

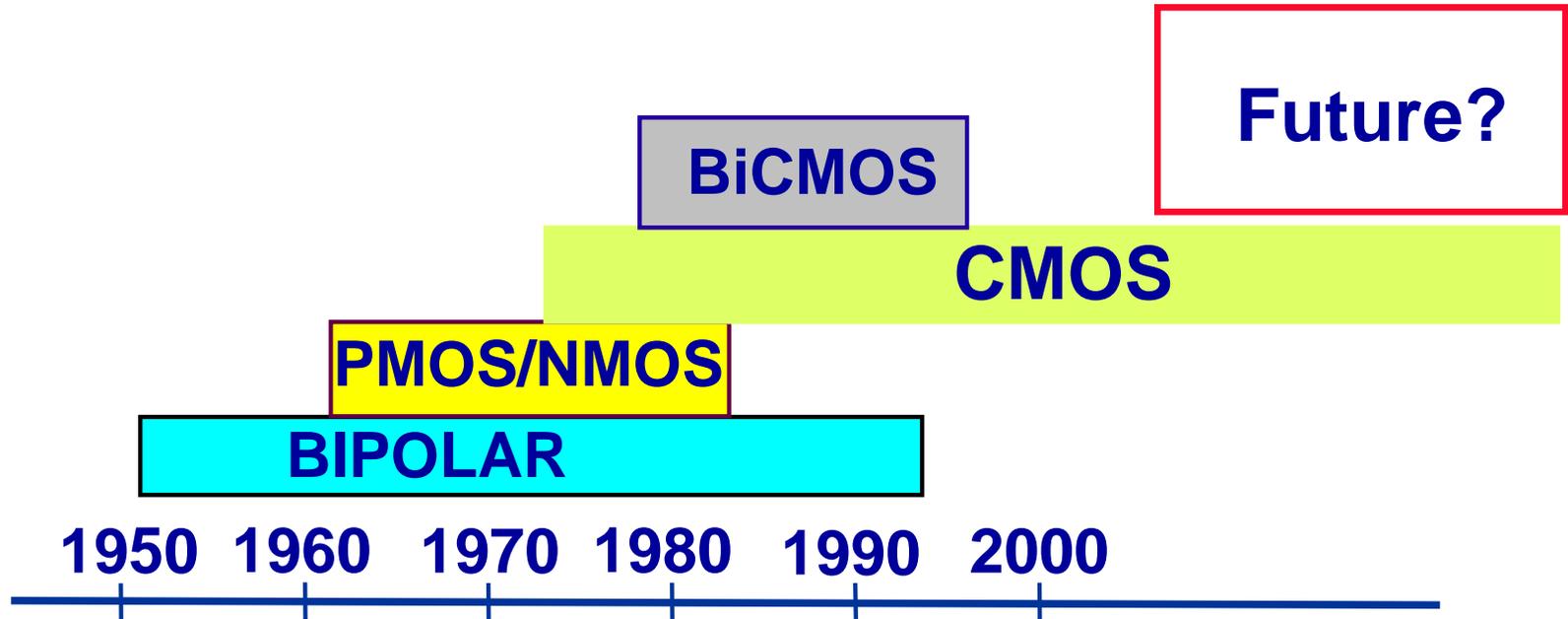
IBM Systems and Technology Group - SRDC

Semiconductor Reliability Topics for Leading Edge CMOS Technologies

Fernando Guarín Ph.D.
Semiconductor Research & Development Center
IBM Systems and Technology Group
2070 Route 52, Hopewell Junction, NY 12533

guarinf@us.ibm.com

Evolution of Integrated-Circuit Technology



First bipolar transistor (1947)

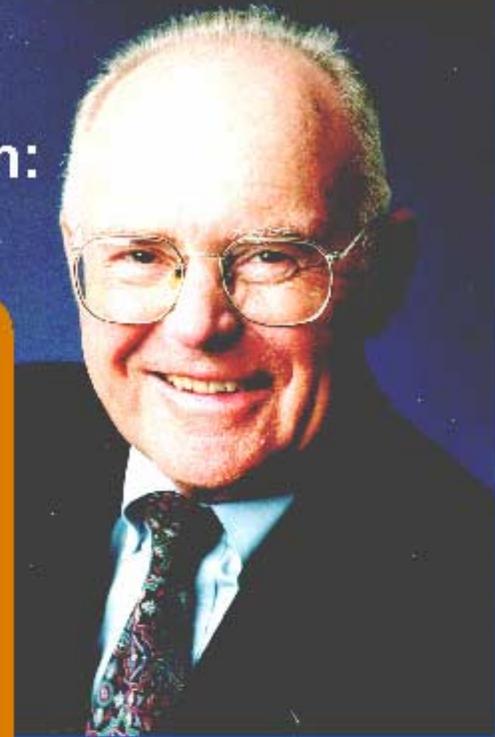
First MOSFET (1960)

CMOS invented (1963)

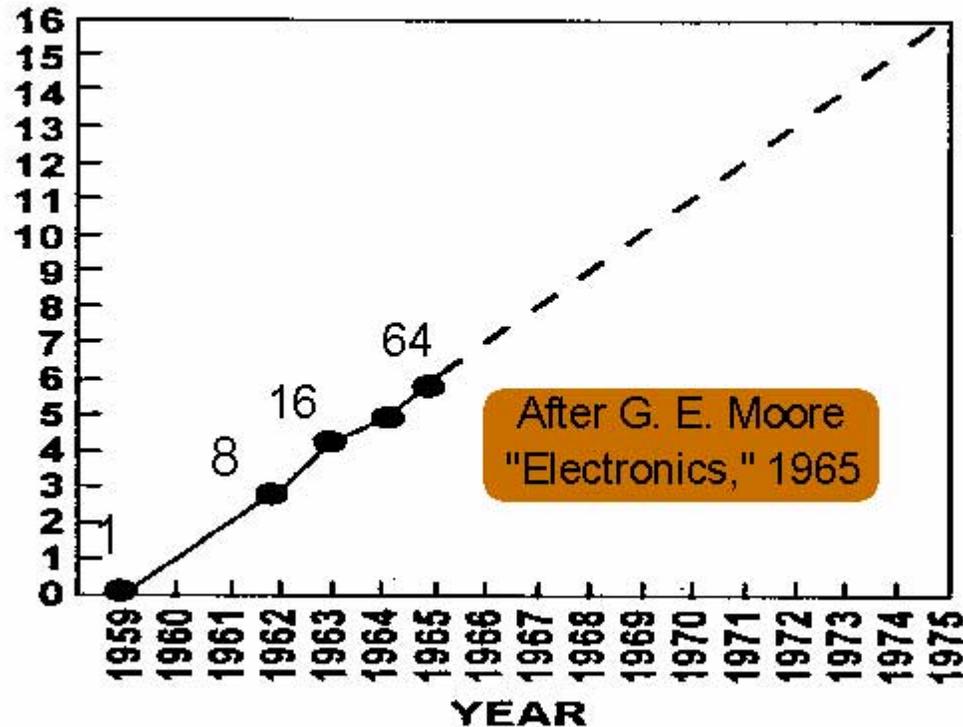
Need a compelling reason to Adopt a new IC Technology



In 1965 Gordon Moore
made an interesting observation:



LOG₂ OF THE NUMBER OF
COMPONENTS PER INTEGRATED FUNCTION



Populist version of Moore's Law:

Any parameter related to semiconductors must form a straight line when plotted on exponential graph paper.



CMOS Scaling

Silicon devices are one of the few manufactured items in the world that perform their function better the smaller you make them.

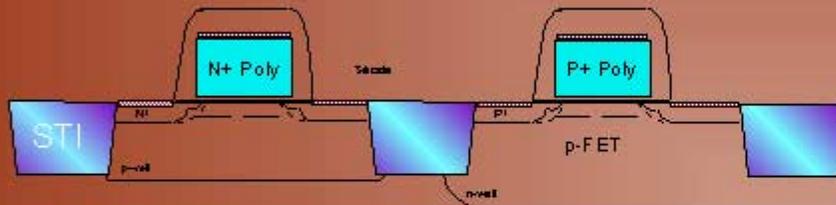


TABLE 1

SCALING RESULTS FOR CIRCUIT PERFORMANCE

Device or Circuit Parameter	Scaling Factor
Device dimension (ax, L, W)	$1/k$
Doping concentration N_d	k
Voltage V	$1/k$
Current I	$1/k$
Capacitance eA/t	$1/k$
Delay time / circuit VC/I	$1/k$
Power Dissipation / circuit VI	$1/k^2$
Power density V/A	1



Dennard, Gaensslen, Yu - 1974

High-Field Effects Limit CMOS Scaling

- Ideally, scaling is to be done at constant field
- In practice, scaling has been achieved by increasing electric fields
- High gate oxide field → large gate tunneling current
 - Solution: High-k gate insulator
 - New reliability challenges PBTI

- Moore's "Law" predicted ONLY continuous size decrease
- Dennard's Scaling theory predicted the performance increases associated with smaller device size
- Moore provide the path and Dennard provided the way to implement it
- The end is near has been a recurring theme in the semiconductor industry for many years.
- First paper predicting the end was published by RCA over 30 years ago
- We are still predicting the end
 - Tunneling Current Increase at larger fields
 - High K to the rescue
- what is next?
 - 3D, Fin Fets, End of Silicon Scaling
- Atoms don't scale !!



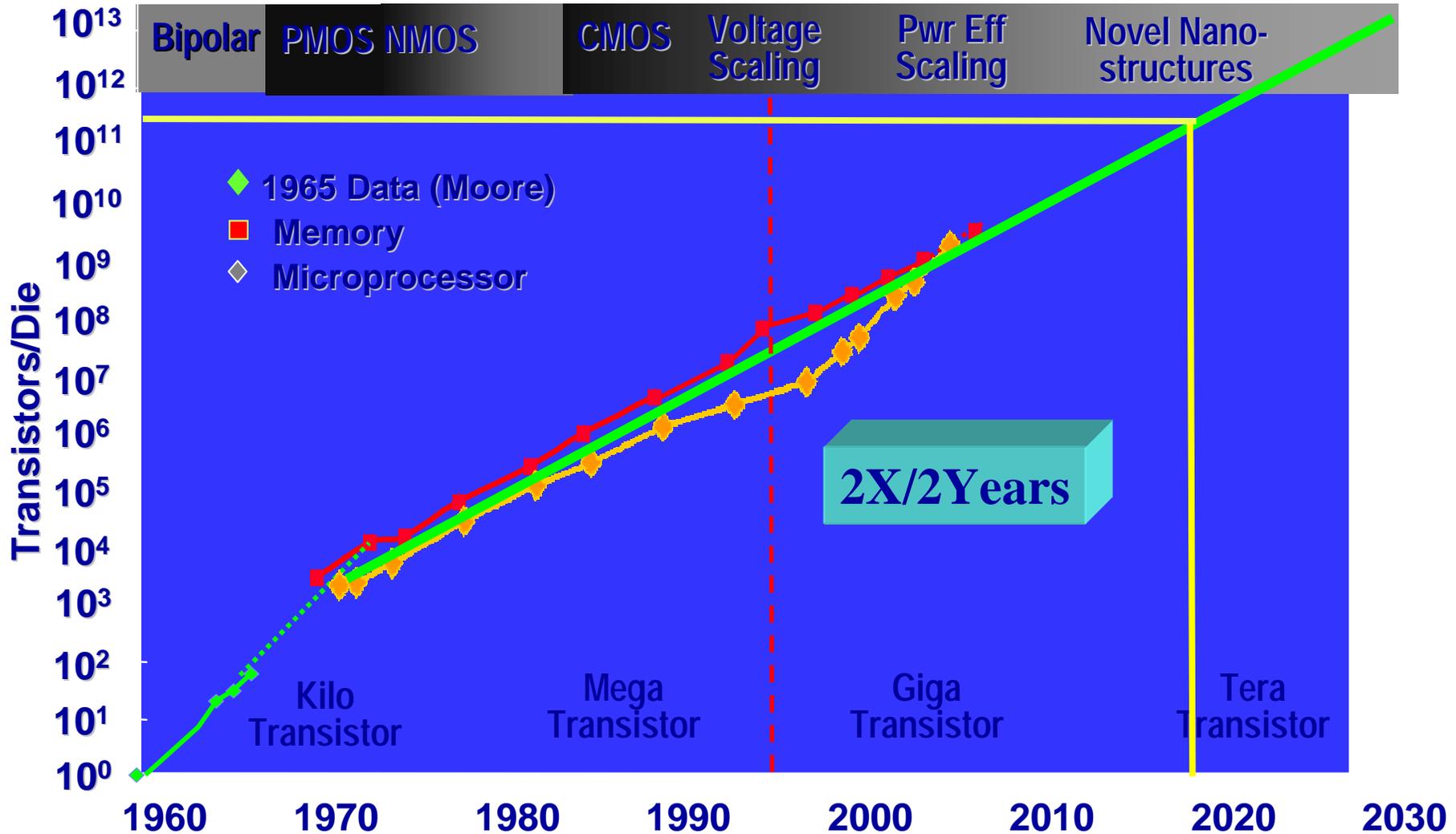
Moore's "Law" Will Continue in the near future

10µm

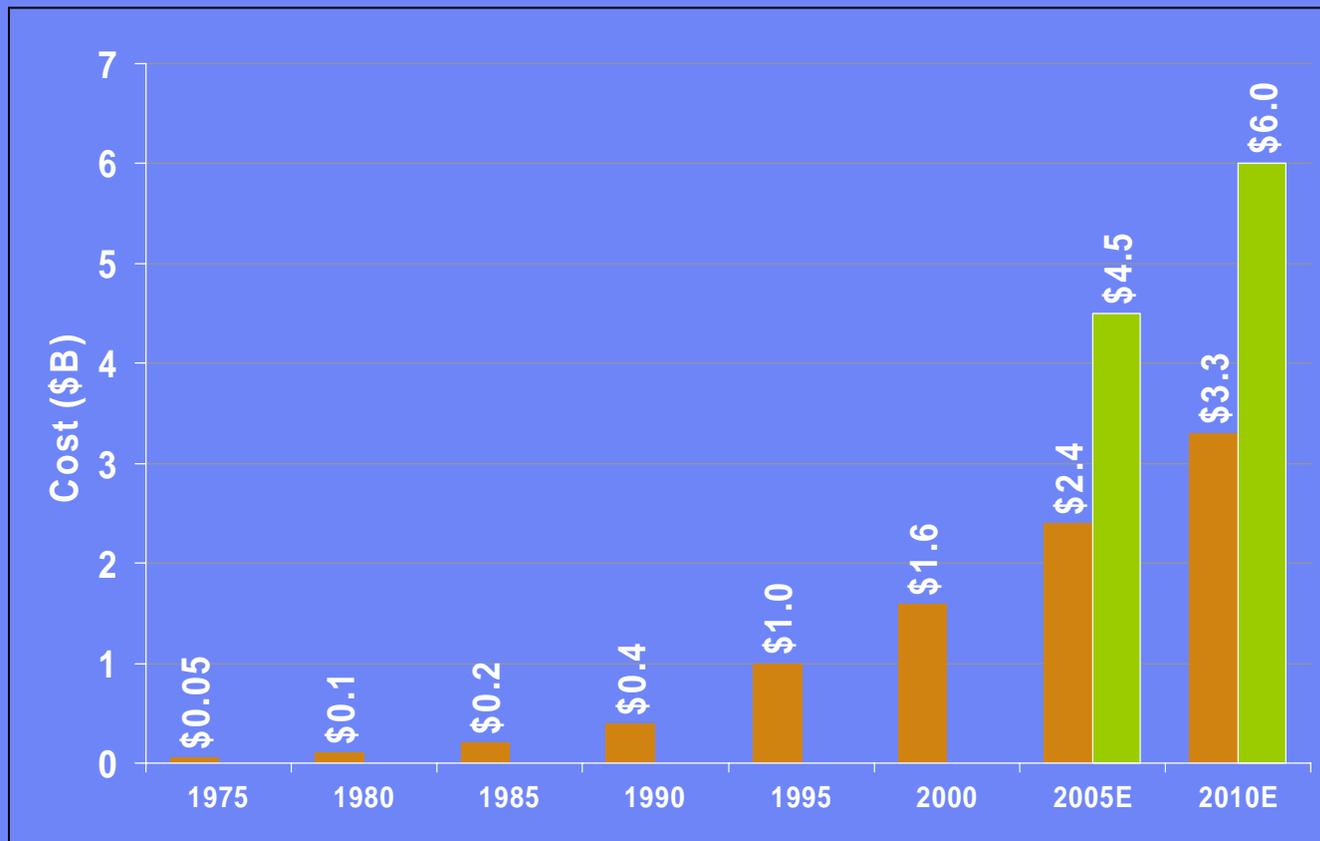
1µm

100nm

10nm



Wafer Fab Cost Trend

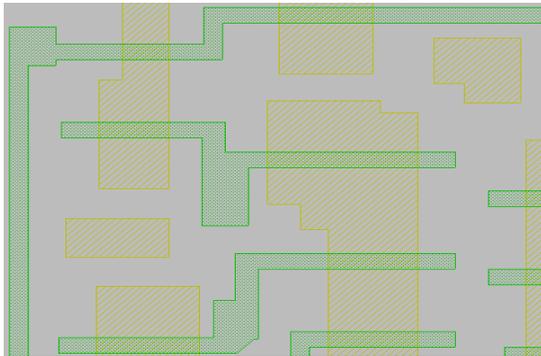


•A dramatic rise in capitalization occurs as technology no longer supports “trivial” linear shrinks, requiring significant innovation in technology and supporting tooling

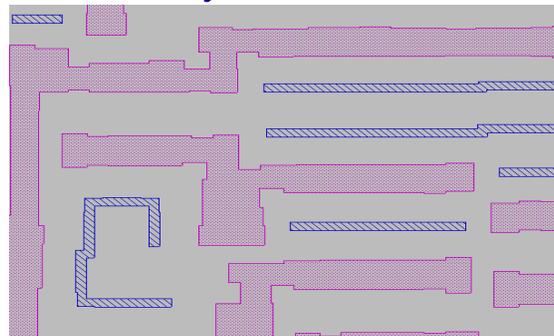
Lithography Challenges

- Scaling now requires you to do the improbable, such as imaging features $1/5^{\text{th}}$ the wavelength of the light you are using to create the image
 - Images are now created by interference, not classical shadowing
 - Computing & checking features required to achieve a given design takes roughly 4000% the time it took in the last technology generation.

