VARIUS-TC: A Modular Architecture-Level Model of Parametric Variation for Thin-Channel Switches

S. Karen Khatamifard, Michael Resch, Nam Sung Kim†, Ulya R. Karpuzcu

University of Minnesota
{khata006, resco059, ukarpuzc}@umn.edu

†University of Illinois
nskim@illinois.edu

10/5/2016

UNIVERSITY OF MINNESOTA

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Motivation

- Manufacturing process becomes less controllable under aggressive scaling.

Motivation

- Manufacturing process becomes less controllable under aggressive scaling.
  - ➢ Deviation of device parameters from nominal becomes more likely.

Motivation

- Manufacturing process becomes less controllable under aggressive scaling.
  - Deviation of device parameters from nominal becomes more likely.
  - Performance and power of manufactured hardware become unpredictable.

Motivation

- Manufacturing process becomes less controllable under aggressive scaling.
  - Deviation of device parameters from nominal becomes more likely.
  - Performance and power of manufactured hardware become unpredictable.
  - We need system-level models of variation at early stages of the design.

Motivation

- Manufacturing process becomes less controllable under aggressive scaling.
  - Deviation of device parameters from nominal becomes more likely.
  - Performance and power of manufactured hardware become unpredictable.
  - We need system-level models of variation at early stages of the design.

- **Contribution: VARIUS-TC**

Motivation

- Manufacturing process becomes less controllable under aggressive scaling.
  - Deviation of device parameters from nominal becomes more likely.
  - Performance and power of manufactured hardware become unpredictable.
  - We need system-level models of variation at early stages of the design.

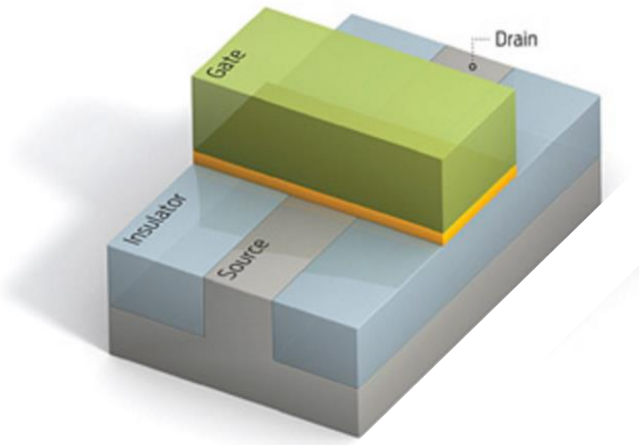
- **Contribution: VARIUS-TC**
  - Modeling process variation in **emerging devices** at **architecture-level**.

Thin Channel (TC) devices

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

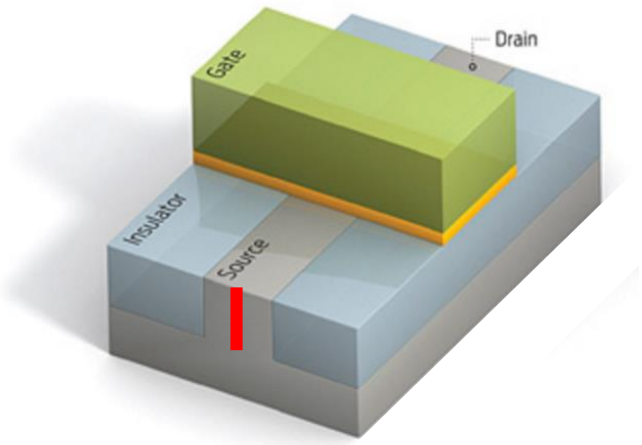
Traditional Planar CMOS



[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS

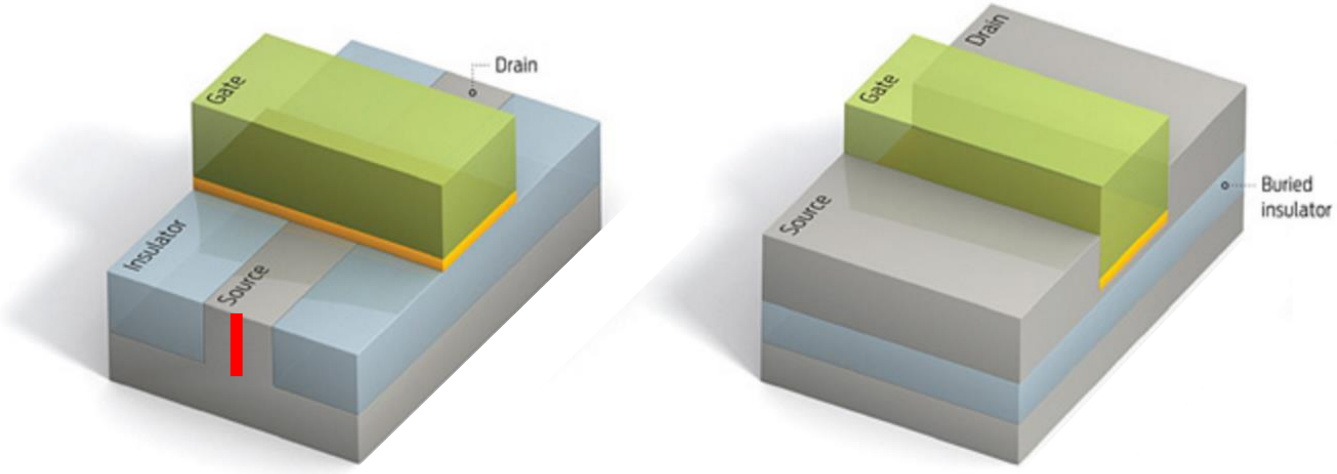


Channel Thickness

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS **SOI**

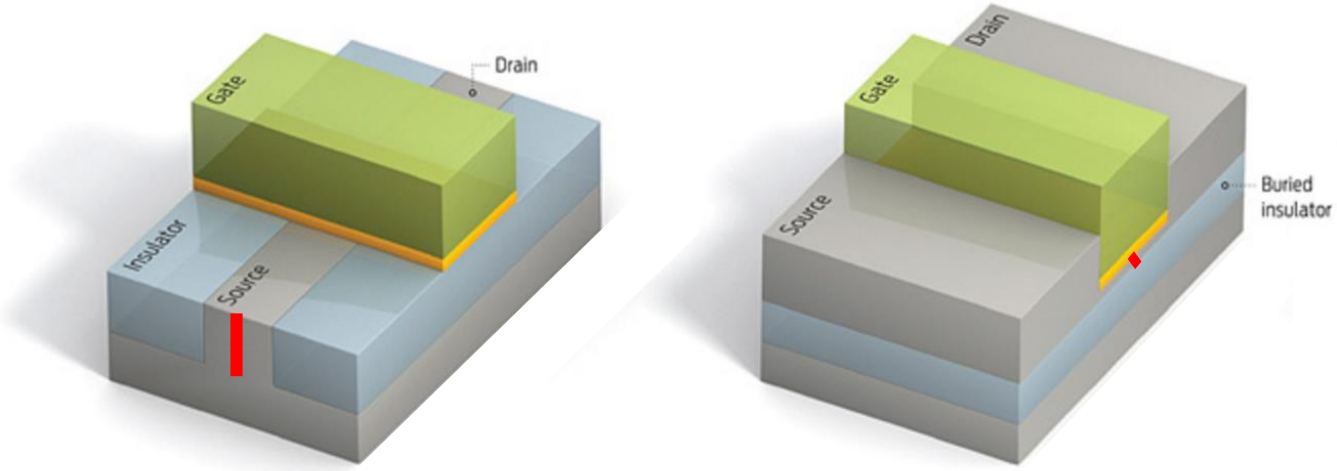


Channel Thickness

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS **SOI**

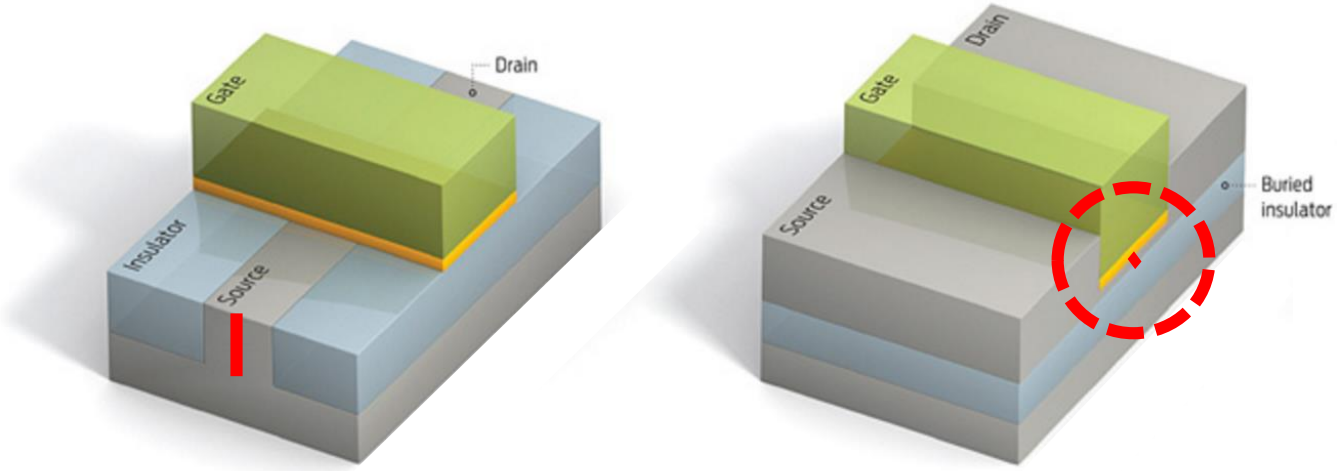


Channel Thickness

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS **SOI**

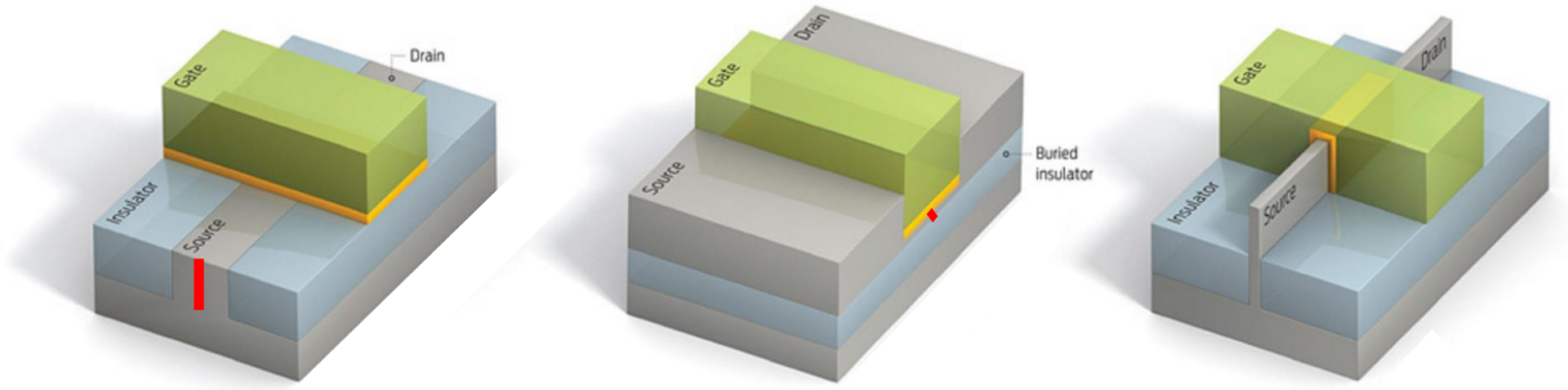


Channel Thickness

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS **SOI** **FinFET**

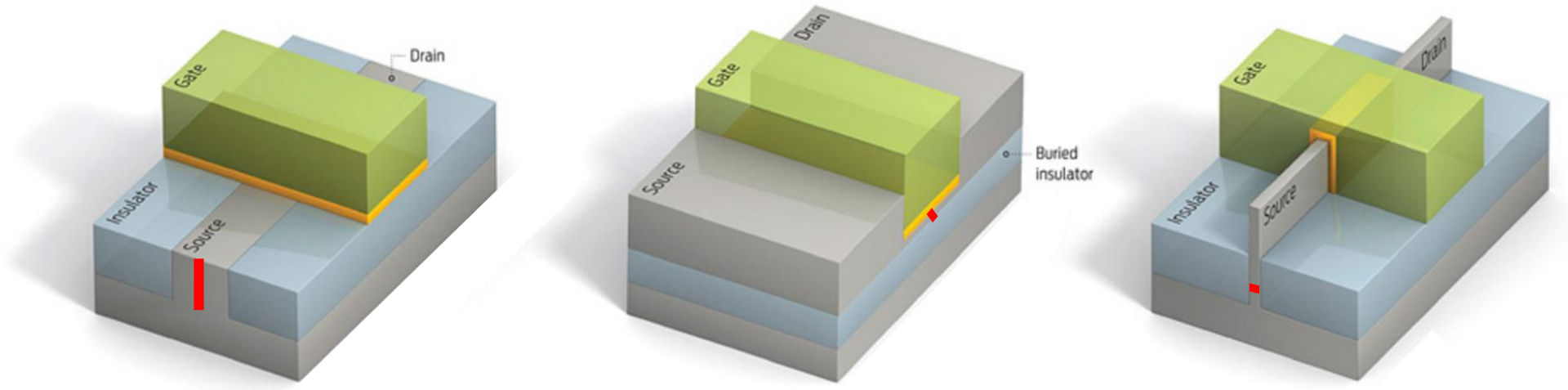


Channel Thickness

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS **SOI** **FinFET**

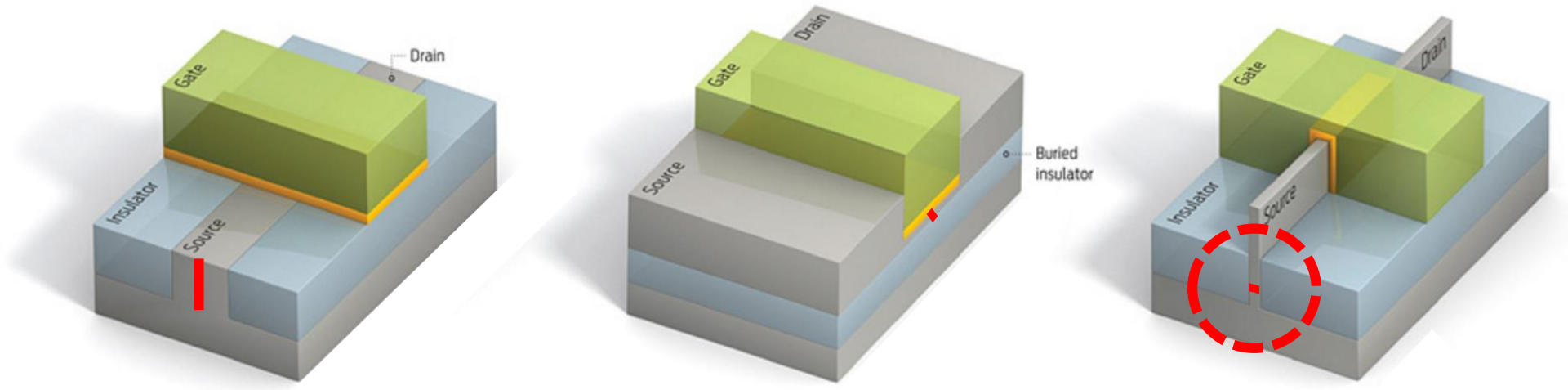


Channel Thickness

[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

Thin Channel (TC) devices

Traditional Planar CMOS **SOI** **FinFET**

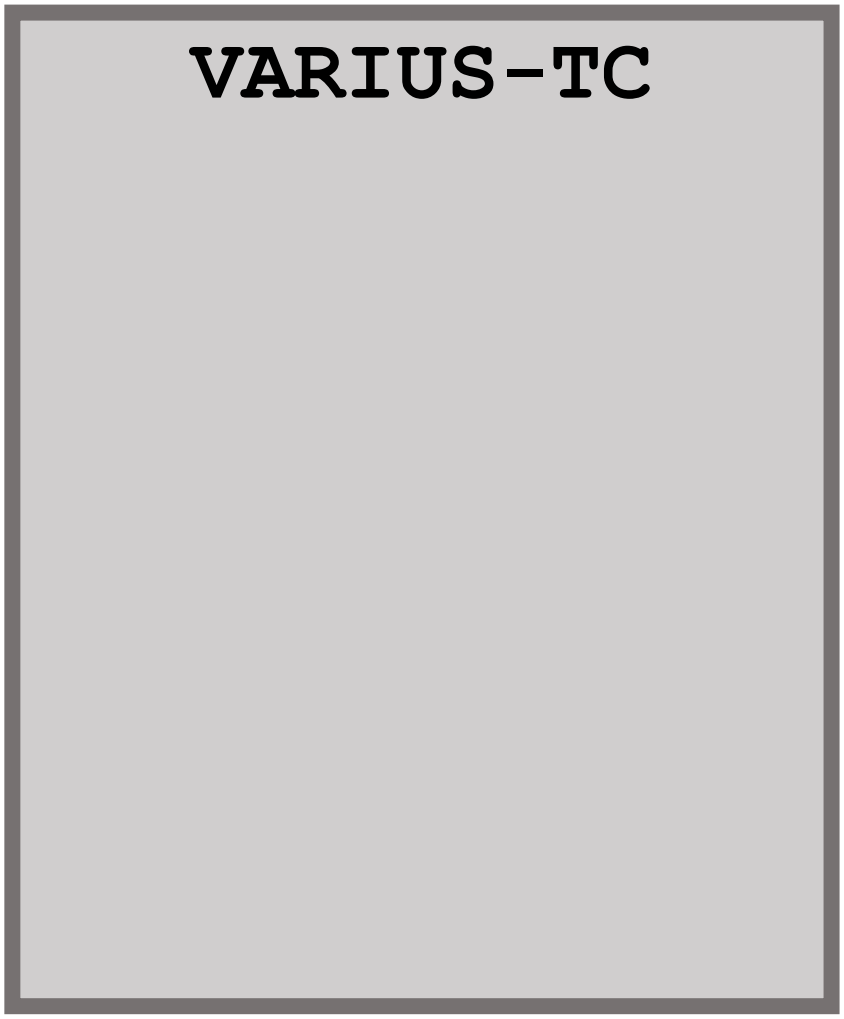


Channel Thickness

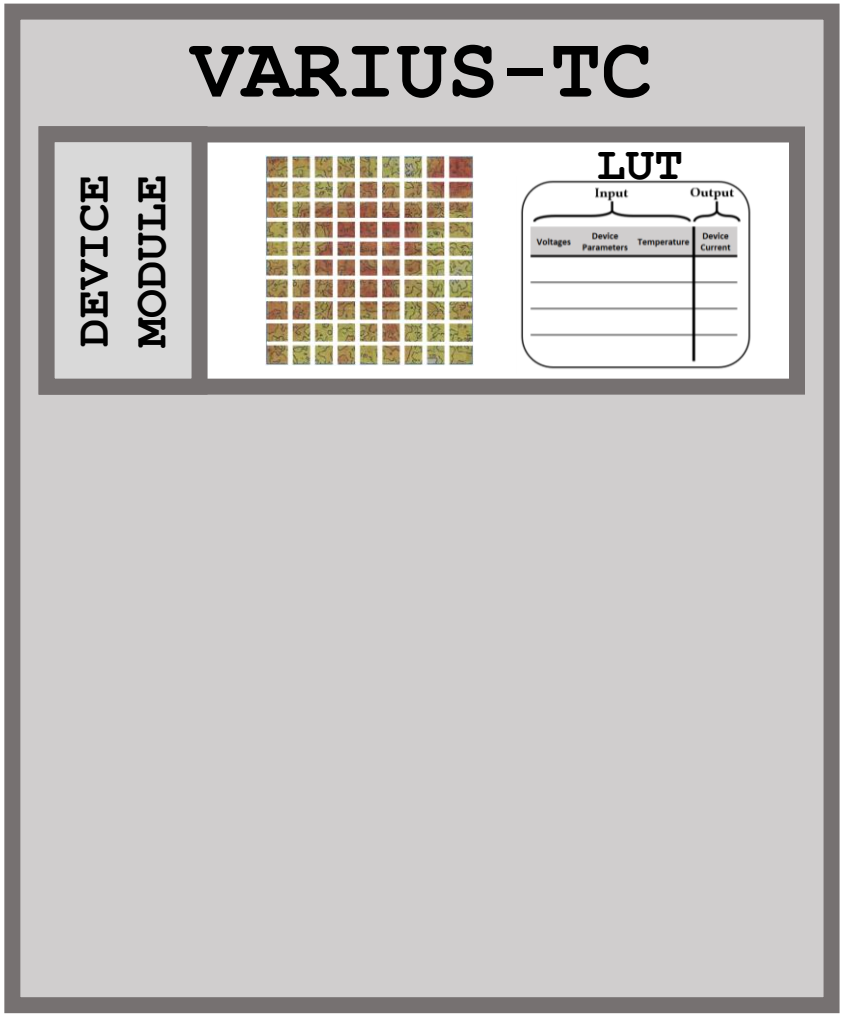
[K. Ahmed, and K. Schuegraf. "Transistor wars." *IEEE Spectrum*48.11 (2011): 50-66.].

