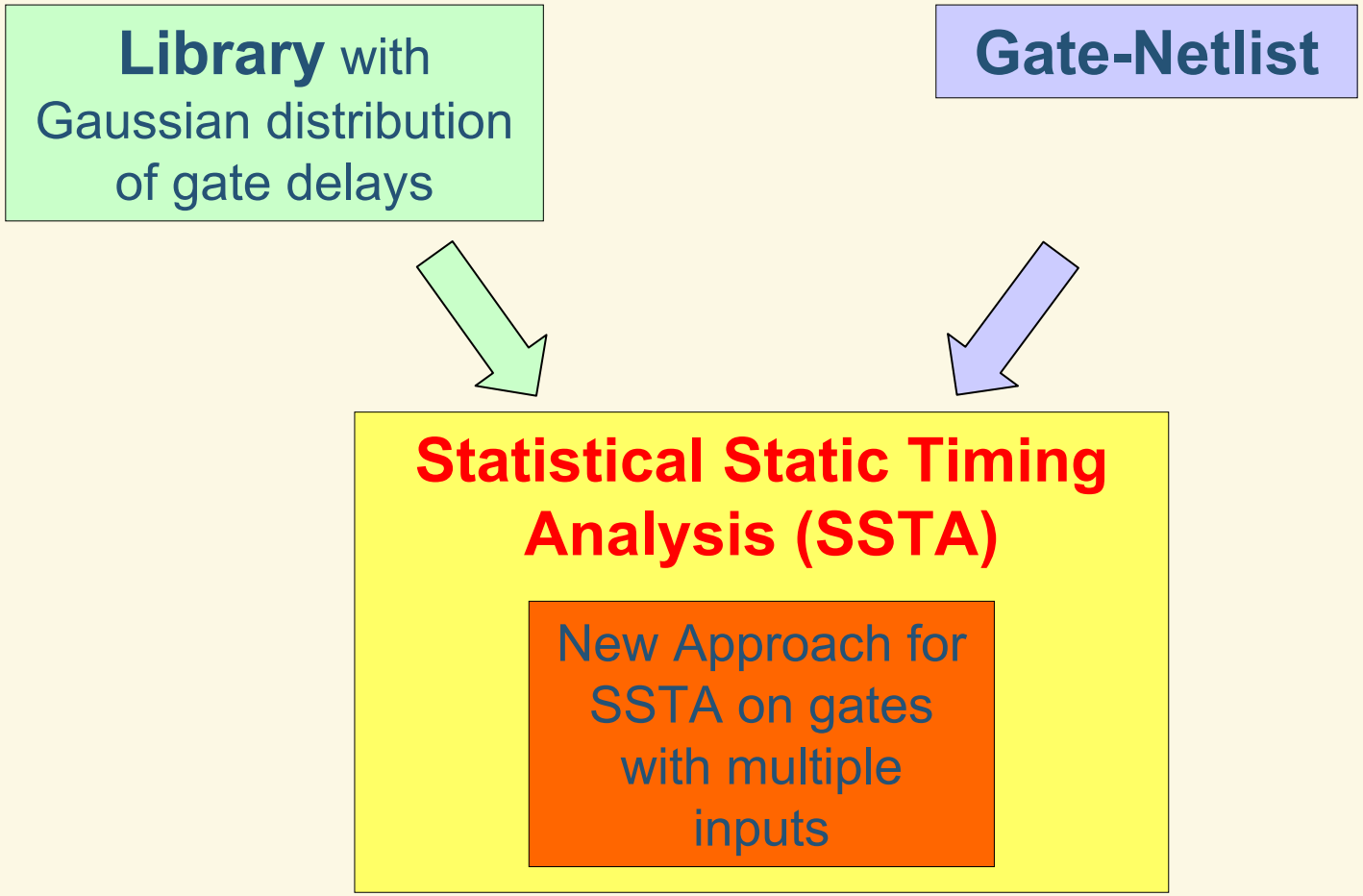
Deterministic STA:
(Worst Case Analysis)



Statistical STA:



Concept

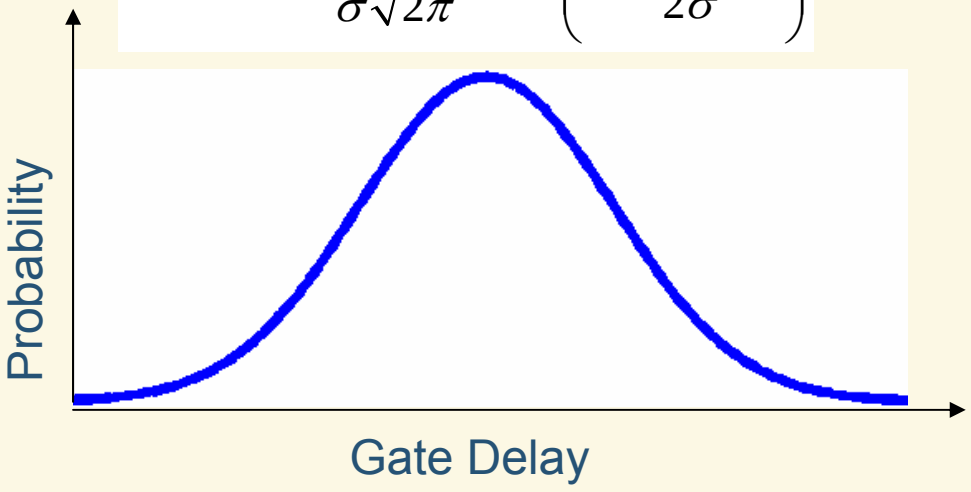

Library with
Gaussian distribution
of gate delays
Gate-Netlist
Statistical Static Timing
Analysis (SSTA)
New Approach for
SSTA on gates
with multiple
inputs