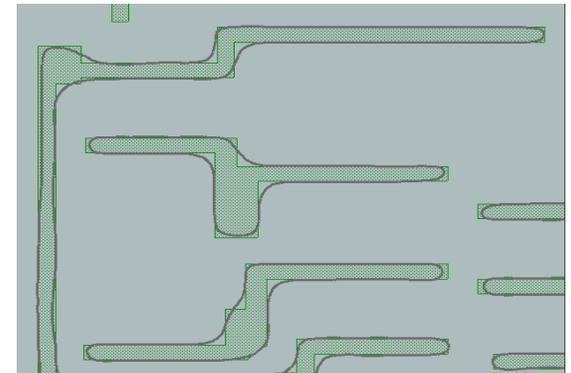
The image you want to print



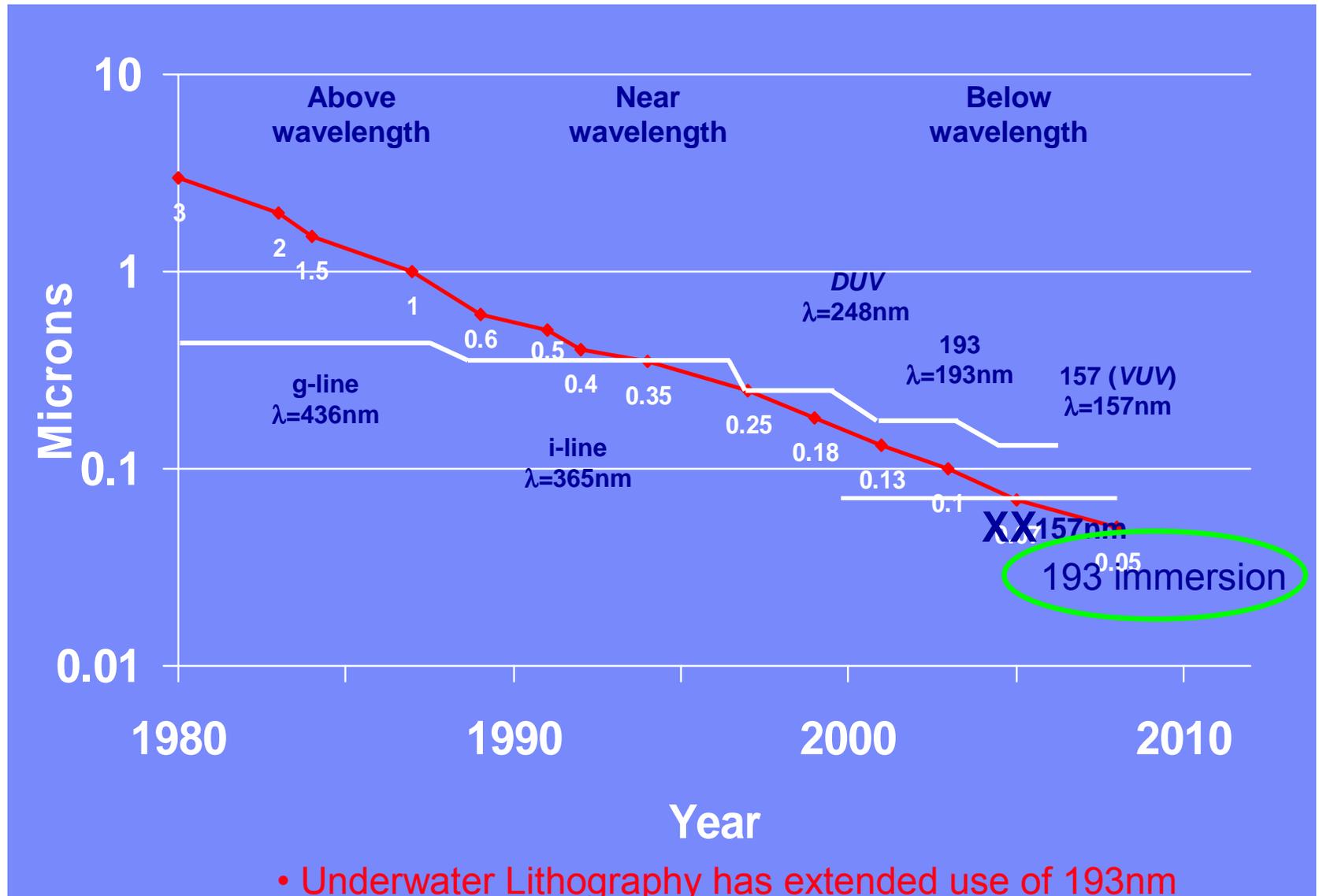
The Mask you shine light through to get what you wanted



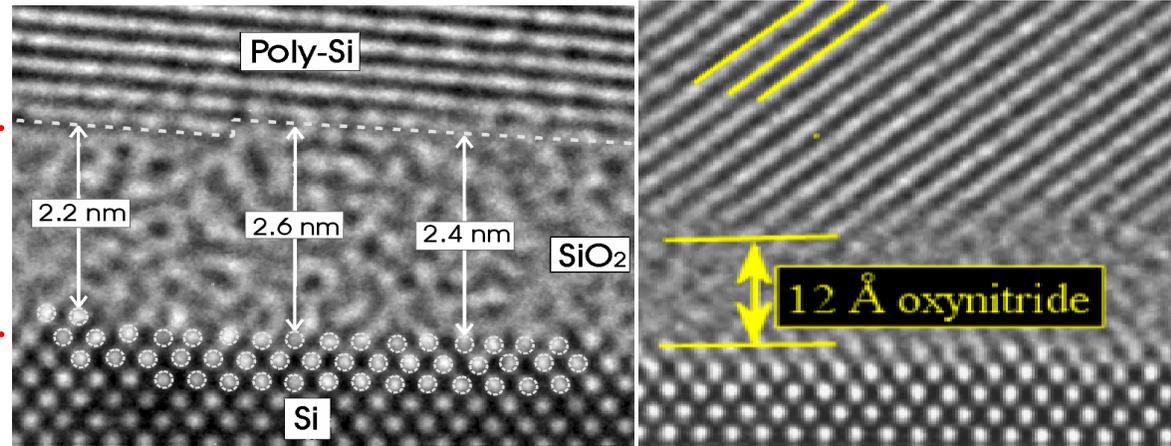
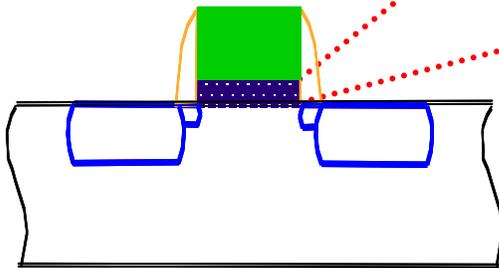
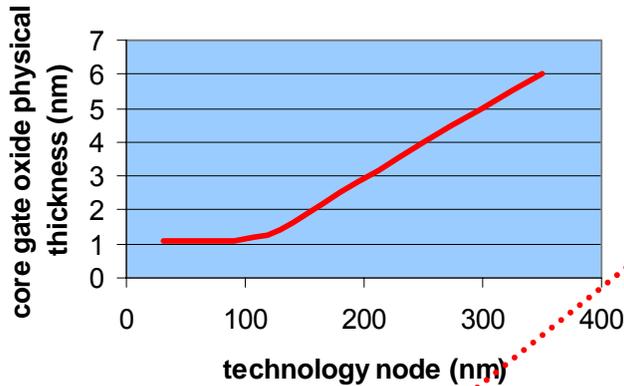
The image you get if everything works perfectly



Lithography Progress



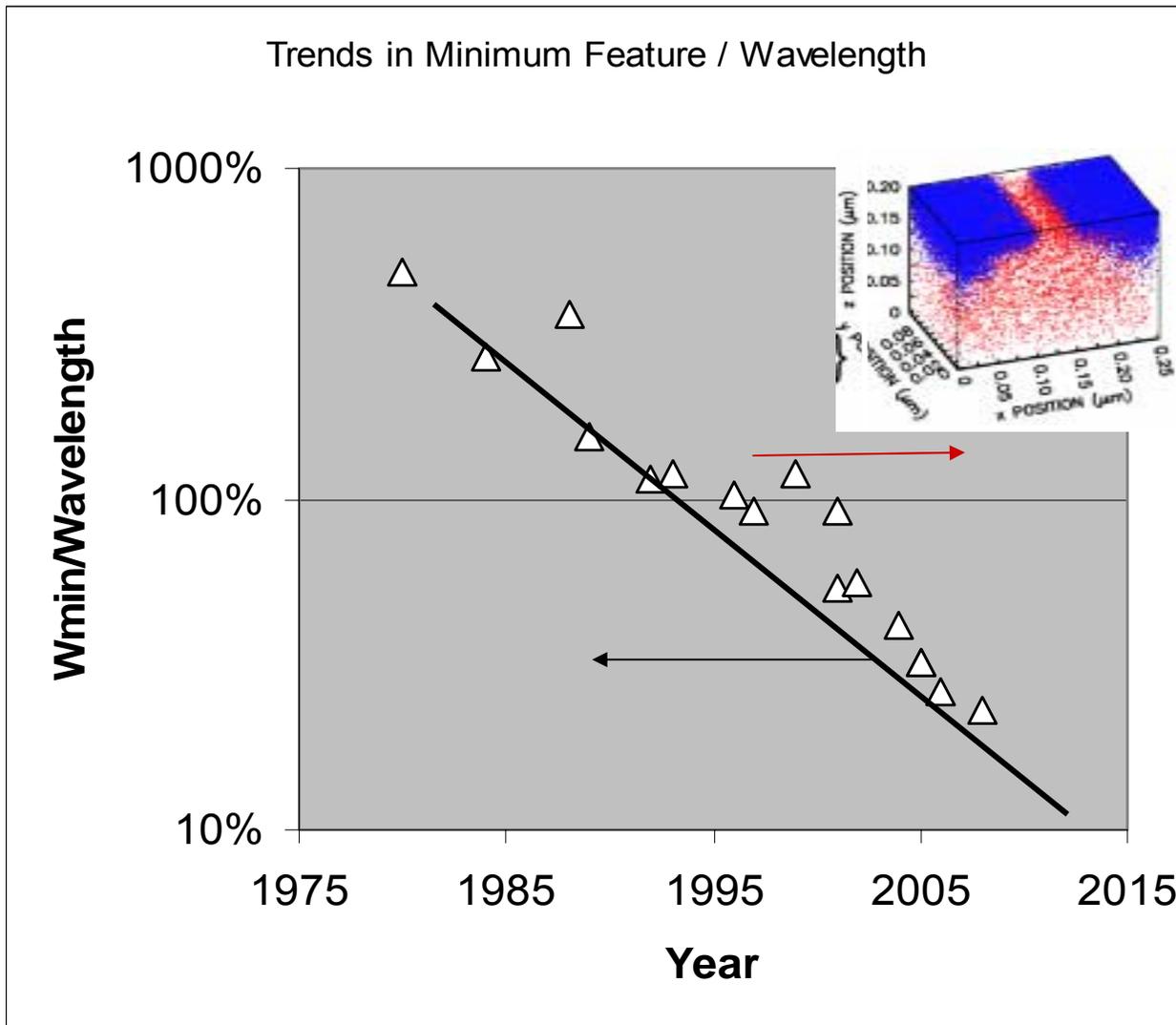
Limitations of Classical Scaling



- We have entered the regime of non-statistical material and device behaviors

→ RELIABILITY IMPLICATIONS

- Oxide thickness is now measured in terms of atomic layers.
 - Atomic level fluctuations have dramatic consequences
- Dopant concentration fluctuations on the atomic scale are now meaningful
 - Devices dimensions are so small as to preclude the use of statistics



Atoms Controlling min device VT

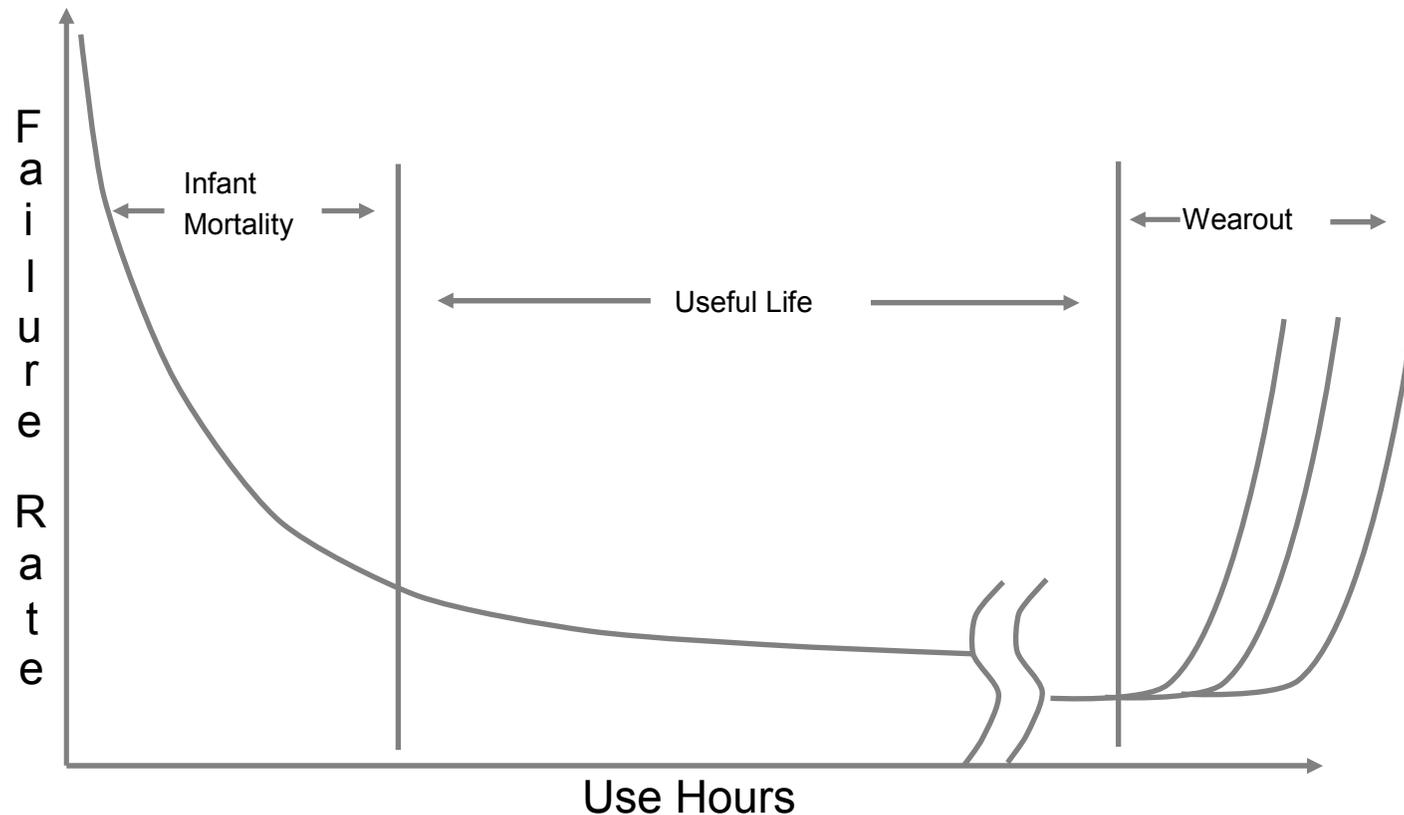
Tolerances of 8 atoms and 8% of light wavelength!

Tim Brunner, Dave Frank – IBM Research

Reliability Topics

Qualification: Reliability Management

- Technology Qualification must cover all reliability aspects removing systematic and managing random defects



Improper Use of “Cumulative Fail”

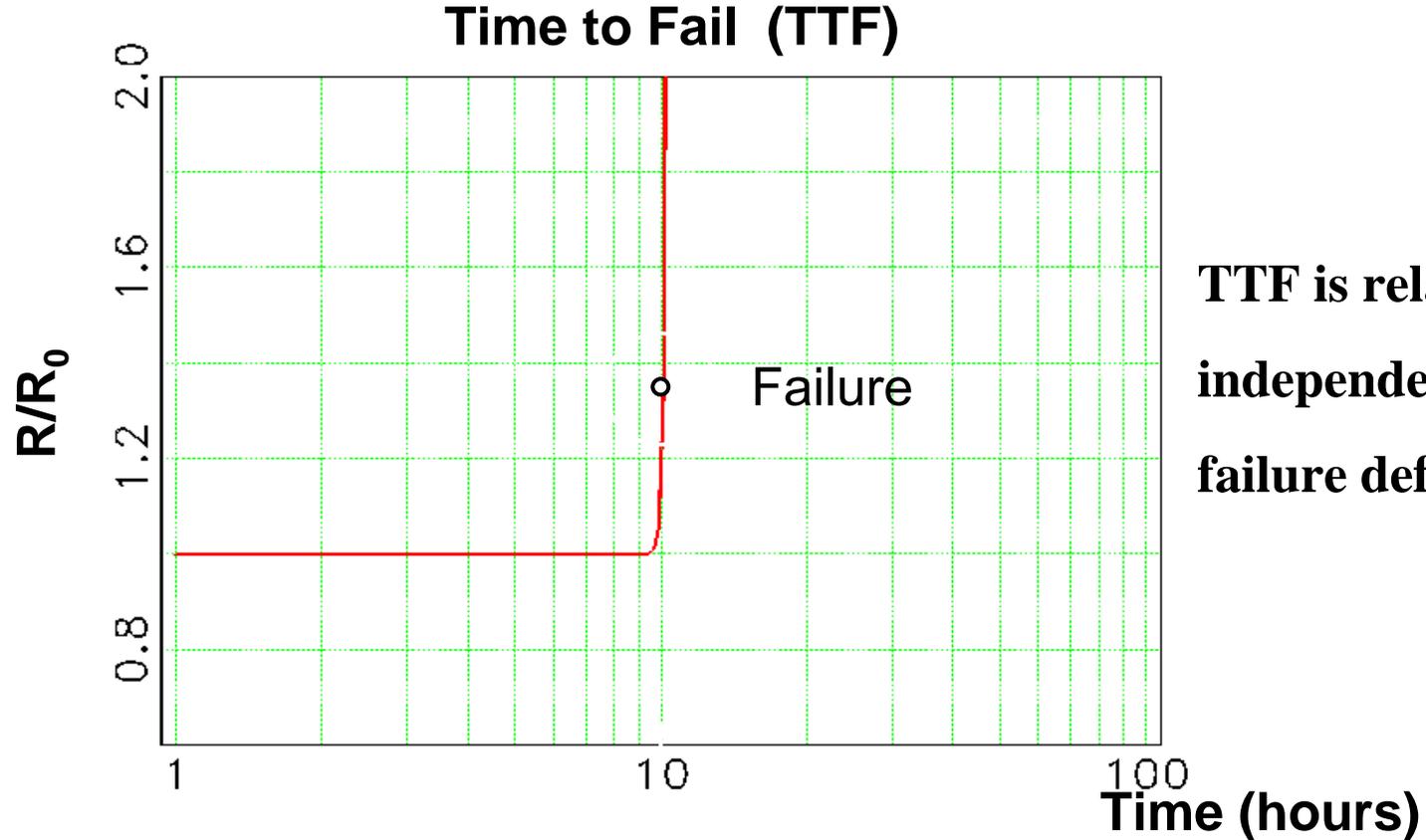
- **First of all, let’s dispense with the improper application of the cum fail concept to parameter drift mechanisms, such as NBTI (HCI)**
- **There seems to be considerable misconception in the industry regarding this point, as for example, this qualification criterion:**

“10% Idsat degradation, 0.1% cum fail”

Implying that the measured statistics of the time for degradation to reach 10% is a valid predictor of product lifetime!