VARIUS-TC: Overview

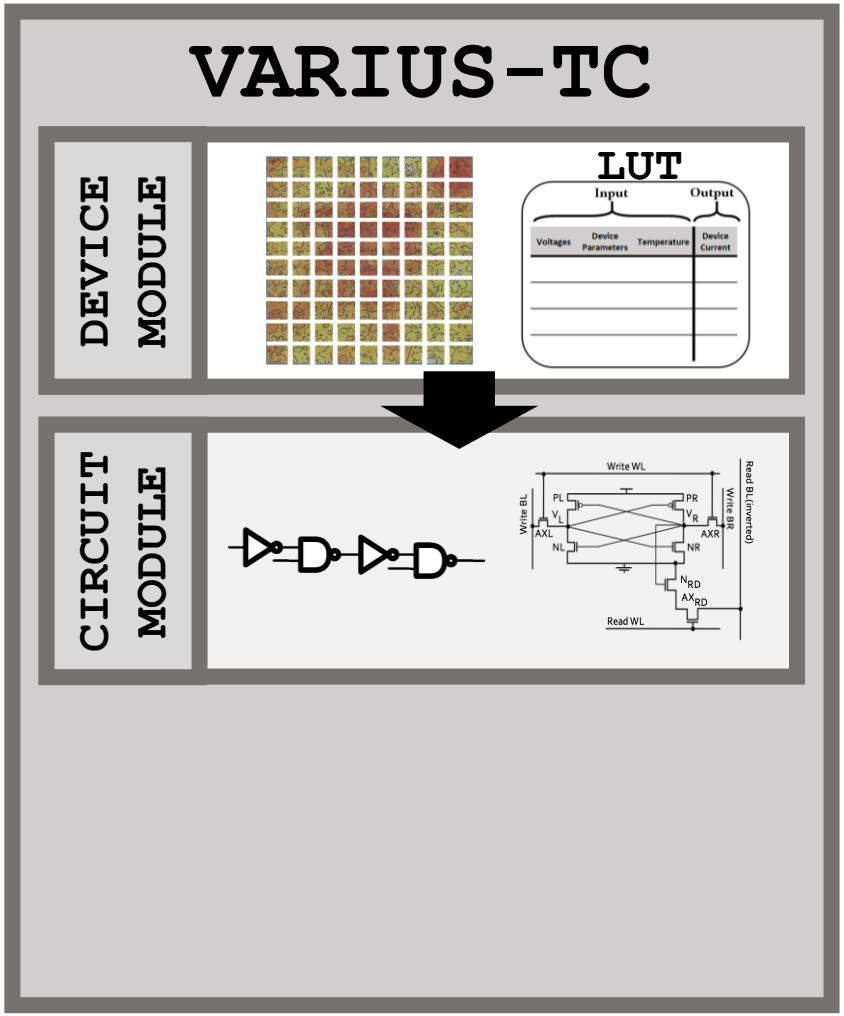
VARIUS-TC: Overview



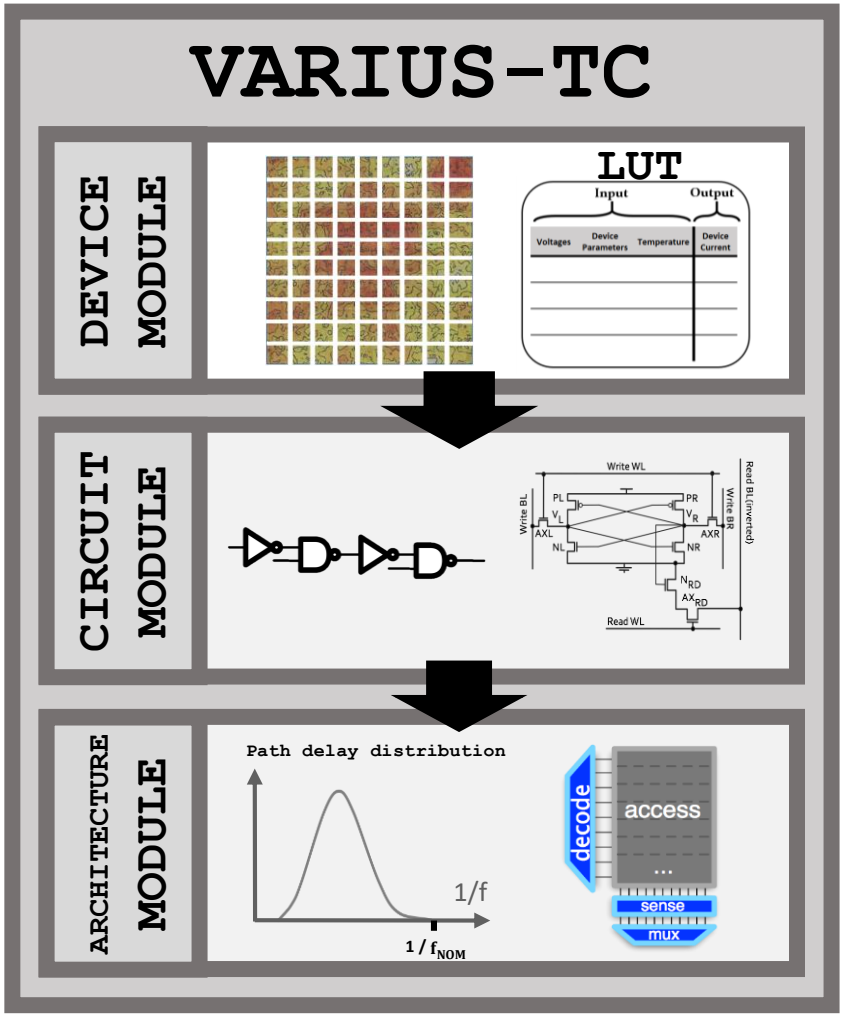
VARIUS-TC: Overview



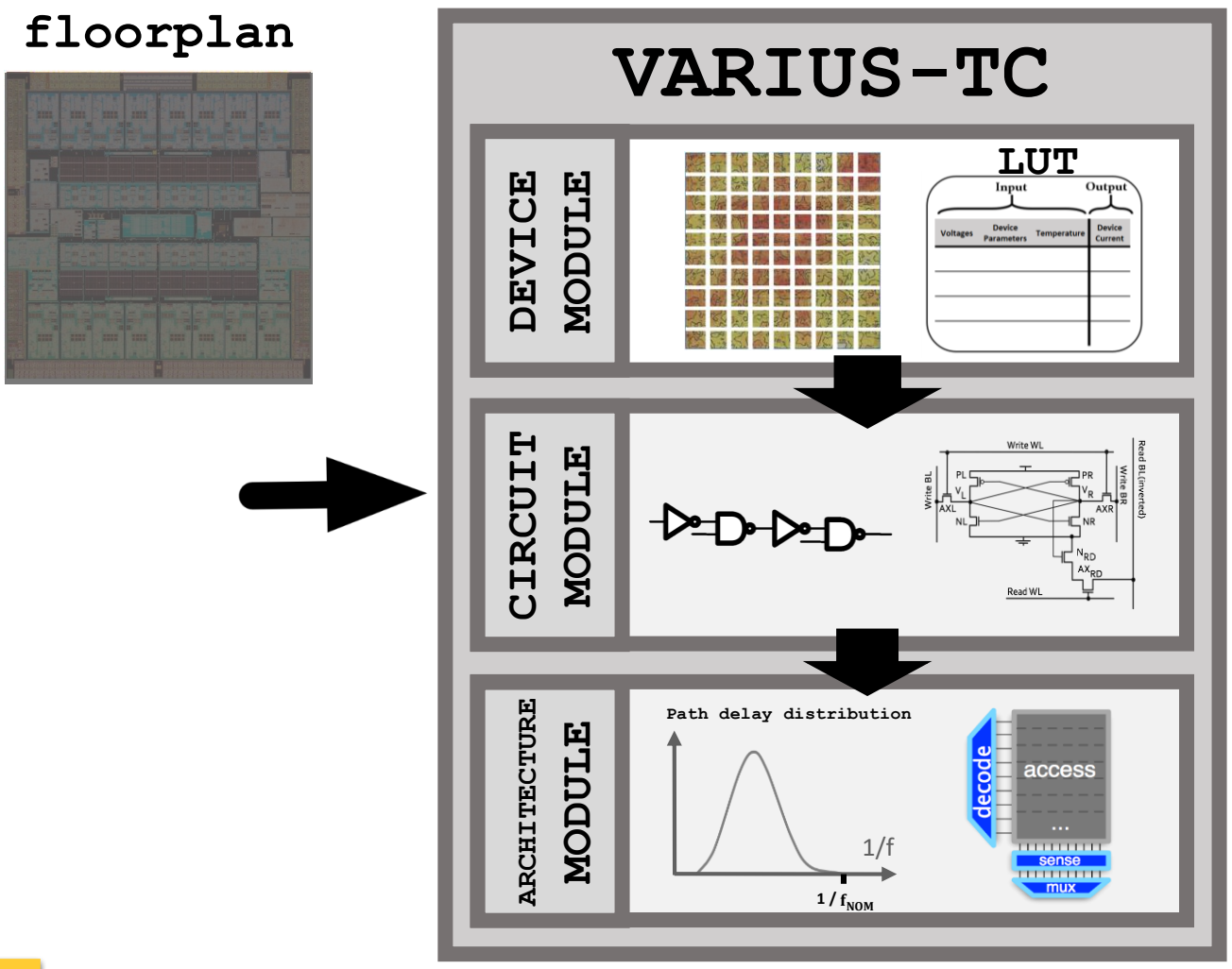
VARIUS-TC: Overview



VARIUS-TC: Overview

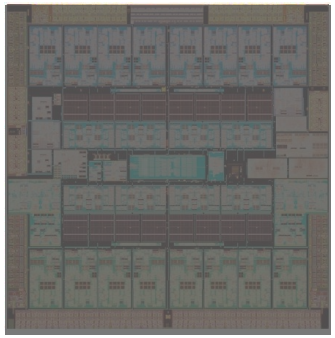


VARIUS-TC: Overview

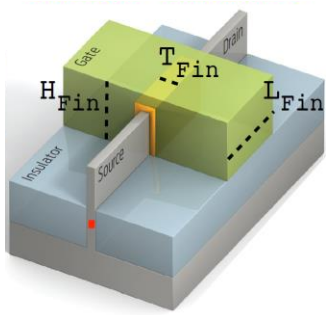


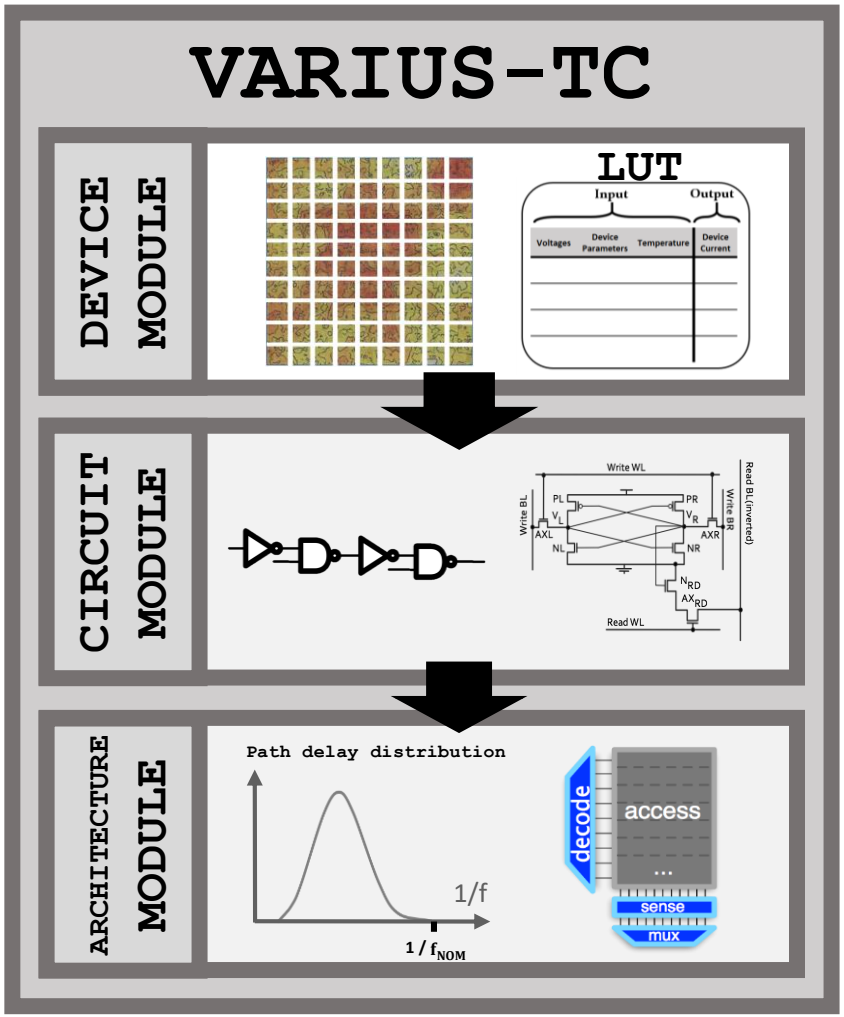
VARIUS-TC: Overview

floorplan



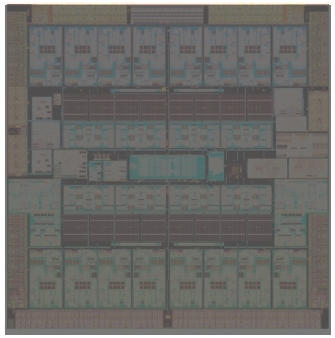
device parameters



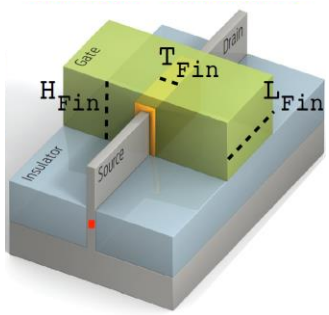


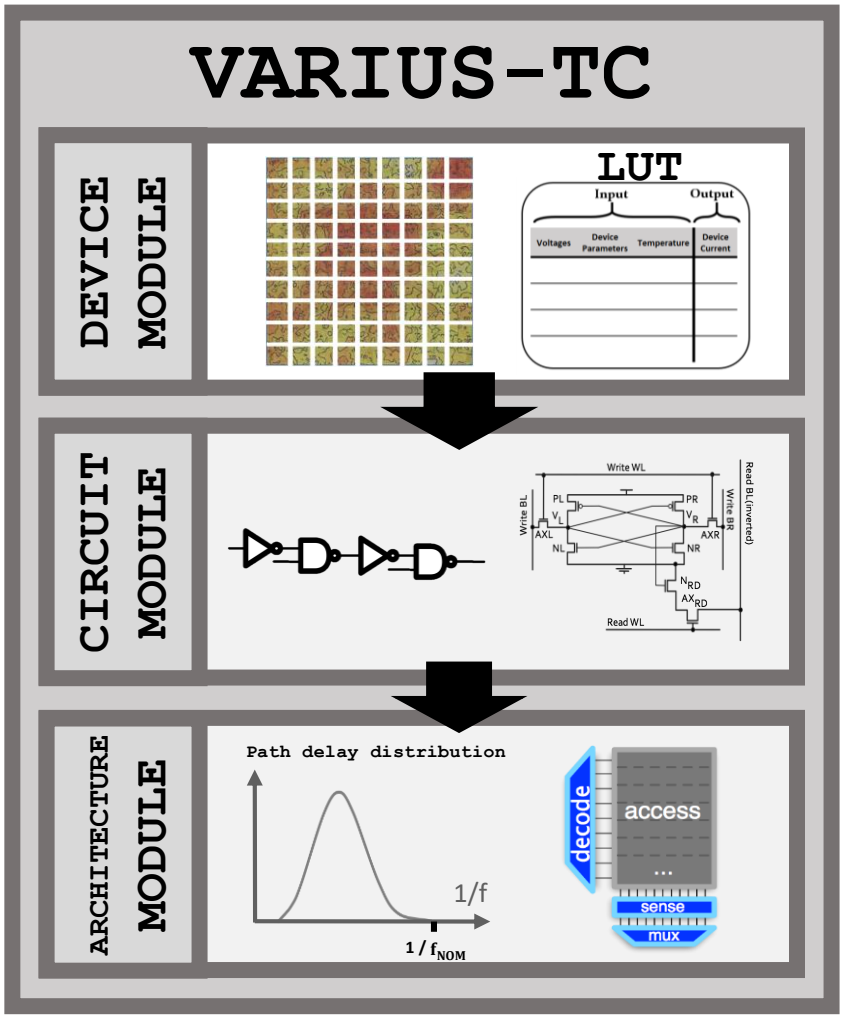
VARIUS-TC: Overview

floorplan

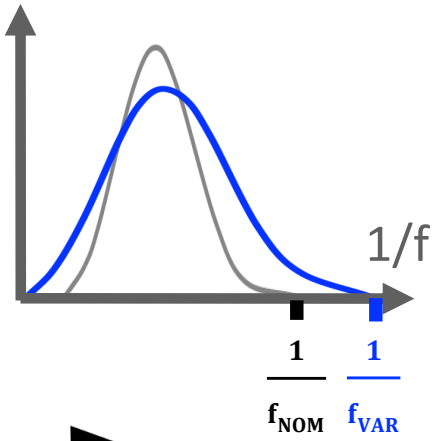


device parameters



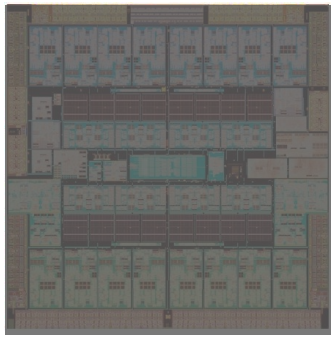


Path delay distribution

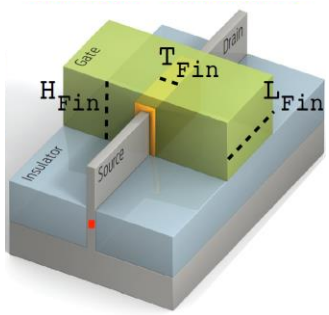


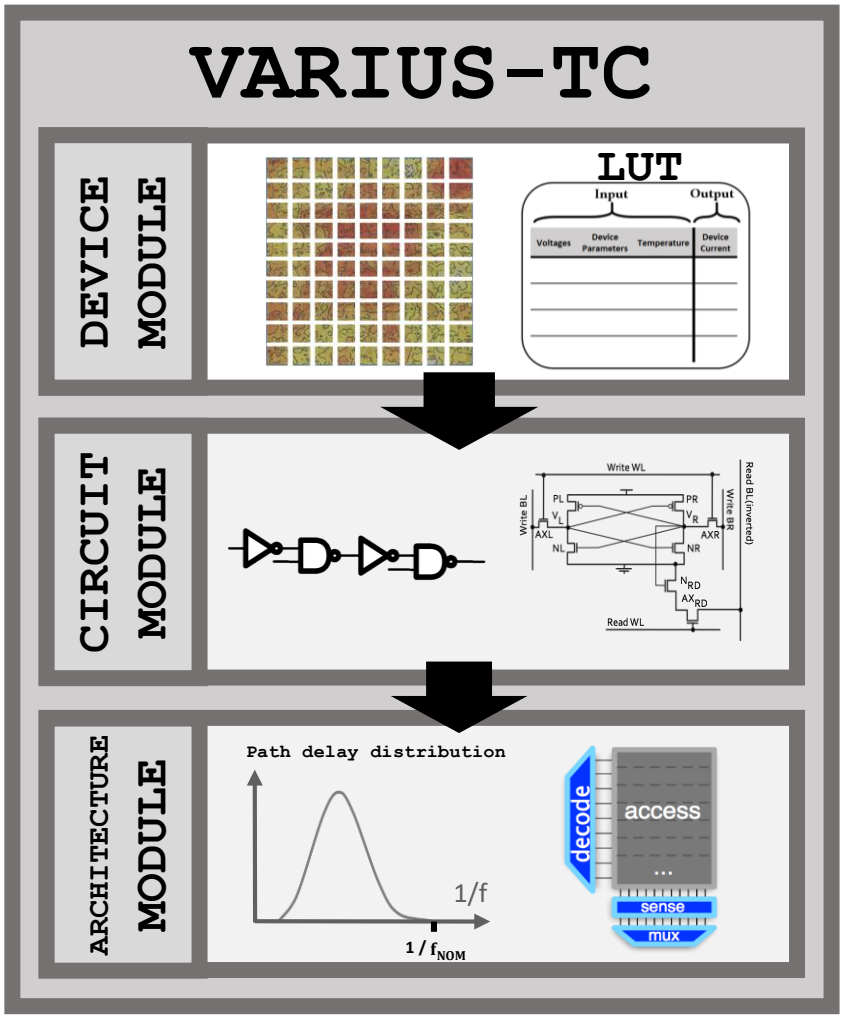
VARIUS-TC: Overview

floorplan

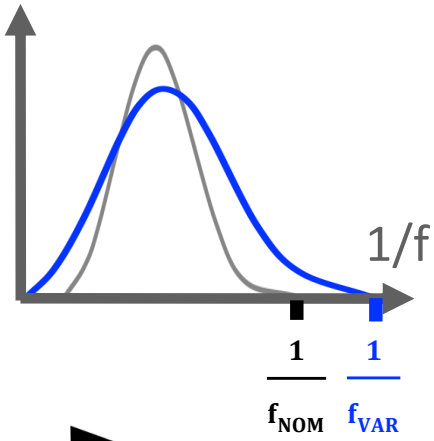


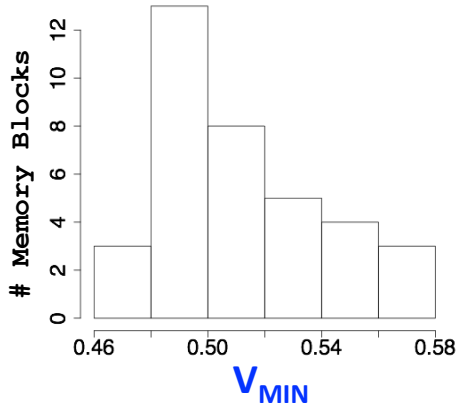
device parameters





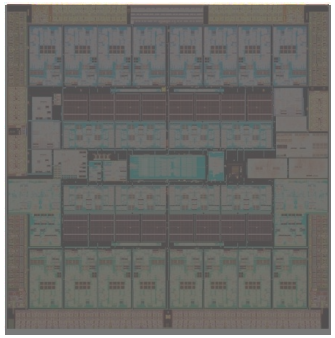
Path delay distribution



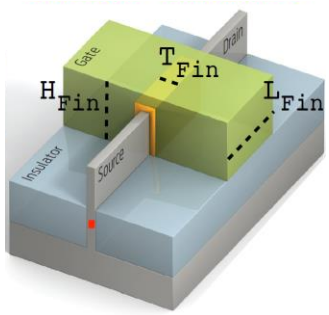


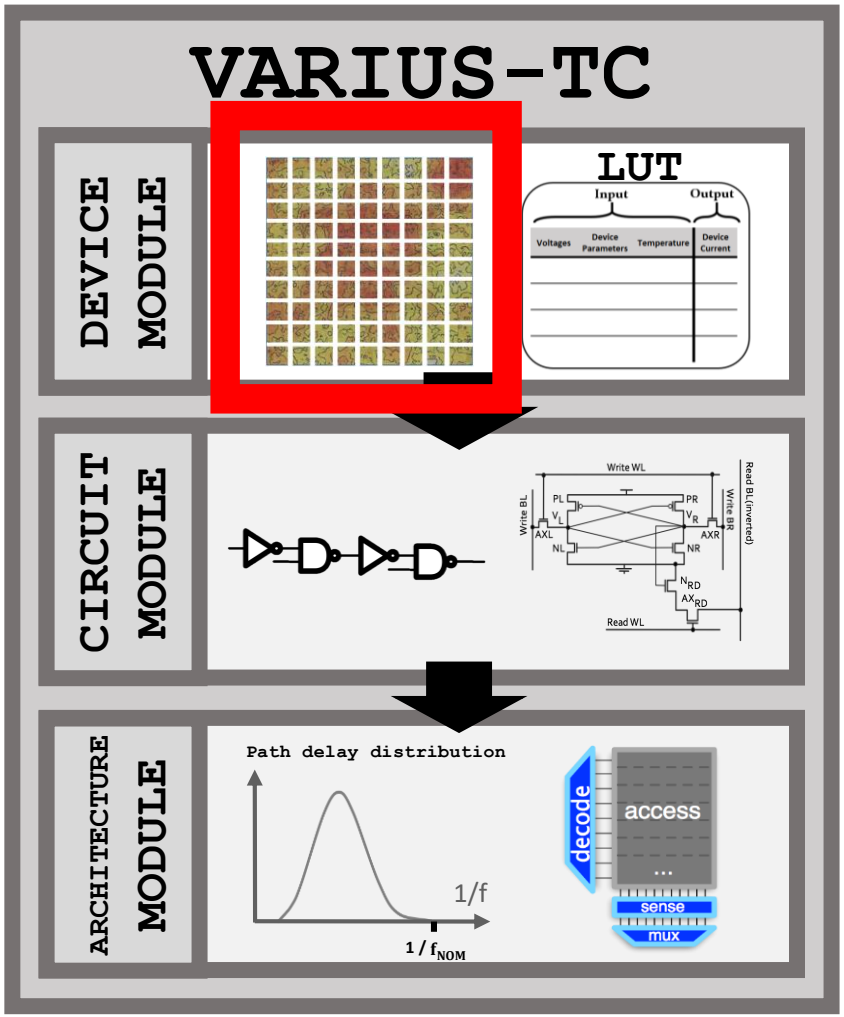
VARIUS-TC: Overview

floorplan

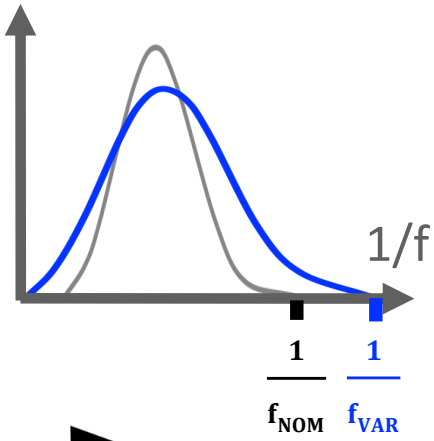


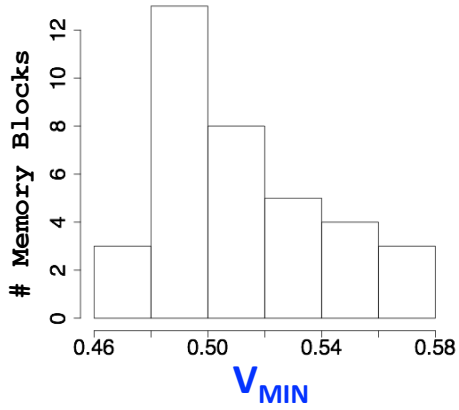
device parameters





Path delay distribution





Hierarchical Variation Modeling

Systematic variation

Hierarchical Variation Modeling

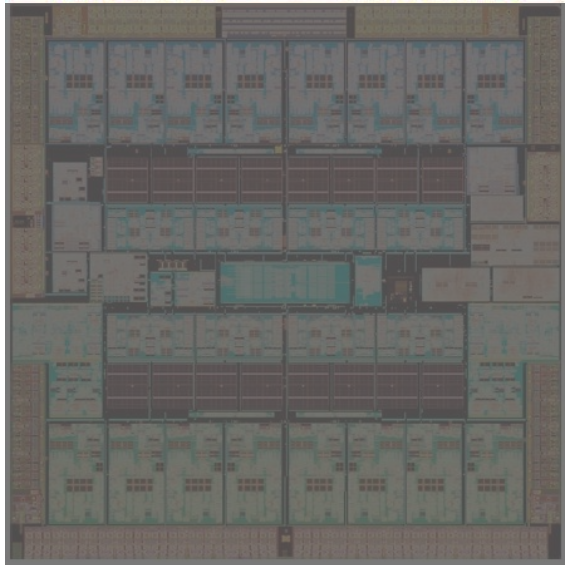
Systematic variation

- Spatial correlation

Hierarchical Variation Modeling

Systematic variation

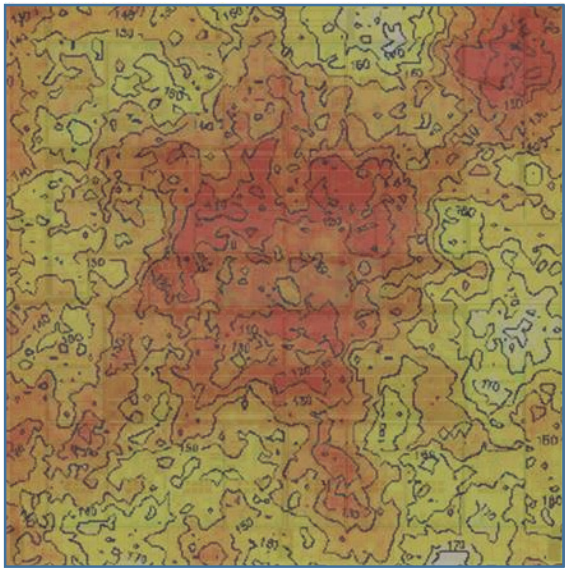
- Spatial correlation



Hierarchical Variation Modeling

Systematic variation

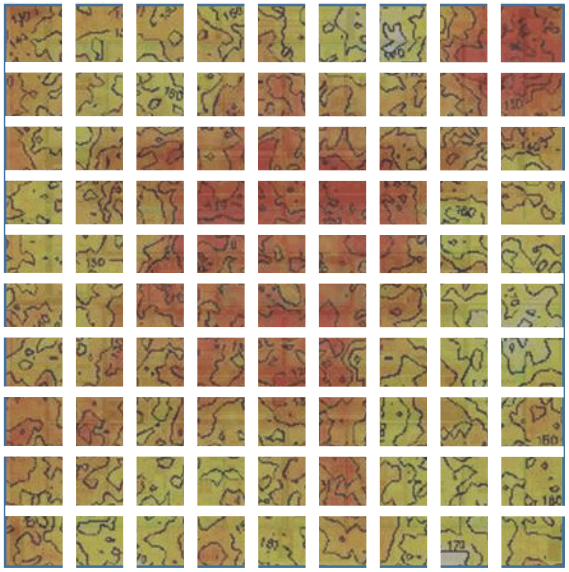