Modeling of Gate Delay

Probability Density Function (PDF)

Probability that gate delay has value x

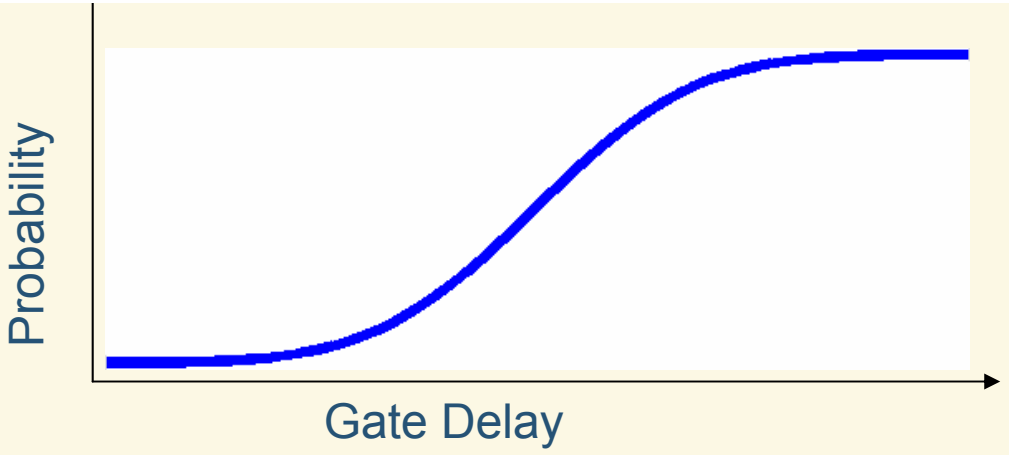
$$PDF(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



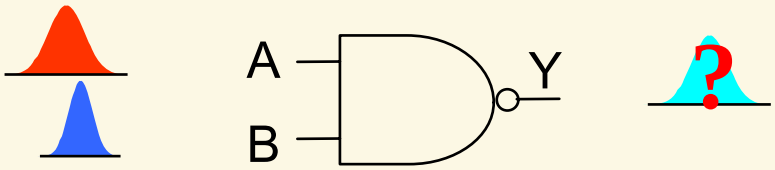
Cumulative probability Distribution Function (CDF)

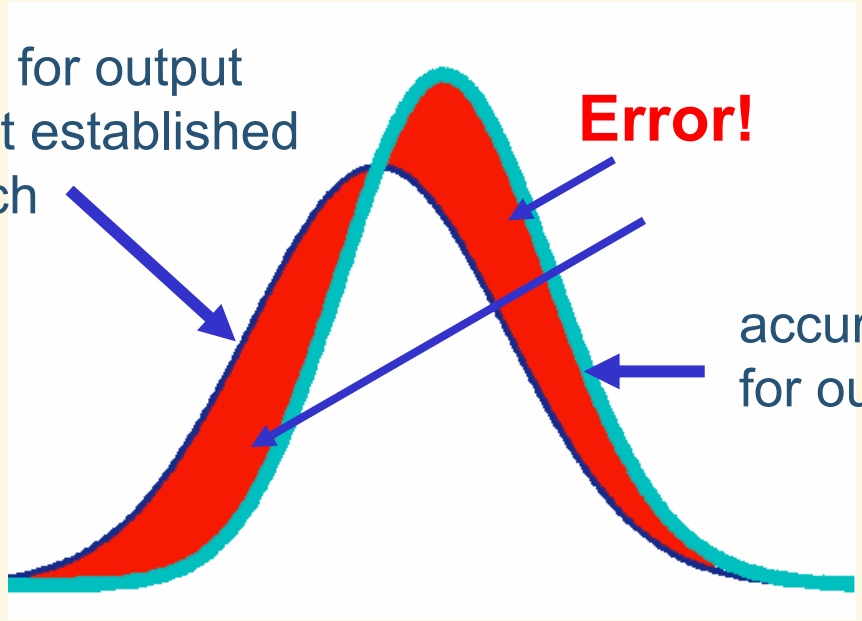
Probability that gate delay is lower than x

$$CDF(x) = \int_0^x PDF(t)dt = \int_0^x \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)dt$$