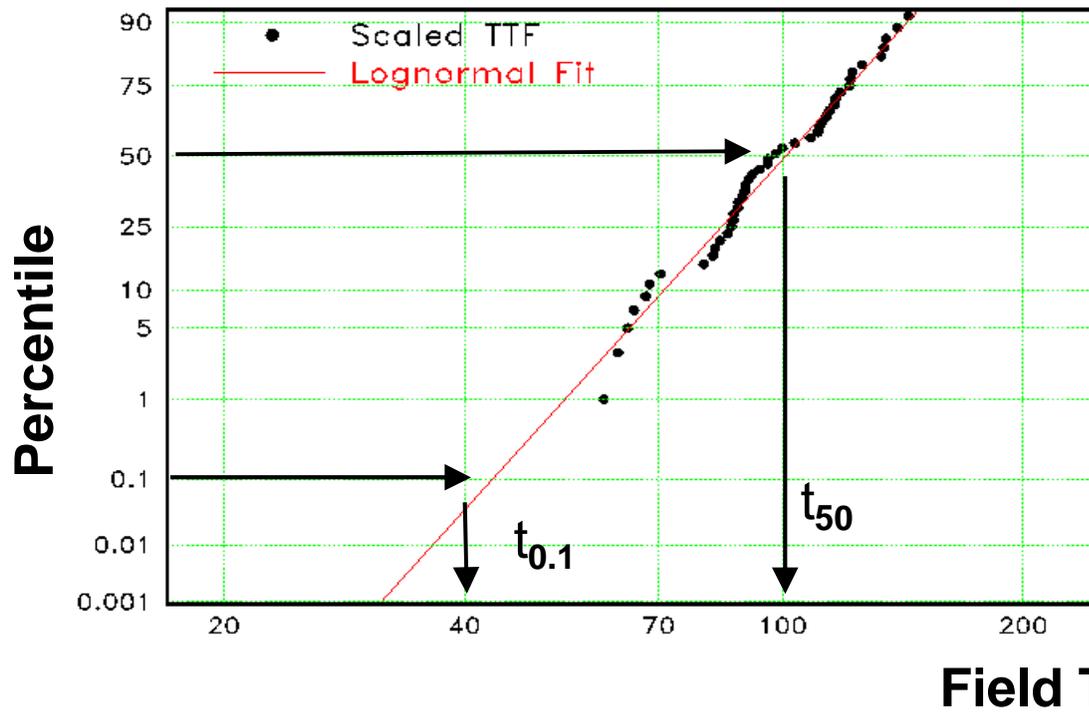
Courtesy S. Rauch

Hard Fail Mechanism (eg. Electromigration)



- Clearly defined TTF that correlates well to product failure.
- Circuit sensitivity or initial resistance does not make any significant difference to the product fail time.

Time to fail Distribution (Electromigration)



$$t_{50} \sim 100 \text{ yr}$$

$$t_{0.1} \sim 45 \text{ yr}$$

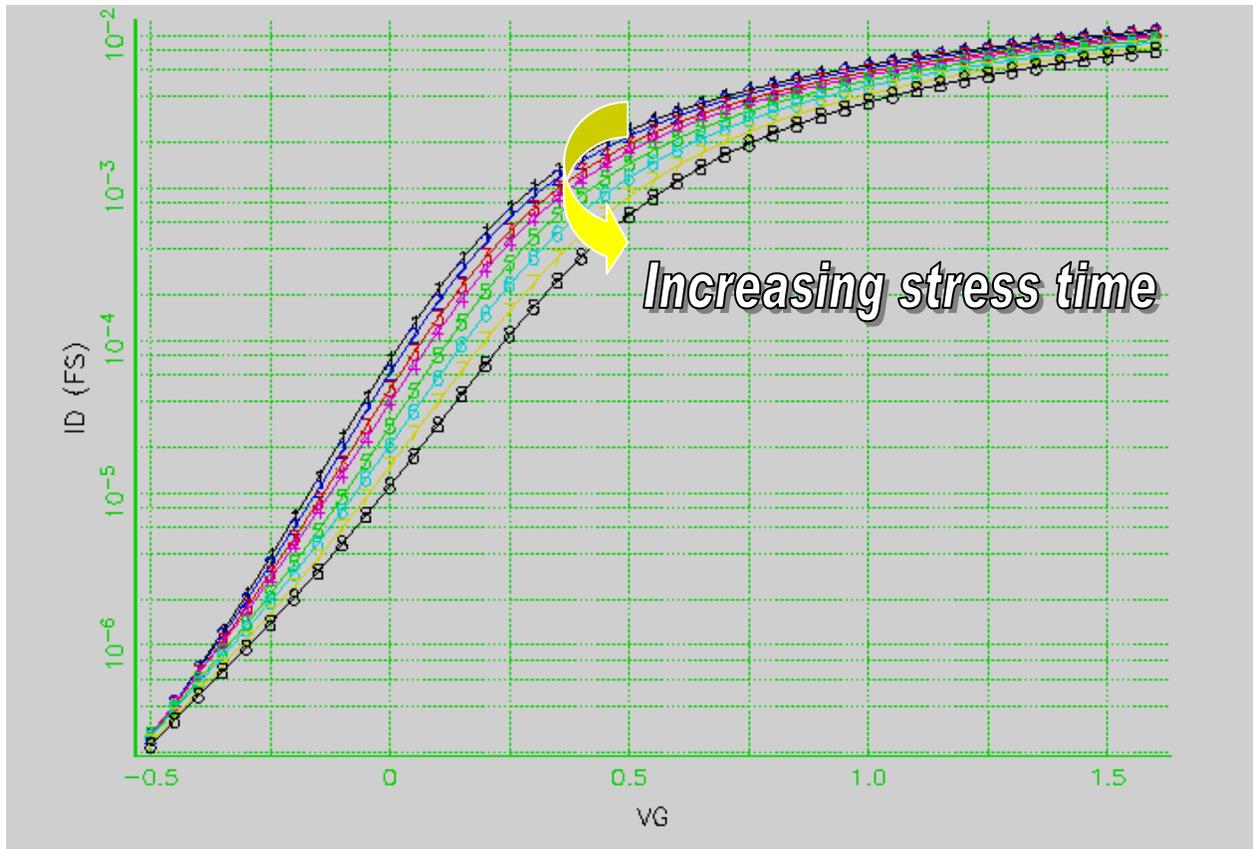
$$t(10^{-11}) \sim 18 \text{ yr}$$

This is useful for predicting product life because:

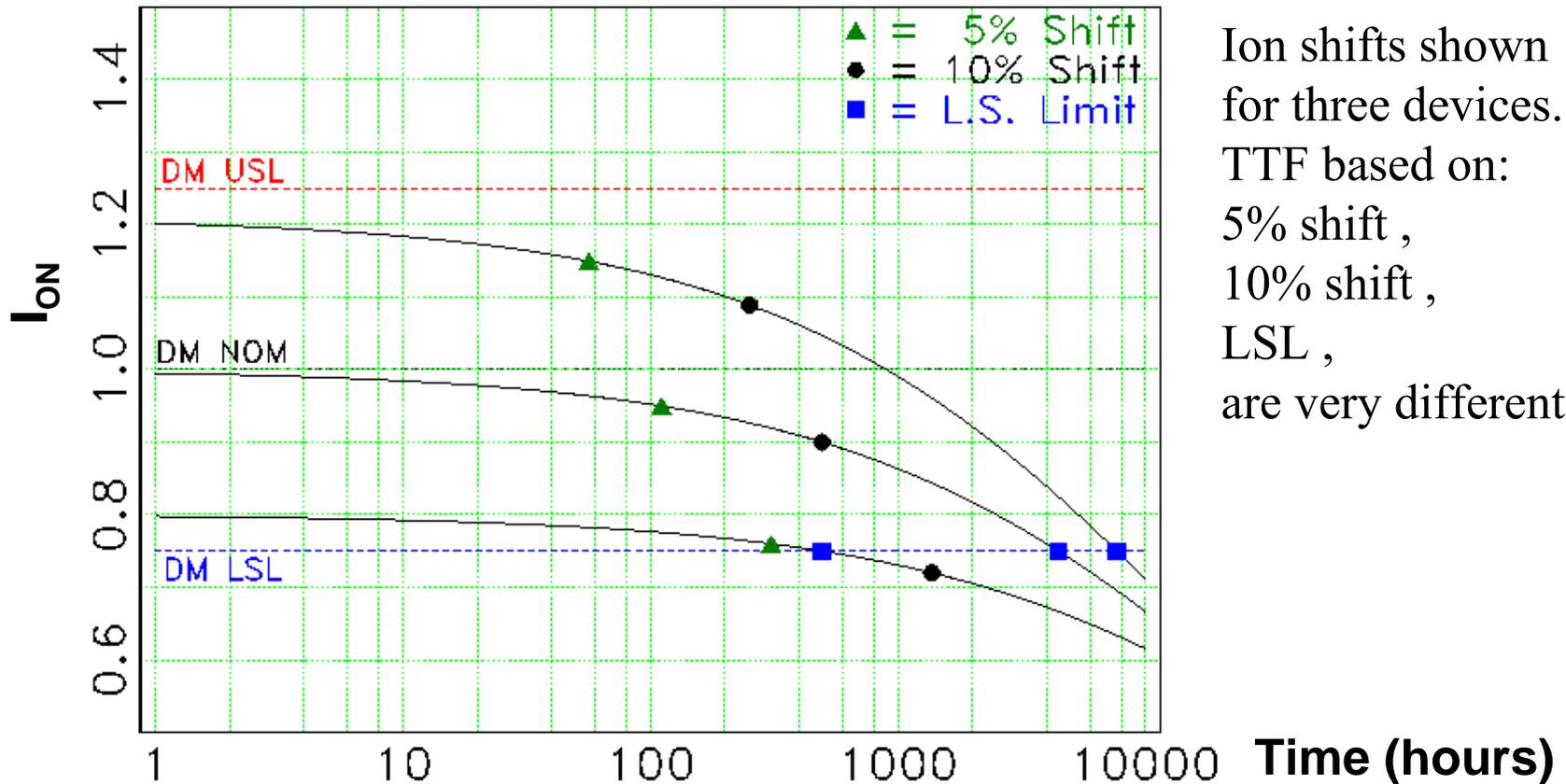
- Product fail statistics are determined by the element fail statistics (metal line in this case.)
- Product Test margin has virtually no effect on product TTF.

Gradual Shift Mechanism (Hot Carriers, BTI)

Ex. I_D V_G for NMOS



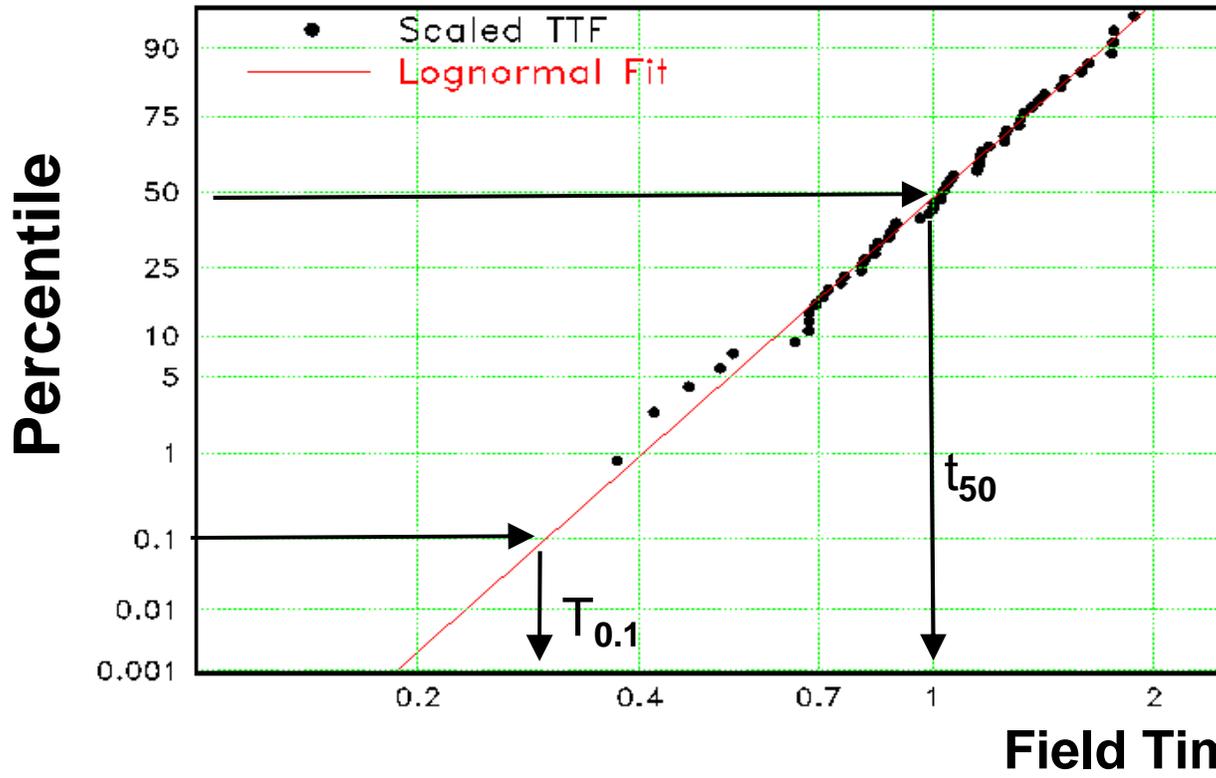
Gradual Shift Mechanism (eg. Hot Carriers)



Ion shifts shown for three devices. TTF based on: 5% shift, 10% shift, LSL, are very different.

- There is no clear and unambiguous TTF that correlates directly with product failure.
- Circuit sensitivity determines the final parameter value that causes failure.
- The amount of shift that causes a product fail depends on the initial parameter value.

Hot Carrier Time to Fail Distribution



Based on
10% shift:

$$t_{50} \sim 1 \text{ yr}$$

$$t_{0.1} \sim 0.33 \text{ yr}$$

This is not useful for predicting product life because:

- Product fail statistics are dominated by product performance distribution.
- Product Test margin has large effect on product Time to Fail.
- Ascribing a fail point to a certain value of parameter shift is obviously not useful.

Modeling Realistic Impact to Product

- Model device shifts as a function of device parameters (L_{poly} , V_T , t_{OX} , etc.) and environment (V_{DD} , T_J)
- Model circuit sensitivity (performance shift vs. device shift.)
- Model product performance before and after degradations, taking into account testing strategy. All device shifts must be considered together (NBTI + HC.)
- A robust strategy for model validation must be in place

Channel Hot Carriers

Channel carriers whose temperature (T_{hc}) is locally (near drain) larger than the lattice temperature.

T_{hc} is mainly determined by:

Large lateral Electric Field in the pinch-off region.

Energy exchange processes (e.g., phonon scattering, impact ionization, e-e scattering, etc.)