- Spatial correlation



Hierarchical Variation Modeling

Systematic variation

- Spatial correlation
- Grid granularity



Hierarchical Variation Modeling

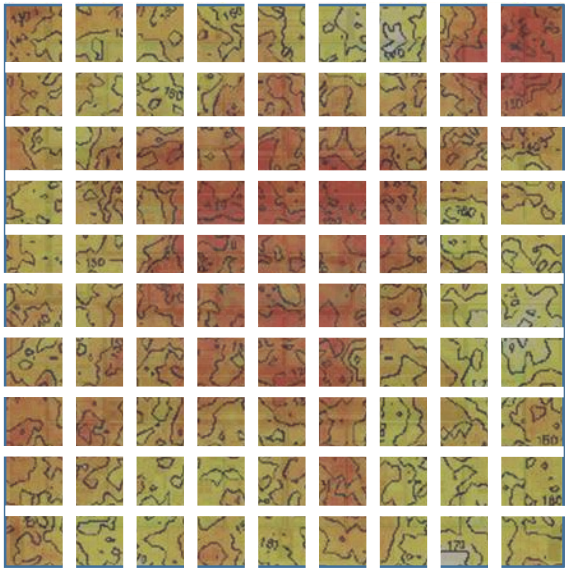
Systematic variation

- Spatial correlation
- Grid granularity

+

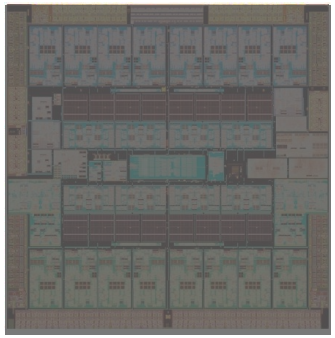
Random variation

- Independent
- Device granularity

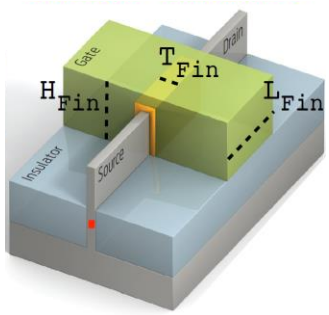


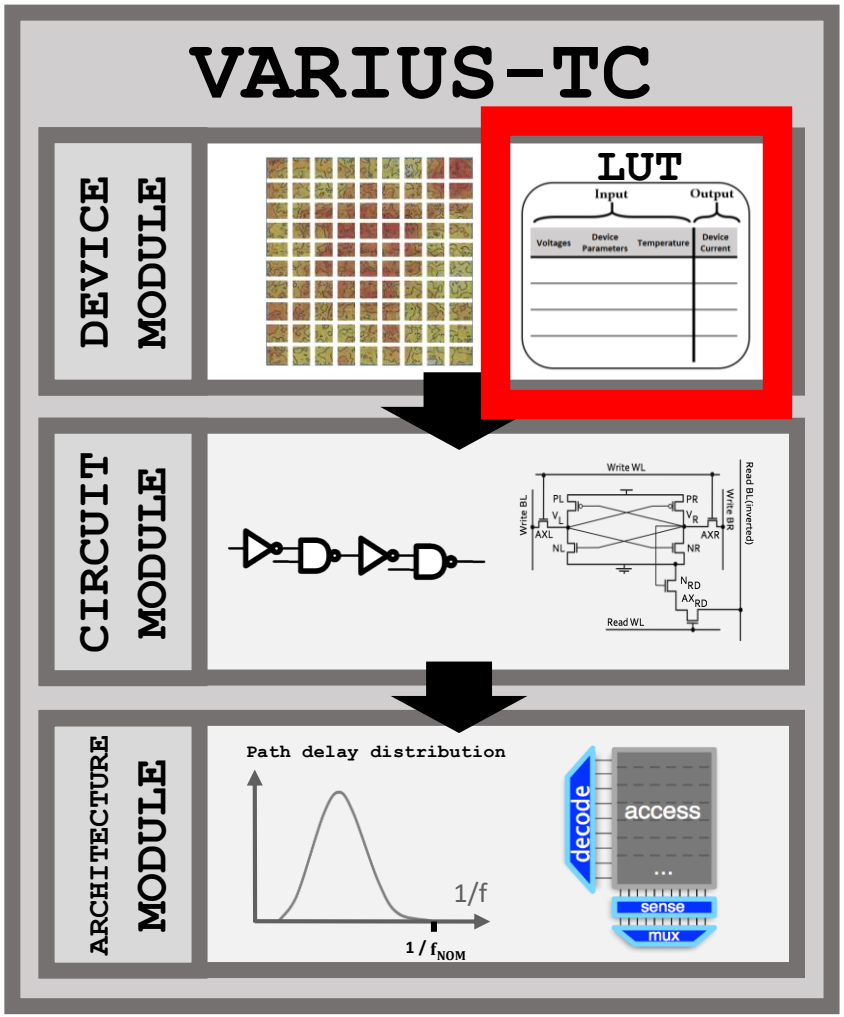
VARIUS-TC: Overview

floorplan

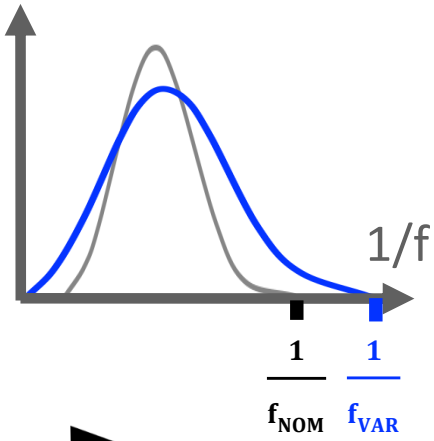


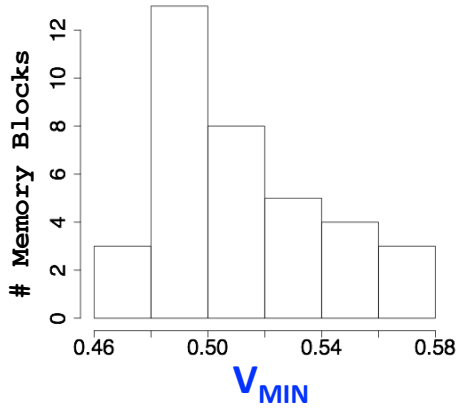
device parameters





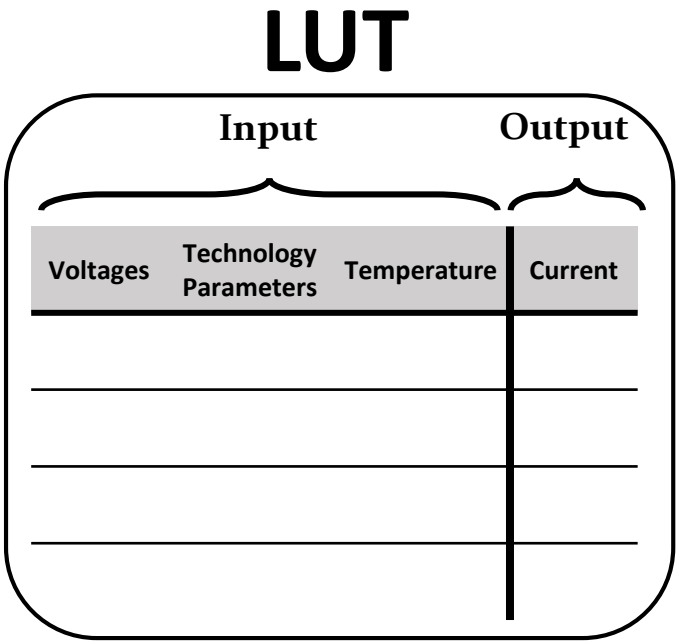
Path delay distribution



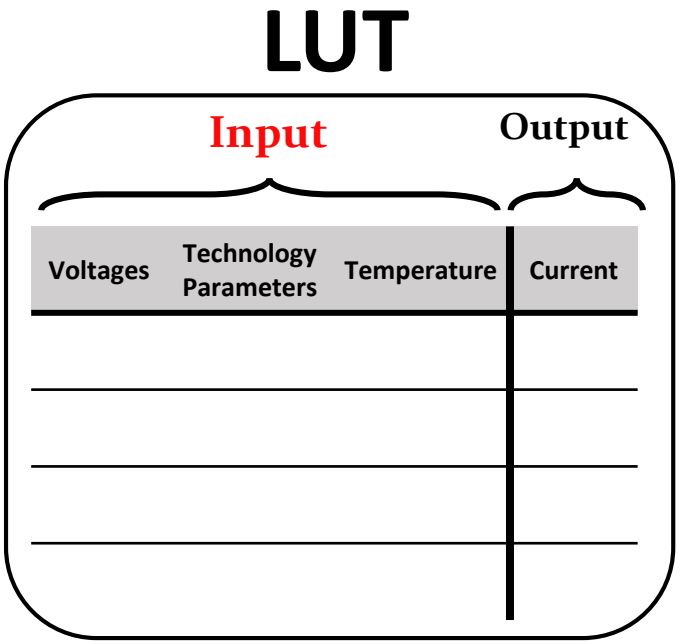


Look-Up Table

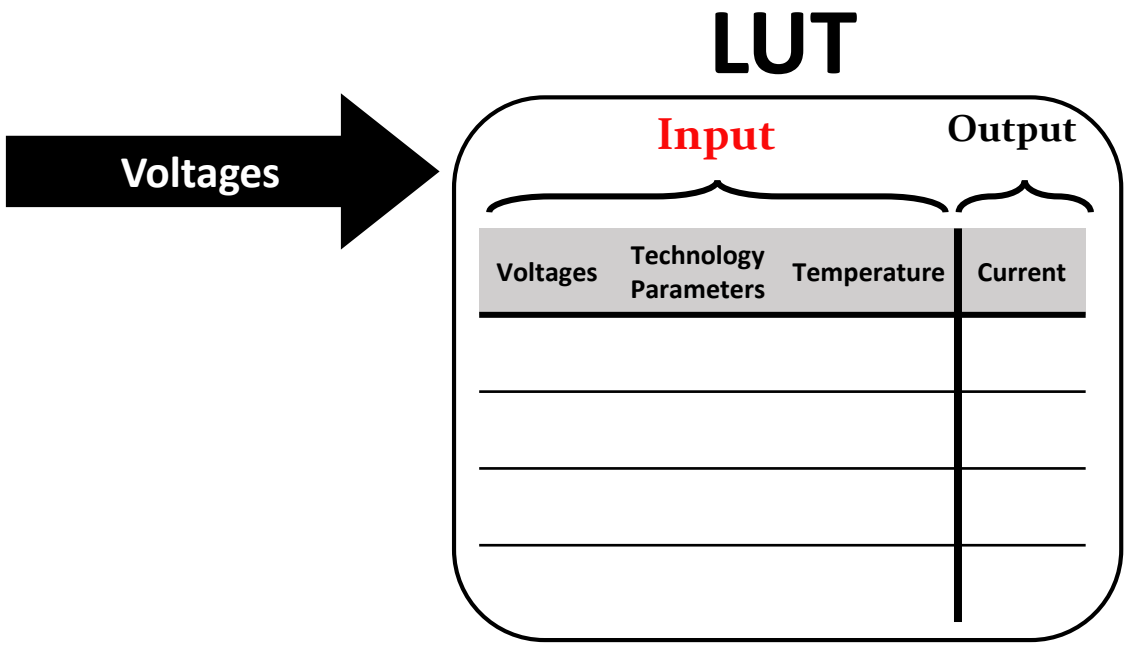
Look-Up Table



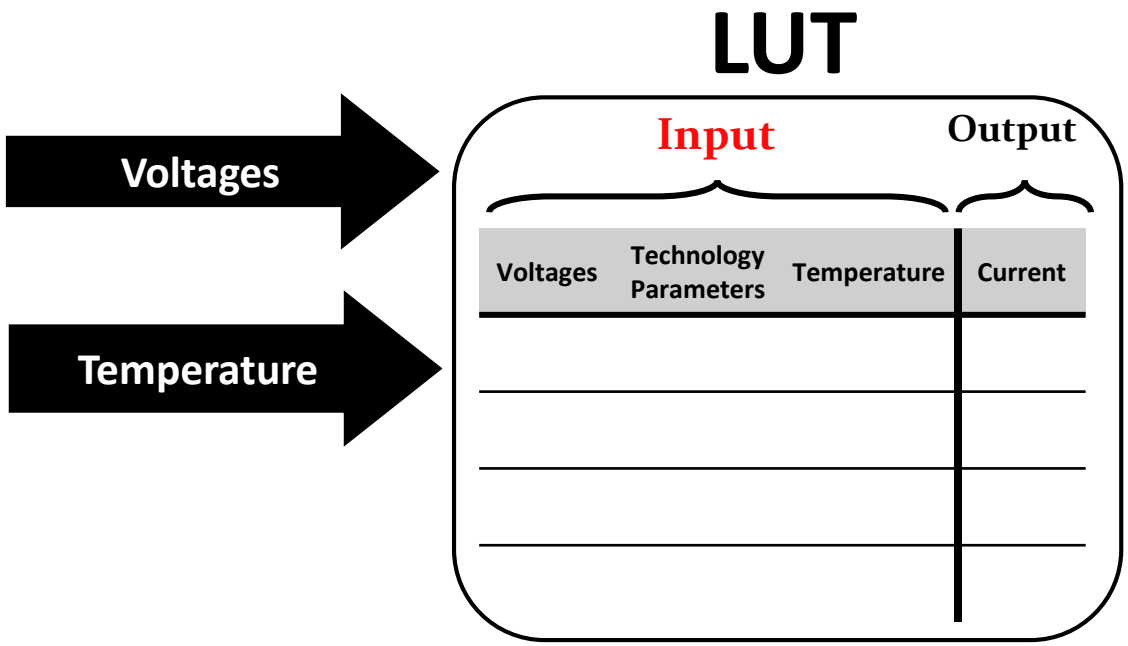
Look-Up Table



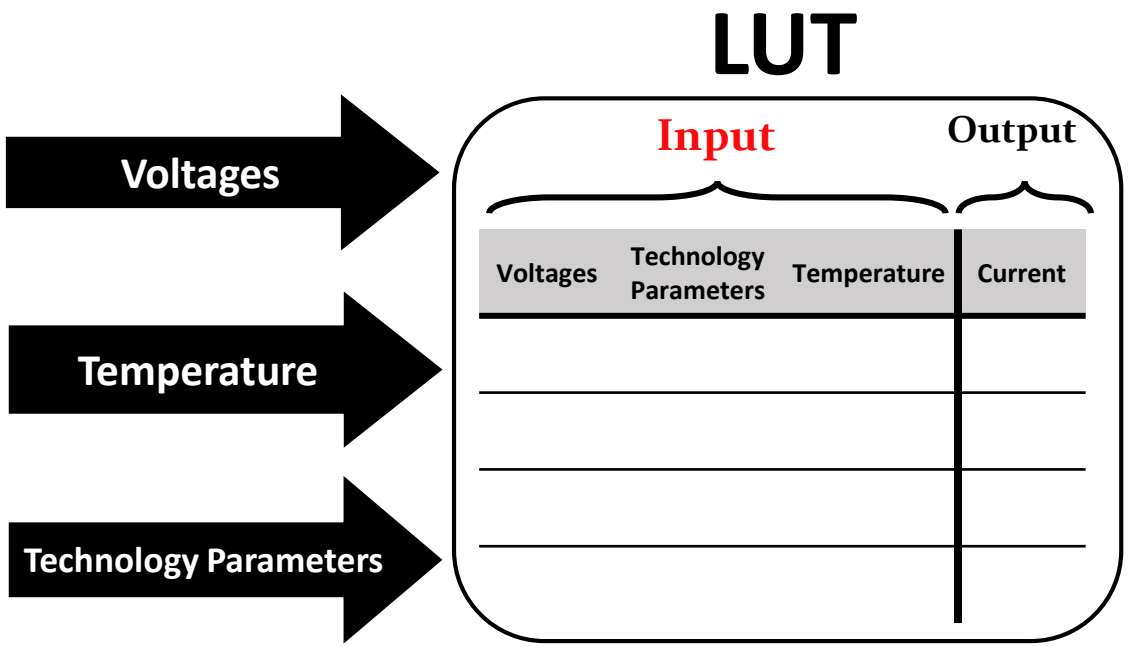
Look-Up Table



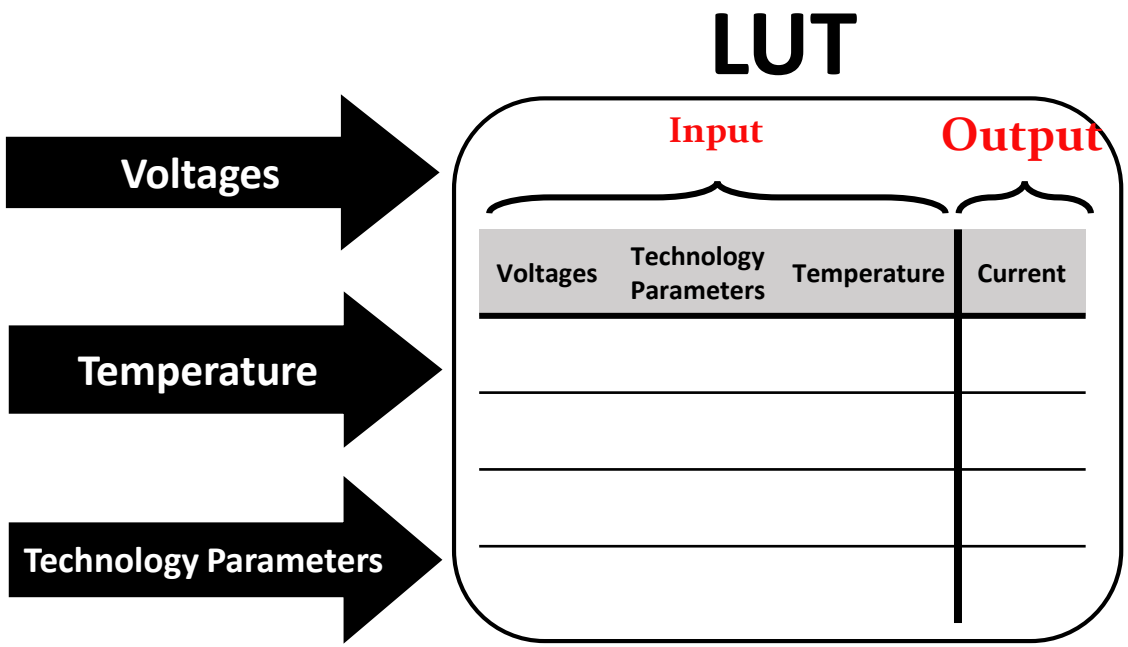
Look-Up Table



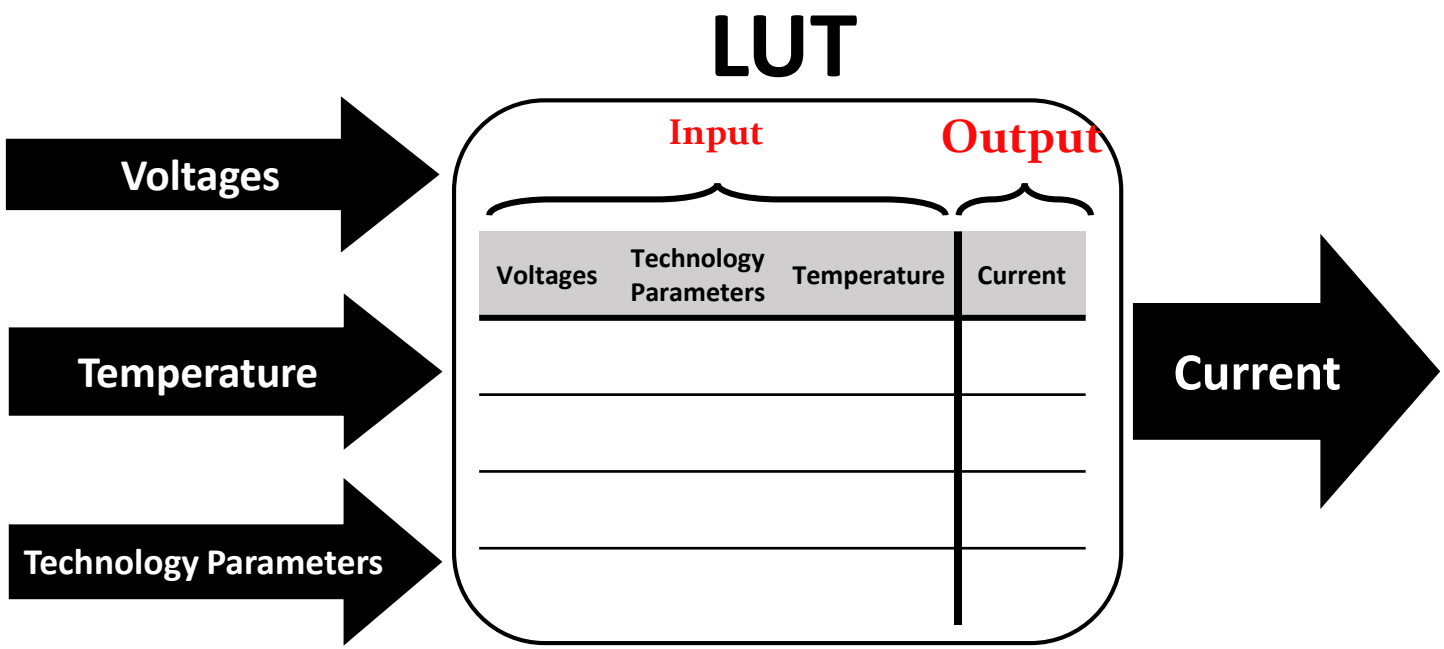
Look-Up Table



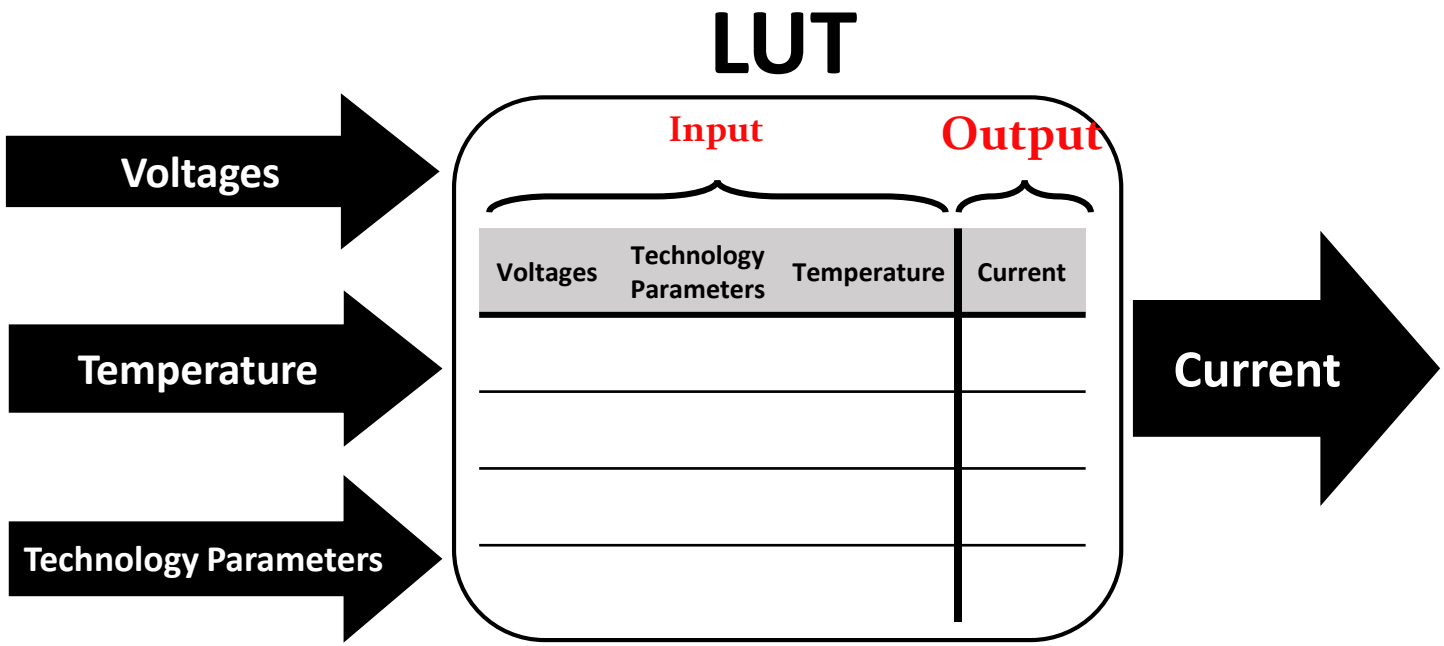
Look-Up Table



Look-Up Table

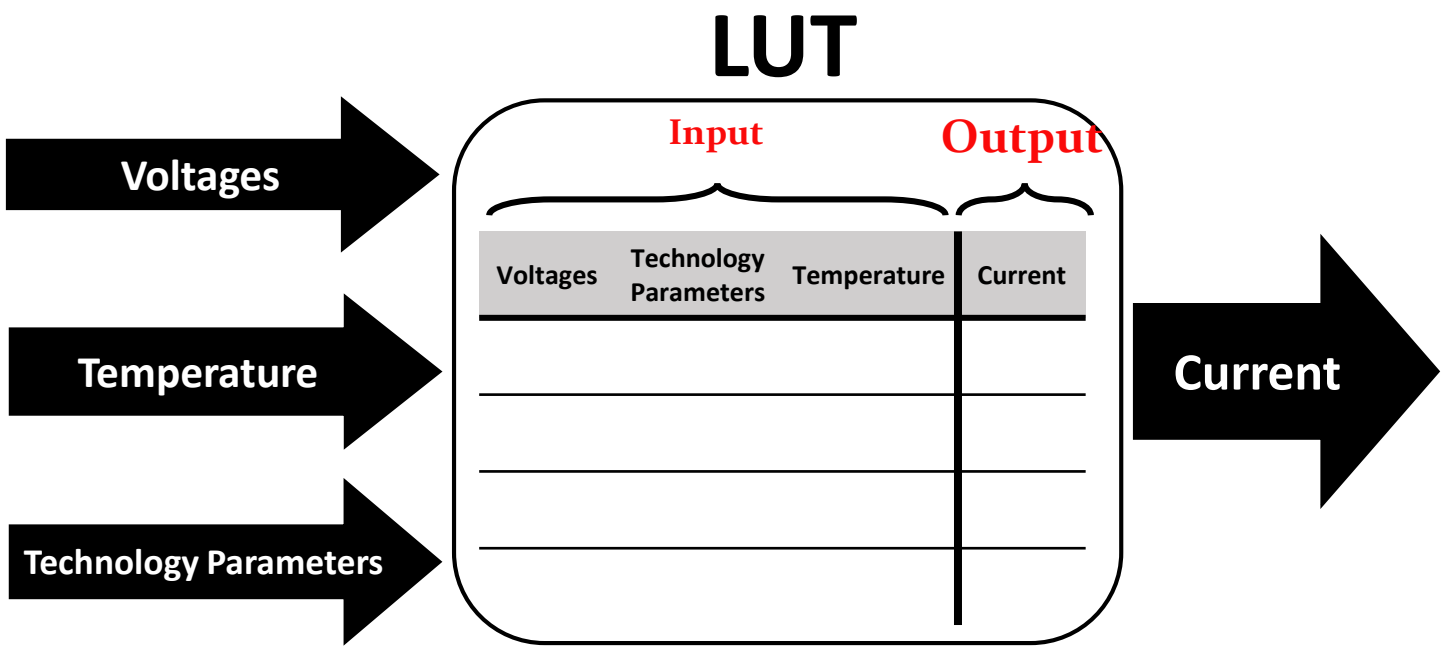


Look-Up Table



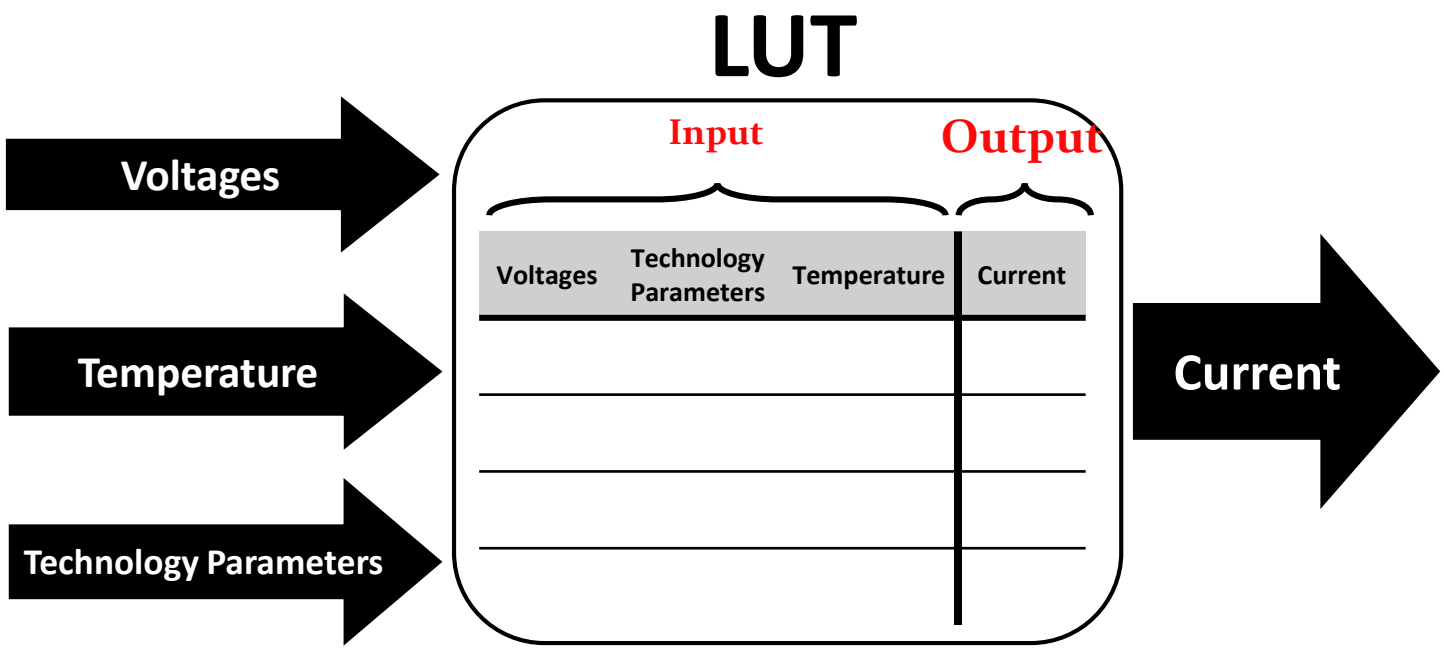
- LUT vs. closed-form formula:

Look-Up Table



- LUT vs. closed-form formula:
  - Modularity eases experimentation with different designs (e.g., SOI variants)

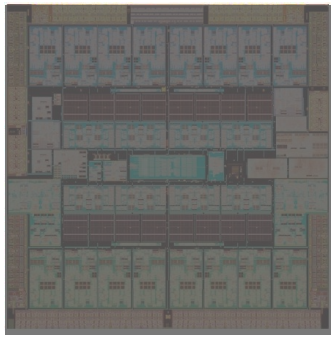
Look-Up Table



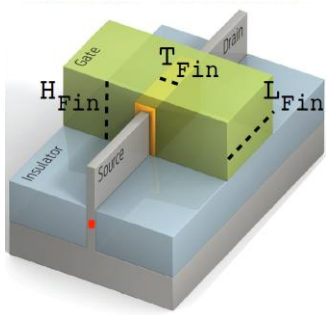
- LUT vs. closed-form formula:
  - Modularity eases experimentation with different designs (e.g., SOI variants)
  - Robust closed-form formula may not always be available for emerging switches