SSTA on Multi-Input Gates





Established Approach:

all but one input are ignored

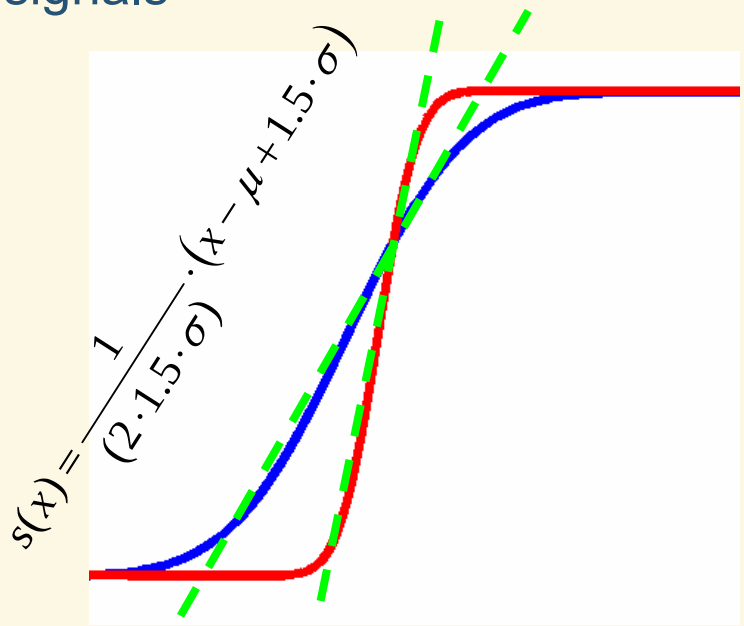
SSTA on Multi-Input Gates cont'd

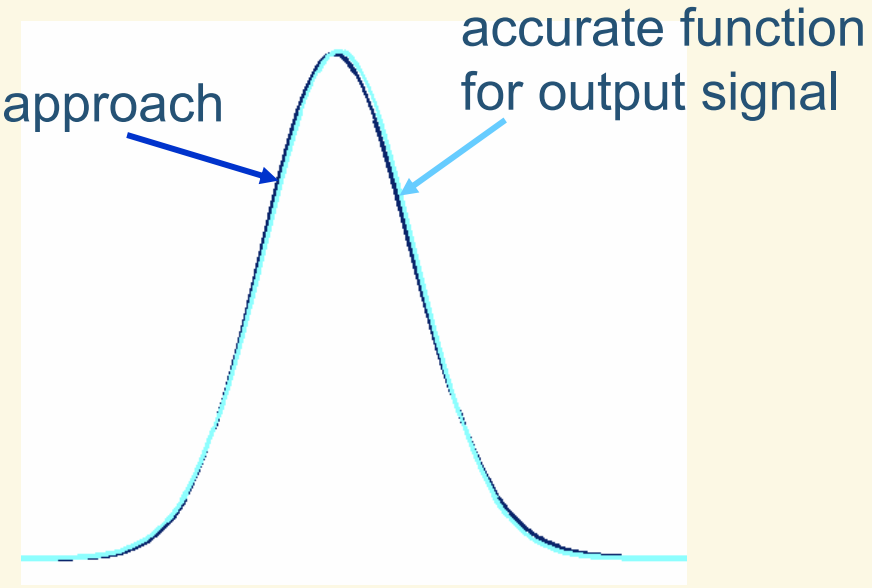
New Approach:

As function of output signal results from multiplication of all input CDFs

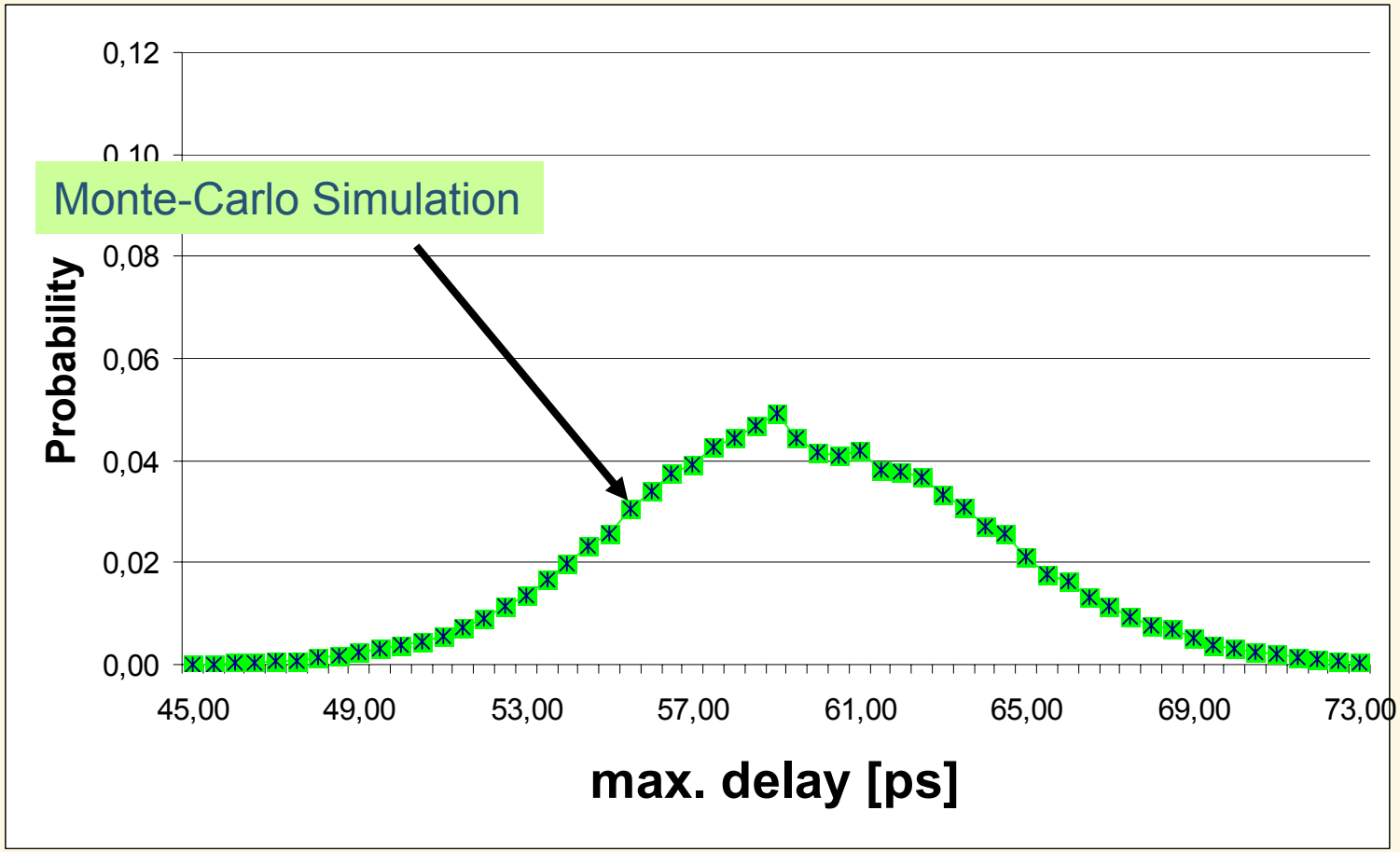
CDF is approximated as straight line

=> new CDF results from multiplication of all approximated straight lines of input signals