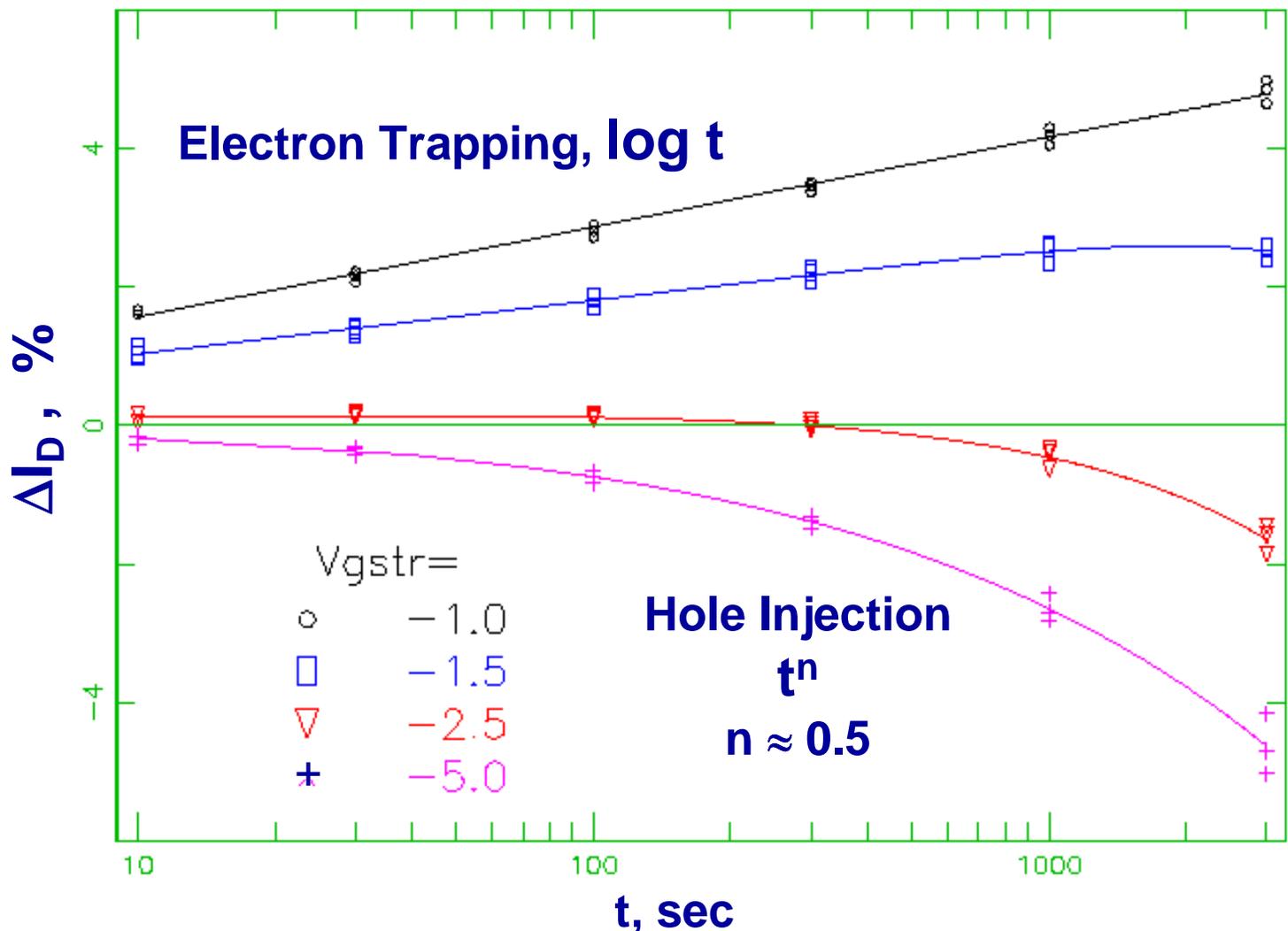
Bad Effects of Channel Hot Carriers

- Carriers with Energy $> E_G$ cause *Impact Ionization* (electron-hole pair generation)
 - Latchup
- Carriers with high Energy impacting the SiO_2 interface cause *Interface State Generation*
 - I_{ON} , V_T degradation
- Carriers injected into the SiO_2 lead to *Charge Trapping*
 - V_T shift

SLOWER CIRCUIT OPERATION

Older Technology PFET HC Behavior

3.3V Appl.: $L_{eff}=0.26 \mu\text{m}$, $t_{ox} = 6.8 \text{ nm}$, $V_{dstr} = -5.0 \text{ V}$

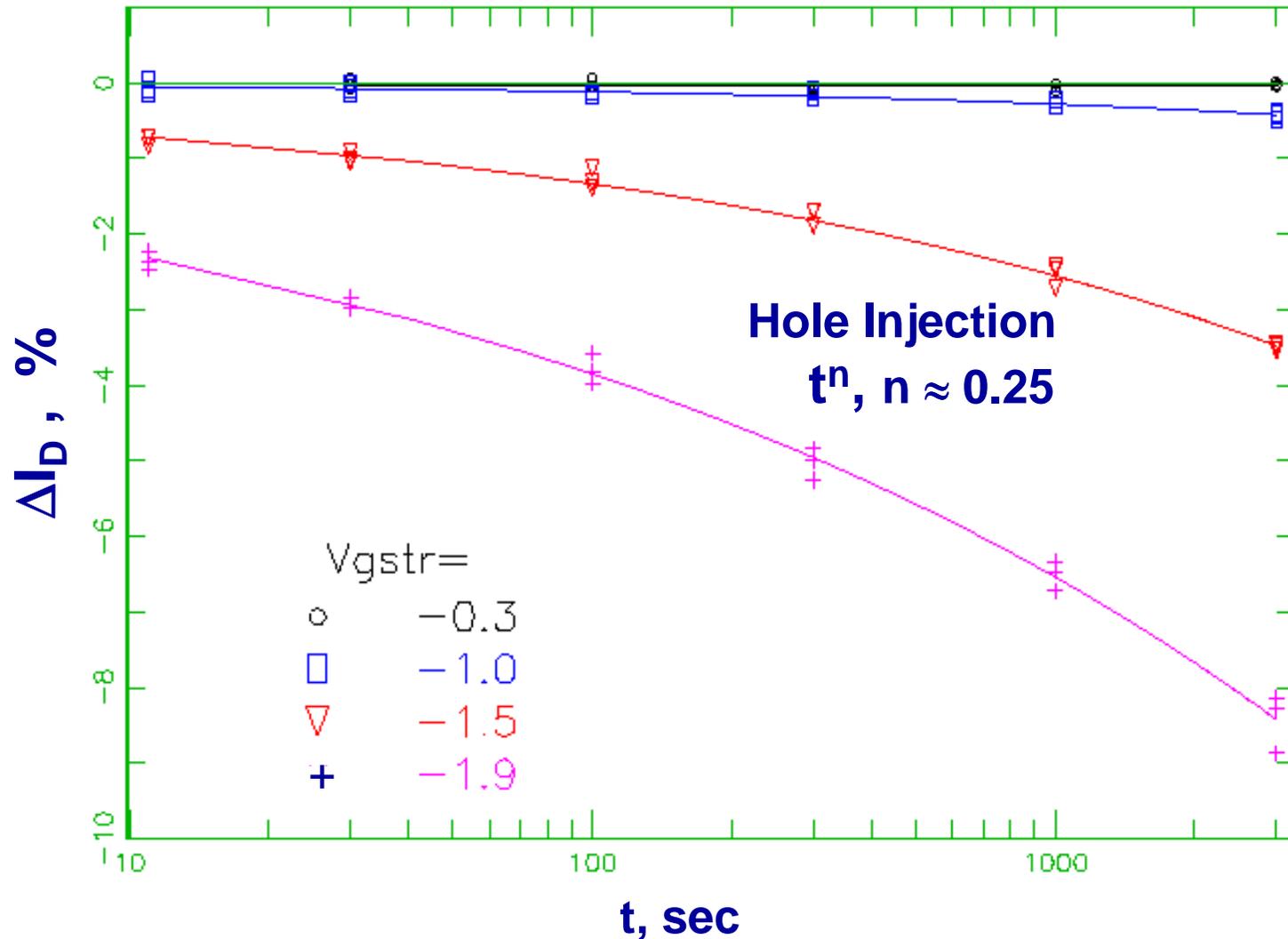


Hole Injection and electron trapping of roughly equal strength.

Note impact of VG stress

Newer Technology PFET HC Behavior

1.2 V Appl.: $L_{\text{eff}}=0.07 \mu\text{m}$, $t_{\text{ox}} = 1.6 \text{ nm}$, $V_{\text{dstr}} = -1.9 \text{ V}$



Hole Injection completely dominant.

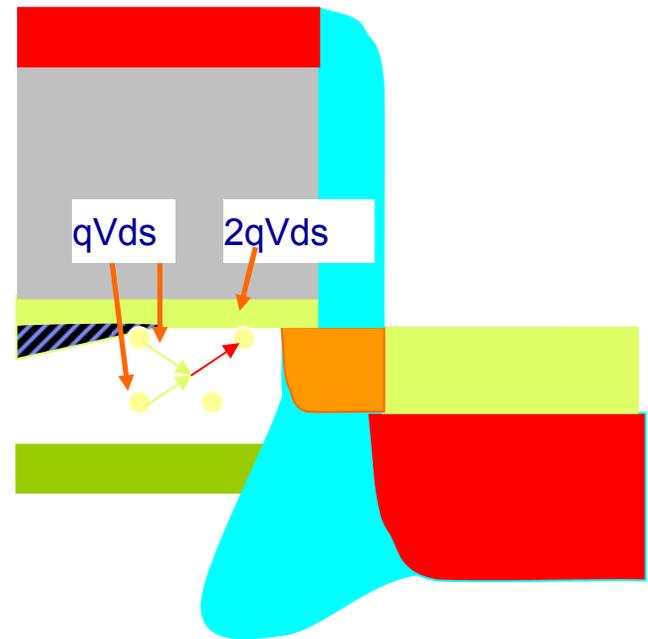
E-E SCATTERING PHENOMENA

In very short channel devices electrons can gain the supply energy qV_{ds} by travelling in the pinchoff region quasi ballistically.

E-E scattering

There is a small probability that two high energy electrons undergo a scattering process so that one electron gains most of the total energy leading to a small electron population of carriers up to about twice qV_{ds}

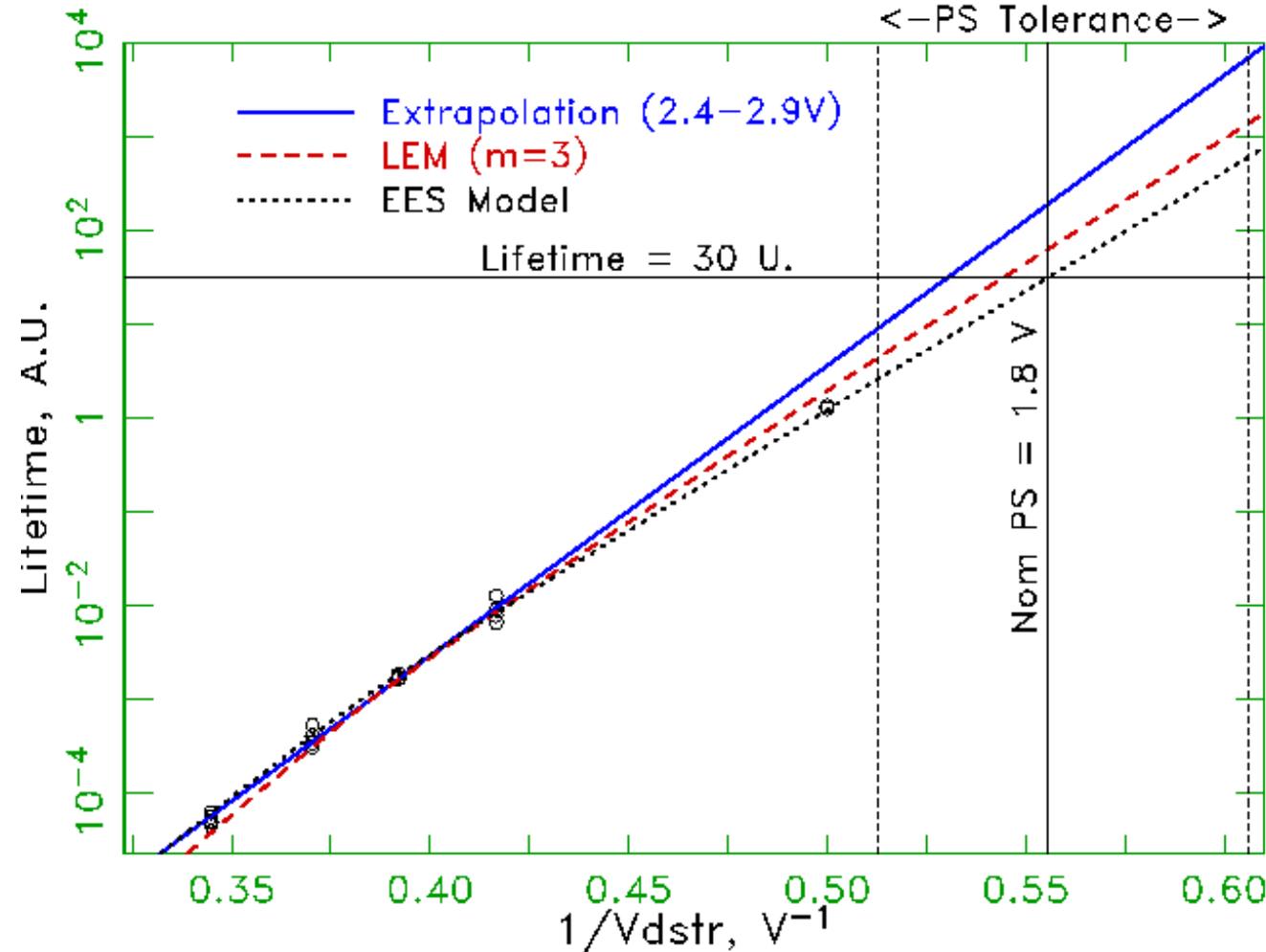
Scattering rate is proportional to n^2
or $\approx I_s^2$



EE scattering event

Non Linearity due to simplification of actual LEM and e-e Scattering

Lifetime vs $1/V_{dstr}$



Linear Extrapolation
 $\tau = 200 \text{ U.}$

Actual LEM
 $\tau = 60 \text{ U.}$
 (Isx not exponential vs $1/V$)

Electron to Electron Scattering
 $\tau = 30 \text{ U.}$

Supported by data collected at voltage closer to use condition

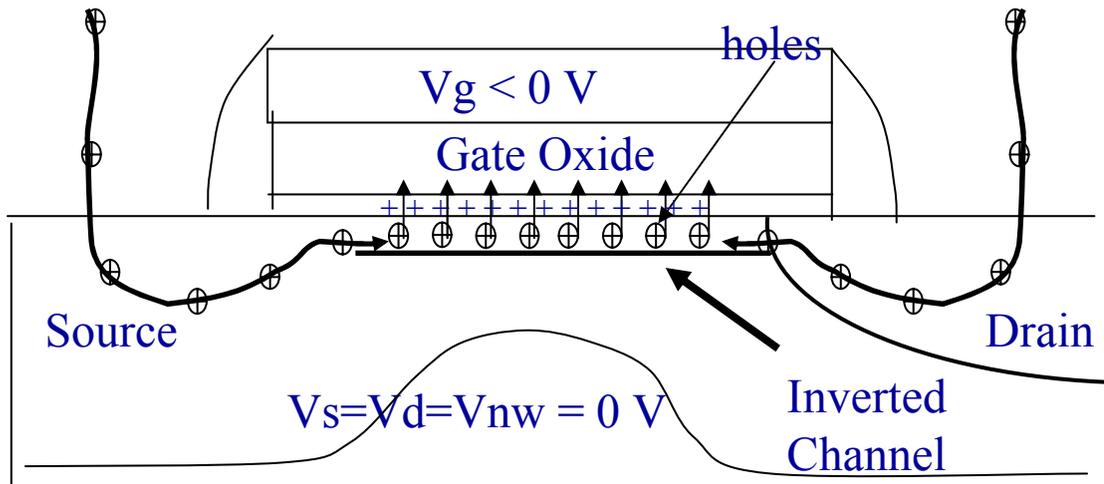
Linear extrapolation and LEM will yield overly optimistic predictions

Bias Temperature Instability NBTI - PBTI

What is NBTI?

PMOSFET wearout mechanism resulting in positive charge buildup – oxide bulk charge and donor type interface states at the SiO₂/Si interface under the influence of applied negative gate voltage

Damage is related to “cold holes” – no drain bias needed.



NBTI damage is

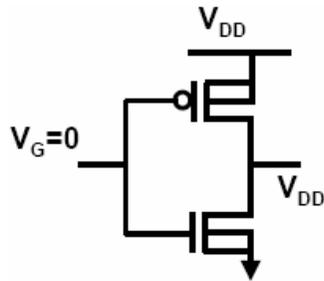
- Not associated with channel carrier transport
- Not related to tunneling gate current in thin oxides

Mechanism typically investigated in a capacitor configuration with channel inverted (symmetric damage)

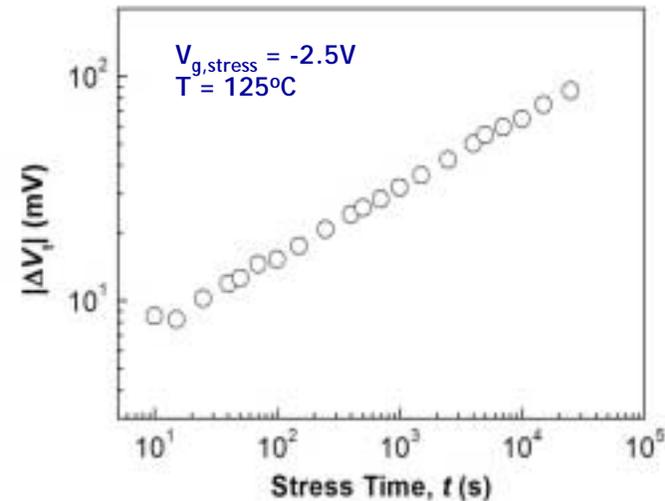
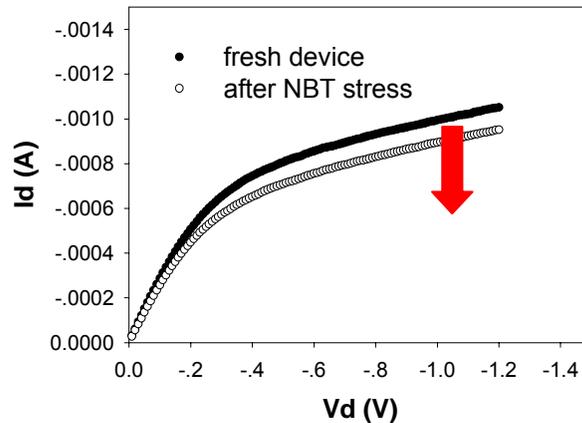
First reported by Deal in 1967. [B. Deal et al., *J. Electrochem. Soc.*, **114**, 266 (1967).]