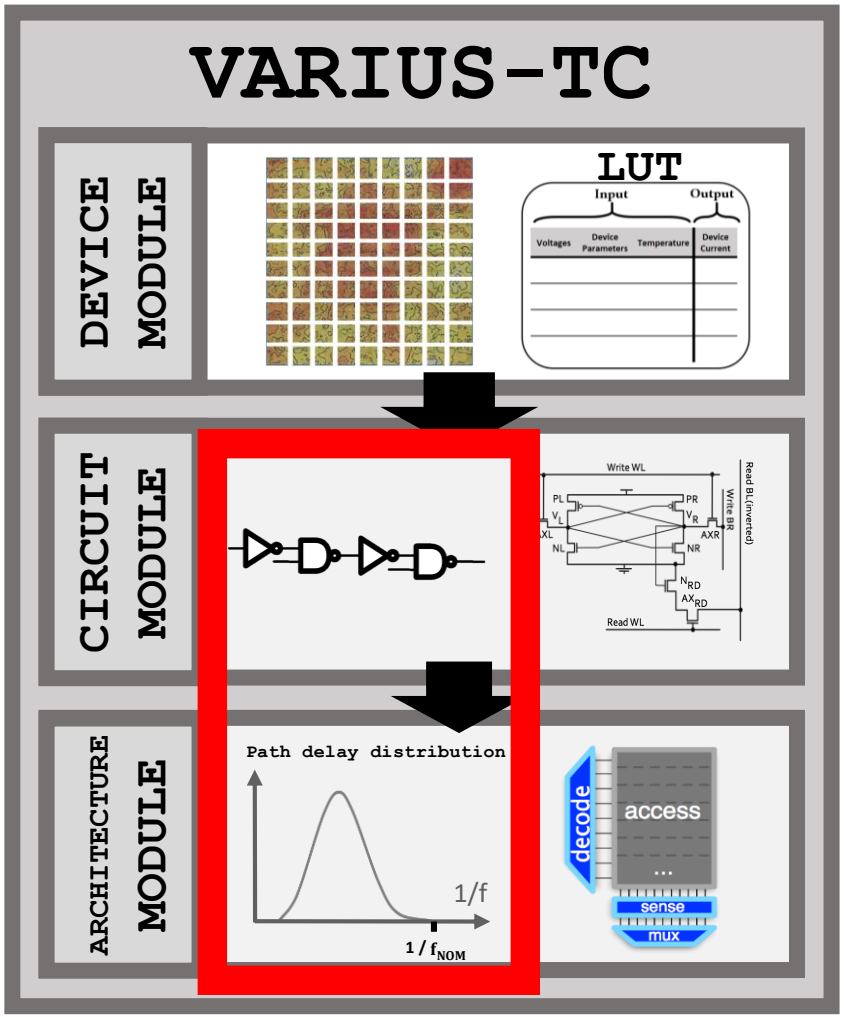
VARIUS-TC: Overview

floorplan

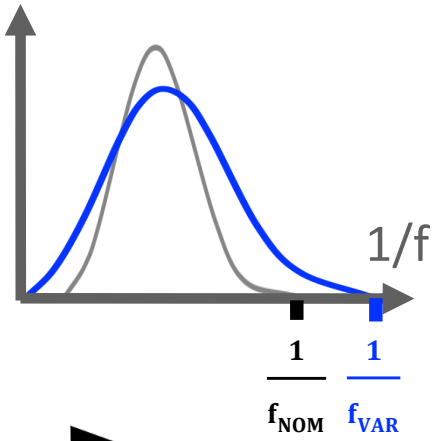


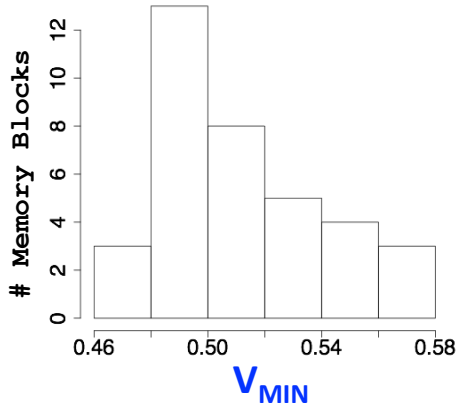
device parameters





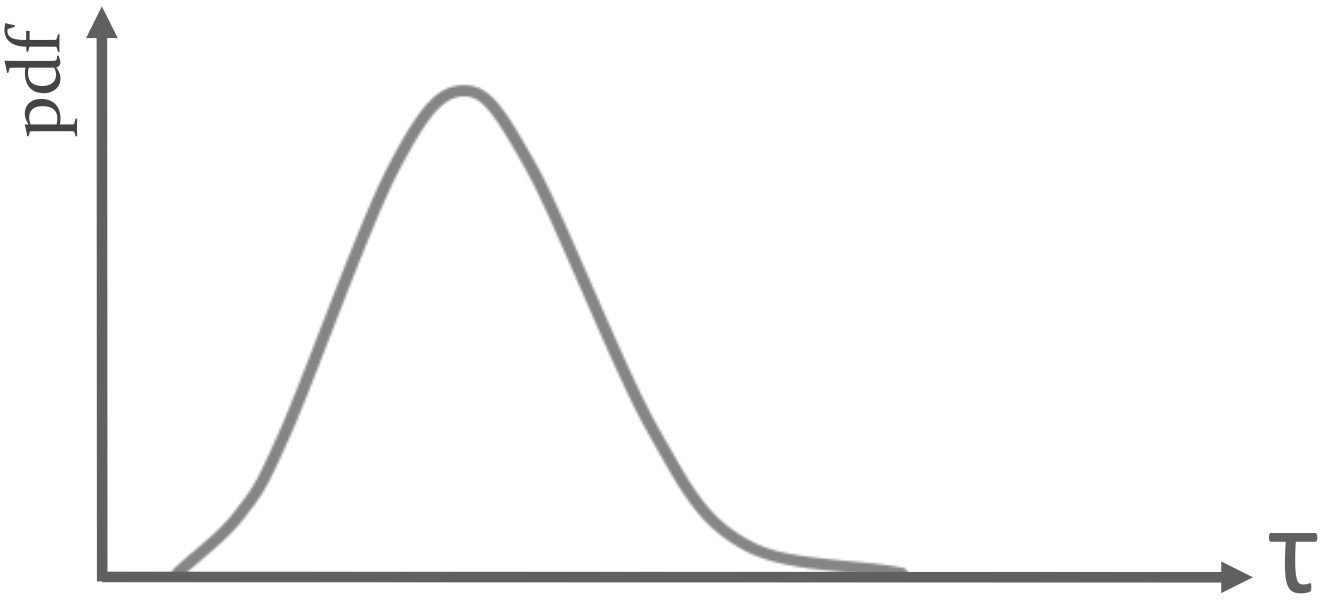
Path delay distribution



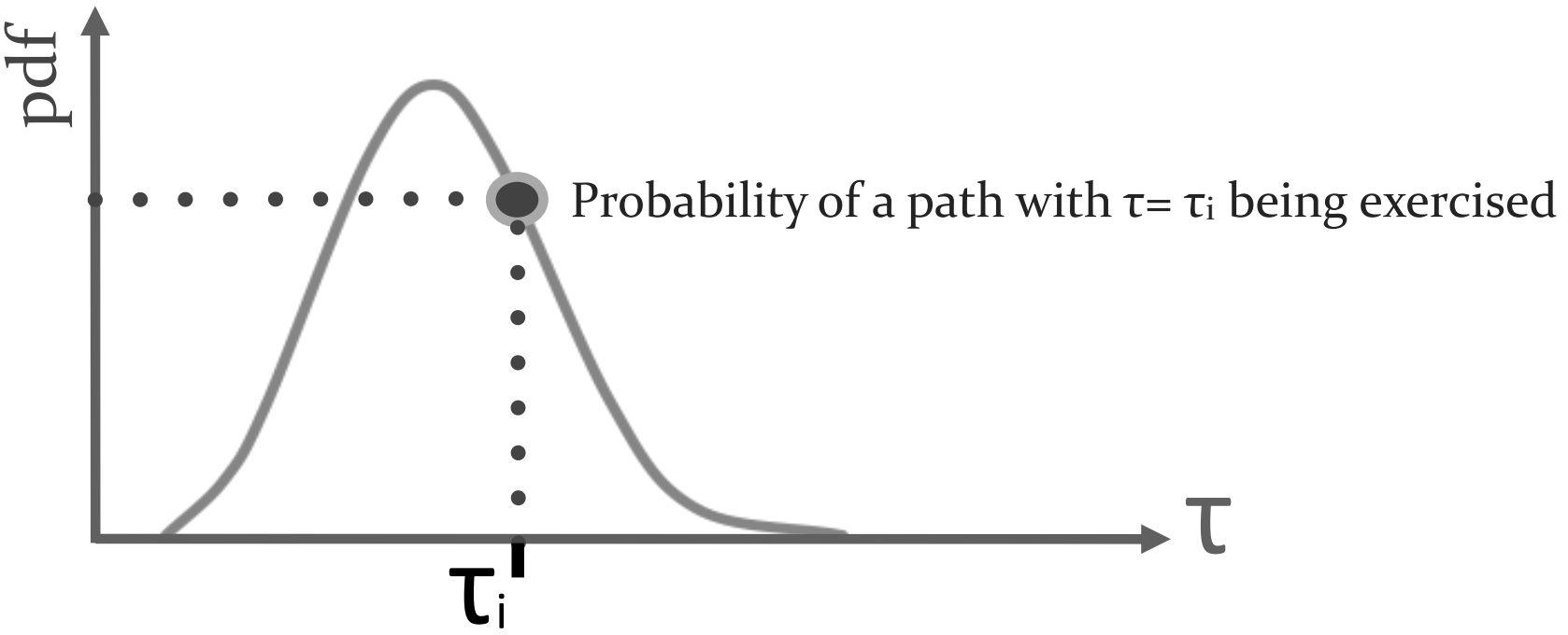


Logic Timing Model

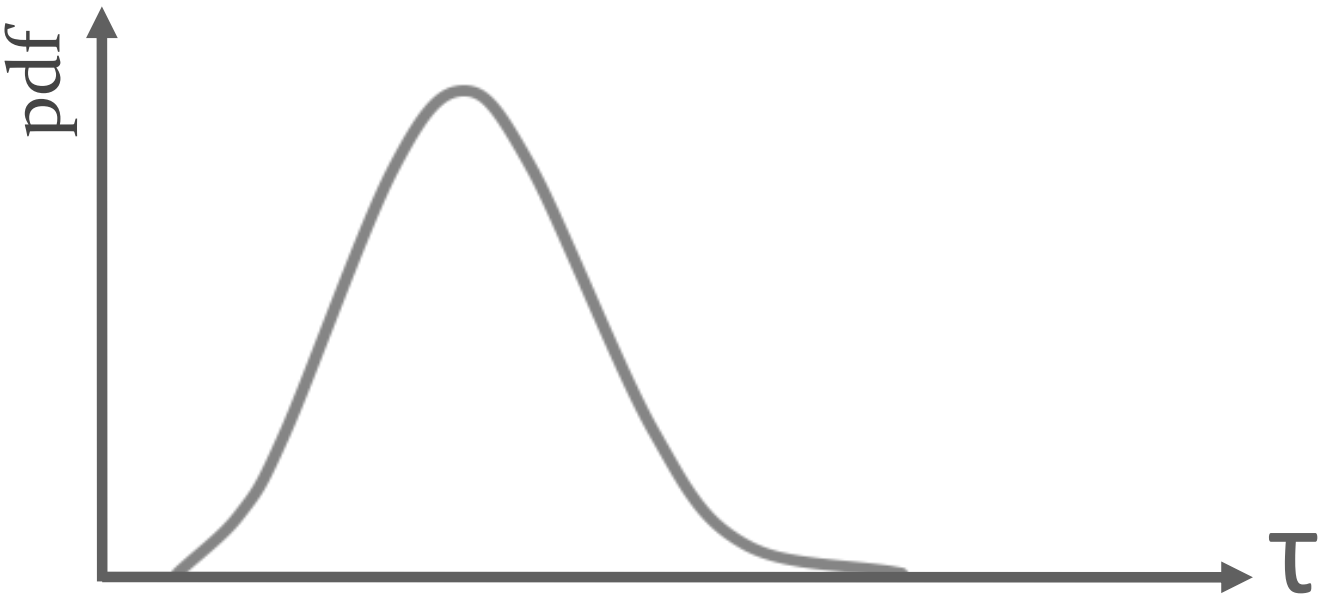
Logic Timing Model



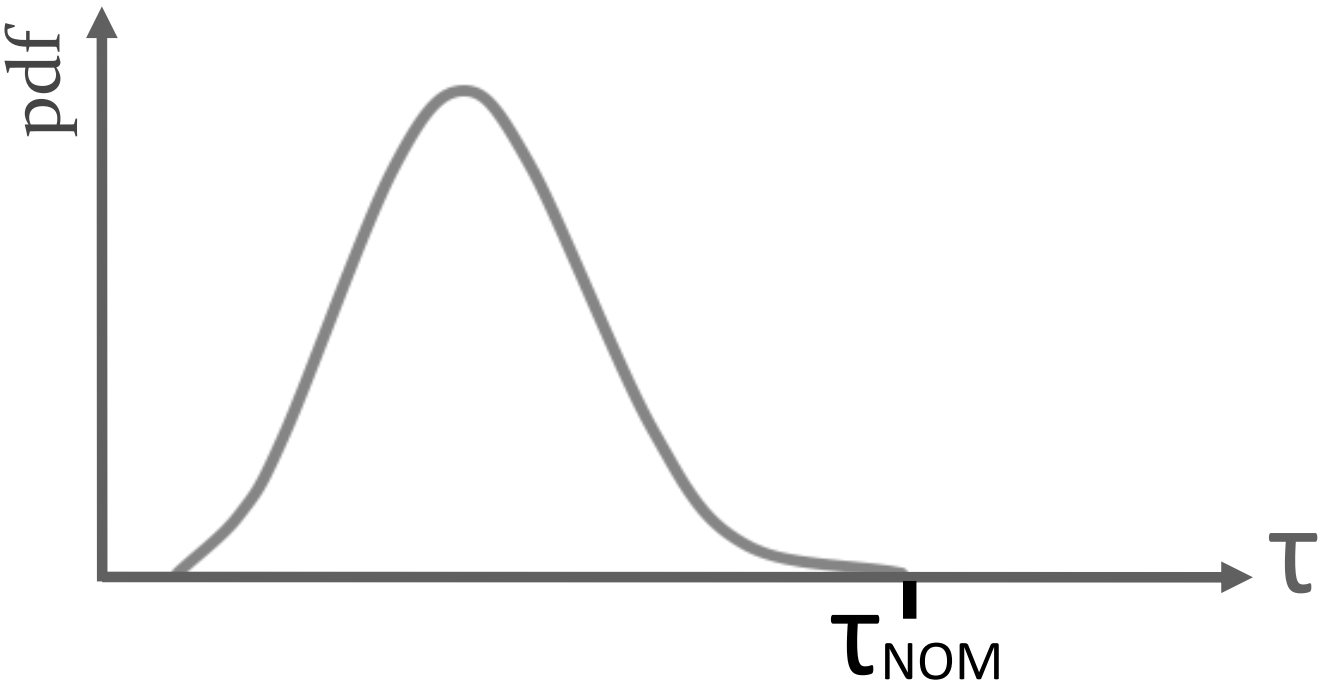
Logic Timing Model



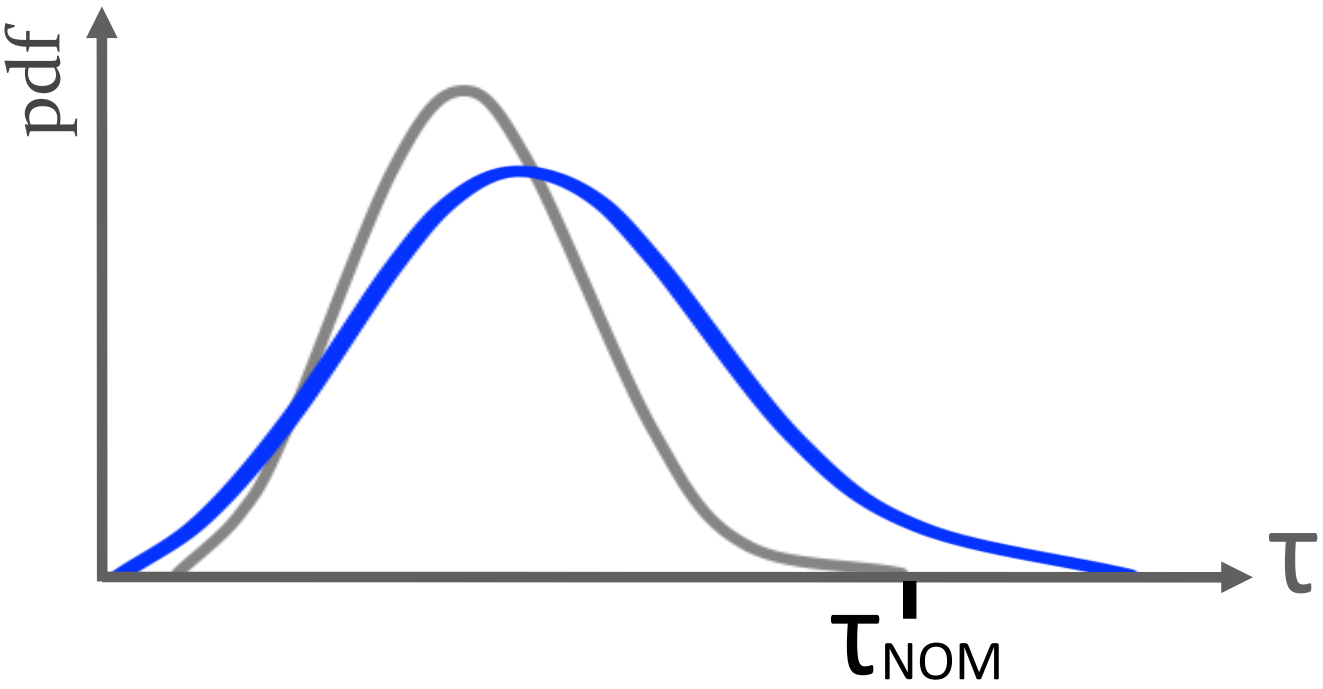
Logic Timing Model



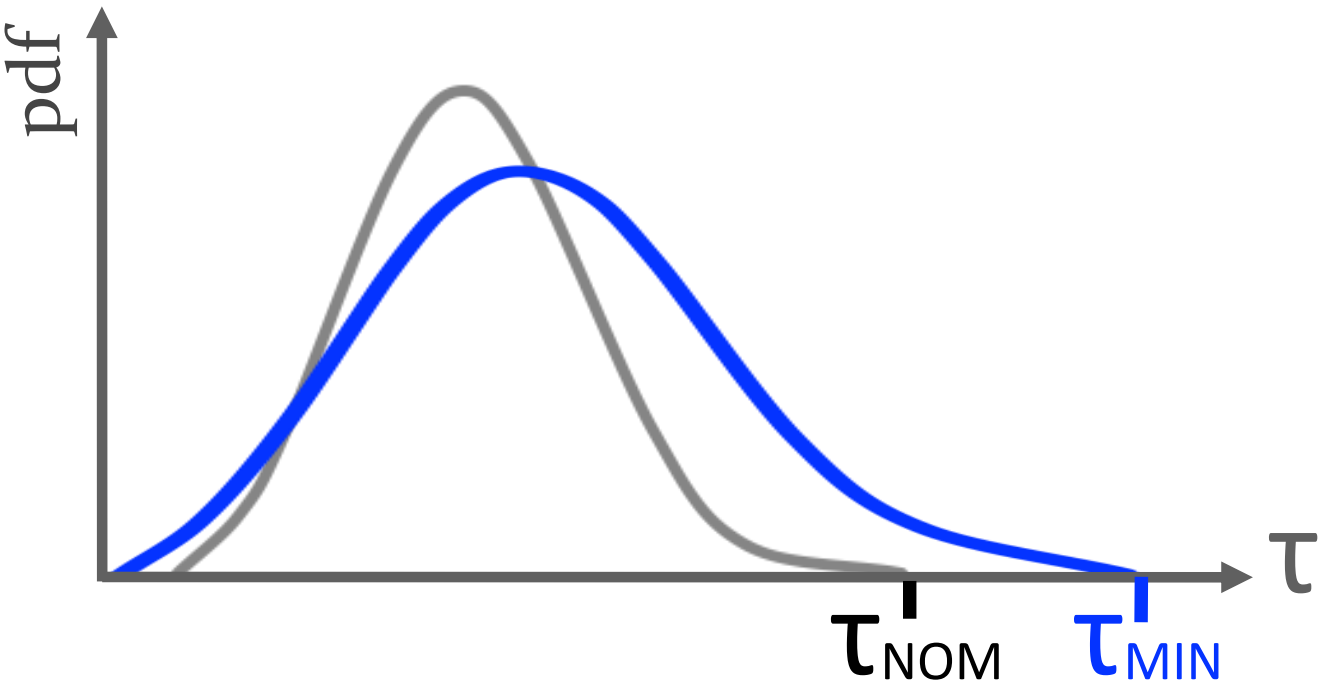
Logic Timing Model



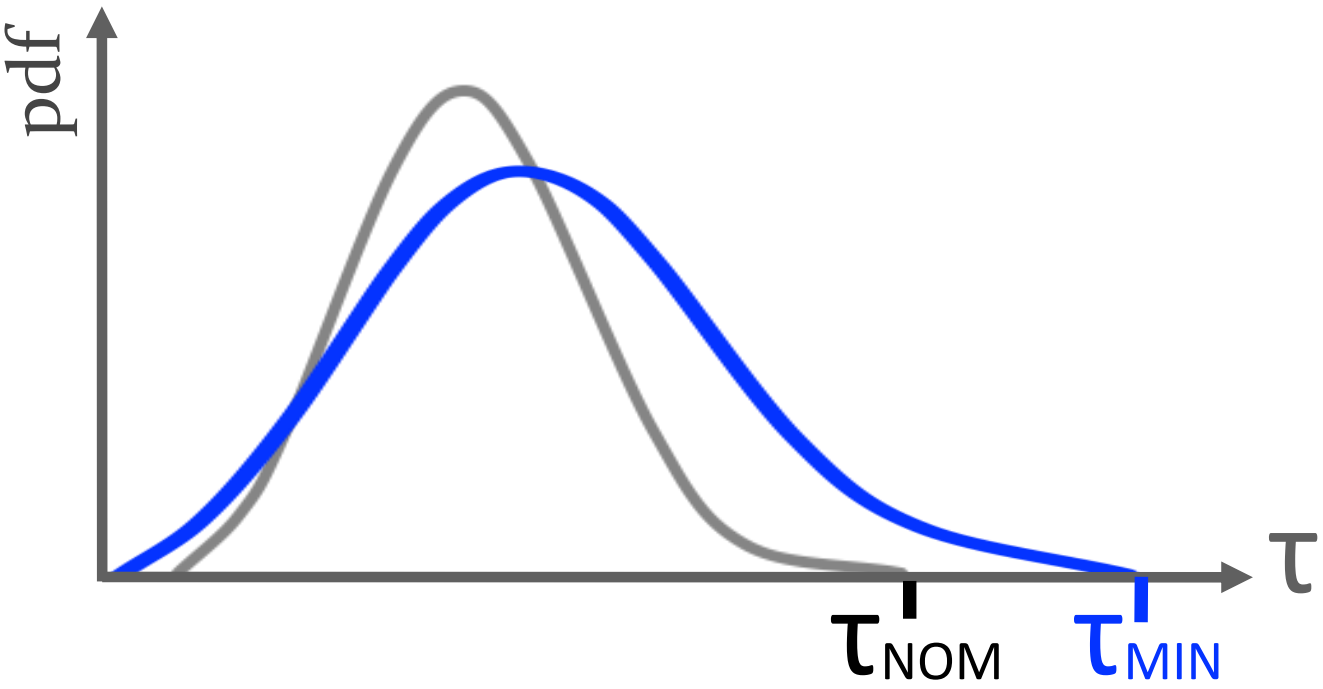
Logic Timing Model



Logic Timing Model

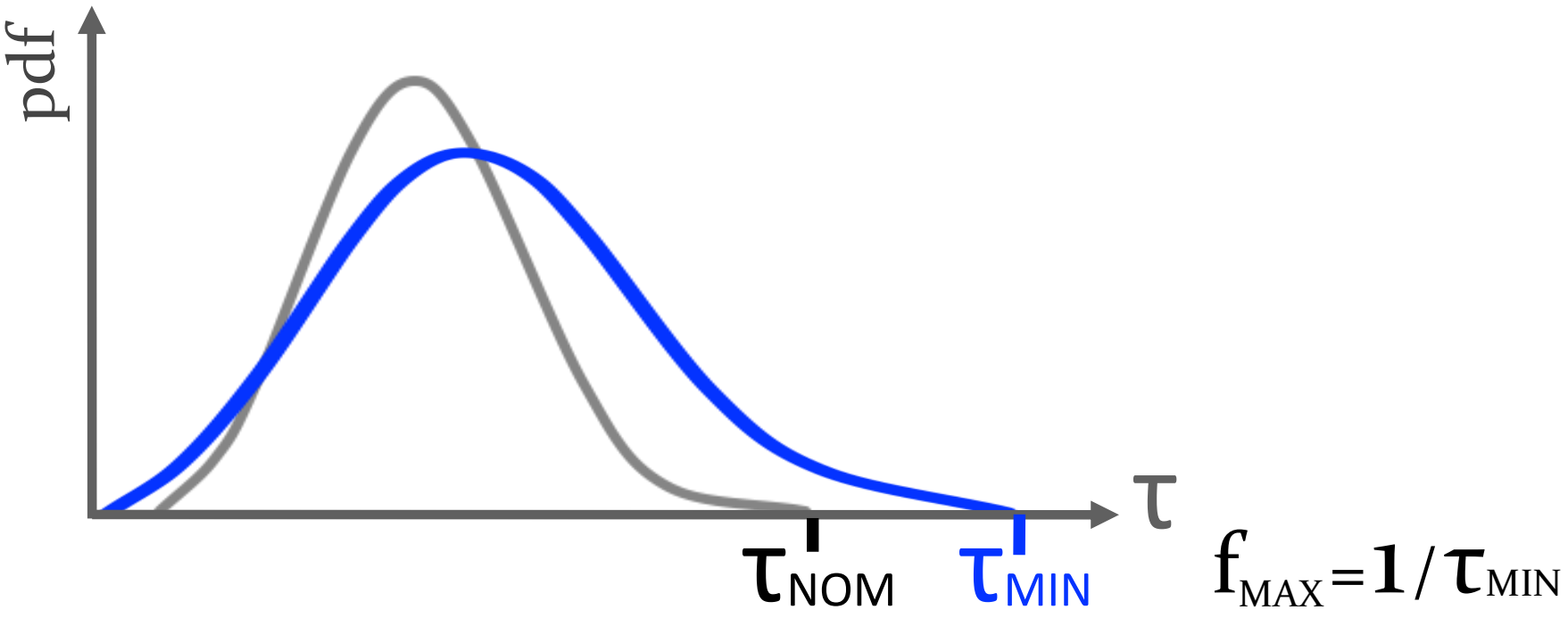


Logic Timing Model

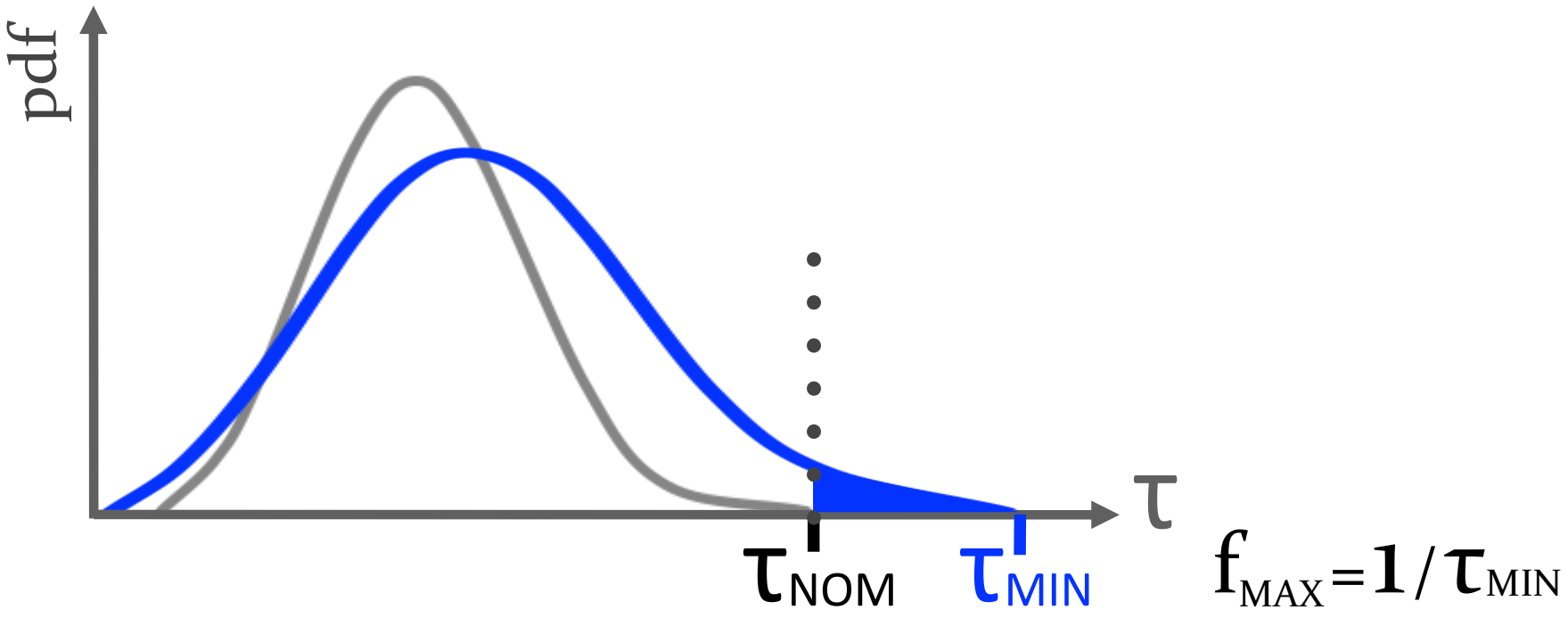


Safe operating point