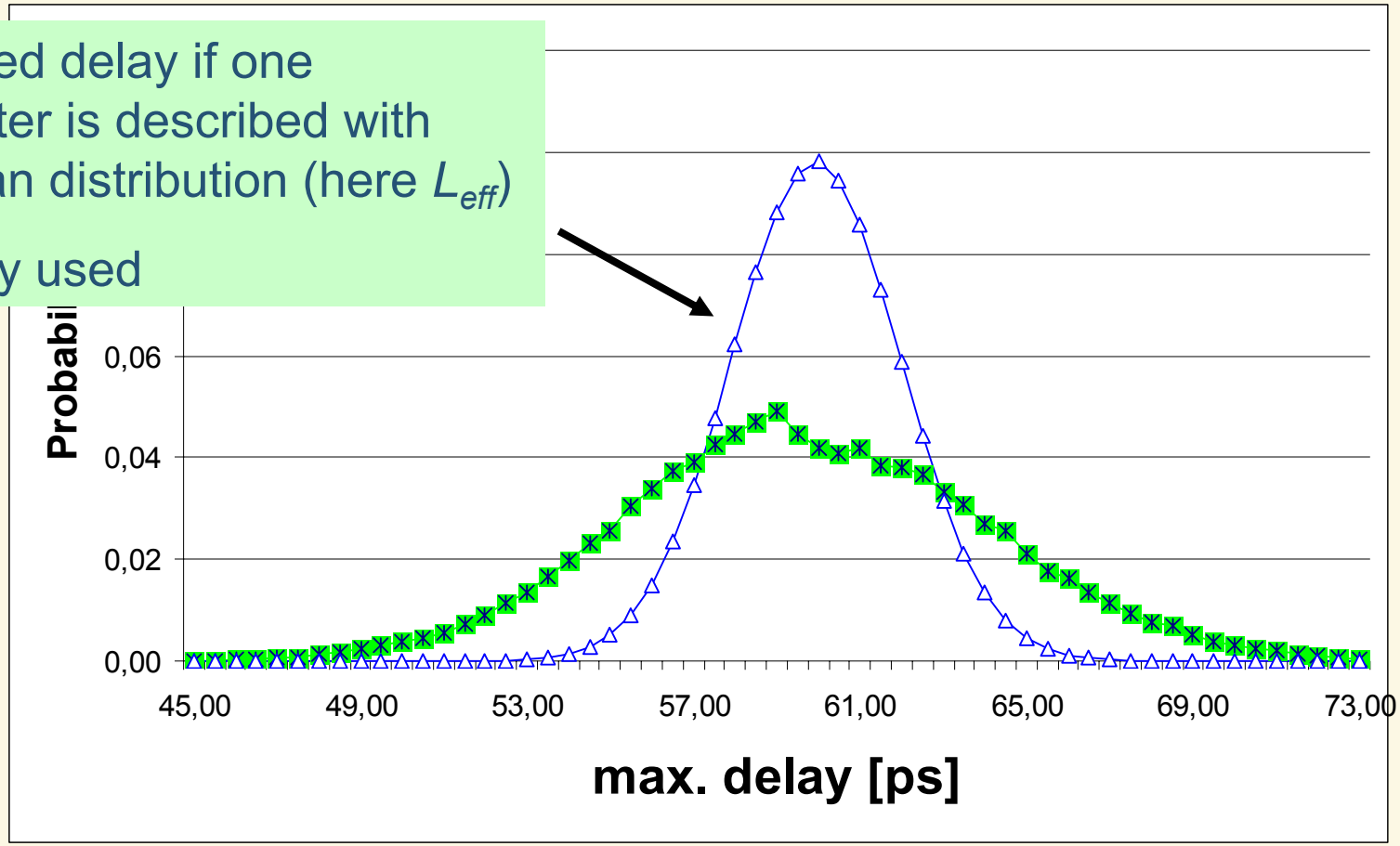
Delay - Extraction



NAND2 in predictive 65nm BPTM technology

Delay - Extraction

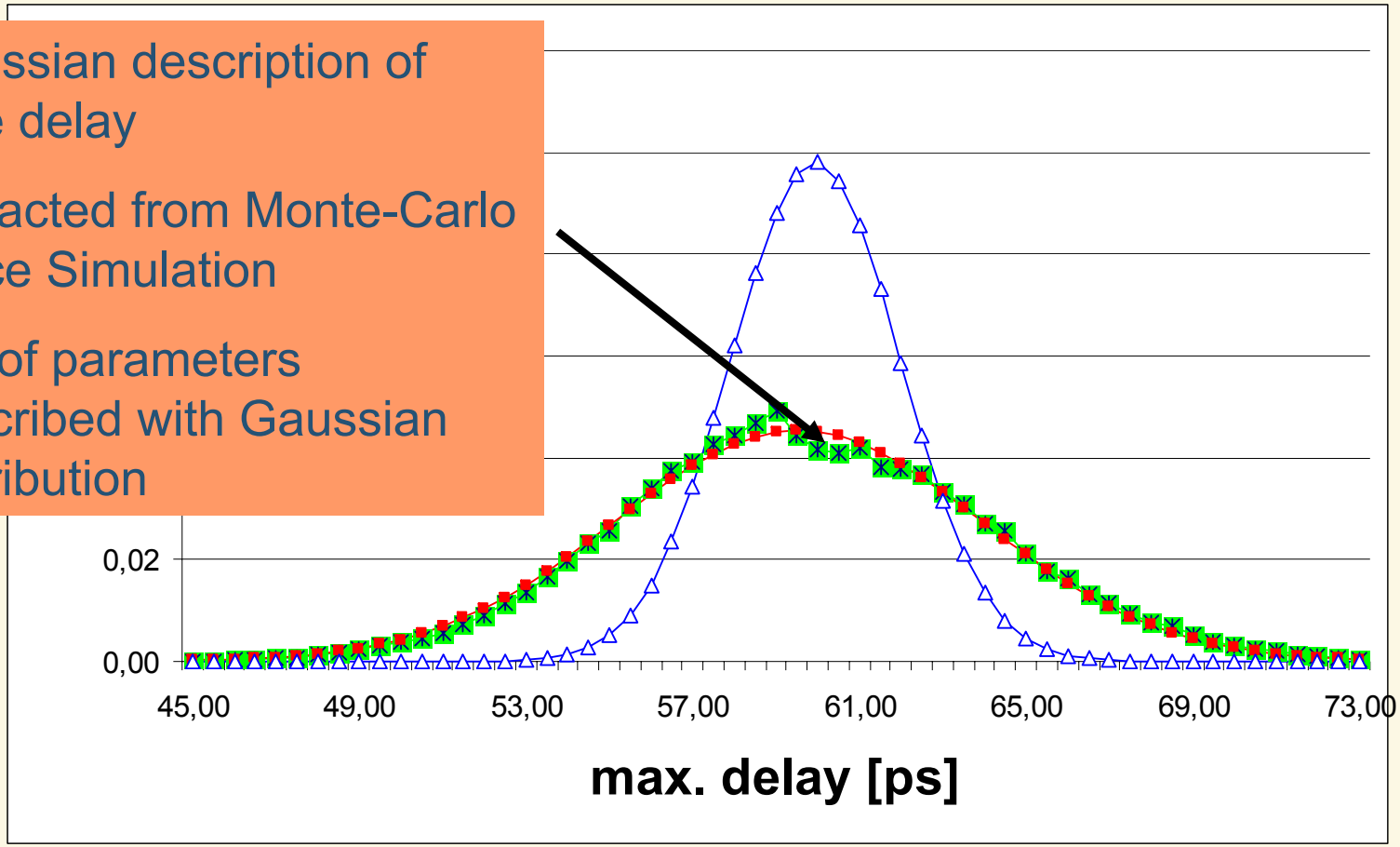
- Evaluated delay if one parameter is described with Gaussian distribution (here L_{eff})
- Normally used



NAND2 in predictive 65nm BPTM technology

Delay - Extraction

- Gaussian description of gate delay
- Extracted from Monte-Carlo Spice Simulation
- Set of parameters described with Gaussian distribution

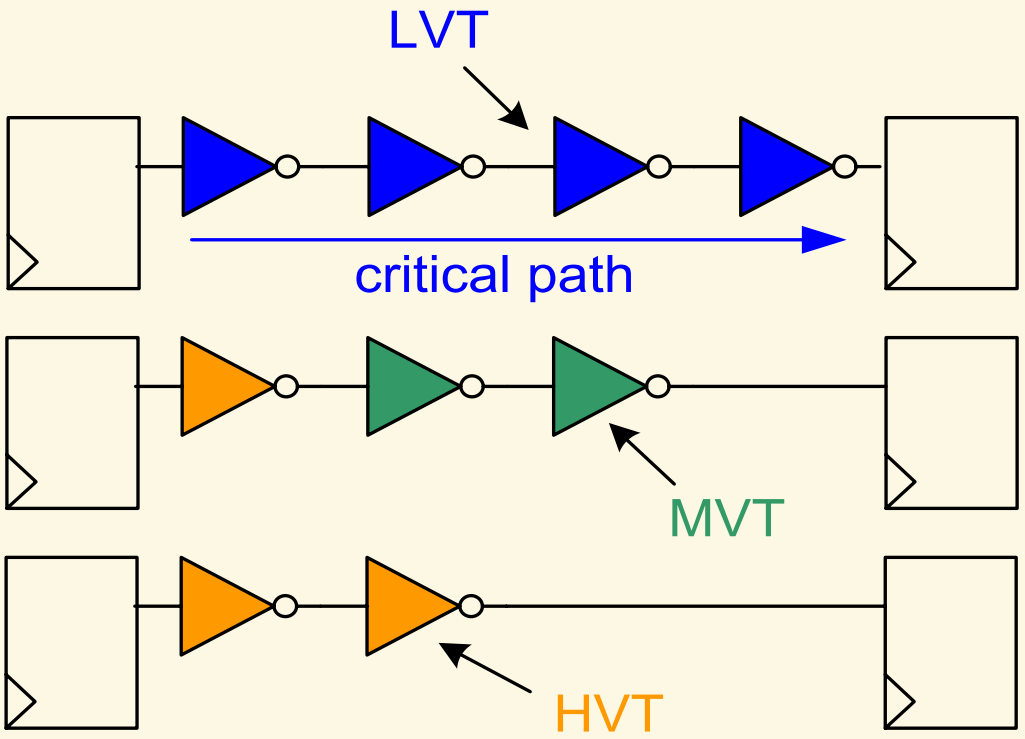


NAND2 in predictive 65nm BPTM technology

Outline

1) Parameter Variations in Upcoming Technologies
2) Static Timing Analysis (STA)
3) Statistical Static Timing Analysis (SSTA)
4) MVT and SSTA
   - Idea of MVT
   - Improvements in delay and leakage
5) Summary

Idea of MVT approach



use different types of Gates:

LVT: fast, high leakage

MVT: medium speed,
medium leakage

HVT: slow, low leakage

Goal: *maximum count of MVT / HVT gates at constant delay*

SSTA and MVT

Two improvements possible:

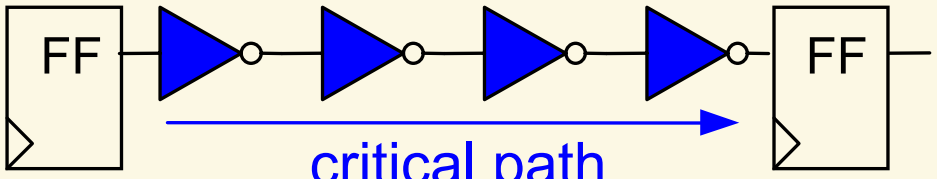
1. Increased *Performance*

   or

2. Reduced *Leakage*

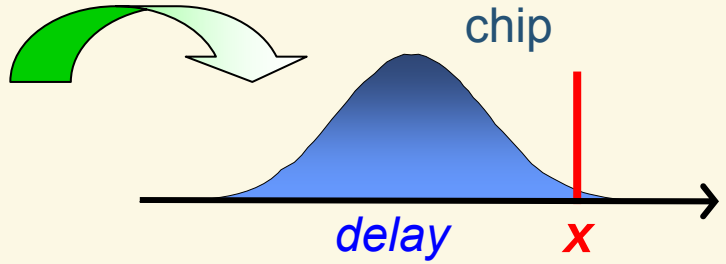
SSTA & MVT: Increased Performance

For a chosen reliability *R* => all critical paths are faster than a specific delay



Probability distribution of delay of output signal

For a reliability *R*, delay is lower than x on every chip



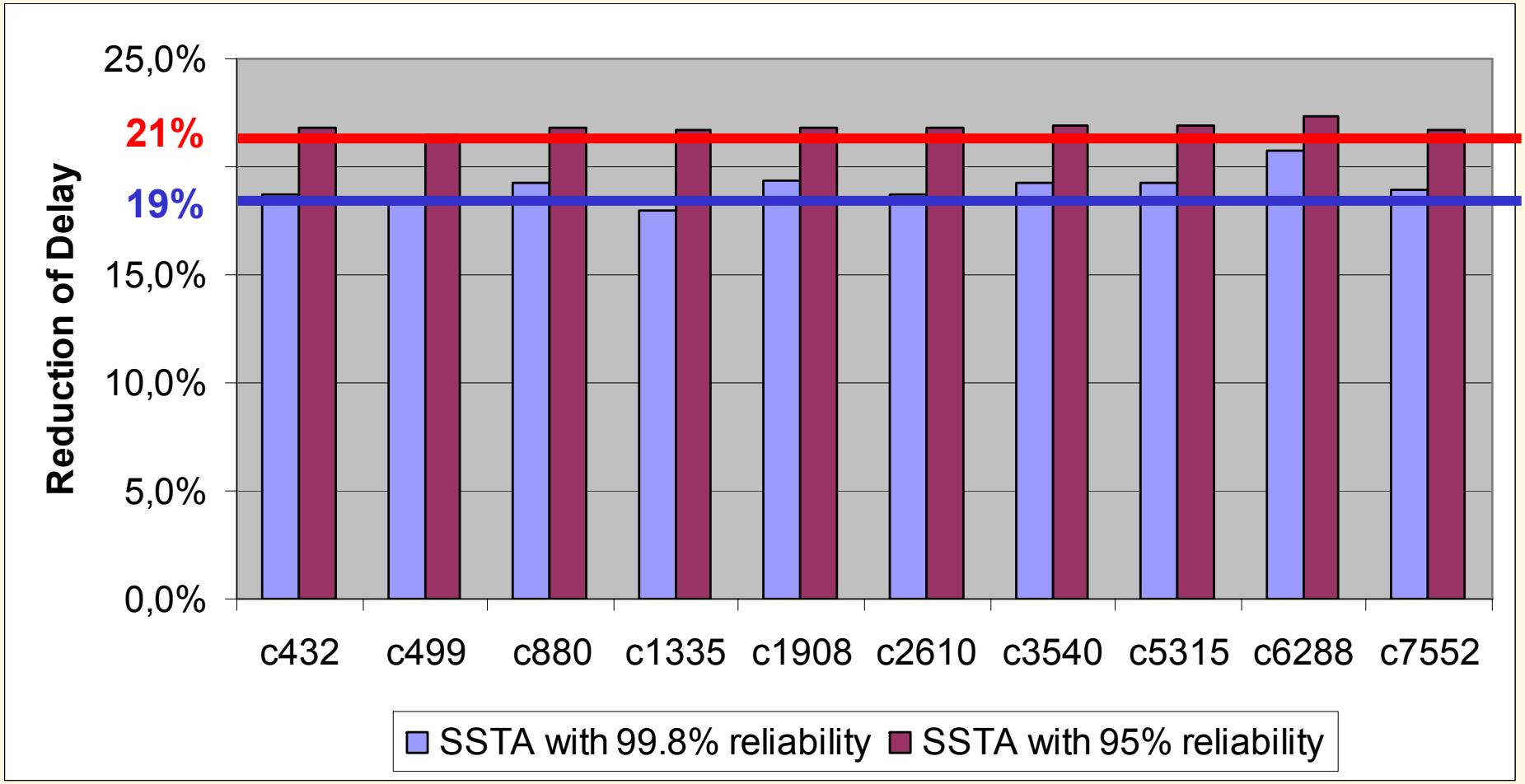
Example:

for ***R = 100 %*** the delay is always lower than ***1400 ps*** (worst case)

for ***R = 99,8 %*** the delay is always lower than ***1200 ps***

for ***R = 95 %*** the delay is always lower than ***1000 ps***

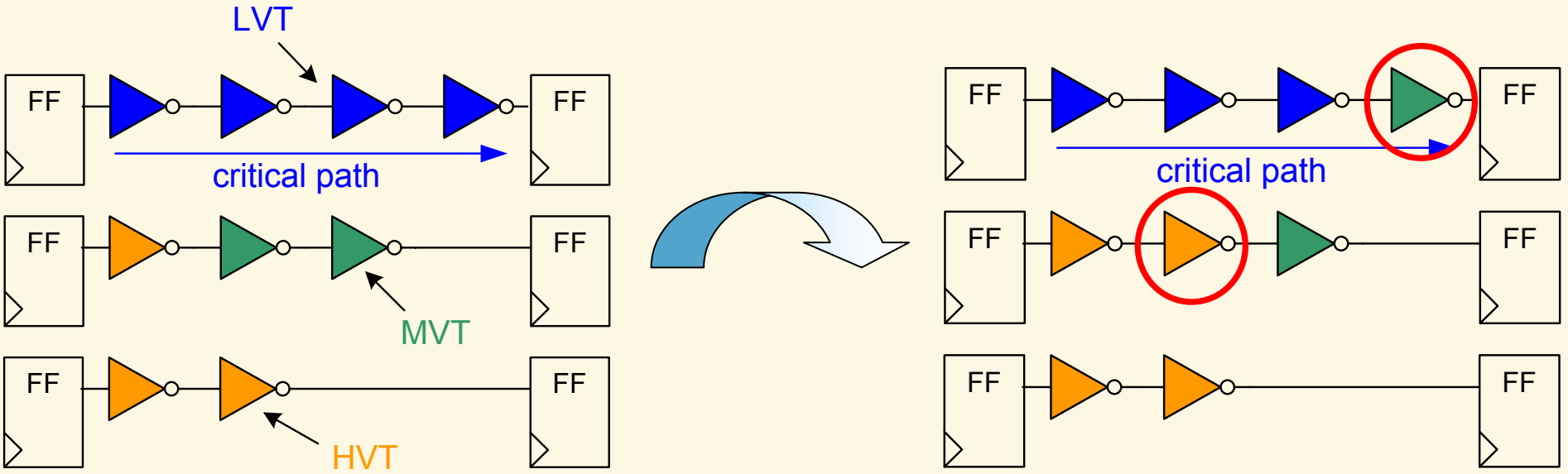
Results: Delay Reduction



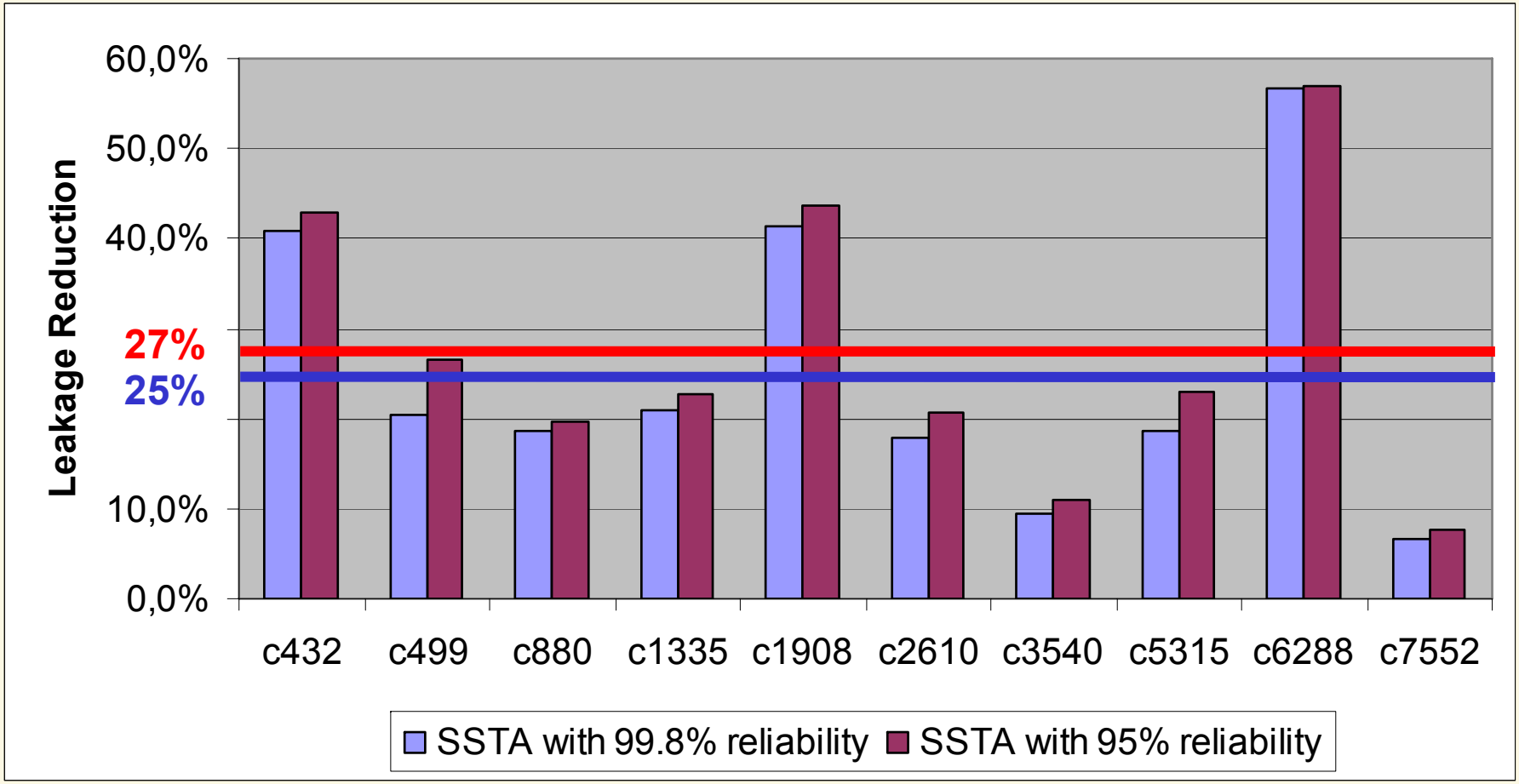
SSTA & MVT: Reduced Leakage

- Estimation of worst case delay (100% reliability)
- SSTA at allocation of HVT / MVT gates

Delay stays the same for a chosen reliability *R*, but count of low leakage gates increases



Results: Leakage Reduction compared to MVT-optimized Circuits



Summary

- Parameter variations, which affect gate behavior, increase in upcoming technologies
- Low performance at 100% reliability
- Statistical Analysis (SSTA) allows higher performance at small decrease of reliability
- SSTA combined with MVT can reduce leakage or can increase performance