Brief overview of NBTI

NBTI (Negative Bias Temperature Instability) refers to the degradation mechanism of p -MOSFETs when the device is stressed with negative gate bias at elevated temperature

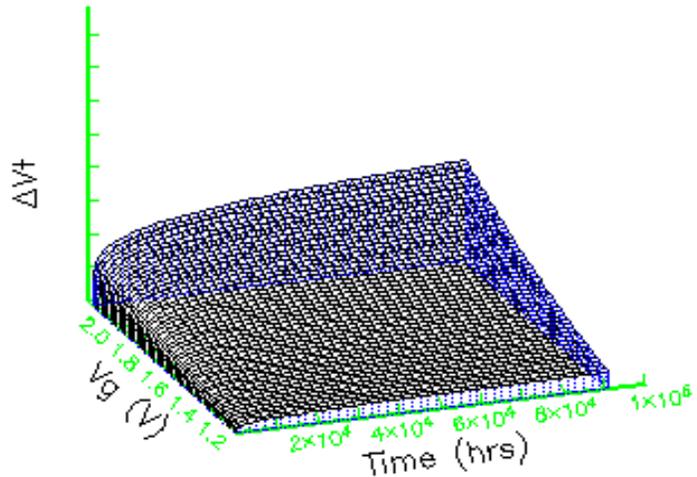


- Negative Bias: Gate is negative to Source, Drain & Bulk.
- Temperature: NBTI is enhanced by high temperature
- Instability: Device parameters (V_t , G_m , $I_{d,sat}$, $I_{d,lin}$, etc.) shift with stress time

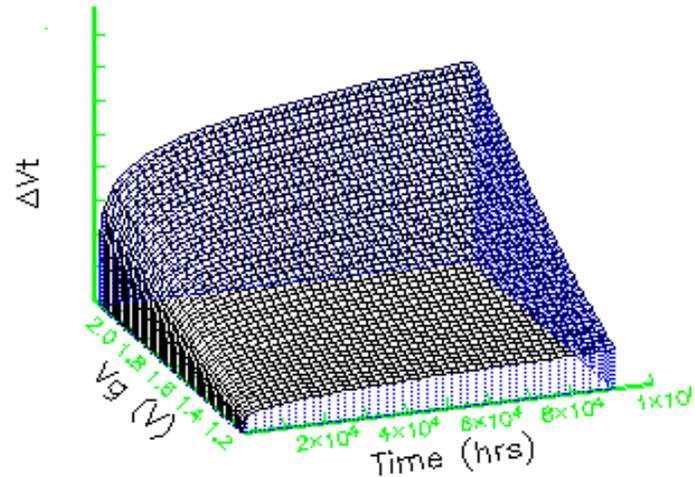


NBTI-Induced V_t Shifts

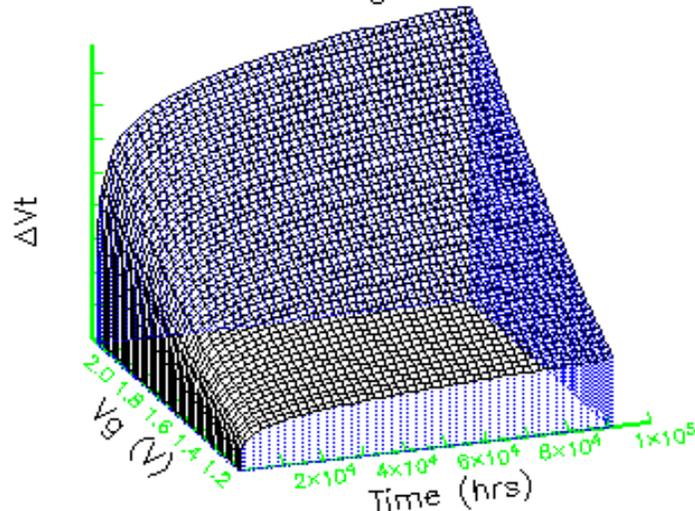
$T=40^\circ\text{C}$ $V_{t0}=0.3$ V



$T=85^\circ\text{C}$ $V_{t0}=0.2595$ V



$T=125^\circ\text{C}$ $V_{t0}=0.2235$ V



Temperature and V_g significantly influence the level of V_t shift

NBTI Degradation Model – Mean Shift

- Typical NBTI Models:

$$1. \Delta V_T = A \left(\frac{|V_{GS}|}{T_{OX}} \right)^m \exp \left(-\frac{\Delta H}{kT} \right) t^n$$

“Power Law”

$$2. \Delta V_T = A \exp \left(a \frac{|V_{GS}|}{T_{OX}} \right) \exp \left(-\frac{\Delta H}{kT} \right) t^n$$

“Exponential Law”

These equations describe the average or mean shift observed.

Often A is higher for narrow PFETs, such as in an SRAM cell.

Typical values: $m \sim 3-4$, $\Delta H \sim 0.1-0.2$ eV, $n \sim 0.15-0.3$

NBTI Implications

NBTI's increased relevance due to:

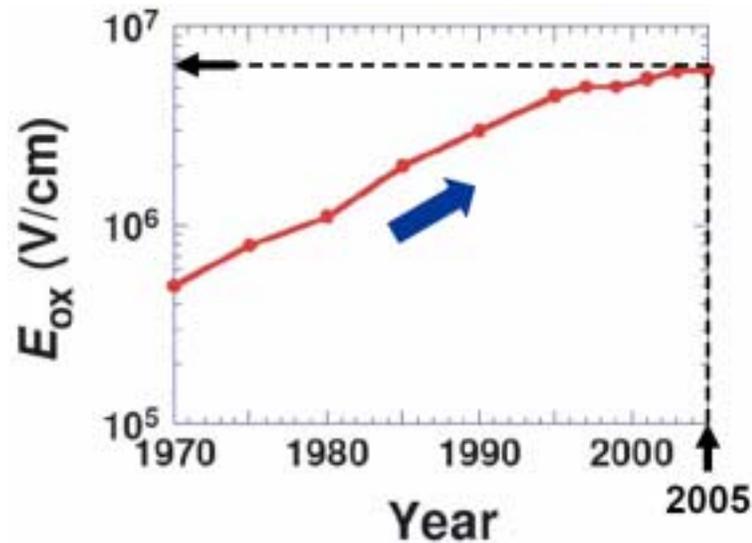
Can not be reduced by increasing Channel Length as in Hot Carrier

Gate electric field has increased as a result of transistor scaling.

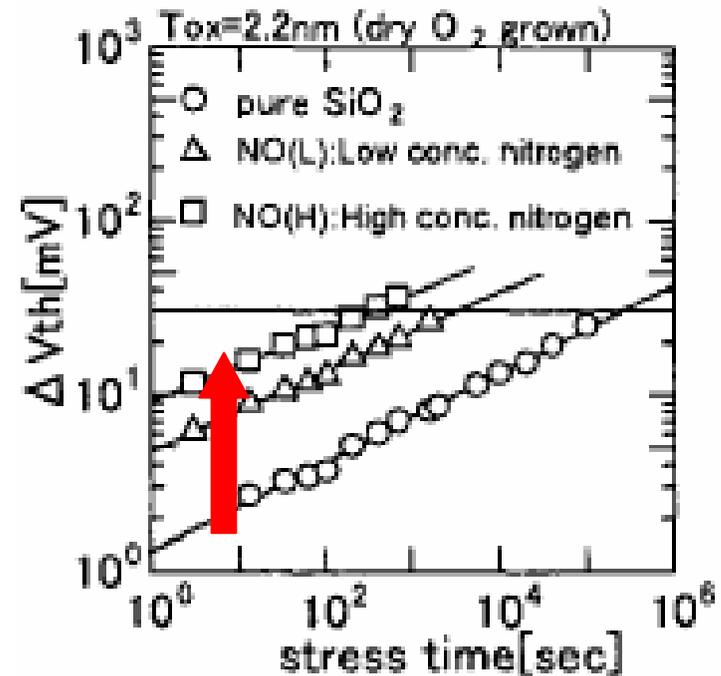
Increased **Nitrogen** levels added into gate oxide.

Chip operating temperature has increased.

Surface p-channel MOSFETs has replaced buried p-channel MOSFETs.



Schroder JAP 2004

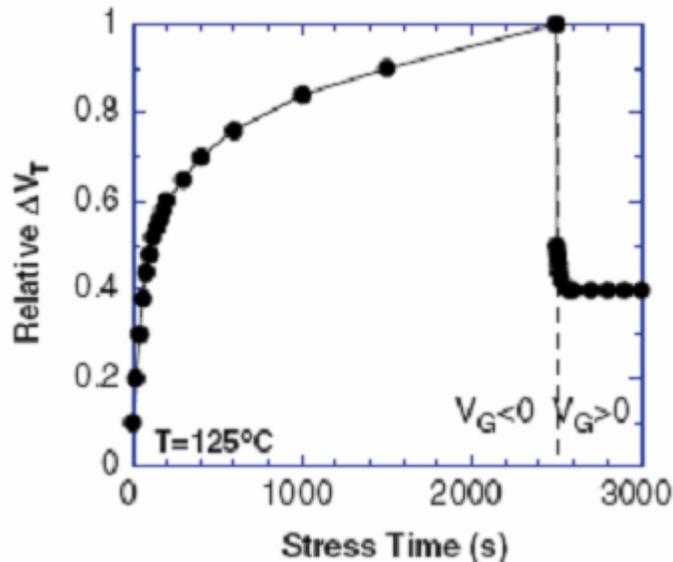


Kimizuka *et al.*, 2000

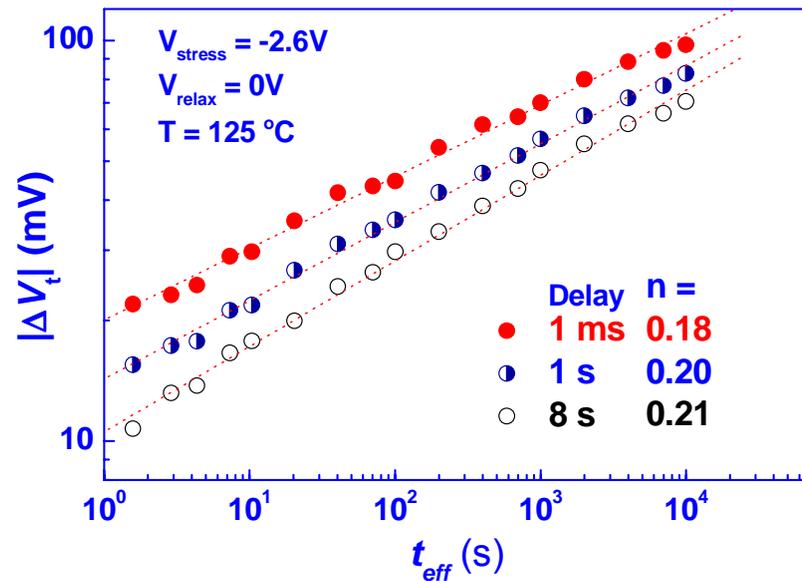
NBTI Recovery – (Measurement Implications)

NBTI measurement results are sensitive to the measurement time, due to the recovery during the measurement.

- Relaxation occurs during the measurement delay
- Power law time-dependence factor is affected

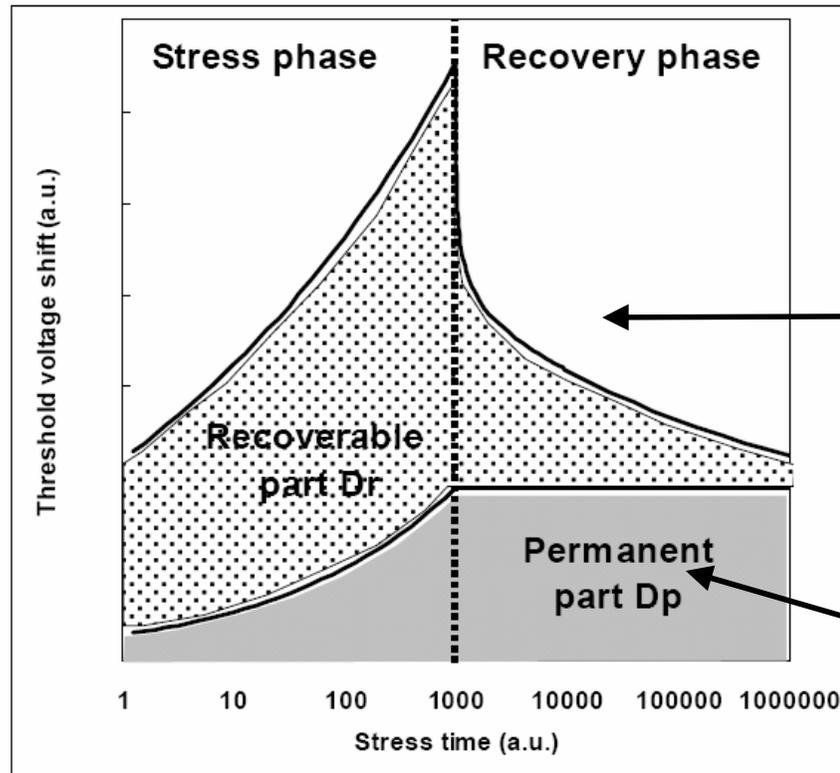


Recovery effect after releasing NBTI stress



ΔV_t recovery due to measurement delay.

NBTI Recovery



2 Components:

-Hole trapping and detrapping
(Recoverable)

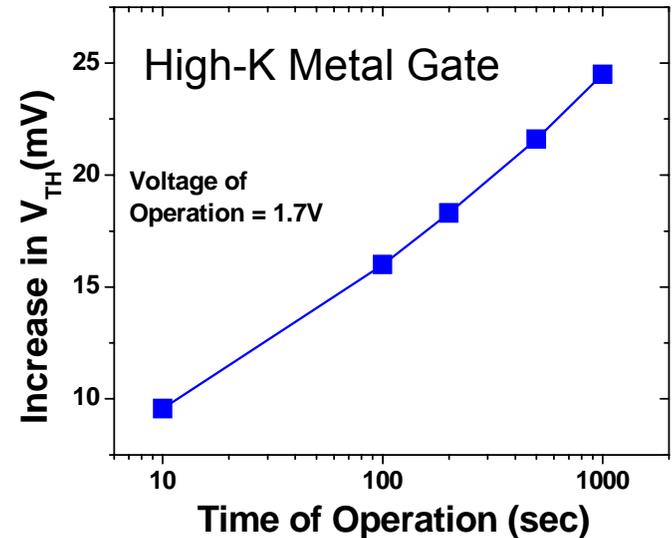
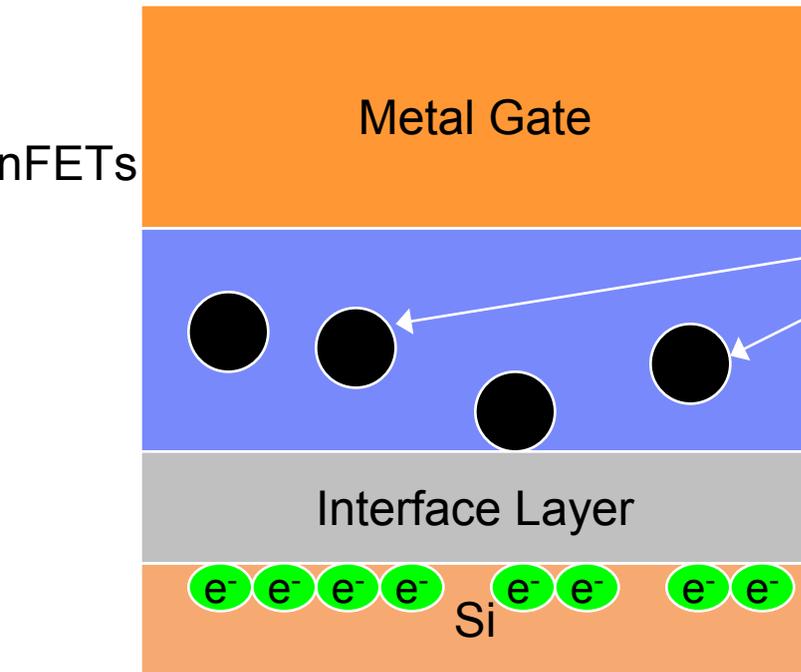
-Interface state trap creation
(Permanent)

Fig. 1: General scheme of NBTI degradation made up of permanent and recoverable parts.

G. La Rosa et al., *IRPS 06*, pp. 274-282.

Huard et al 2007 IEDM

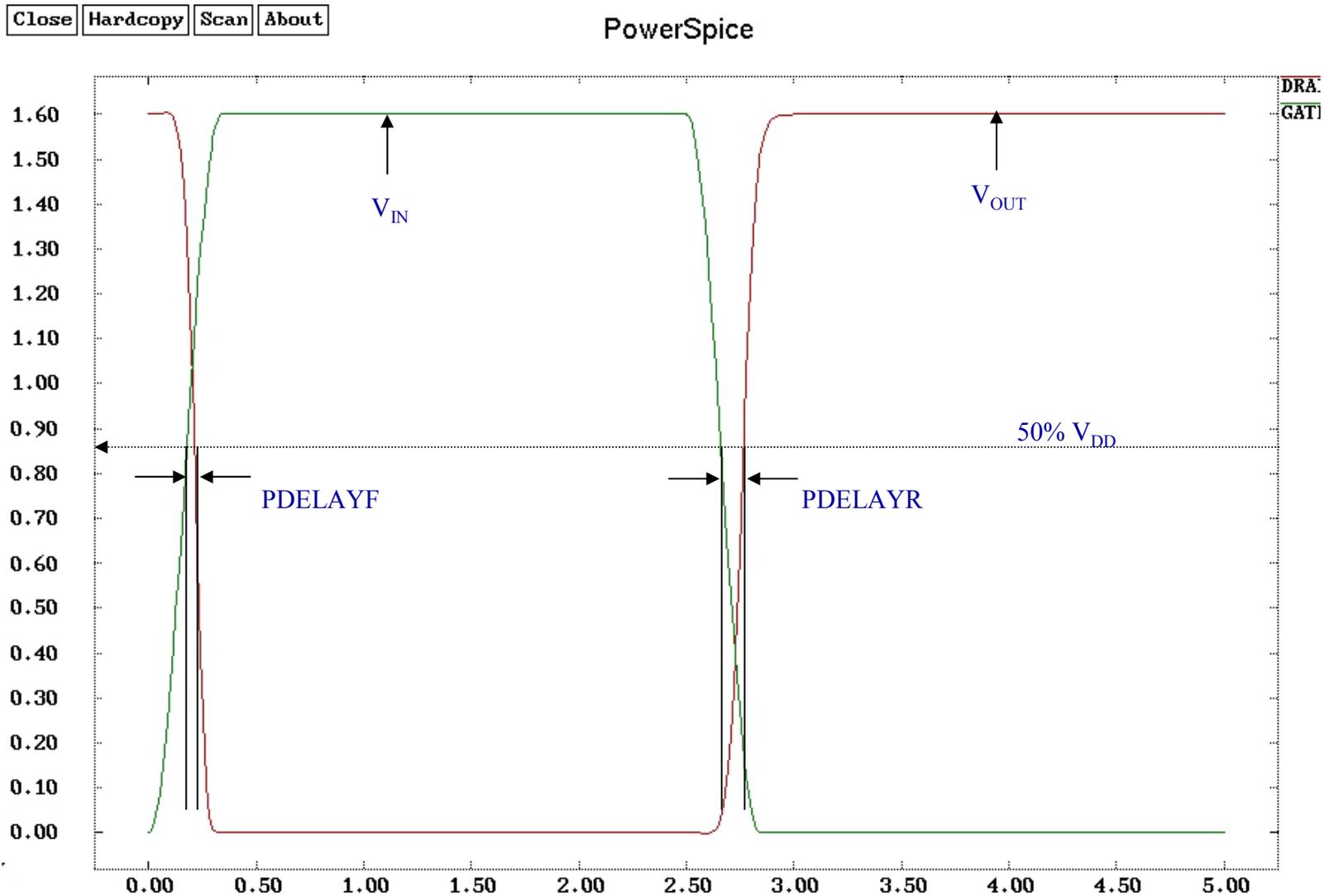
Charge Trapping in High-K



- HfO_2 contains pre-existing Defects (or traps)
 - Electron injection from Si in nFETs in On-State
 - Traps “trap” electrons
 - Trapped charge causes Threshold Voltage (V_{TH}) – (Need V_{TH} to be low)
 - Performance ↓
 - Known as Positive Bias Temperature Instability (PBTi) or “Charge Trapping”
 - Did not exist in SiON based technologies → RELIABILITY IMPLICATIONS