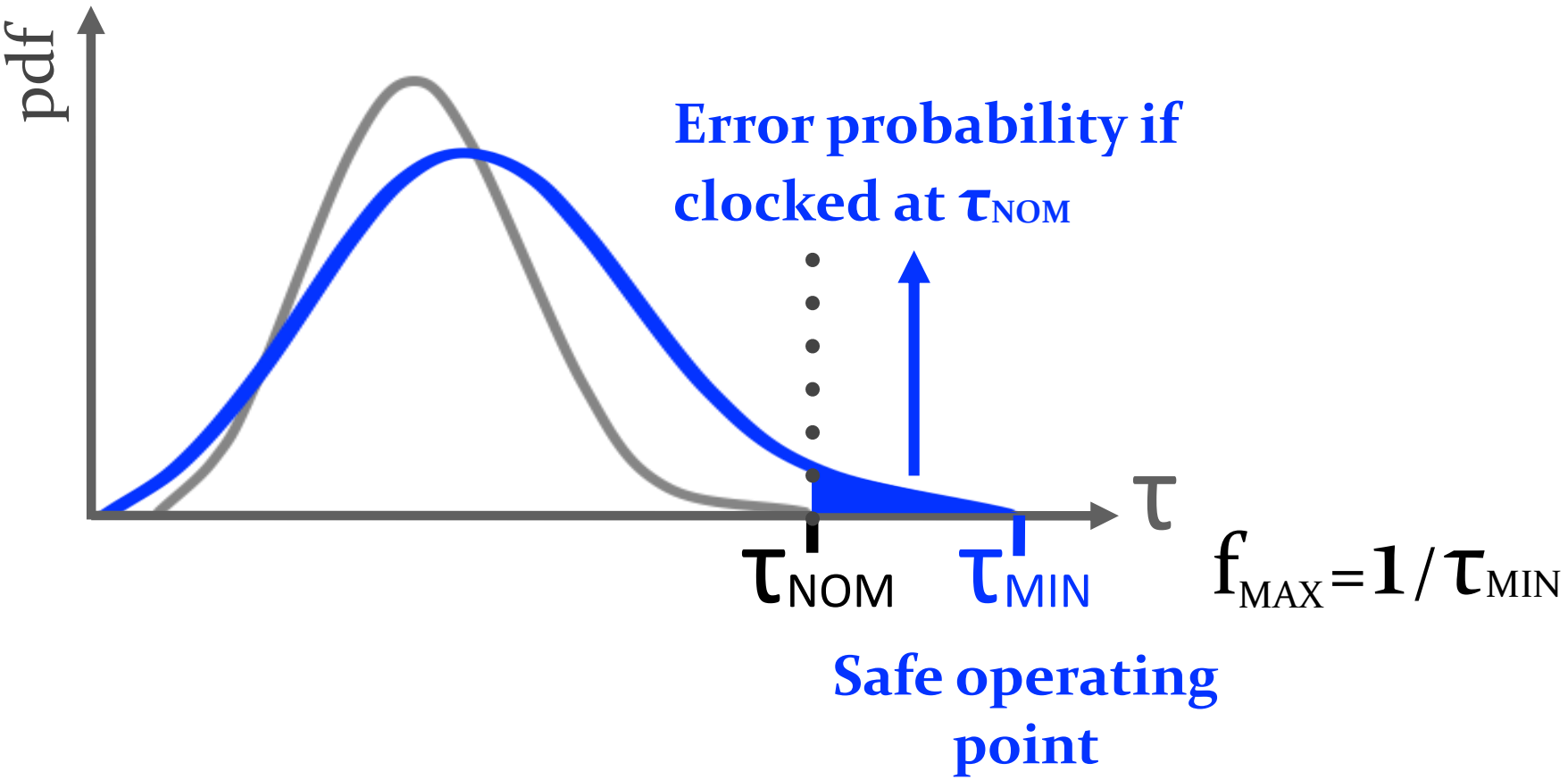
Logic Timing Model



Logic Timing Model

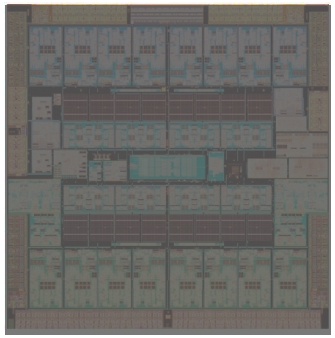


Logic Timing Model

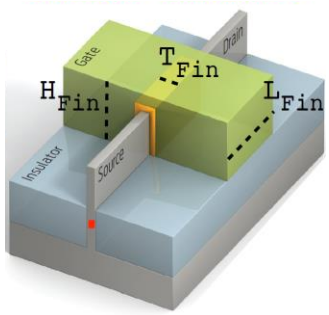


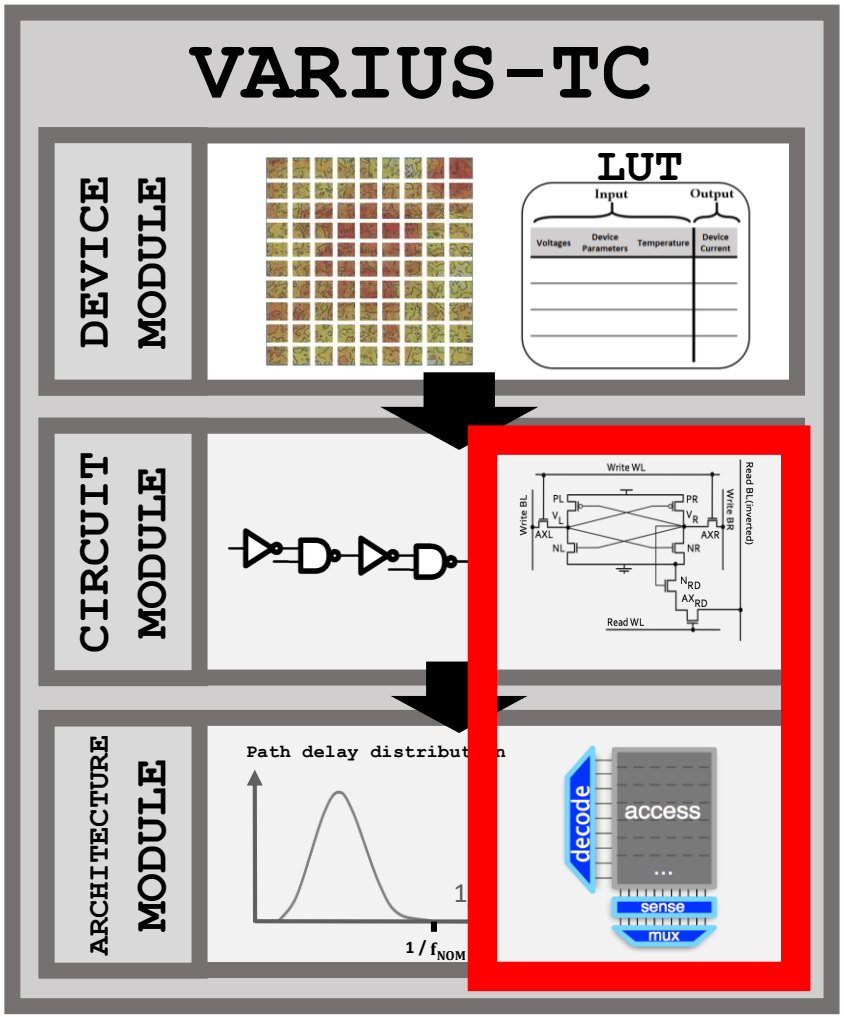
VARIUS-TC: Overview

floorplan

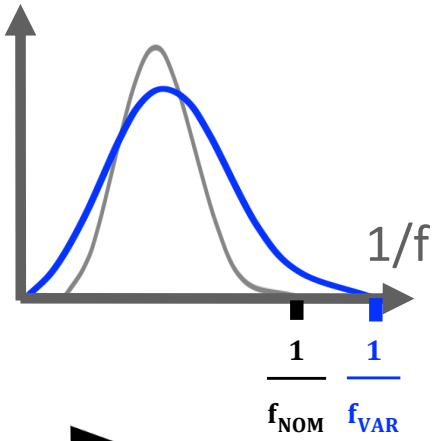


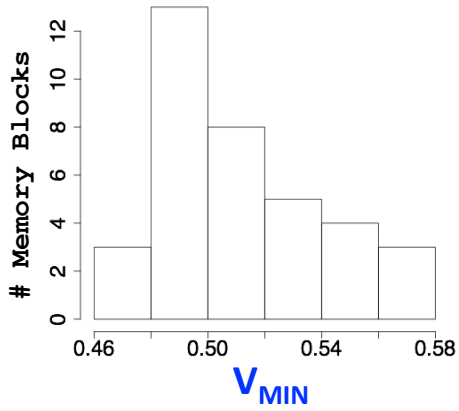
device parameters





Path delay distribution





Memory Model

Memory Model

- Supports 6T and 8T memory cell

Memory Model

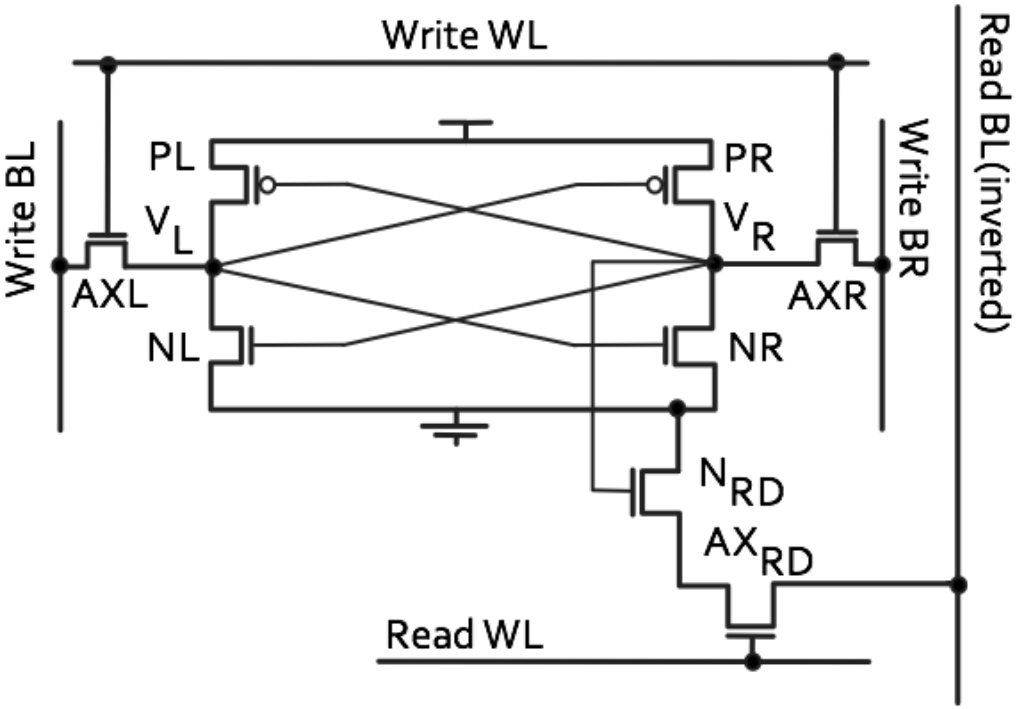
- Supports 6T and 8T memory cell
- Timing errors
  - Write timing
  - Read timing

Memory Model

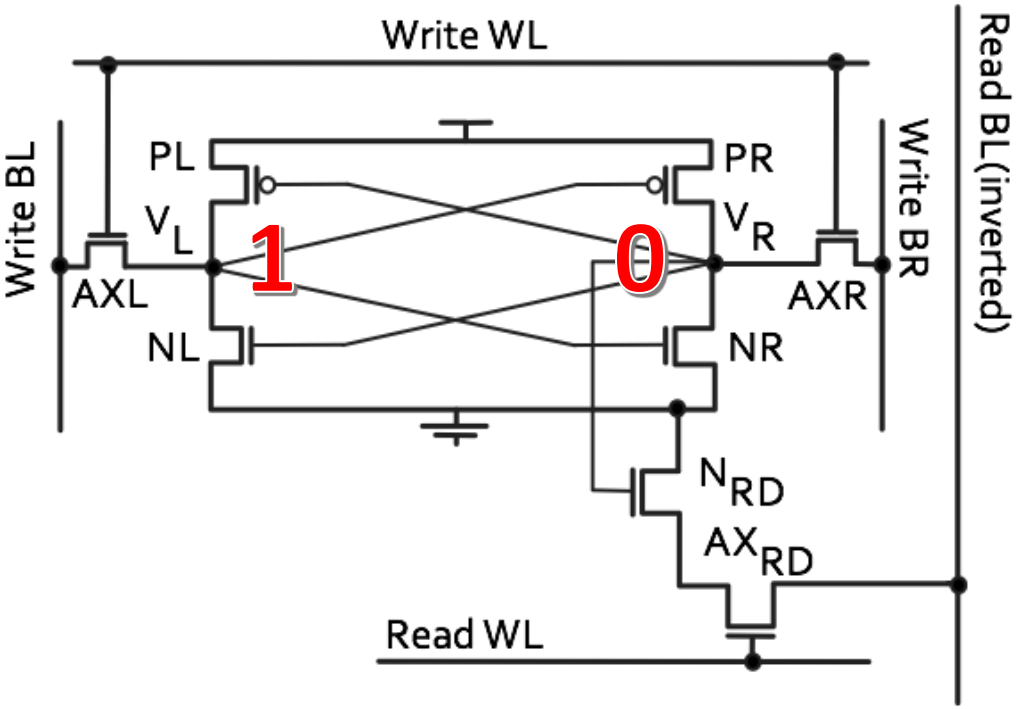
- Supports 6T and 8T memory cell
- Timing errors
  - Write timing
  - Read timing
- Stability errors
  - Hold error
  - Write Stability