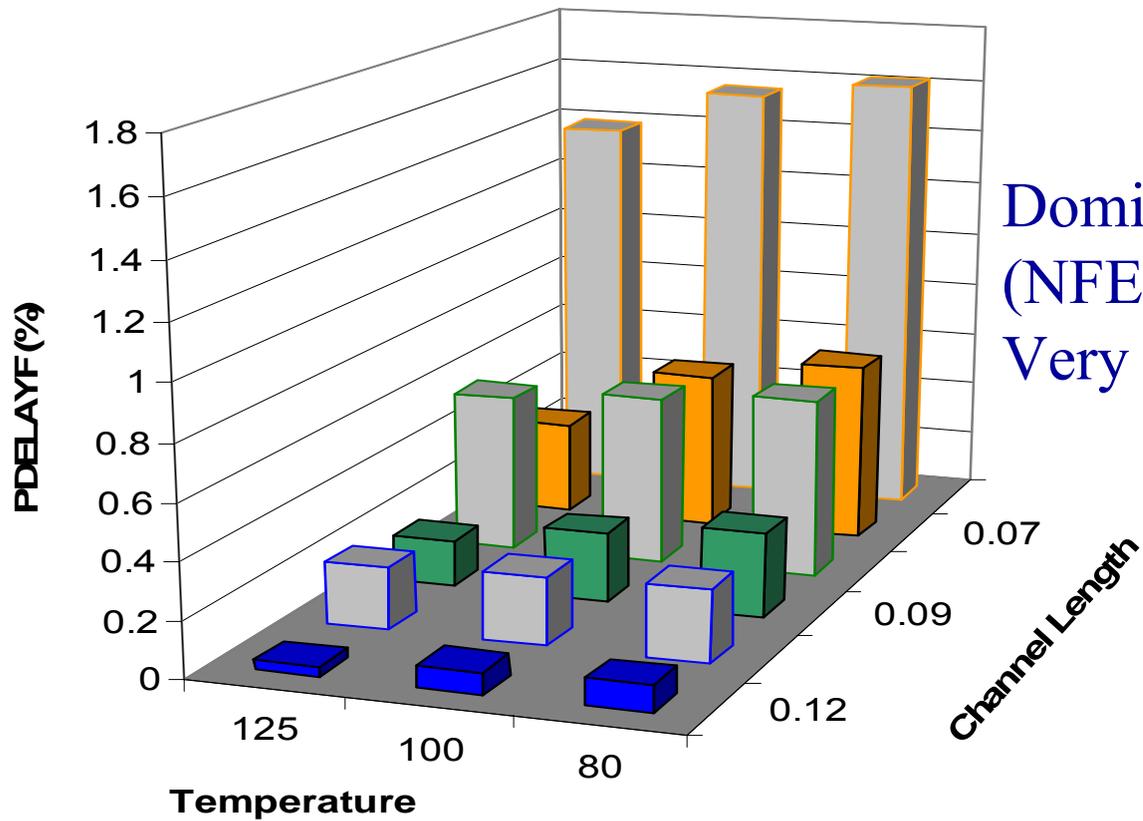
AC Considerations

Typical Voltage vs. Time plot generated from PowerSPICE for inverter



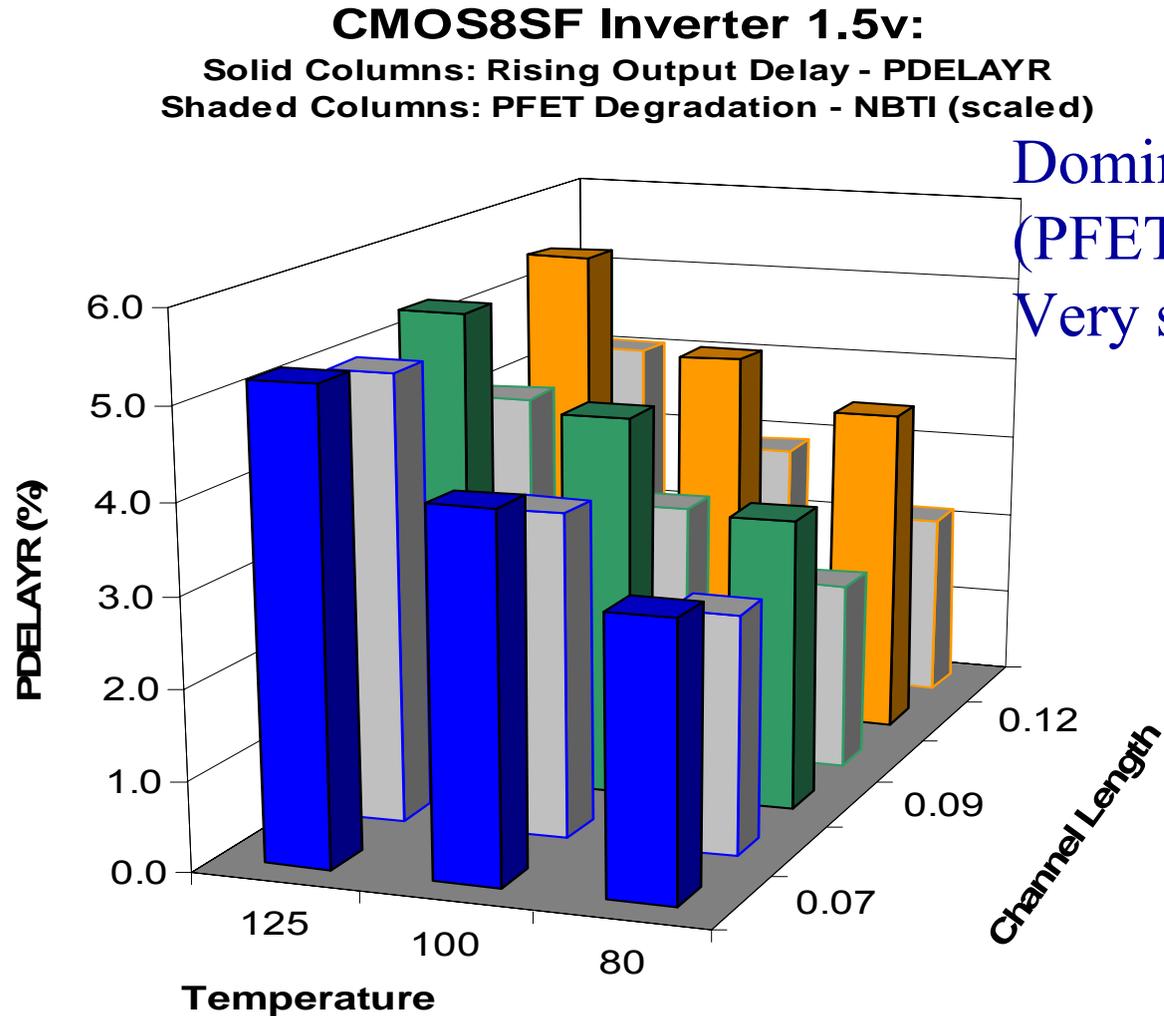
PDELAYF vs Temp/Leff @ 1.5V

CMOS8SF Inverter 1.5v:
Solid Columns: Falling Output Delay - PDELAYF
Shaded Columns: NFET Degradation

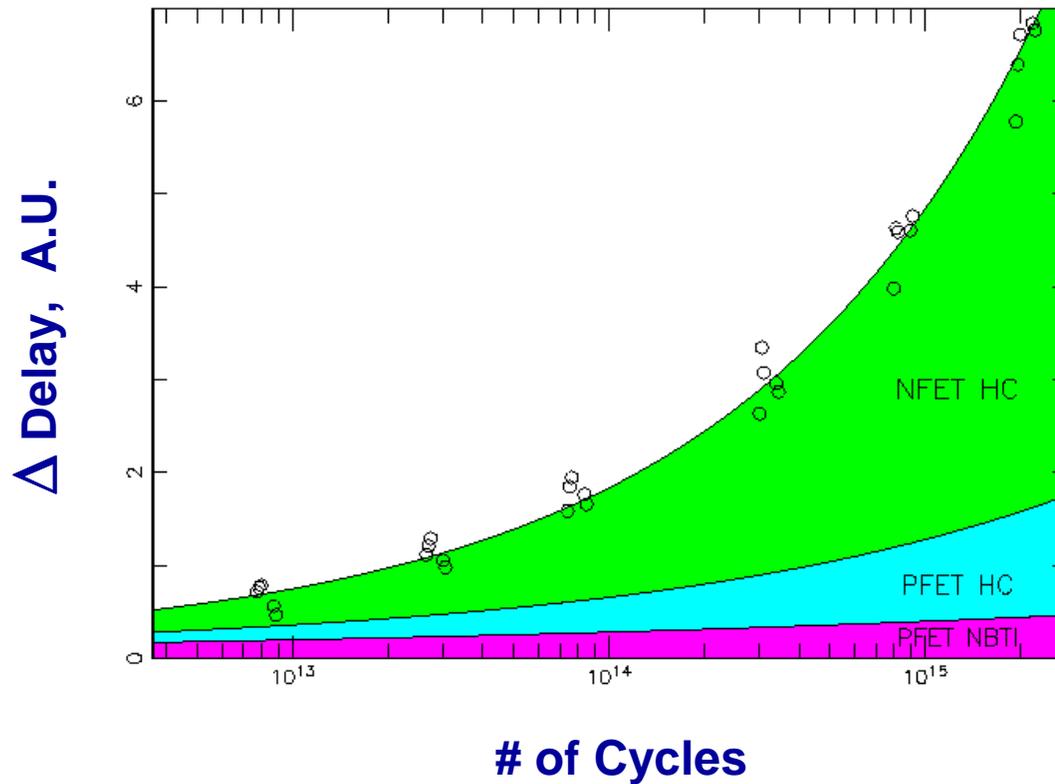


Dominated by H.C
 (NFET)
 Very sensitive to Leff

PDELAYR vs Temp/Leff @ 1.5V

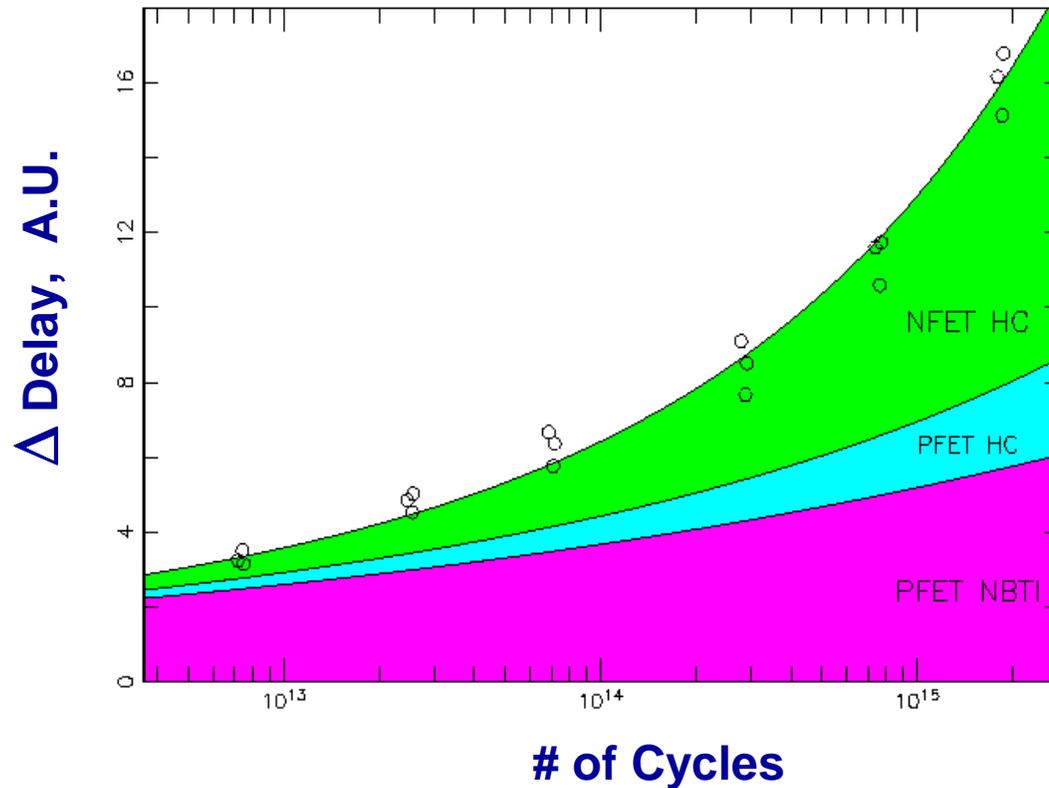


Ring Oscillator Δ Delay vs # Cycles @ 30°C



At room temperature
for a high # cycles
NFET is dominant.
but PFET contribution
H.C + NBTI
is not negligible

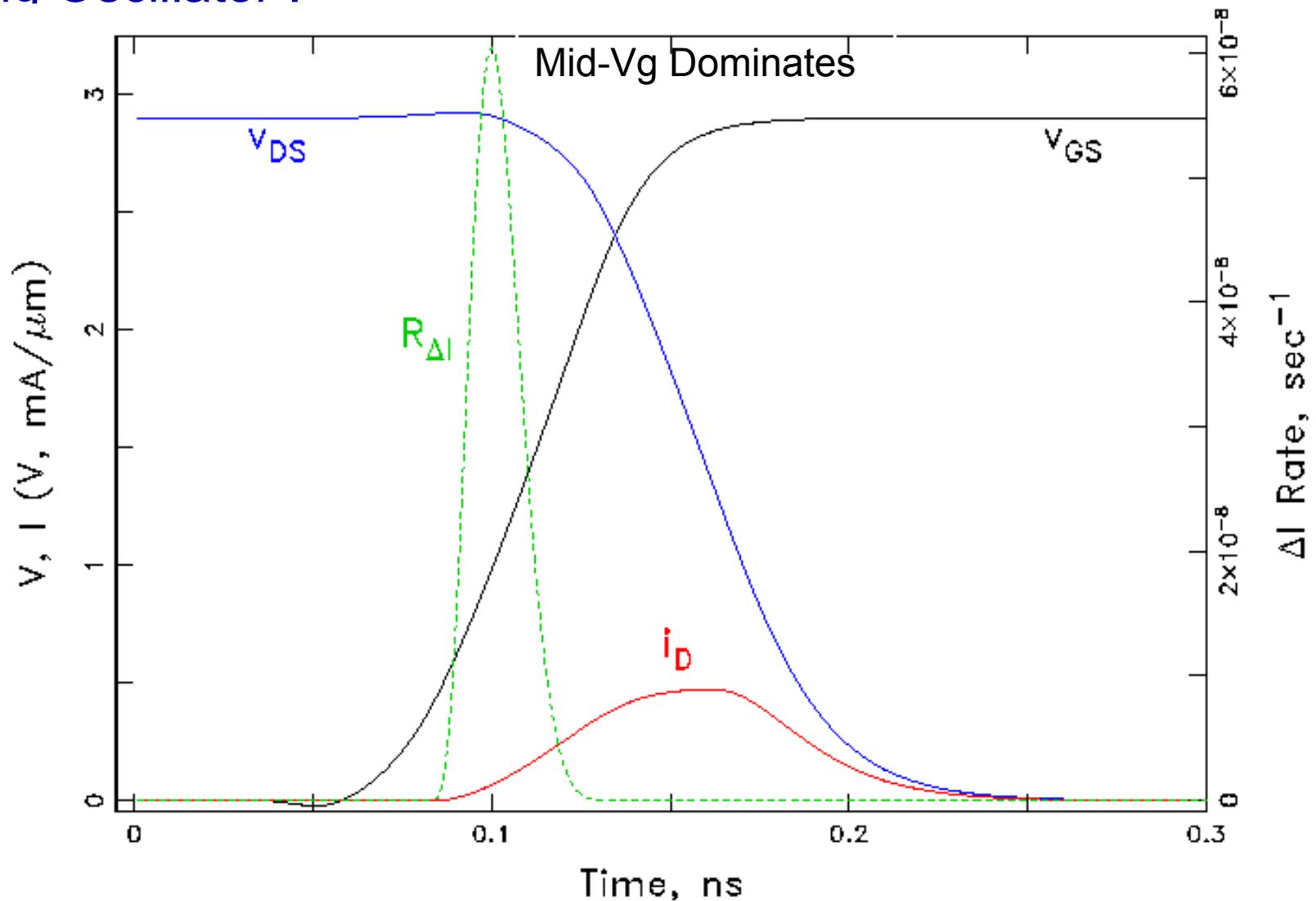
Ring Oscillator Δ Delay vs # cycles @ 125°C



At high temperature for # cycles $< 4 \times 10^{14}$ NBTI contribution from PFET is clearly dominant.

Quasi-static calculation for the PFET- Energy Driven HC [37]

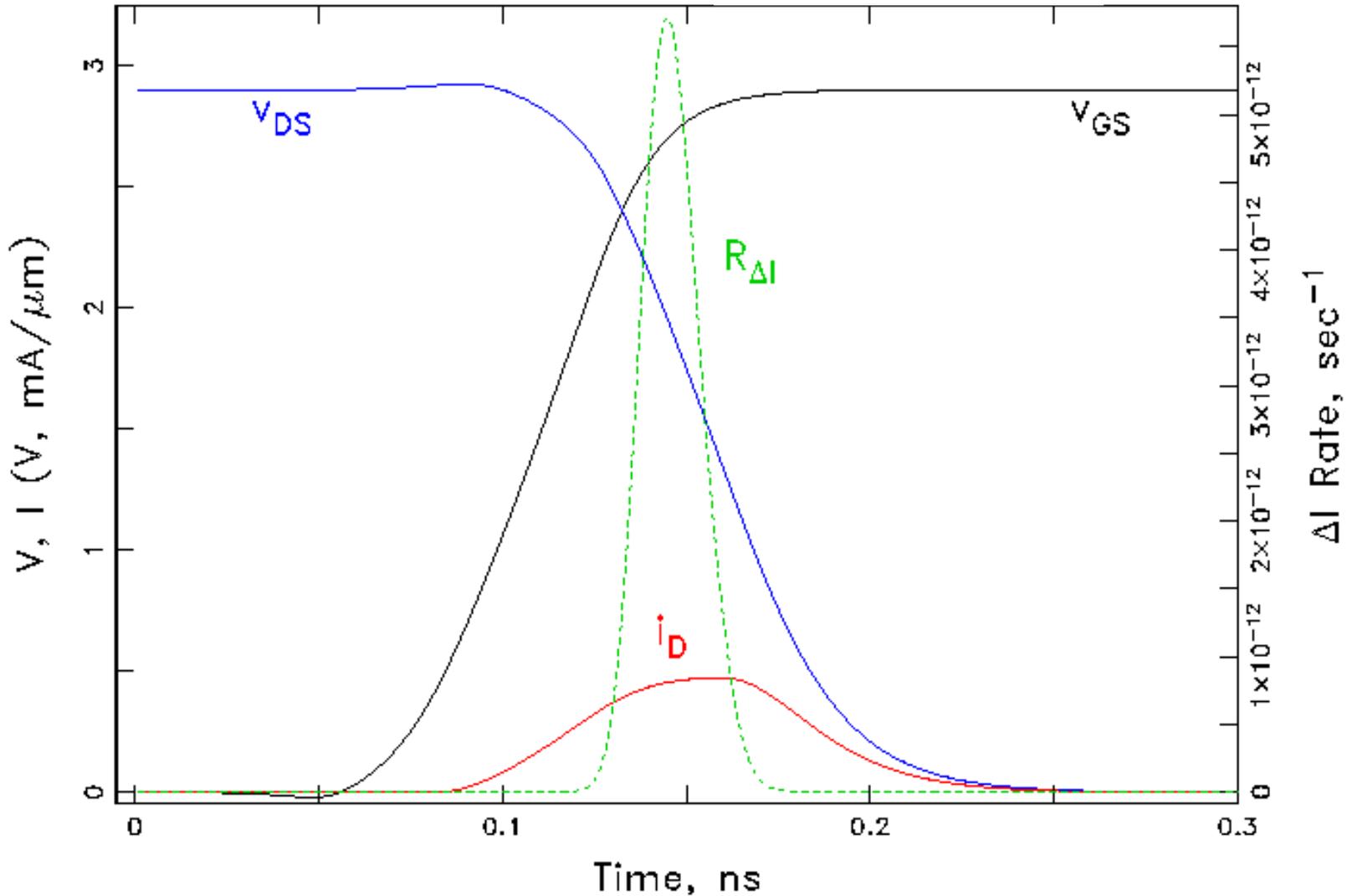
Ring Oscillator I



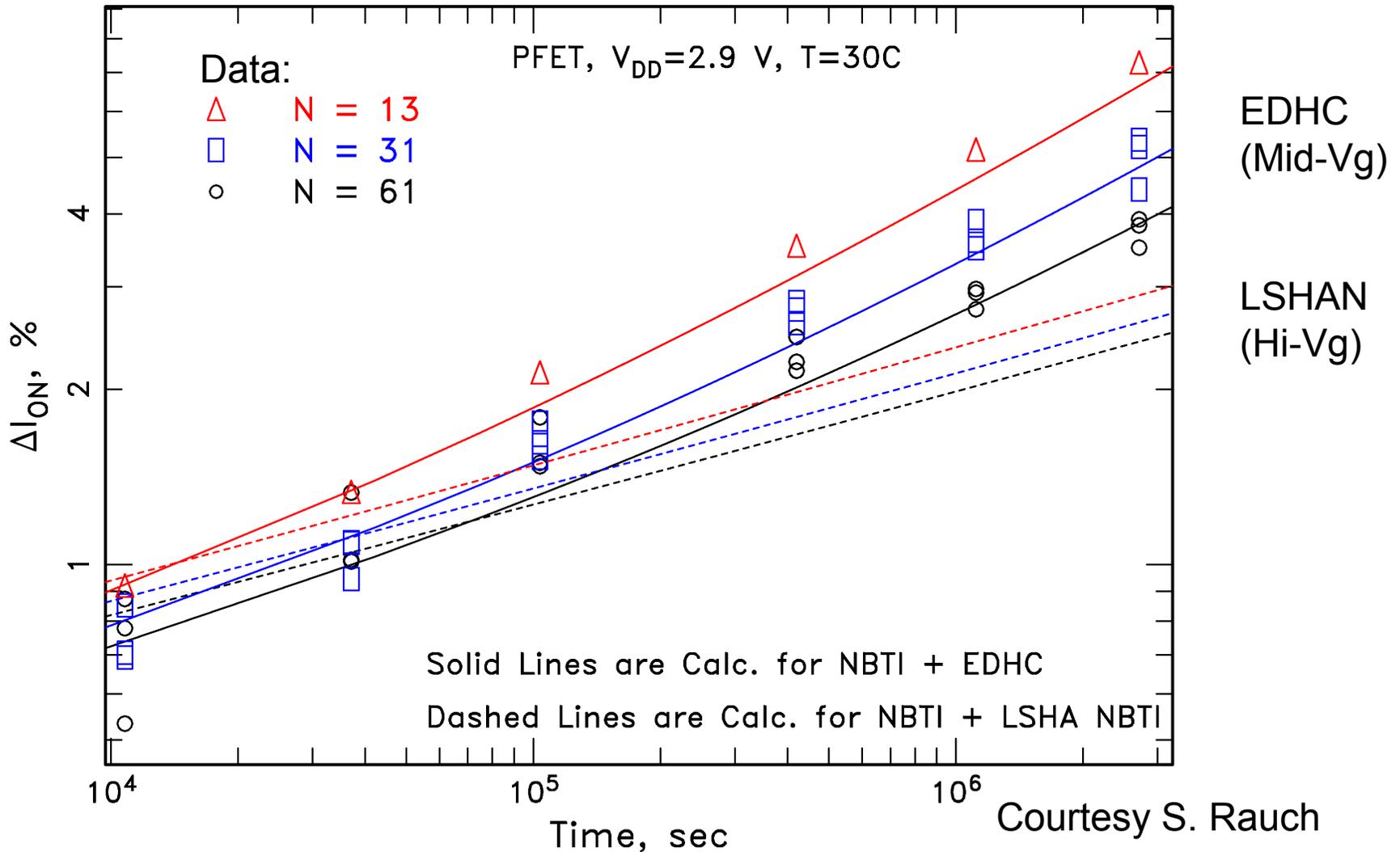
Quasi-static calculation for PFET- LSHAN [37]

Ring Oscillator

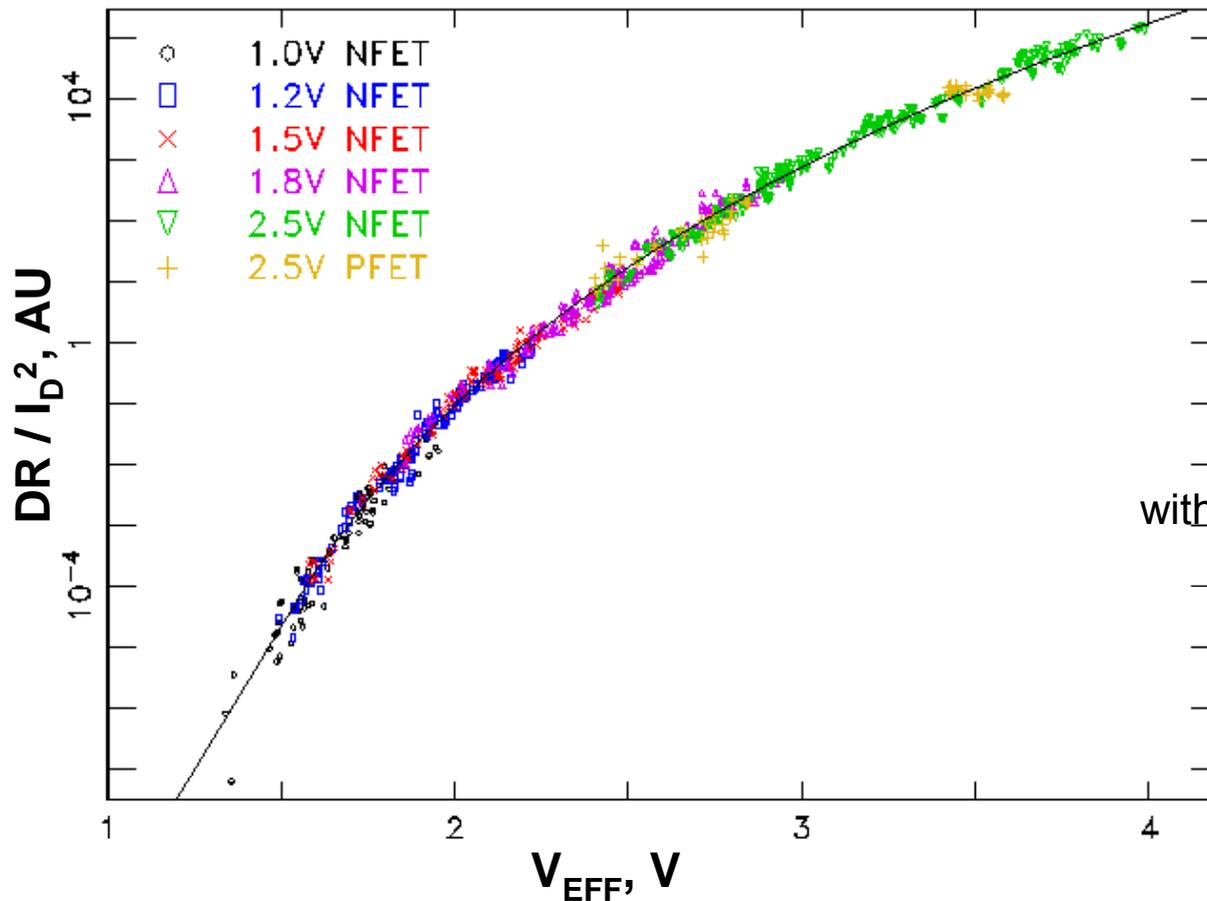
Hi-Vg Dominates



PFET Results (Data Points), and Predictions (Lines)
 (pure NBTI added to each quasi-static calculation) [37]



Universal Energy Driven Hot Carrier – Mid Vg



> 900 Devices
 $V_{DS} = 1.2 - 4.25V$
 6 Device Types
 Various L_{DES} , V_{GS}
 ("Mid Vg")

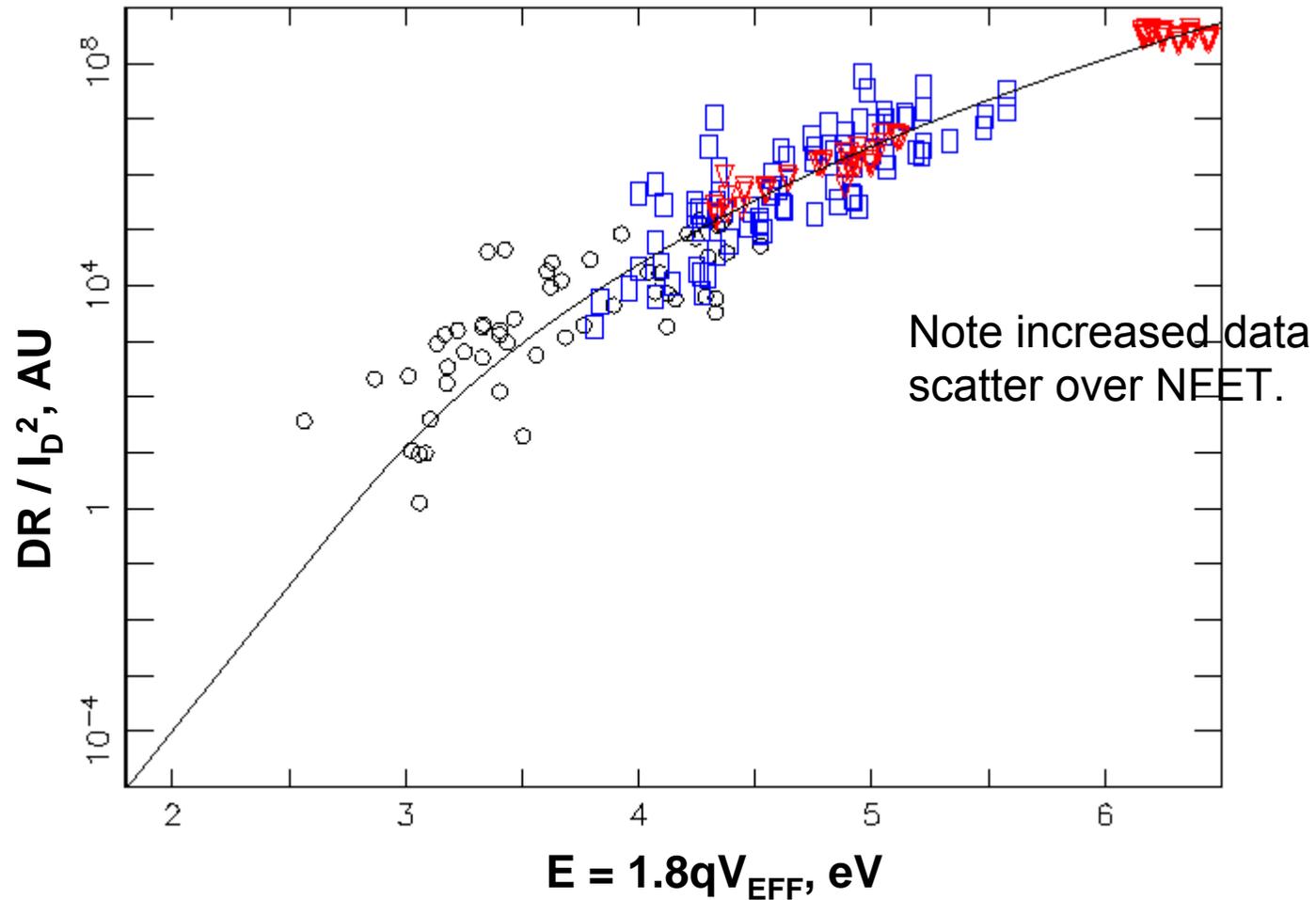
$$S_{IT} \propto \exp(aE), \quad E \leq \phi_{IT} + p/a$$

$$\propto (E - \phi_{IT})^p, \quad E > \phi_{IT} + p/a$$

with $\phi_{IT} = 1.6 \text{ eV}$, $p = 14$, $a = 12 \text{ eV}^{-1}$

The Damage Rate, DR, follows a universal curve: $DR(V_{EFF}) \propto I_D^2 S_{IT}(q m_{EE} V_{EFF})$ irrespective of scaling, proving that the available energy, not electric field, is driving the damage rate. [S. Rauch]

Universal Energy Driven Hot Carrier – PFET Mid Vg

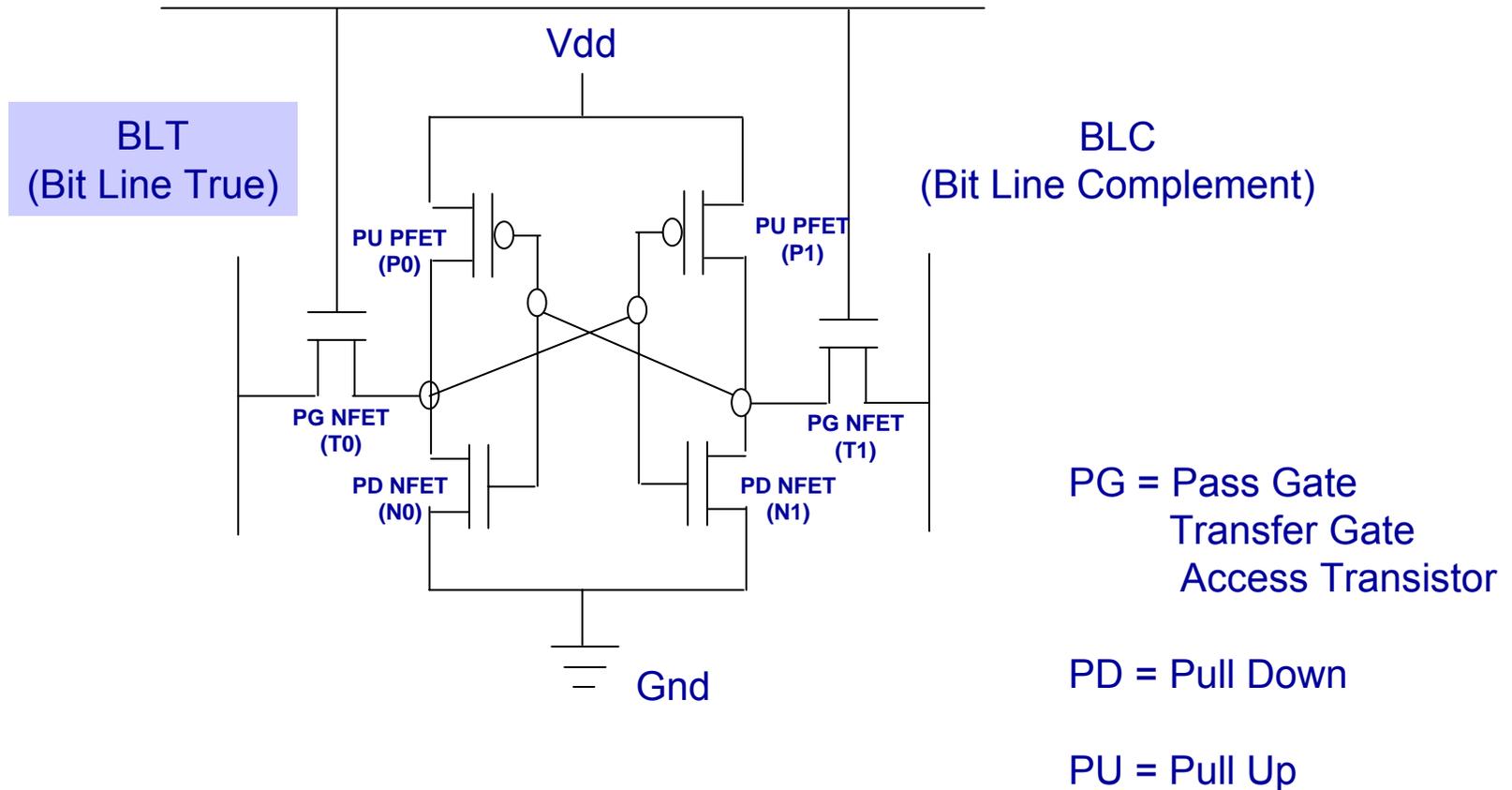


PFET Mid Vg HC follows the same universal curve. [37]

SRAM Reliability Topics

6 Transistor SRAM bit cell

Word Line (WL)



Threshold Voltage Variation

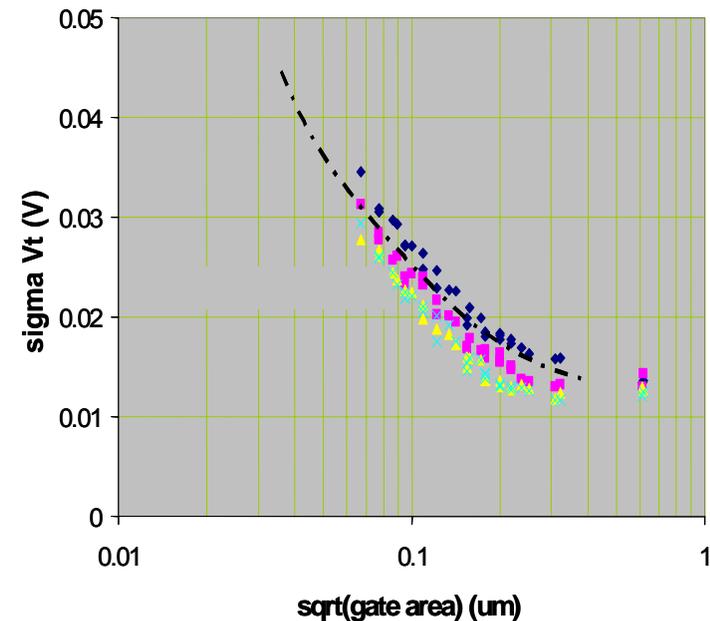
Statistical Nature of SRAMs:

- Large number of bits on a chip
- SRAM yield is determined by the weakest bit(s)
- Cell operation is sensitive to PG,PD, PU ratios
- The larger the variation, greater chance of a failing bit

Smaller cells \rightarrow smaller FETs \rightarrow greater V_{th} variation

V_{th} variation inversely proportional to $\sqrt{(L * W)}$
due to random dopant fluctuation

Sigma V_t vs $\sqrt{\text{Area}}$



R.W. Keyes, "The effect of randomness in the distribution Of impurity atoms on FET thresholds," *App. Phys.* vol. 8, no. 3: 251-9, 1975

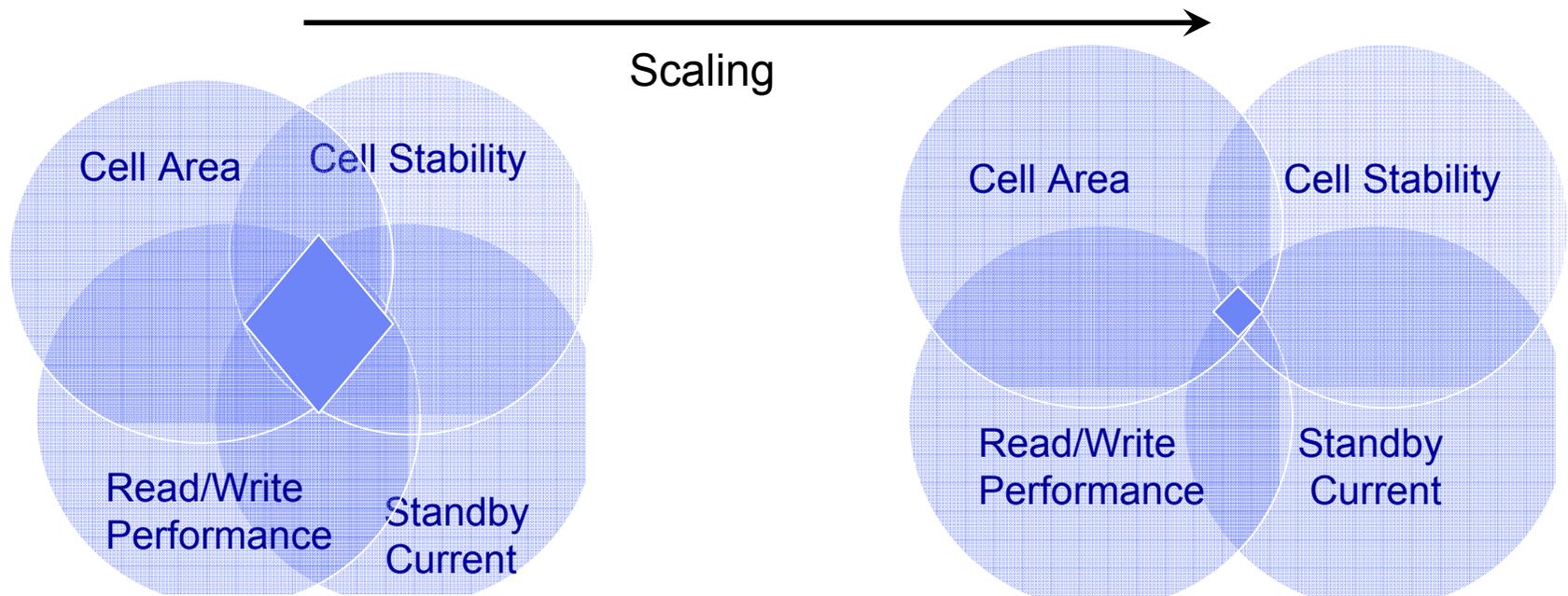
SRAM Cell Design

Bit cell physical layout and cell transistor characteristics chosen to meet

- Cell area
- Cell stability
- Speed of a read and write and
- Standby current

As technologies scale, cell sizes shrink and more bits are in a chip

- FET variations increase
- Balancing the PG, PD, PU ratios to meet all criteria is more difficult



SRAM Cell Reliability

Cell must meet requirements over product lifetime

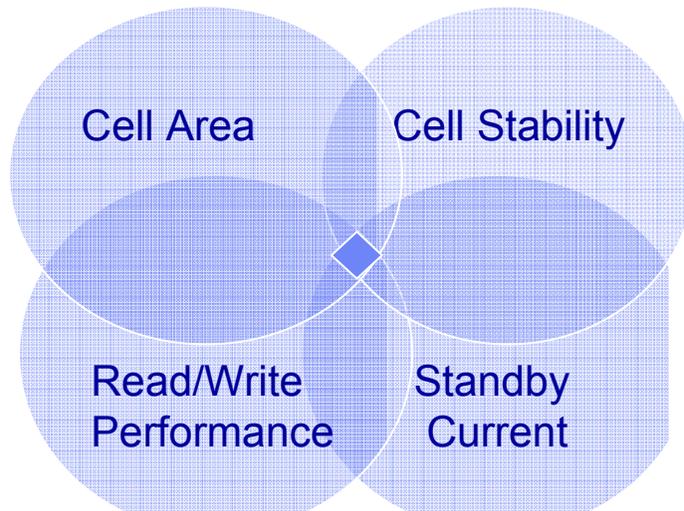
Transistor changes over a cell's lifetime can degrade cell stability, performance, and/or standby current

If cell changes are large enough, SRAM may fail in the field

In addition, physical layout must ensure no shorts or opens are induced in cell operation
Example: gate to contact short due to oxide breakdown

Time 0:

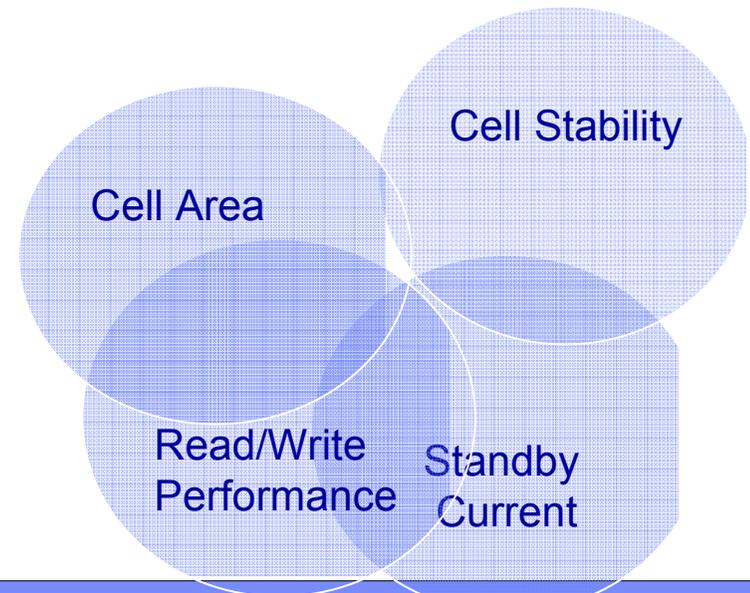
Cell meets criteria, but
Design for reliability not considered



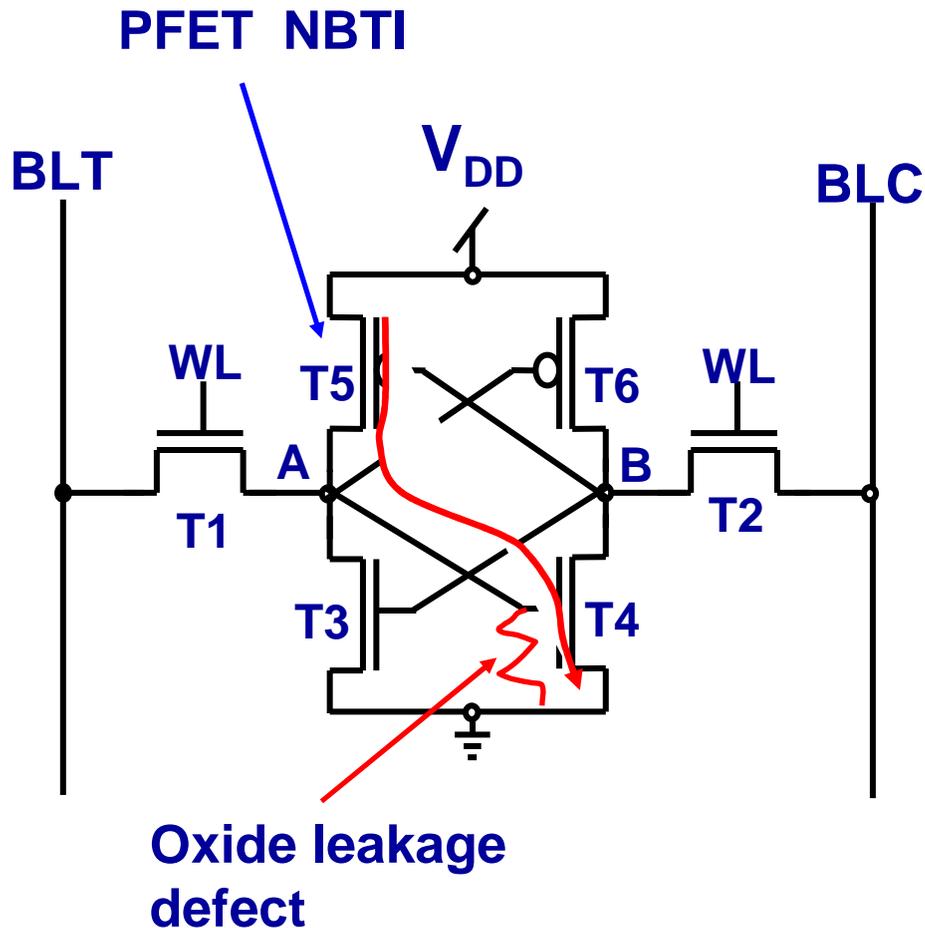
After N hours of operation:

Cell no longer meets stability criteria

→
Chip Operation



SRAM V_{\min} Failure Mechanism



- Gate oxide leakage degrades the pull-down N-FET
- P-FET degrades due to NBTI
- P-FET fights to keep high node at V_{DD}
- This mechanism degrades stability of the cell
- Causing V_{\min} to be out specification

Mueller *et al.* IRPS 2004

Implications

SRAM Stability degradation can cause chip failures after burn-in or operation in the field due to so-called “ V_{\min} ” or “ V_{ccmin} ” fails (primarily read disturb).

Major root mechanisms:

A. NBTI

B. Gate Dielectric Soft Breakdown

- 1. Intrinsic**
- 2. Extrinsic**

A. Haggag et al., *IRPS 06*, pp. 541-544.

K. Mueller et al., *IRPS 04*, pp. 426-429.

NBTI Degradation Model – Distribution

Even for identical use conditions and devices, NBTI will cause V_T mismatch shifts due to random variations in the number and spatial distribution of the charges/interface states formed.

This is similar to random dopant fluctuation induced mismatch, and obeys similar statistics.

This means that small gate area devices will experience more NBTI induced mismatch.

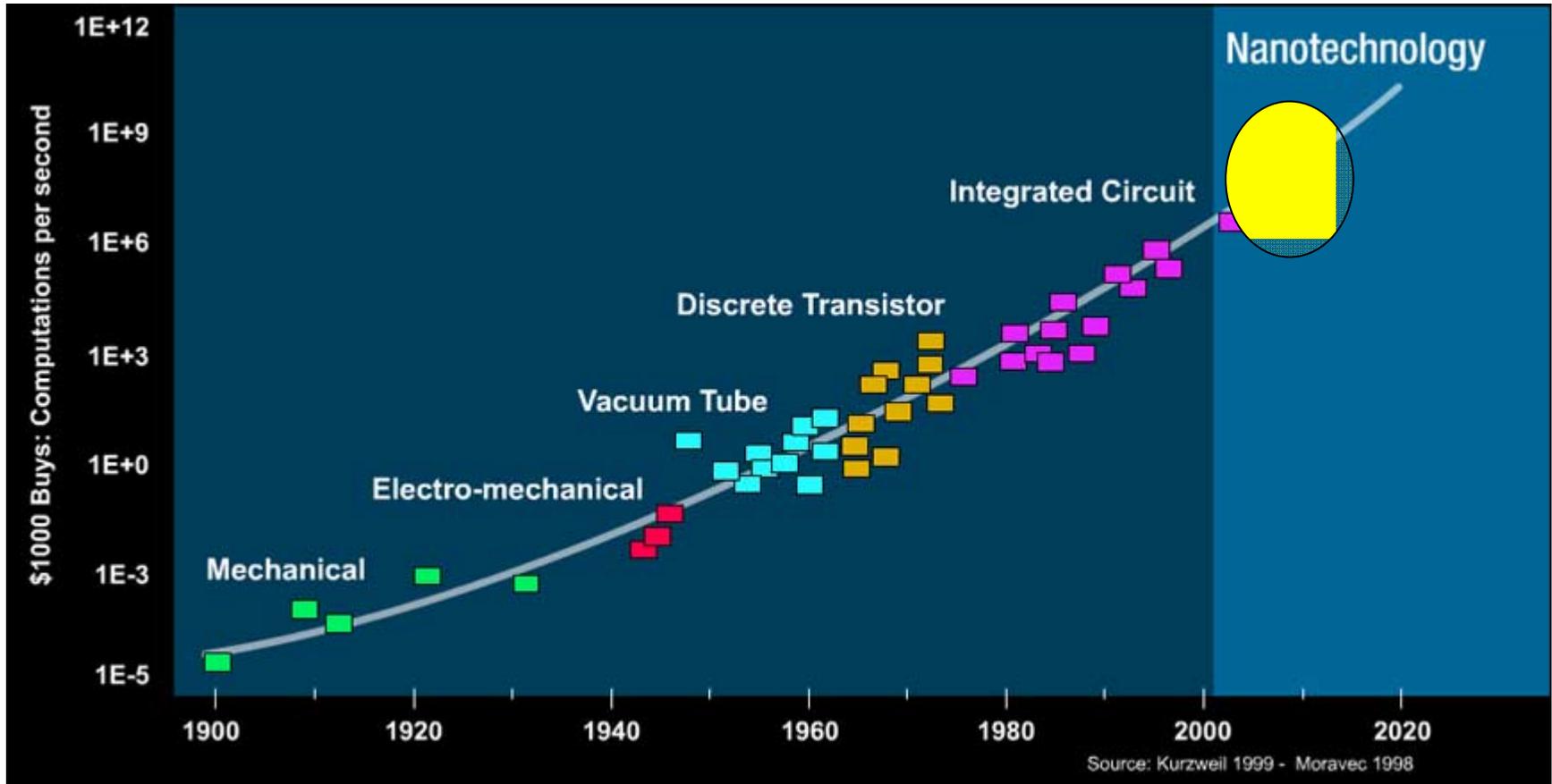
The effect is relatively unimportant for typical CMOS digital logic, since path delays tend to average out individual device shifts, and device widths are fairly large. But for SRAM, it must be considered.

S. Rauch, *IEEE TDMR*, Vol. 2, No.4, pp. 89-93, (2002).

Review and Perspective

Accelerating Advances in Technology

In 100 years 15 order of magnitude improvement !



Perspectives - 2010

- Years since 1st transistor (1947) – 63 years
- Estimated number of transistors produced – 1.5×10^{19}
(If rice, about a trillion tons of long grain rice -
about 1,300 x the world wide production*)
- Transistors consumed per person – 2×10^9
How many transistors in your pocket?
- \$'s per person spent on transistors - \$40
- Price of a transistor – 20 nanodollars
(Is that why its called nanotechnology?)
- Data Explosion:
 - In 2010 digital information will grow to 988 Exabytes
 - 1 Exabyte = 10^{18}
 - 1 Zettabyte = 10^{21}

Courtesy R. Lange

Conclusions

- CMOS scaling fast approaching its fundamental limits
However, it remains a CMOS world; no alternative on horizon
- The end of scaling is in sight → Atoms don't scale !
- Silicon Technology will continue to dominate
- As we scale further Reliability challenges are greater
 - Oxide integrity, Bias Temperature Instabilities, Hot Carriers, electromigration.....
- SRAM stability compromised by reliability shifts
- Reliability mechanisms must be taken into account for successful designs as technology scales
- Reliability Issues gate the viability of many new promising technologies
- Thank you for attending