Hold Error

Hold Error

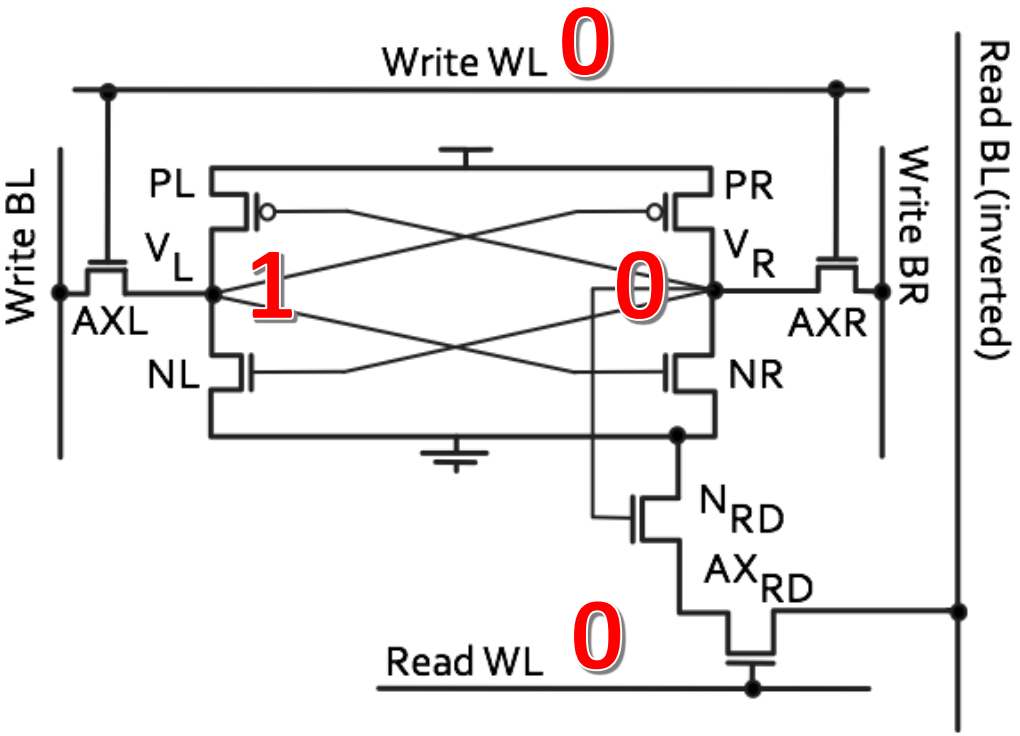

Write WL
Read BL(inverted)
Write BL
Write BR
PL
PR
V_L
V_R
AXL
AXR
NL
NR
N_{RD}
AX_{RD}
Read WL

Hold Error


Write WL
Write BL
PL
PR
V_L
V_R
1
0
AXL
AXR
Write BR
NL
NR
N_{RD}
AX_{RD}
Read BL(inverted)
Read WL

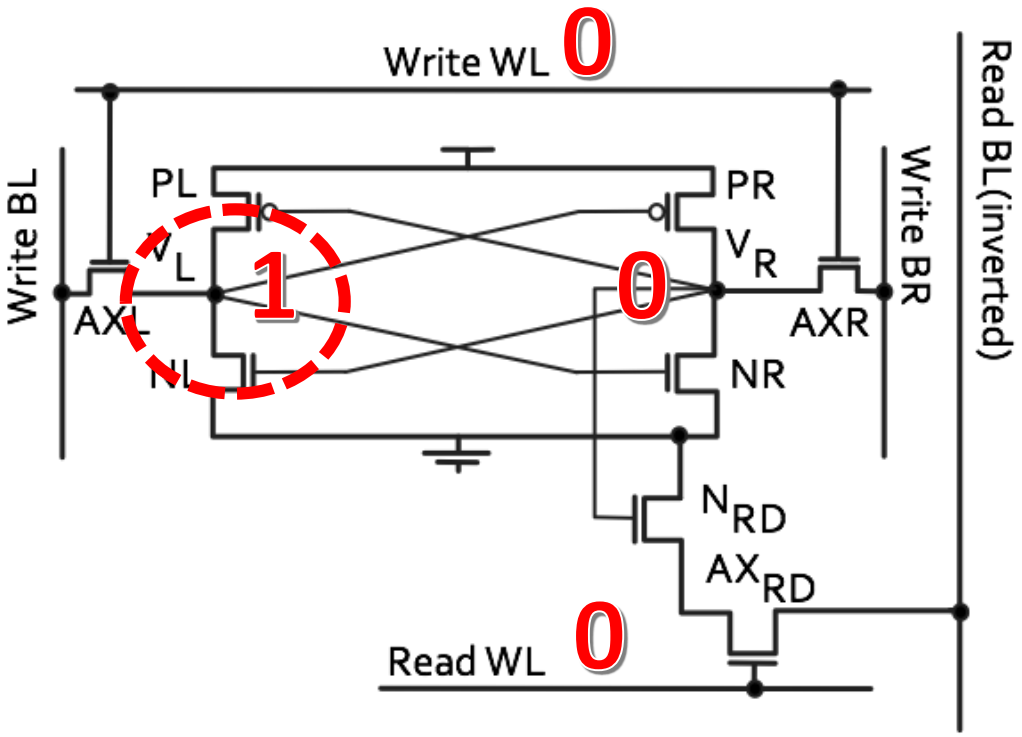
Hold Error

- The cell is not accessed.



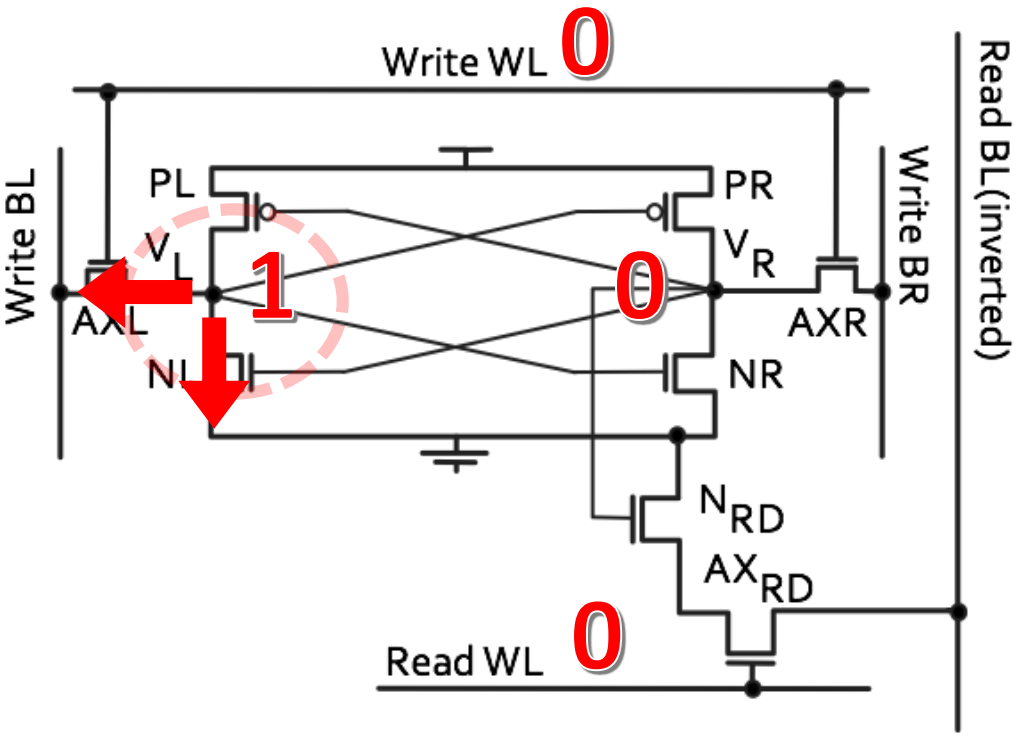
Hold Error

- The cell is not accessed.
- Node V_L looses its state.



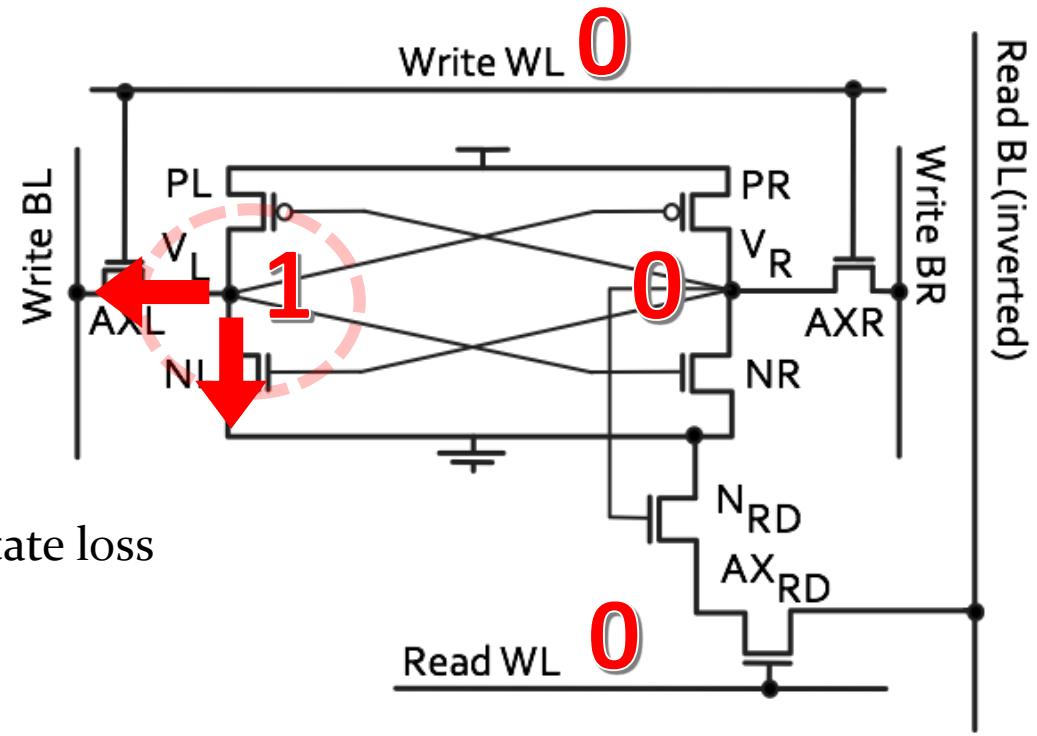
Hold Error

- The cell is not accessed.
- Node V_L looses its state.
  - Excessive leakage



Hold Error

- The cell is not accessed.
- Node V_L looses its state.
  - Excessive leakage



VARIUS-TC

- Minimum Vdd (V_{MIN}) to exclude state loss

Evaluation Setup

Evaluation Setup

- Device parameters
  - PTM, FinFET, 16nm

Evaluation Setup

- Device parameters
  - PTM, FinFET, 16nm
- Parametric sweep
  - L_{Fin}, T_{Fin}, ϕ_g

Evaluation Setup

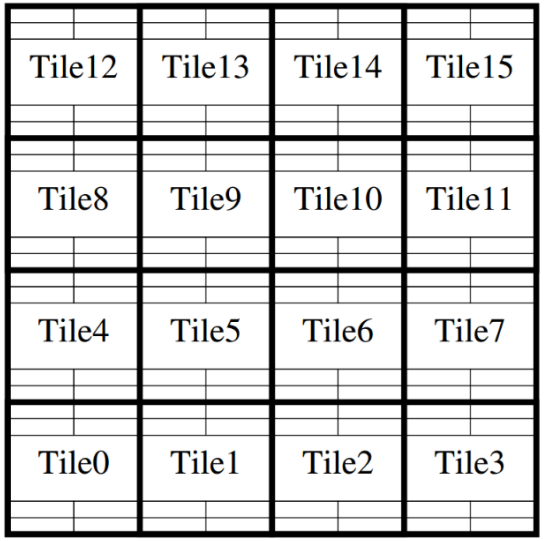
- Device parameters
  - PTM, FinFET, 16nm
- Parametric sweep
  - L_{Fin}, T_{Fin}, ϕ_g
- 3 levels of variation
  - low, medium, high

Parameter	Low var.	Medium var.	High var.
L_{Fin}	3.5%	7%	10.5%
T_{Fin}	3.5%	7%	10.5%
ϕ_g	0.16%	0.32%	0.48%

Evaluation Setup

- Device parameters
  - PTM, FinFET, 16nm
- Parametric sweep
  - L_{Fin}, T_{Fin}, ϕ_g
- 3 levels of variation
  - low, medium, high
- Many-core system
  - 16 tiles

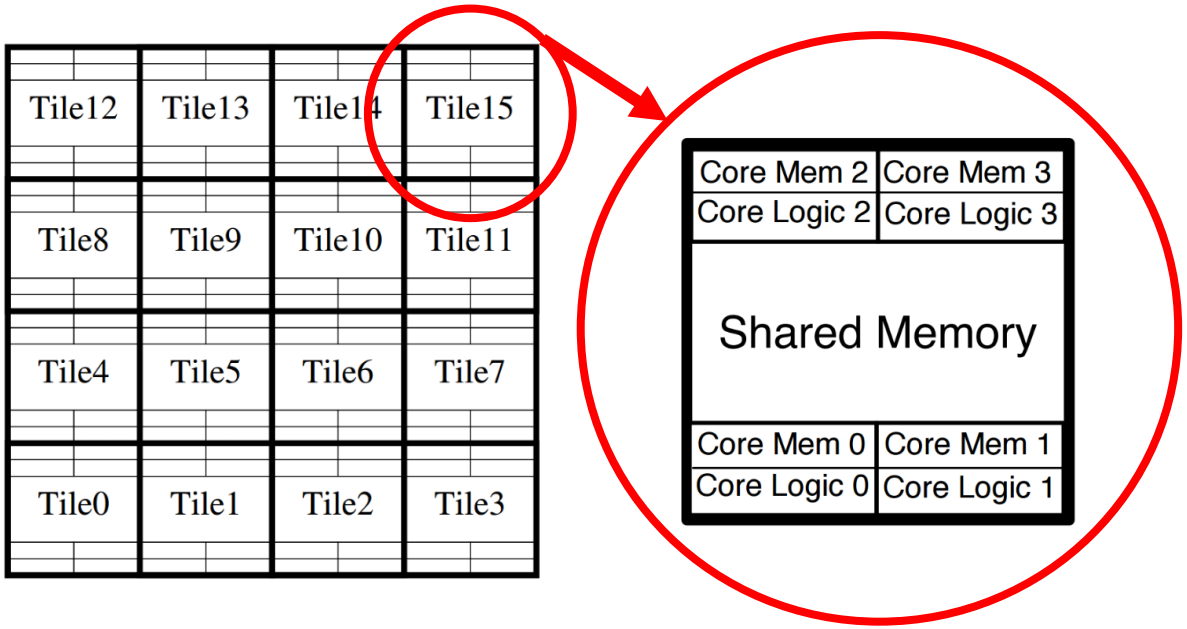
Parameter	Low var.	Medium var.	High var.
L_{Fin}	3.5%	7%	10.5%
T_{Fin}	3.5%	7%	10.5%
ϕ_g	0.16%	0.32%	0.48%



Evaluation Setup

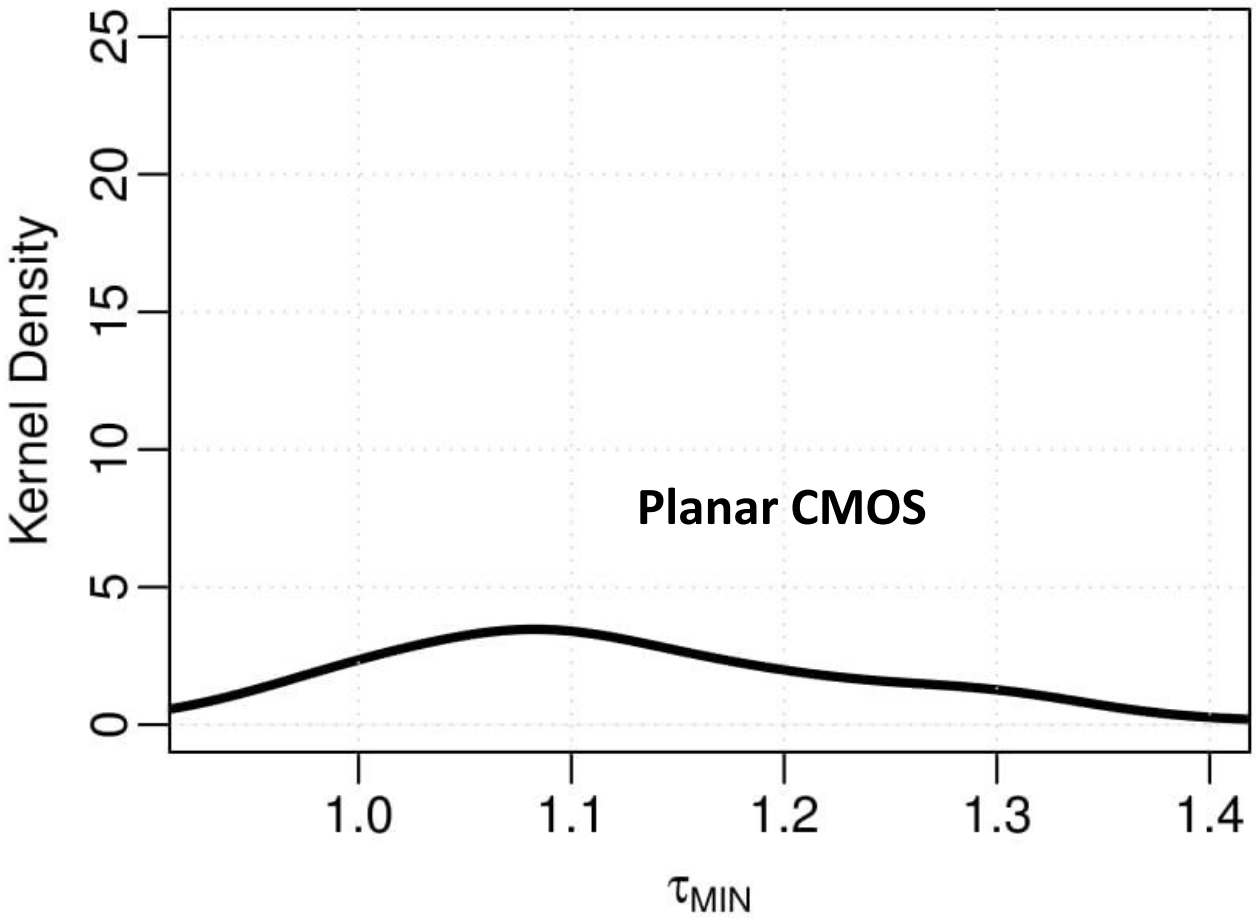
- Device parameters
  - PTM, FinFET, 16nm
- Parametric sweep
  - L_{Fin}, T_{Fin}, ϕ_g
- 3 levels of variation
  - low, medium, high
- Many-core system
  - 16 tiles
  - 4 core per tile
  - Private L1, Shared L2

Parameter	Low var.	Medium var.	High var.
L_{Fin}	3.5%	7%	10.5%
T_{Fin}	3.5%	7%	10.5%
ϕ_g	0.16%	0.32%	0.48%

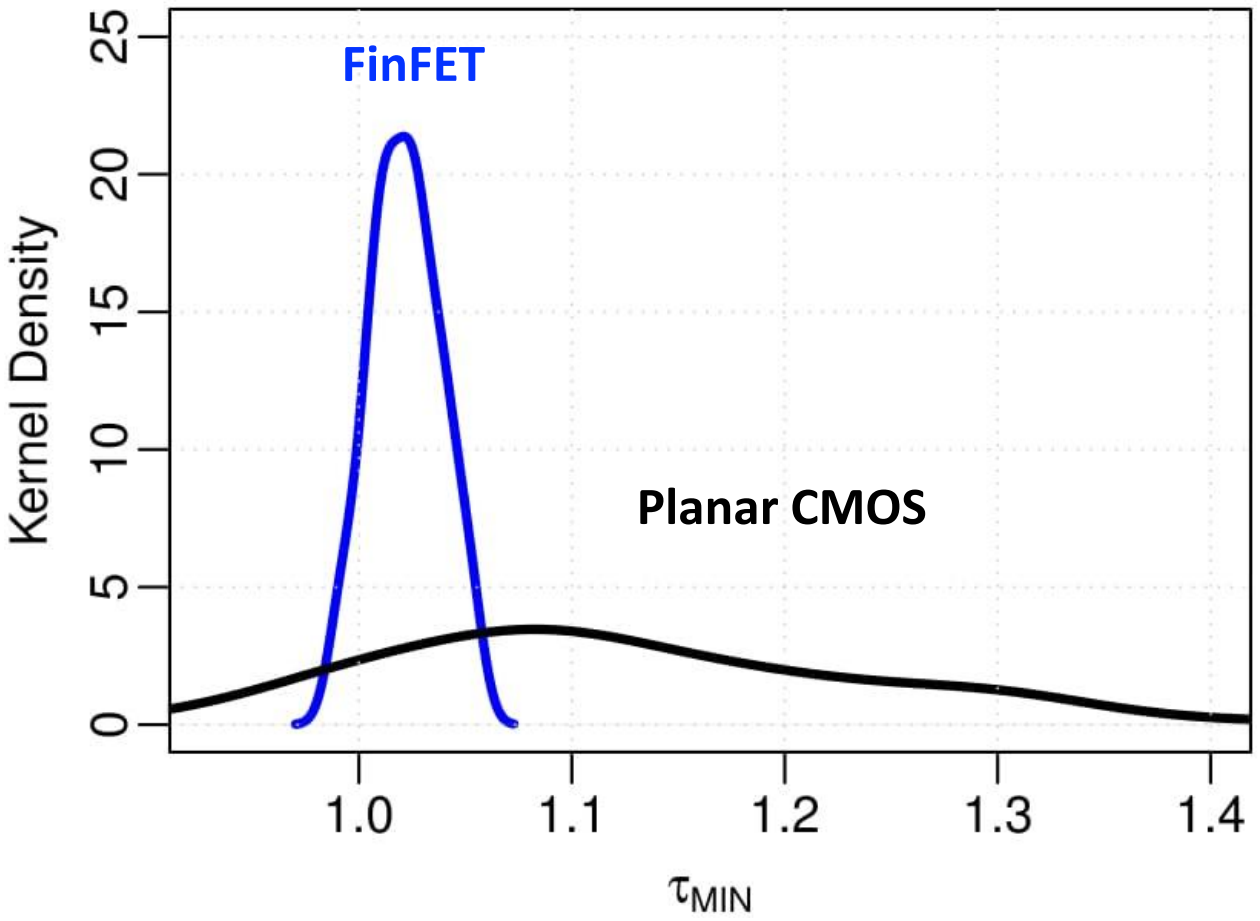


Impact on Logic Timing

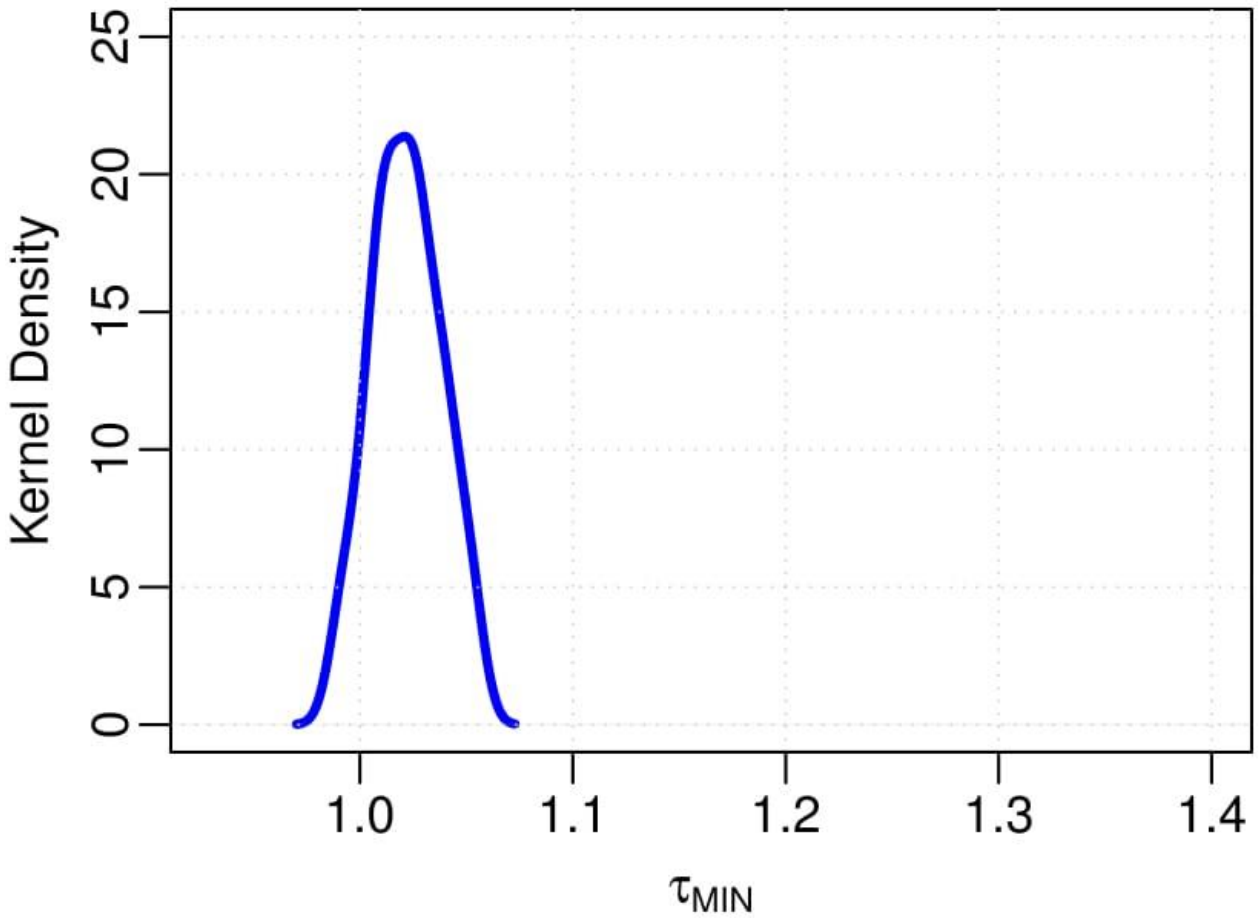
Impact on Logic Timing



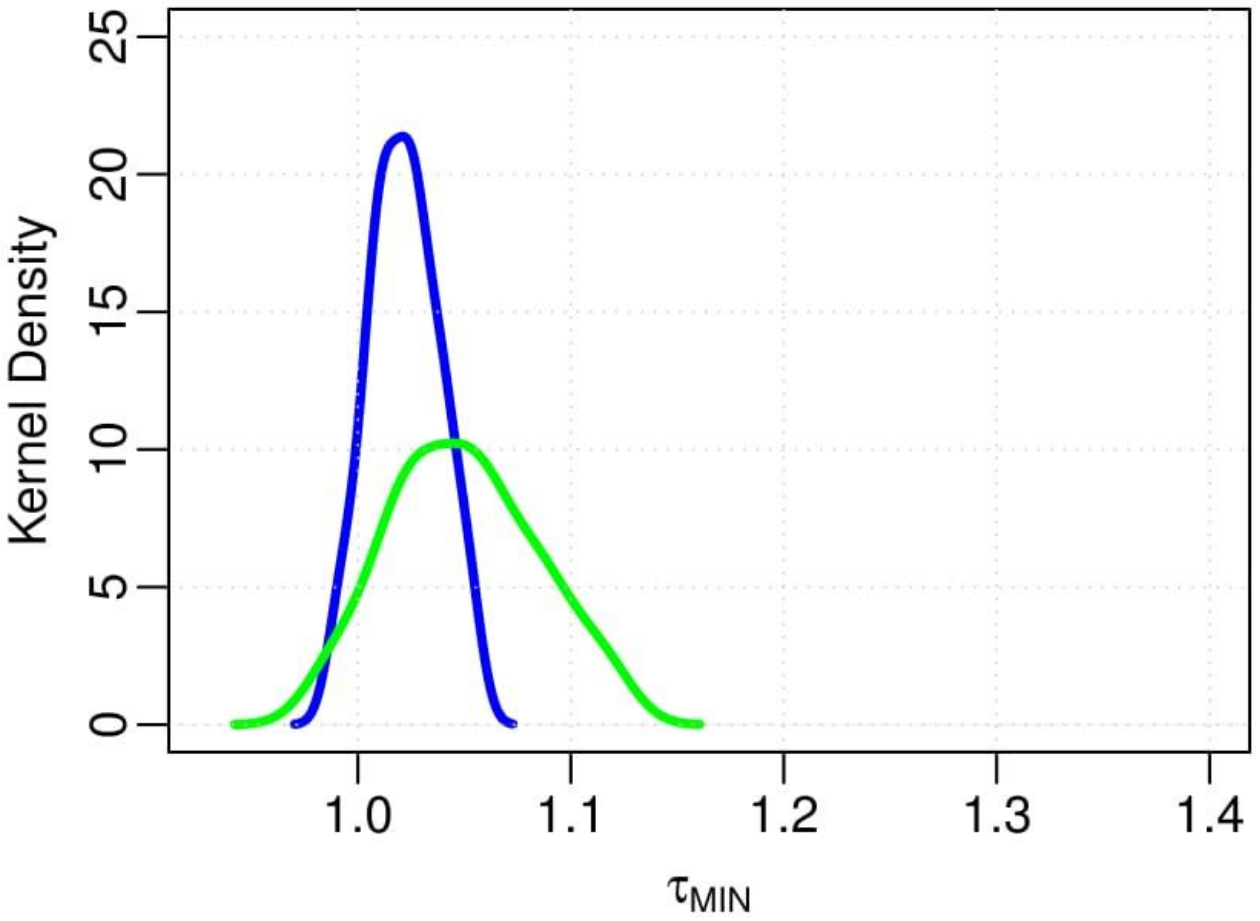
Impact on Logic Timing



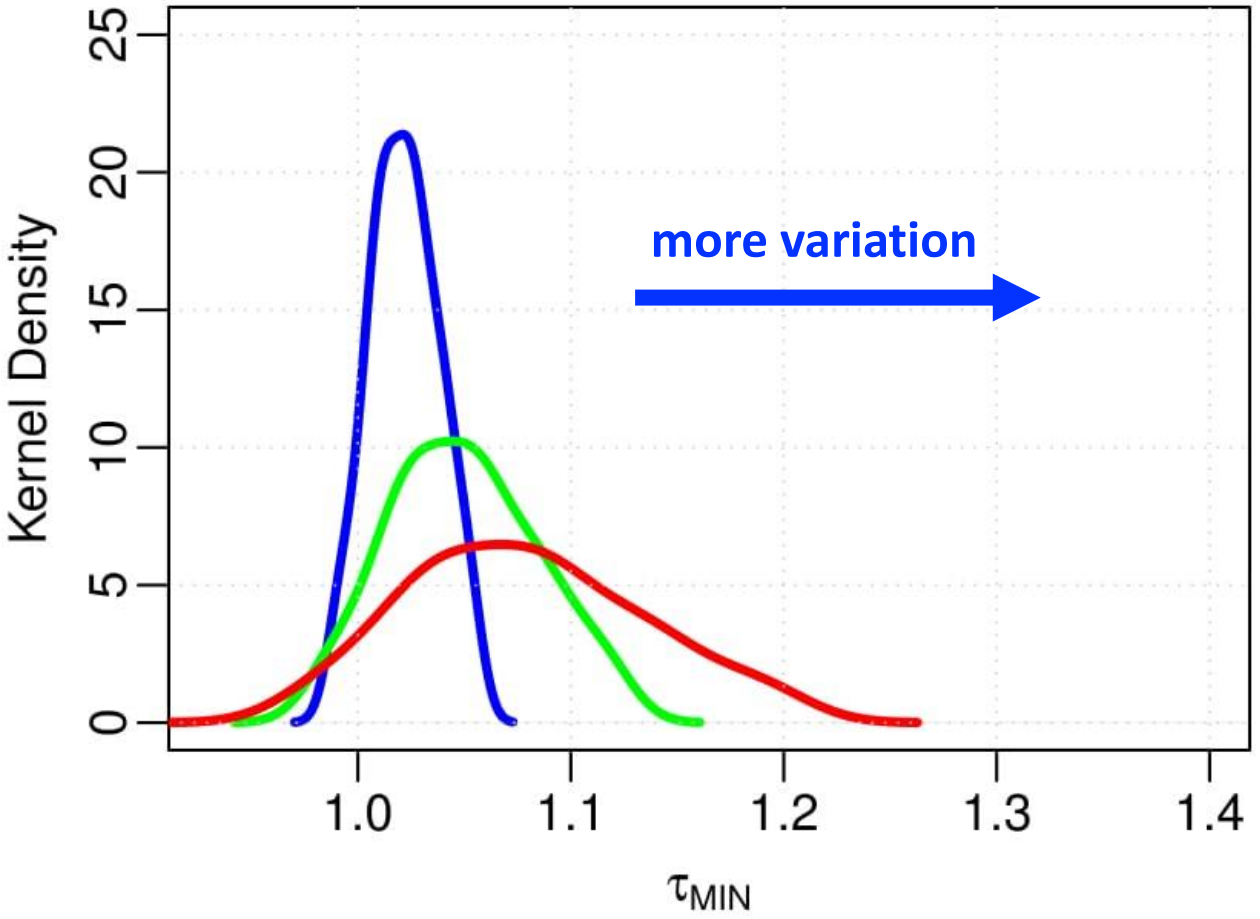
Impact on Logic Timing



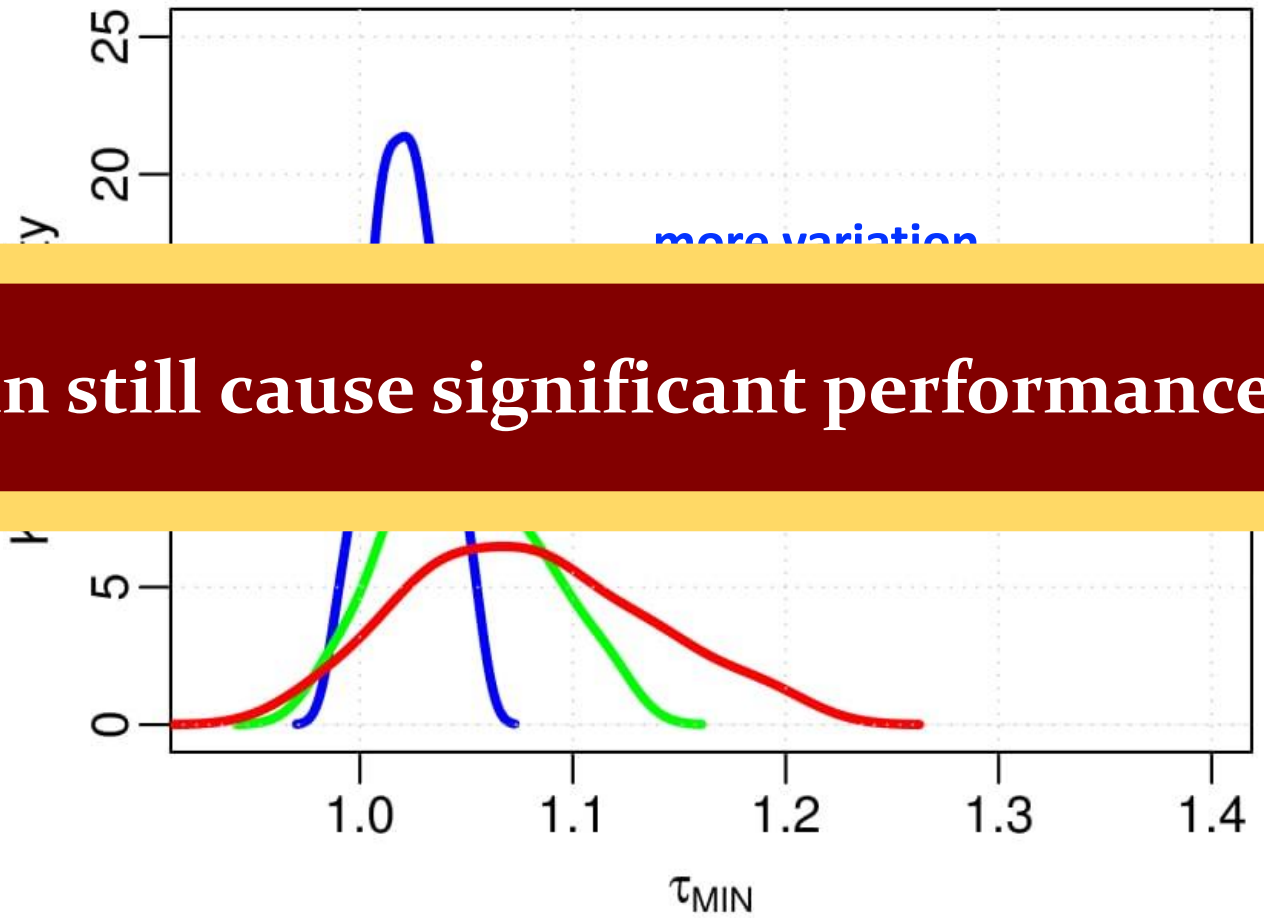
Impact on Logic Timing



Impact on Logic Timing



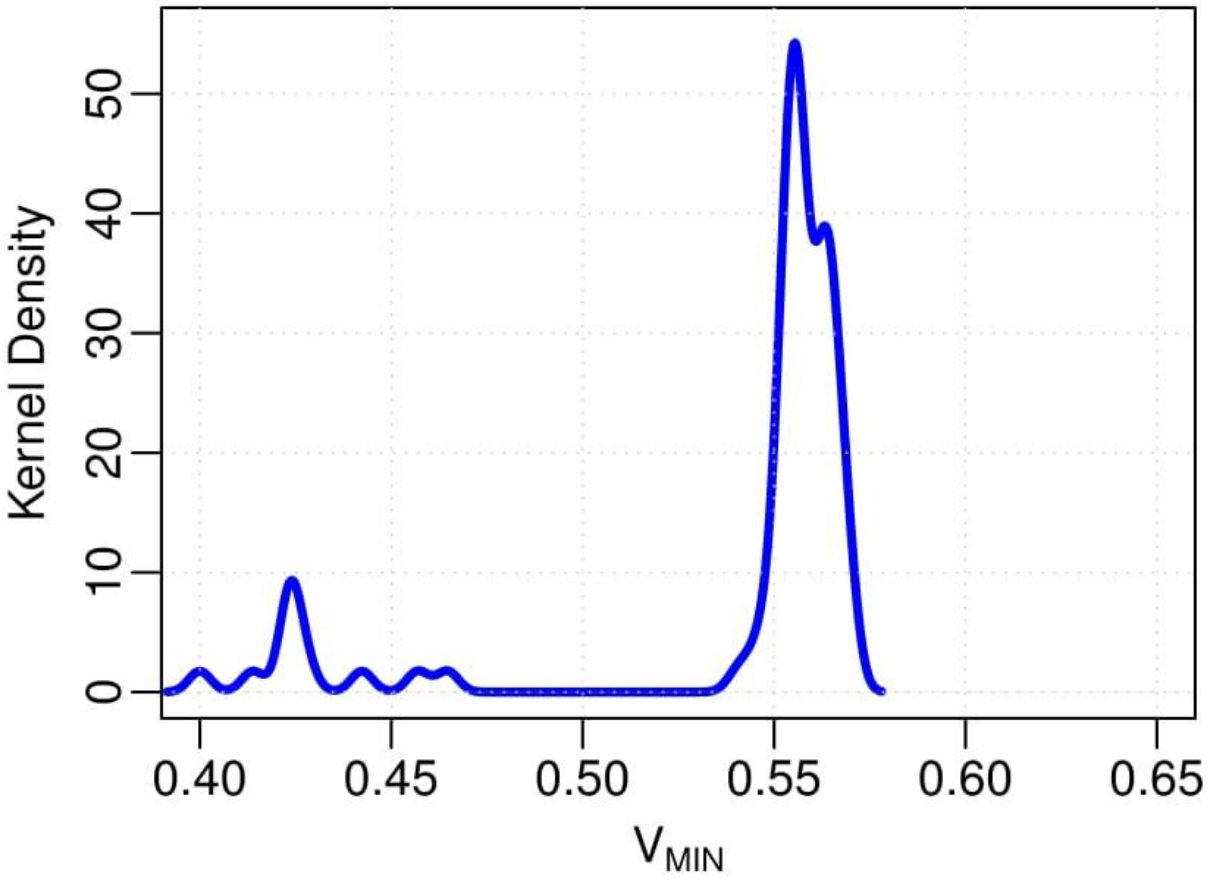
Impact on Logic Timing



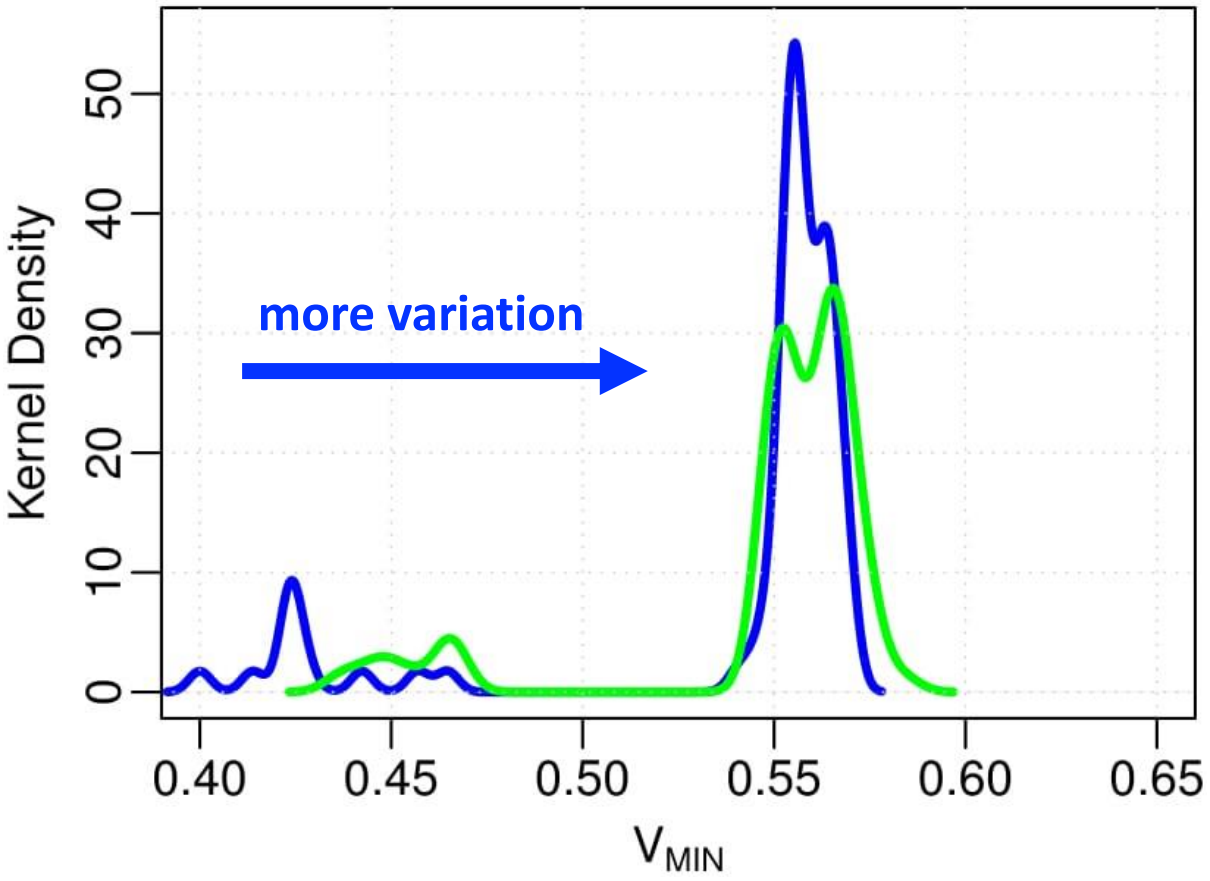
PV can still cause significant performance loss.

Impact on V_{MIN}

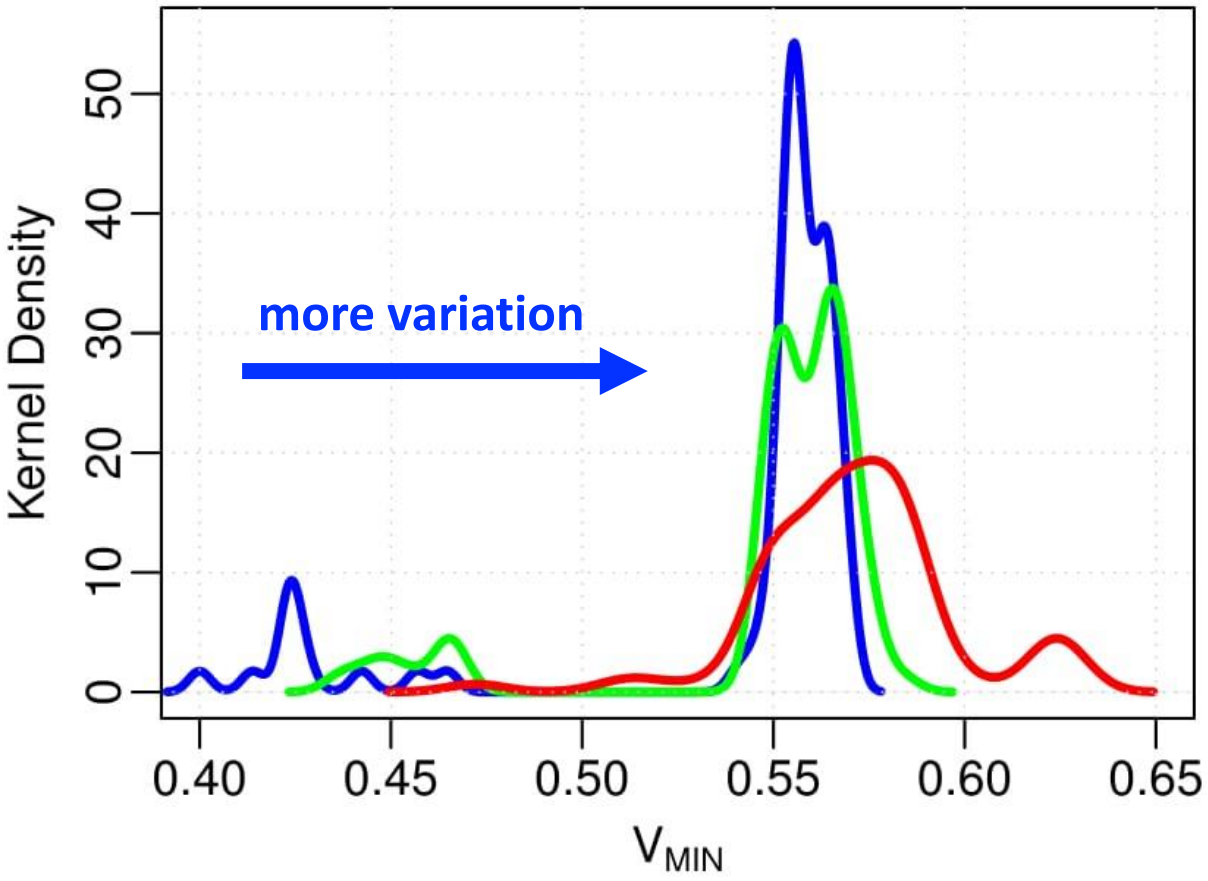
Impact on V_{MIN}



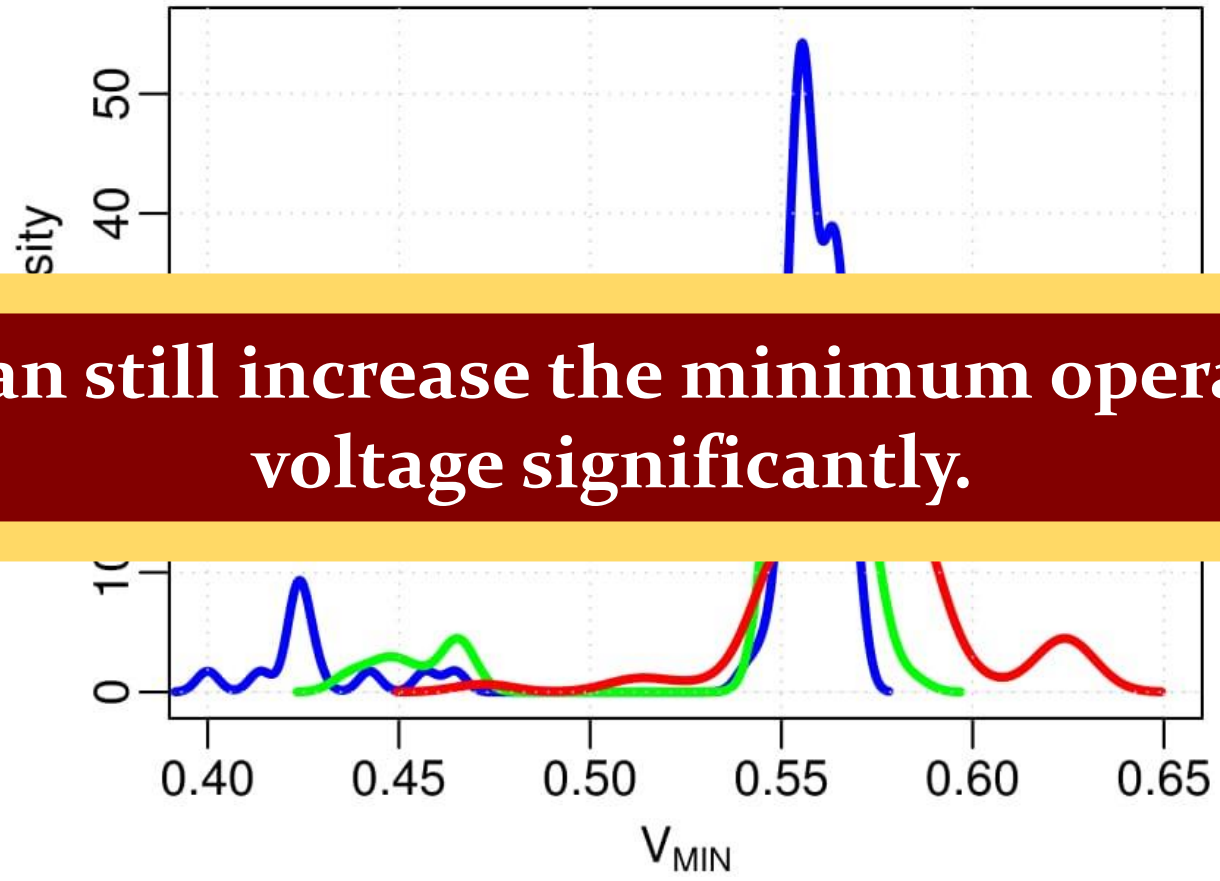
Impact on V_{MIN}



Impact on V_{MIN}



Impact on V_{MIN}



PV can still increase the minimum operating voltage significantly.

Example Use Case

Example Use Case

- Reducing operating voltage reduces power consumption.

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\textbf{Throughput} \propto N_{cores} \times f_{cores}$$

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\textbf{Throughput} \propto N_{cores} \times f_{cores}$$

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

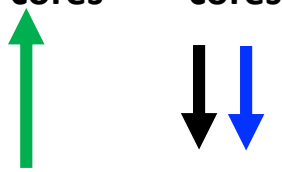
$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

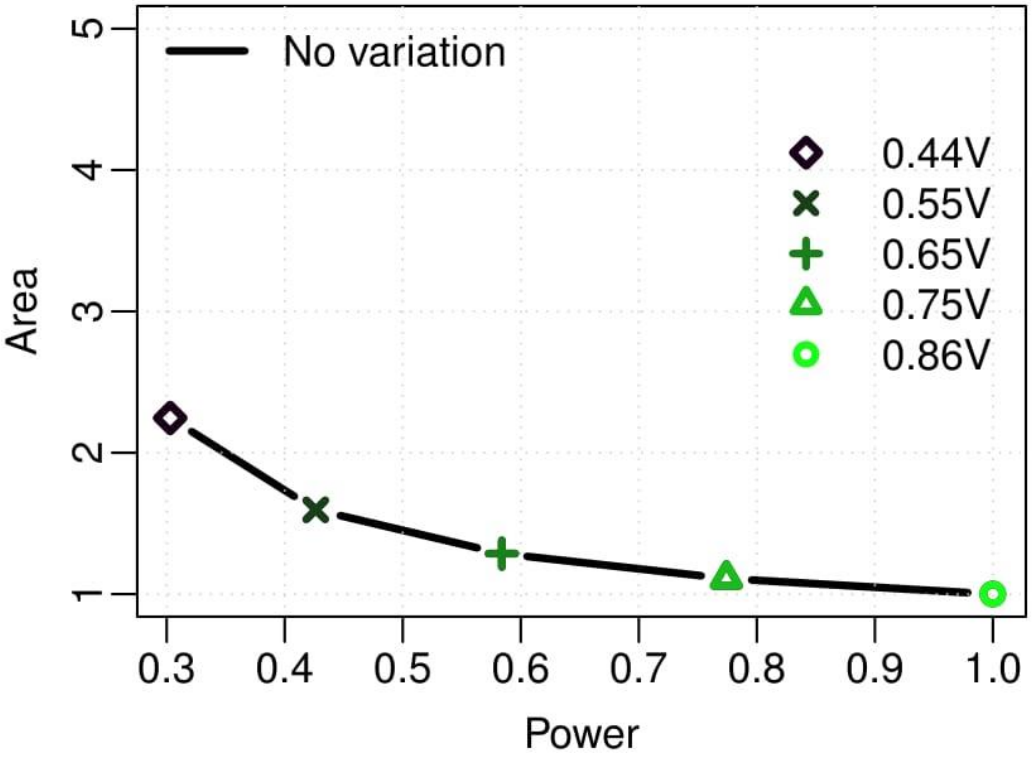


- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

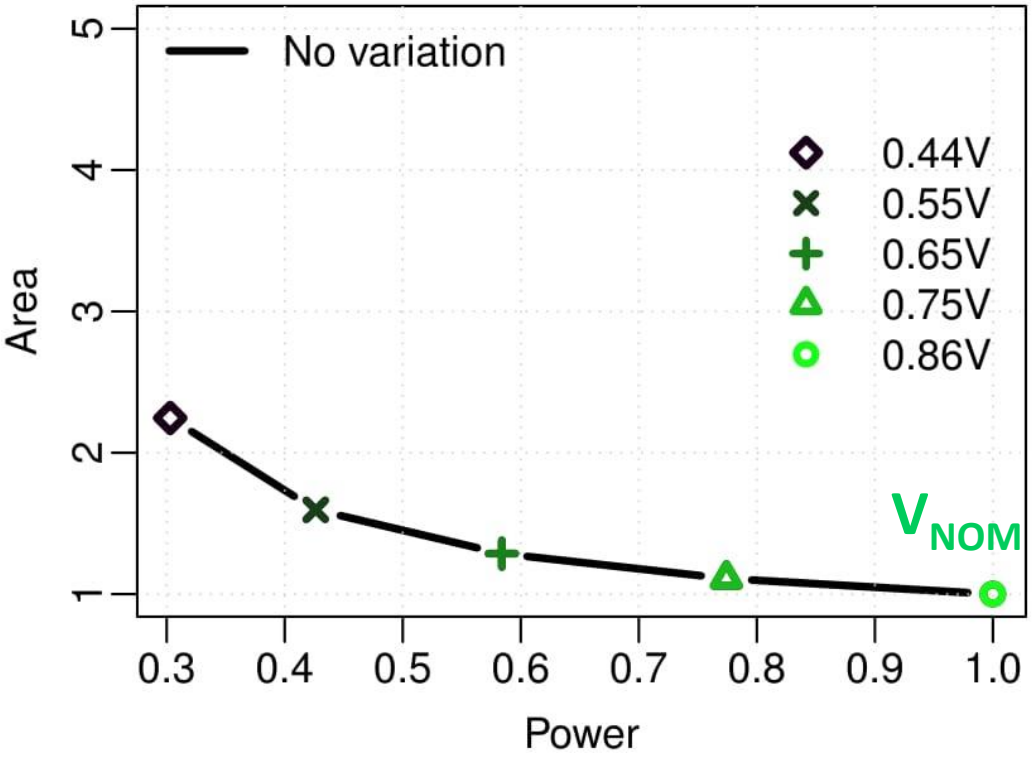


- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

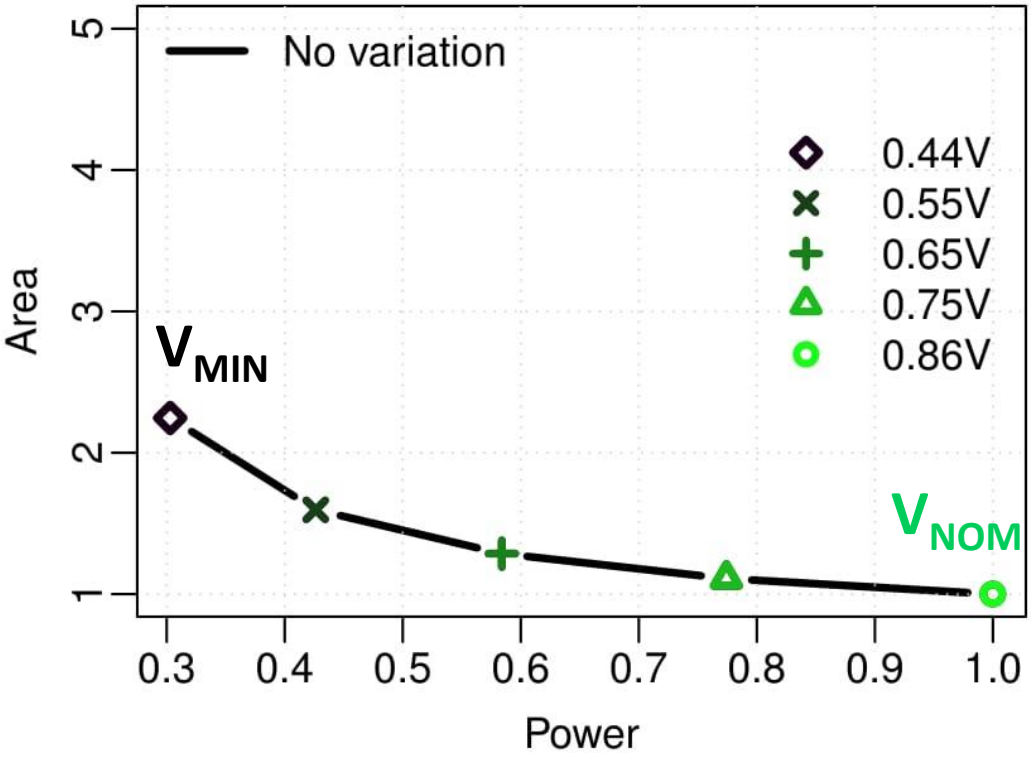


- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

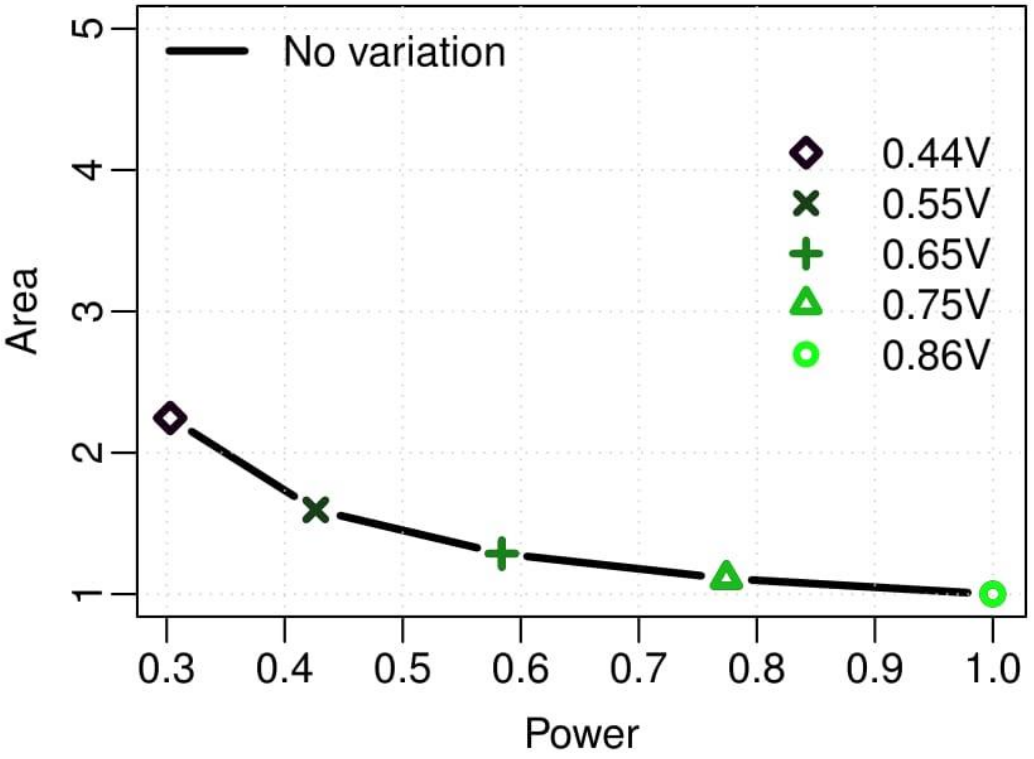


- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

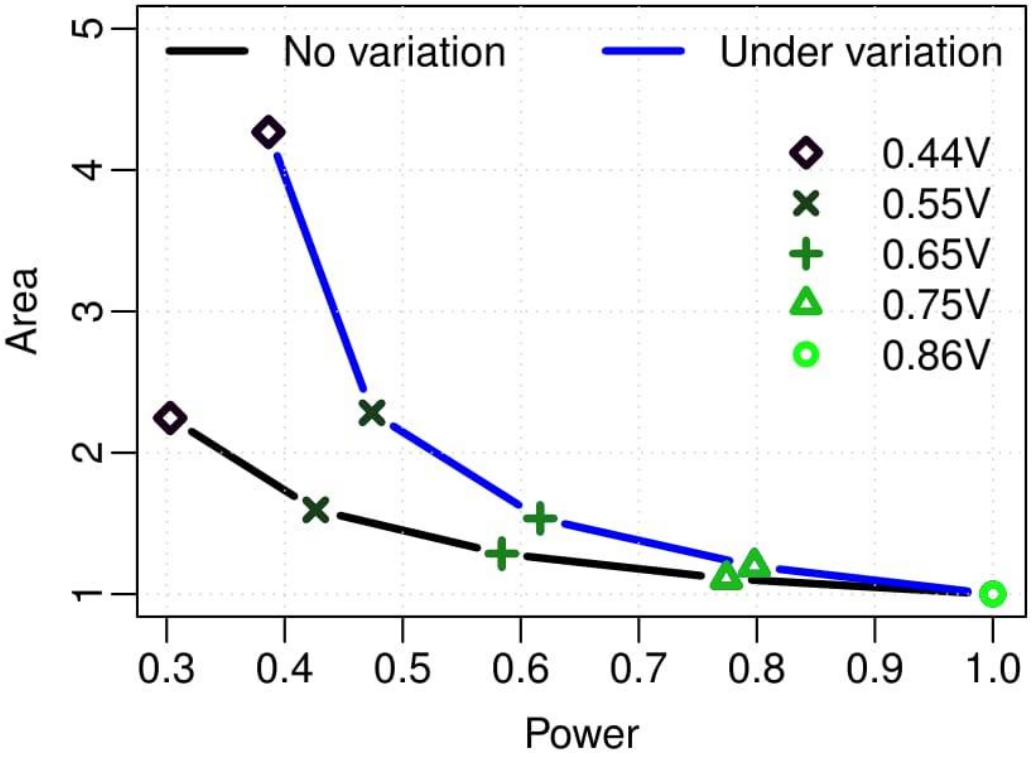


- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

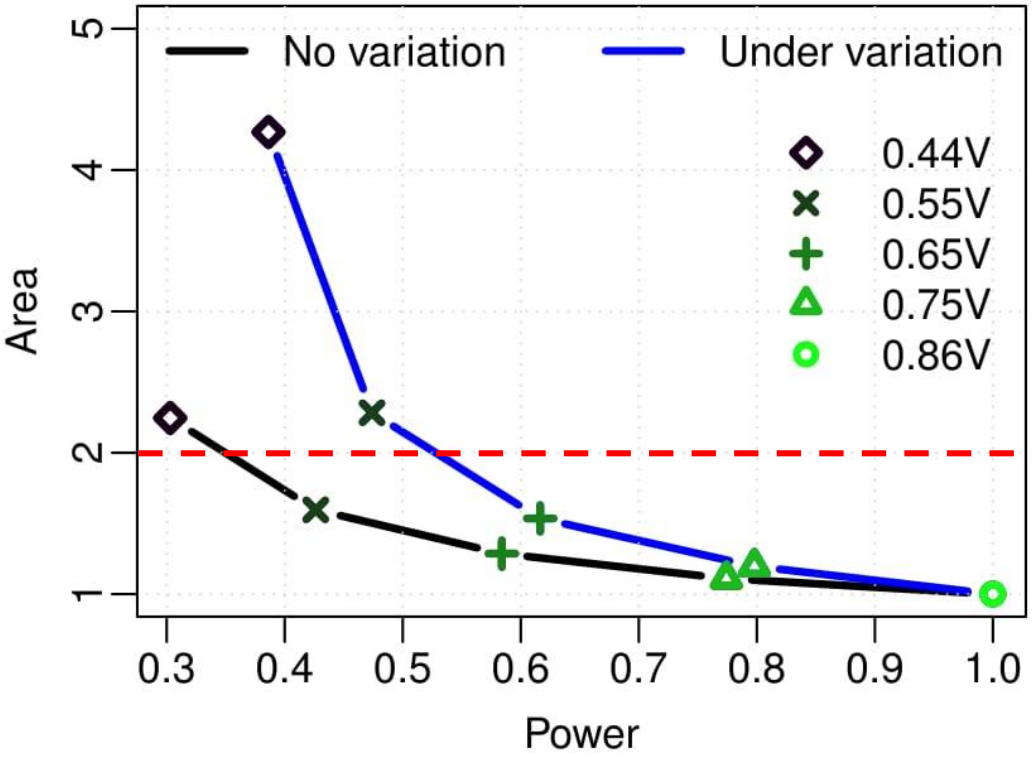


- Process Variation …

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

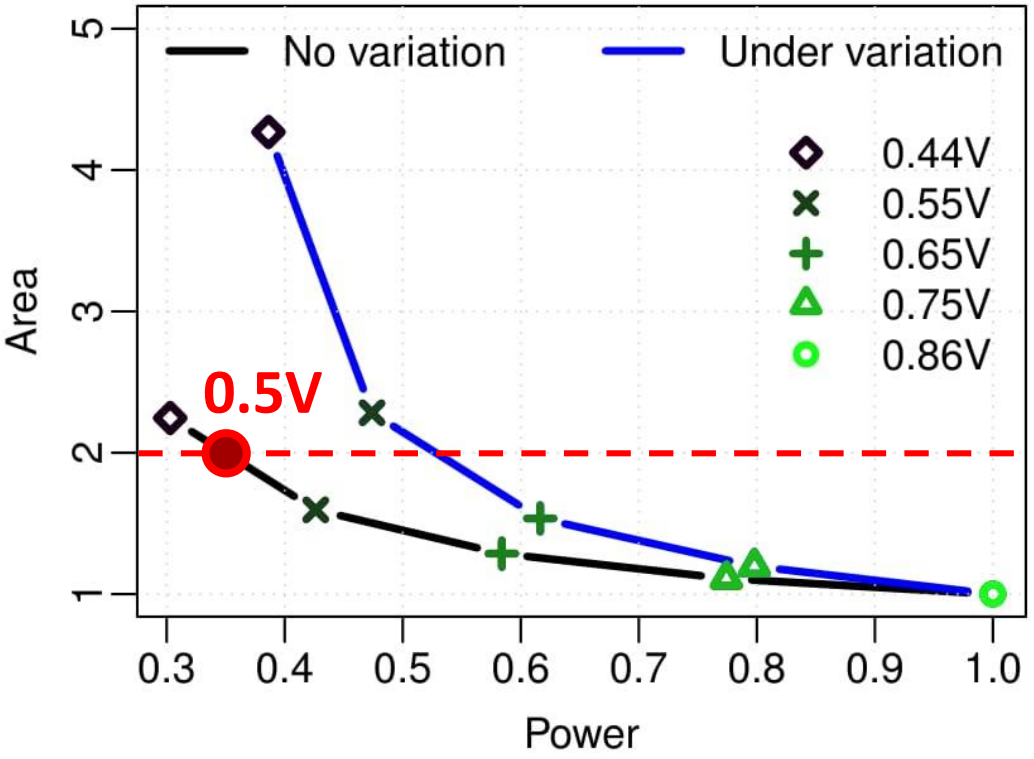


- Process Variation …
  - If > 2x area increase is not affordable…

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

$$\text{Throughput} \propto N_{cores} \times f_{cores}$$

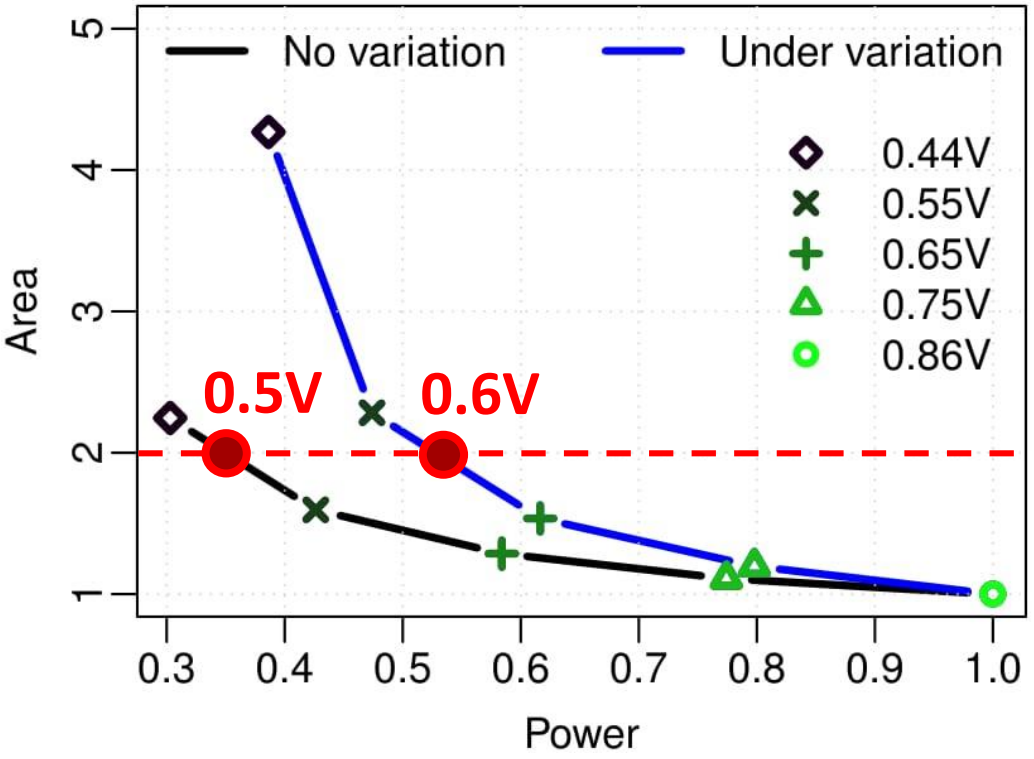


- Process Variation …
  - If > 2x area increase is not affordable…

Example Use Case

- Reducing operating voltage reduces power consumption.
  - Frequency reduces, too.

Throughput $\propto N_{cores} \times f_{cores}$



- Process Variation …
  - If > 2x area increase is not affordable…

Related Work

Related Work

- **VARIUS-NTV** is tailored for planar CMOS only.

Related Work

- **VARIUS-NTV** is tailored for planar CMOS only.
- **FinCANON** and **McPAT-PVT** are architecture-level FinFET-based models.

Related Work

- **VARIUS-NTV** is tailored for planar CMOS only.
- **FinCANON** and **McPAT-PVT** are architecture-level FinFET-based models.
  - Modular

Related Work

- **VARIUS-NTV** is tailored for planar CMOS only.
- **FinCANON** and **McPAT-PVT** are architecture-level FinFET-based models.
  - Modular
  - TCAD-based device-level simulations

Related Work

- **VARIUS-NTV** is tailored for planar CMOS only.
- **FinCANON** and **McPAT-PVT** are architecture-level FinFET-based models.
  - Modular
  - TCAD-based device-level simulations
- **VARIUS-TC**'s strength

Related Work

- **VARIUS-NTV** is tailored for planar CMOS only.
- **FinCANON** and **McPAT-PVT** are architecture-level FinFET-based models.
  - Modular
  - TCAD-based device-level simulations
- **VARIUS-TC**'s strength
  - Probabilistic model to analyze **processor logic** and **error modes of memory**

Conclusion

Conclusion

- **VARIUS-TC**

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage
  - Generation of critical path delay, power, and V_{MIN} distributions

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage
  - Generation of critical path delay, power, and V_{MIN} distributions
  - Calculation of error probabilities for

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage
  - Generation of critical path delay, power, and V_{MIN} distributions
  - Calculation of error probabilities for
    - logic (timing)

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage
  - Generation of critical path delay, power, and V_{MIN} distributions
  - Calculation of error probabilities for
    - logic (timing)
    - memory (timing and stability) error modes

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage
  - Generation of critical path delay, power, and V_{MIN} distributions
  - Calculation of error probabilities for
    - logic (timing)
    - memory (timing and stability) error modes
  - Design space exploration

Conclusion

- **VARIUS-TC**
  - Models process variation in **emerging devices** at **architecture-level**.

- **VARIUS-TC** facilities
  - Extraction of a safe operating frequency and voltage
  - Generation of critical path delay, power, and V_{MIN} distributions
  - Calculation of error probabilities for
    - logic (timing)
    - memory (timing and stability) error modes
  - Design space exploration

- Modularity eases experimentation with different designs (e.g., SOI variants)

VARIUS-TC: A Modular Architecture-Level Model of Parametric Variation for Thin-Channel Switches

S. Karen Khatamifard, Michael Resch, Nam Sung Kim†, Ulya R. Karpuzcu

University of Minnesota
{khatami, resco059, ukarpuzc}@umn.edu

†University of Illinois
{nskim}@illinois.edu

http://altai.ece.umn.edu/varius

10/5/2016



UNIVERSITY OF MINNESOTA

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN