IS RELIABILITY IMPORTANT?

Some famous examples of the consequences of insufficient reliability
IS RELIABILITY IMPORTANT?

Reliability is an essential aspect of all products and technologies.
OUTLINE

1. Introduction

2. Basic reliability concepts

3. Time-dependent Dielectric Breakdown (TDBB)

4. Hot carrier degradation

5. Bias-Temperature Instability (NBTI)

WHY IS RELIABILITY IMPORTANT?

Reliability is becoming a major bottleneck in the further downscaling of VLSI technologies

- Increasing fields and current densities
- Increasing power density and chip temperatures
- Increasing number of interfaces
- Increasing process complexity
- Introduction of new materials (high k, low k, Cu, metal gate)
- Increasing failure rate requirements
MARKET CONTINUOUSLY DEMANDS FOR HIGHER RELIABILITY LEVELS

Maximum failure rate and defect density requirements are continuously increasing

Power supply voltage evolution

Power supply voltage is saturating at about 1V due to non-scaling of subthreshold slope
ELECTRIC FIELDS IN OXIDE AND IN SILICON STARTED TO INCREASE AGAIN!

Third scaling period: renewed constant voltage scaling due to non-scaling subthreshold slope will have new implications on reliability!

POWER DENSITY INCREASES

Power density will soon be dominated by static power!

D. Cox, IRPS Tutorial 2004

Almost all reliability problems are accelerated by Temperature!
TECHNOLOGY TRENDS

Performance requirements push intrinsic reliability of materials to the limit
- Gate oxide fields > 5MV/cm
- Interconnect current densities > 5x10^5 A/cm^2

New materials introduced
- 70’s: Si, Al, SiO_2 (oxides), Si_3N_4 (nitrides)
- 80-90’s: Ti-silicide, Co-silicide, W, TiN, Al-Cu
- 2000: high k gate dielectrics (Hf-oxide, Zr-oxide, Ta-oxide), low k dielectrics, Cu metallization
- 2010: Ge, III-V, CNT, graphene ???

HUGE RELIABILITY CHALLENGES AHEAD OF US!
- High k gate dielectrics for logic and memory applications (HfO_2, HfSiON, Al_2O_3, etc.)
- Metal gates (TaN, TiN, RuO, etc.)
- Ferroelectric materials (PZT, BST, etc.)
- Phase change materials, Resistive RAM materials
- Nanodevices (MuGFET, Finfet, Ge, III-V devices, TFET’s, CNT, nanowires, graphene devices)
- Low k dielectrics/barrier materials/airgaps
- Cu metallization, CNT interconnects, Cu contacts, ....

Unknown reliability behavior and/or new failure mechanisms!

Speed of introduction exceeds capabilities to explore reliability performance
MAIN ON-CHIP RELIABILITY FAILURE MECHANISMS

1) Time-dependent dielectric breakdown
2) Hot carrier degradation
3) Bias-Temperature Instability
4) Electromigration
5) Stress voiding
6) Interconnect dielectric breakdown/ stability
7) Electrostatic Discharge sensitivity

MAIN OFF-CHIP RELIABILITY FAILURE MECHANISMS

1) Creep
2) Fatigue
3) Stiction
4) Corrosion
5) Electromigration
6) Charging

Reliability needs a broad field of disciplines and expertises going from device physics, material science, surface chemistry, electrical and physical characterisation
OUTLINE
1. Introduction
2. Basic reliability concepts
3. Time-dependent Dielectric Breakdown (TDBB)
4. Hot carrier degradation
5. Bias-Temperature Instability (NBTI)

DEFINITIONS

Reliability = probability that an item will
- perform a required function
- under stated conditions
- for a stated period of time

- probability: probabilistic models/statistical methods
  - random variables, sample populations
  - density functions (p.d.f., c.d.f.)
  - mean time, median time-to-failure, failure rate

- required function: includes definition of failure
  - specification of satisfactory operation

- stated conditions: total physical environment
  - mechanical, thermal, environmental, electrical conditions

- stated period of time: concept of lifetime
  - required lifetime depends on application: telecom, consumer, space, automotive
DEFINITIONS

Failure = termination of the ability of an item to perform a **required function** (within previously specified limits)

Failure mechanism = the physical, electrical, chemical, thermal or any other **process** which leads to a failure.

Examples:
- **Devices**
  - hot carrier injection
  - NBTI
  - Gate oxide breakdown
- **Interconnects**
  - electromigration
  - stress migration
  - low k dielectric BD
- **Packages**
  - humidity/corrosion
  - moisture penetration
  - die cracking
  - bond wire fracture

DEFINITIONS

Failure mode = consequence of the mechanism through which the failure occurs **or**

that characteristic of the device for which it has been named a failure

Examples:
- **Devices/interconnects**
  - internal short or open
  - threshold voltage shift
  - parametric variation
  - Junction/gate leakage
- **Circuits**
  - shift in access time
  - power supply current increase
  - functional failure
  - stuck-at fault
DEFINITIONS

Intrinsic failures:
▸ inherent in the design and materials used
▸ managed by insuring that they occur beyond useful life of a product, by limiting loads that drive the failure mechanisms (e.g. maximum field, maximum current density, thermal management)

Extrinsic failures:
▸ due to process or manufacturing defects or to misapplications such as overload, EOS, ESD etc.
▸ managed by improving manufacturing process, reducing defect density
▸ usually product reliability is determined by extrinsic failures

RELIABILITY AND FAILURE FUNCTIONS

Random variable x: time t

a) Cumulative Failure Function: F(t)
   - t < 0 → F(t) = 0
   - 0 ≤ t ≤ t' → 0 ≤ F(t) ≤ F(t')
   - t → ∞ → F(t) → 1
   F(t) = probability that a device will fail at or before time t
   = percentage of failures at time t

b) Reliability function: R(t)
   R(t) = Probability that a device will survive to time t without failure
   = percentage of good devices at time t
   \[ R(t) = 1 - F(t) \]
RELIABILITY AND FAILURE FUNCTIONS

c) Failure density function: \( f(t) : = p.d.f. \)

\[ f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt} \]

\[ F(t) = \int_0^t f(t) dt \]

\[ R(t) = 1 - \int_0^t f(t) dt = 1 - \int t_1^t f(t) dt \]

d) Mean time to failure: MTTF

\[ \text{MTTF} = \text{mean value to the p.d.f.} \]

\[ \text{MTTF} = \int_0^\infty t f(t) dt \]

e) Median time to failure: MTF

\[ \text{MTF} = \text{median value to the p.d.f. } f(t) \]

\[ \text{MTF} = \text{time when } F(t) = 50\% = t_{50} \]

f) Failure rate

- instantaneous failure rate or hazard rate:

\[ \lambda_i(t) = \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{R(t) \times \Delta t} \]

\[ = \frac{1}{R(t)} \cdot \frac{dF(t)}{dt} \]

\[ \lambda_i(t) = \text{probability that a device will fail in the time interval } (t, t+\Delta t), \text{ given that it has survived at time } t \]

- cumulative failure rate = average hazard rate

\[ \lambda_c(t) = \frac{1}{t} \int_0^t \lambda_i(t) dt \]

\[ H(t) = \int_0^t \lambda_i(t) dt = \text{cumulative hazard function} \]

Units of failure rate: \# Failures/(devices x hours)
Failure unit: 1 FIT = 1 failure/10^9 device hours = 8.76 ppm/year
HAZARD RATE AND BATHTUB CURVE

If hazard rate \( \lambda(t) \) is
- Decreasing
- Constant
- Increasing
With time
- Infant mortality
- Useful life
- Wear-out

Bath tub curve = sum of 3 curves
1) **quality failures**: weaker parts (manufacturing defects) are getting removed from the population
2) **stress related failures**: failures occur occasionally at a constant rate
3) **wear-out failures**: accumulated damage takes its toll, devices start to fail with increasing frequency

OUTLINE
1. Introduction
2. Basic reliability concepts
3. Time-dependent Dielectric Breakdown (TDDB)
   - Basics
   - Statistics
   - Acceleration model
   - Impact on circuits
4. Hot carrier degradation
5. Bias-Temperature Instability (NBTI)
TUNNELING BASICS

\[ \begin{align*}
V & \quad \text{Si substrate} \\
\phi & \quad \text{Gate} \\
\text{Si} & \quad \text{SiO}_2 \\
\text{Si} & \quad \text{Si} \\
\text{SiO}_2 & \quad \text{SiO}_2 \\
\phi_B & \quad \text{direct tunneling} \\
\epsilon^- & \quad \text{FN-tunneling} \\
\end{align*} \]

\[ \begin{align*}
J_{\text{gate}} & \quad (A/cm^2) \\
V_{\text{gate}} & \quad (V) \\
t_{\text{ox}} & = 9.7 \text{ nm} \\
\end{align*} \]

(GIST OF) GATE OXIDE BREAKDOWN


\[ \begin{align*}
V & \quad \text{Gate} \\
\text{Si substrate} & \quad \text{Si substrate} \\
I & \quad \text{Gate oxide breakdown = catastrophic failure} \\
\end{align*} \]
RAMPED VOLTAGE TEST

- Apply a fast **voltage ramp** (typically 1-100 V/s) and measure current
- **Breakdown voltage** is defined as the gate voltage at which a sudden current increase occurs
- Fixed **trigger current** is set (100uA to 5mA)
- Sensitive to series resistance effects!

CONSTANT VOLTAGE STRESS

Oxide breakdown is a time-dependent phenomenon

TDDB = Time dependent Dielectric Breakdown
CONSTANT VOLTAGE STRESS (CVS)

Dielectric breakdown = (sudden) loss of the insulating property

Oxide degradation = current change as a consequence of degradation

CVS SEEN BY THE EMISSION MICROSCOPE

Uniform tunnel current

Localized breakdown spot

Breakdown is a localized event
**TDDB: CVS AND CCS**

**Constant Voltage Stress (CVS)**
- average field is constant
- current increase: net positive charge; current decrease: net negative charge
- at fixed Vg: tbd decrease with thinner oxides

**Constant Current Stress (CCS)**
- net positive charge => decrease of average field (and voltage)
- net negative charge => increase applied voltage
- at fixed J: tbd increases with thinner oxides

---

**CHARGE-TO-BREAKDOWN**

Degradation scales with charge fluence: for higher currents, lower t_{BD}

\[
Q_{BD} = \int_{t_{BD}}^{t_{break}} J_{bd} \, dt
\]

\[
Q_{BD} = \int_{0}^{t_{bd}} J_{bd} \, dt
\]

- for CCS: \( Q_{BD} = J_{bd} \cdot t_{bd} \)

**Q_{BD}** is one of the main BD figure-of-merits of a technology: typically between 1-100 C/cm²

When oxide thickness scales down: Q_{bd}-value becomes less relevant
TIME-TO-BREAKDOWN = STATISTICALLY DISTRIBUTED PARAMETER

• Time-to-breakdown is Weibull distributed

• Monomodal Weibull distribution is described as:

\[ F(t) = 1 - \exp \left( -\left( \frac{t}{\eta} \right)^\beta \right) \Rightarrow \ln(-\ln(1-F)) = \beta \ln(t) - \beta \ln(\eta) \]

BREAKDOWN STATISTICS

Oxide breakdown \((t_{BD}, Q_{BD})\) is Weibull distributed

Extrapolation towards low percentile from lognormal distribution yield overoptimistic results!
TWO BREAKDOWN MODES

Intrinsic BD = caused by electrical stress-induced defect generation (creation of neutral electron traps)

Extrinsic BD = caused by process-induced imperfections (impurities, roughness, thinning, pinholes, structural weakness)

EXTRINSIC BREAKDOWN

Example of extrinsic breakdown

- Process-induced damage at the gate isolation edge causes multiple breakdowns after continued stress
EXTRINSIC RELIABILITY

Influencing factors:
- substrate quality (epi vs Cz substrates)
- crystalline defects in the silicon: stacking faults!
- wafer cleaning procedures
- oxidation procedures (wet, dry, post-oxidation anneal)
- purity of chemicals
- cleanliness of oxidation furnace
- metal contamination (e.g. heating elements in furnace)
- mobile ion contamination (Na)
- photolithographic related defects
- mechanical stress buildup
- post-oxidation processing: e.g. plasma induced damage

AREA DEPENDENCE OF TBD

For \( A_1 > A_2 \): series rule applies:

\[
F_{A_i}(t) = 1 - \prod_n (1 - F_{A_i}(t)) \quad \text{with } n = \frac{A_1}{A_2}
\]

\[
\ln(1 - F_{A_1}(t)) = \sum_n \ln(1 - F_{A_2}(t))
\]

\[
= n \times \ln(1 - F_{A_i}(t))
\]

\[
\ln(-\ln(1 - F_{A_i}(t))) = \ln\left(\frac{A_1}{A_2}\right) + \ln(-\ln(1 - F_{A_i}(t)))
\]

The area effect: changing from \( A_1 \) to \( A_2 \): distribution shift vertically over a distance \( \ln(A_1 / A_2) \)

The area dependence of breakdown: schematically
THE AREA SCALING EFFECT

• Changing from $A_1$ to $A_2$ gives a vertical shift of $\ln(A_1/A_2)$
• For Weibull: $\eta_1 = \eta_2 (A_2/A_1)^{1/\beta}$
• Scaling of $\eta$ is $\beta$-dependent
• Extrinsic breakdown can only be measured if gate area is large

Normalized area = $10^{-6}$ cm²
Measurement area =
- $10^{-6}$ cm²
- $2.10^{-5}$ cm²
- $5.25.10^{-4}$ cm²
- $2.45.10^{-2}$ cm²

- Normalization of the breakdown distribution to a reference area is always possible
- Depending on test structure area different parts of the distribution are measured
- Allows to extract Weibull slope with higher accuracy
**BREAKDOWN MODEL**

Breakdown model:
- electrons are tunneling
- gain energy in the oxide conduction band
- at anode: energy is transferred to holes

1) **anode hole injection model**: holes tunnel back and generate traps
2) **hydrogen release model**: H is released and moves through the oxide, generate traps

At critical trap density: breakdown occurs!

---

**OXIDE DEGRADATION**

NBD = critical defect density

Decreases with Tox!

Breakdown occurs when a critical density of neutral electron traps is reached

Degraeve et al, TED, p. 904 1998
PERCOLATION MODEL

Electron traps are randomly generated in the oxide
Trap is characterized by a sphere with constant radius \( r \)
Breakdown when conducting path is formed from anode to cathode
Critical trap density = total number of traps / testvolume

STATISTICAL MODELLING OF BREAKDOWN

Simulation fits experimental \( Q_{BD} \)-distribution very well

- \( t_{ox} = 8.9 \) nm
- \( C = 1.9 \times 10^{21} \) C \( \cdot \) cm \( ^{-0.6} \) cm \( ^{-1.8} \)
- \( r = 0.45 \) nm

experiment
simulation

Simulation fits experimental \( Q_{BD} \)-distribution very well
Statistics of percolation model explains decreasing Weibull slope (larger spread of $t_{BD}$) for decreasing oxide thickness.

The decrease of the Weibull slope $\beta$ for thin oxides is an intrinsic statistical property of the degradation mechanism.

Degraeve et al, TED, p. 904, 1998
Comparison of $Q_{BD}$ for different oxide thicknesses can be very misleading!!
Thick oxide: large Weibull slope: small area dependence
Thin oxide: small Weibull slope: large area dependence

$$Q_{BD,1} = Q_{BD,2} \left( \frac{A_2}{A_1} \right)^{1/\beta}$$

Nigam et al. IRPS 1998

Between $10^{-6}$ and 0.1 cm$^2$ lifetime drops with 3-4 orders of magnitude for 4 nm oxides, only factor of 2 for 11 nm oxides
STATISTICAL MODELING OF BREAKDOWN

Critical electron trap density \( N_{BD} \) is reduced with oxide thickness scaling, which is also explained by percolation model.

Trap size: 0.9 nm diameter (IMEC) to 2.5 nm (IBM)

FIELD ACCELERATION MODELS

**E-model**

\[
t_{BD} \approx t_0 \exp(-\gamma E_{ox})
\]

- Breakdown is not fluence driven
- Based on Thermochemical model
- Worst case?

**1/E-model**

\[
t_{BD} \approx t_0 \exp\left(\frac{G}{E_{ox}}\right)
\]

- Breakdown is fluence driven
- Based on Anode hole injection
- Very optimistic: \( 1/E \) probably not valid at lower voltages
IMPACT OF SCALING ON ACCELERATION MODEL

- Electron energy at the anode determines oxide degradation and breakdown.
- Decreasing $t_{ox}$:
  - FN tunneling
  - Ballistic tunneling
  - Direct tunneling

- Thick oxide: $E_{ox}$ determines electron energy at anode
- Thin oxide: $V_g$ determines electron energy at anode

VOLTAGE ACCELERATION MODEL

- Power law model:
  - $t_{BD} \approx V_g^{-n}$
  - Acceleration increases at low $V_g$
  - Power factor $n \sim 44$

- Area Scaling Exp. p+/n Capacitors
- NFET set 1
- NFET set 2
- NFET set 3

E. Hu et al, TED p 2244, 2002
PRINCIPLE OF EXTRAPOLATION TO SPECIFIED CONDITIONS

63% TBD at test structure, test voltage and test temperature

- Temperature acceleration NO Arrhenius-law
- Use spec. T = test T to avoid acceleration issue

63% TBD at test structure, test voltage and spec. temperature

- Voltage acceleration to operating voltage
- Use power law $t_{BD} = V_{G}^{n}$

63% TBD at test structure, operating voltage and spec. T

- Area scaling law to total gate oxide area on chip

63% TBD at spec. area, operating voltage and spec. T

- Weibull distribution

low percentile TBD at spec. area, operating voltage and spec. T

---

PRINCIPLE OF EXTRAPOLATION TO SPECIFIED CONDITIONS

- Distribution extrapolated to operating voltage

\[ \ln(-\ln(1-F)) \]

- 63% for area
- Temp.
- low percentile

- 0.01% for 0.1 cm² after 10 years
- DC operation at operation voltage and temperature
- Area scaling and percentile scaling require correct Weibull slope $\beta$
DETERMINATION OF THE MAXIMUM APPLICABLE GATE VOLTAGE

Combined effect of temperature acceleration and statistical scaling (depends on $\beta$!)

This extrapolation is only valid for soft breakdown

Voltage power-law more accurate than exponential also in high-k

Wu & Suñe, TED 2009

PHASES OF GATE OXIDE BREAKDOWN

Electrical stress = Additional gate oxide leakage paths (TDDB)
**STRESS-INDUCED LEAKAGE CURRENT**

SILC: macroscopically uniform over the area
- tunneling mechanism = trap assisted tunneling

**TWO COMPENSATING EVOLUTIONS IN ULTRA-THIN OXIDE SILC**

Effect 1: trap generation rate drastically drops (figure)
- Less conductive paths through the oxide

Effect 2: conductivity per trap increases drastically
- Tunnel distance decreases and current rises exponentially
- Each trap contributes more to the total SILC

(Stathis et al, IEDM 1998)
THE CHANGING APPEARANCE OF SILC

Thick / large capacitor  thin / small capacitor

SILC and individual trap components

Capacitor top view

many small contributions

few large contributions

Capacitor cross section

Discrete SILC events in very thin layers

Abrupt current increase of about 250nA
= creation of single trap conduction path
= ‘discrete’ SILC event
This is NOT a breakdown
TWO-TRAP CONDUCTION PATH

Percolation path with trap-assisted tunneling
- this is **Anomalous Stress-Induced Leakage Current** = a-SILC (non-volatile memories) (Okada et al. IEDM 2001, Degraeve et al., IEDM 2001)
- this is **Micro-Breakdown** (Cellere et al. TED 49(8), 2002) or **pre-Breakdown** (Degraeve et al, IEDM 2001)

THIN OXIDE: MICRO-BREAKDOWNS

- No significant change of the gate current noise after micro-BD
- Is this a breakdown or not?
  - Application dependent specifications needed
**SOFT BREAKDOWN**

Soft breakdown makes classical detection of breakdown problematic: alternatives have been introduced, based on current noise increase.

\[ t_{ox} = 4.3 \text{ nm} \]

**NO SIGNIFICANT EFFECT ON FET CHARACTERISTICS AT THE MOMENT OF SBD**

Circuits will generally work after SBD!

Gerrer et al., ESSCIRC 2009

Kaczer et al., IRPS, p. 79 (2004)

Weir et al., IEDM, p. 73 (1997)
OVERVIEW

<table>
<thead>
<tr>
<th>NAME</th>
<th>Conduction model</th>
<th>physical damage</th>
<th>Observable properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>SILC (=A-mode SILC)</td>
<td>trap-assisted</td>
<td>single trap</td>
<td>- no noise change</td>
</tr>
<tr>
<td></td>
<td>tunneling</td>
<td></td>
<td>- RTS behavior</td>
</tr>
<tr>
<td>Micro BD (=pre-BD,</td>
<td>trap-assisted</td>
<td>cluster of two</td>
<td>- no noise change</td>
</tr>
<tr>
<td>anomalous SILC)</td>
<td>tunneling</td>
<td>traps</td>
<td>- RTS behavior</td>
</tr>
<tr>
<td>Soft BD (=quasi BD,</td>
<td>localized</td>
<td>percolation path</td>
<td>noise increase</td>
</tr>
<tr>
<td>partial BD,</td>
<td>conduction over</td>
<td>with limited</td>
<td></td>
</tr>
<tr>
<td>B-mode SILC)</td>
<td>lowered barrier</td>
<td>wear out</td>
<td></td>
</tr>
<tr>
<td>Hard BD</td>
<td>resistor (+diode)</td>
<td>percolation path</td>
<td>very large current increase</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with thermal</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>damage</td>
<td></td>
</tr>
</tbody>
</table>

PHASES OF DEGRADATION AND WEAR OUT IN THIN OXIDES

1. SBD: Weibull ($\beta_{SBD}, \eta_{SBD}$)
2. Wearout: assume Weibull ($\beta_{wo}, \eta_{wo}$)
3. Runaway $R_S$-limited
CONDUCTION PATH + THERMAL DAMAGE

- Mass transport and Si regrowth in the breakdown spot (Lei Jun Tang et al., TDMR 4(1), 2004)
- Localized current flow through resistor (+diode) (Kaczer et al. TED 49(3) 2002)
- This is a hard breakdown

HARD BREAKDOWN AND MASS TRANSPORT CONFIRMED BY AFM AND TEM

Porti et al., TED 50, p. 933 (2003)

Tang et al., TDMR 4, p. 38 (2004)
SEARCHING MORE RELIABILITY MARGIN

At breakdown: leakage current is still low (soft breakdown)
Leakage path is wearing out as a function of time
Delay time is available until the leakage current reaches unacceptable levels

F. Monsieur et al, IRPS2002

PROGRESSIVE BREAKDOWN

Wearout of breakdown path progresses gradually, rate is lower at low stress voltages

Hosoi et al, IEDM p. 155, 2002
**IMPACT OF PROGRESSIVE BD ON RELIABILITY MARGIN**

Post-SBD degradation rate depends on voltage:

- >1mA/s at 4V
- <1nA/s at 2V
- Extrapolations to 10fA/s at 1.2V

More than 10 years to grow BD-path from SBD to >10 µA at 1.2V: gives us again reliability margin

---

**OXIDE BD IS OBSERVABLE WITH EMISSION MICROSCOPE**

“Satellite” spots are observed, identified as hot carrier emission by spectral and SPICE analysis
Logical functions are maintained even after many FET oxide (hard) breakdowns

Kaczer et al, IEDM 2000

Border case

For $R_{bd} > 1.5\,k\Omega$, logic function is maintained

Soft BD has ~ $\text{M}\Omega$ resistance $\rightarrow$ BD effect much smaller
WIDE SRAM AFTER HBD AT NFET DRAIN: STILL STABLE

$L = 0.13 \text{ um}$
$W_n = 1.3 \text{ um}$
$W_p = 2.5 \text{ um}$
$R_{GD} = 3.2 \text{ k}\Omega$
$R_{GS} = \infty \text{ k}\Omega$
$V_{DD} = 1.2 \text{ V}$

“NARROW” SRAM AFTER HBD: DISABLED!

$L = 0.13 \text{ um}$
$W_n = 0.25 \text{ um}$
$W_p = 0.5 \text{ um}$
$R_{GD} = 3 \text{ k}\Omega$
$R_{GS} = \infty \text{ k}\Omega$
$V_{DD} = 1.2 \text{ V}$
“NARROW” SRAM AFTER SBD: STABLE

Good news: SBD is expected prevalent failure mode

EFFECT OF BD ON SRAM STATIC NOISE MARGIN

50% reduction in Static Noise Margin (SNM) for BD current >50μA

Reliability specification becomes strongly dependent on application: requires interaction with designers!
OUTLINE

1. Introduction
2. Basic reliability concepts
3. Time-dependent Dielectric Breakdown (TDDB)
4. Hot carrier degradation
   - Hot carrier injection
   - Effects of hot carrier injection
   - Acceleration models and examples
   - Dynamic degradation
5. Bias-Temperature Instability (NBTI)

DEFINITION OF “HOT CARRIERS”

“Hot carriers” are channel electrons and/or holes with high energy

- Generation: high lateral electric field at drain side
- The distribution of the carrier energy can be approximated by a Fermi-Dirac or Maxwell-Boltzman distribution with a high energy tail (high electron temperature)
- Hot carriers are always present when a transistor operates in saturation!
HOT CARRIER INJECTION

- Hot carrier = carrier that has gained large energy in a **high electric field**
- Hot carrier is **NOT** in equilibrium with the crystal: energy distribution is **NOT** Maxwell-Boltzmann distribution
- This means that hot carriers can be injected at much **smaller oxide fields** than tunneling injection
- Once the electrons are in the oxide conduction band they travel with a relatively high mobility to the anode

![Diagram of hot carrier injection]

SUBSTRATE HOT ELECTRON INJECTION

- Electrons injected from an external injector junction or by illumination are heated in depletion layer and injected into the gate
- Separate control of oxide current density, oxide field and silicon field (i.e. carrier field)
  - Oxide field controlled by $V_g$
  - Electron energy controlled by $V_{sub}$
  - Gate current controlled by substrate current
- Allows to measure oxide degradation and breakdown at fields considerably lower than 9MV/cm (i.e. closer to normal operation conditions)
- Use p-channels for hole injection
HOT CARRIER INJECTION

High-energetic carriers are generated in the high lateral electric field at the drain side of the MOS transistor

\[ E_m = \frac{V_d - V_{dsat}}{l} \]

\[ V_{dsat} = \frac{(V_g - V_i) \times L_{eff} \times E_{sat}}{(V_g - V_i) + L_{eff} \times E_{sat}} \]

\[ l = 0.22 \times t_{ox}^{1/3} \times x_j^{1/m} \]

\[ l = 0.017 \times t_{ox}^{1/8} \times x_j^{1/3} \times L_{eff}^{1/5} \]

Large electric fields at drain lead to high energetic carrier generation
Impact ionization causes electron-hole pairs, electrons are flowing to the drain, holes are flowing to the substrate. Hot electron generation and injection causes gate current.

**SIMPLIFIED EQUATIONS SUBSTRATE CURRENT**

Substrate current: 
\[ \text{I}_{\text{sub}} = I_s \cdot C \cdot \exp\left( -\frac{\phi_i}{q \lambda_e E_m} \right) \]

Multiplication factor: 
\[ M = \frac{I_{\text{sub}}}{I_s} = C \cdot \exp\left( -\frac{\phi_i}{q \lambda_e E_m} \right) \]

\[ \ln\left( \frac{I_{\text{sub}}}{I_s} \right) = \ln C - \frac{\phi_i}{q \lambda_e E_m} = \ln C - \frac{I \cdot \phi_i}{q \lambda_e (V_d - V_{\text{dsat}})} \]

Multiplication factor \( I_{\text{sub}}/I_s \) is a measure for the lateral electrical field \( E \).

\( \phi_i \) = impact ionization energy threshold (typically 1.5-1.6 eV)

\( \lambda_e \) = mean free path of electron in Si (typically 9.2 nm), T-dependent

For increasing T: \( \lambda_e \) becomes smaller: less hot carriers generated!
Measured gate currents:
At large $V_g (\sim V_d)$:
  negative gate current: electron injection

For small $V_g$ (near $V_t$):
  positive gate current: hot hole injection

Currents scale exponentially with drain voltage

Injection current vs. gate current: influence of gate voltage
HOT CARRIER INJECTION: ROLE OF OXIDE FIELD (GATE VOLTAGE)

When $V_g < V_d$: repulsive field for electrons exists at the drain. Large injection current is repelled back into the substrate.

When $V_g > V_d$: all injected electrons are traveling to the gate: e-current lateral field is usually too low for hole injection, holes are pushed back into channel.

When $V_g < V_d$: most injected electrons are pushed back into channel: no e-current, all injected holes will travel to the gate: h-current.

HOT CARRIER INJECTION FOR PMOS

Drain, substrate and gate currents for a pMOS device.
HOT CARRIER BASICS
A. Bravaix, IRPS Tutorial 2014

Substrate current is determined by product of drain current and multiplication current

Measured gate current is determined by generated hot carriers and attractive or repelling oxide field at the drain

Both electron and holes can be injected, depending on the oxide (gate voltage) and lateral silicon (drain voltage) fields.

HOT CARRIERS (HC) CAUSE MANY EFFECTS

HC generate electron-hole pairs by impact ionization
- carrier energy > 1.6eV
- generated current is measured as a bulk current (I_b)

HC can be injected into the gate oxide
- damage to the channel - gate oxide interface and electron trapping in the gate oxide → transistor degradation
- Carrier energy > 3.1eV (for electrons), > 4.8eV (for holes)
- I_{inj} \sim I_b
- Can also be measured as I_g (under specific conditions)

HC cause also other phenomena
- Light emission
- Secondary effects (e.g. too high I_{sub} can cause CMOS latch-up, …)
Change in Id-Vg characteristics and transconductance

After stress asymmetric Id-Vd characteristics (different when S-D are reversed):
- when damage is located at drain side: damaged area is depleted in saturation
- S-D reversed: damage is located at source side, influence I-V characteristic for all voltage

In linear region: characteristics are symmetrical but degraded
**NMOS DEGRADATION**

Gate voltage dependence of the nMOS hot carrier degradation

- $V_g$ (V)
- $V_{dd}$ = 6.5 V
- $t_{stress}$ = 3600 s

Maxima:
- $\Delta I_{CP}$ = 250 %
- $\Delta V_{CP}$ = 1.24 V
- $\Delta V_t$ = 15 mV
- $\Delta M$ = 59 %

**PMOS DEGRADATION**

Gate voltage dependence of hot carrier degradation for p-MOSFET's
INFLUENCE ON PMOS PERFORMANCE

Fixed negative oxide charge gives rise to a channel shortening effect (magnitude dependent on short channel effect) ⇒ increase of drain current

Interface trap generation causes a decrease of $I_d$

Since $\lambda_e > \lambda_h$ and because of smaller electric fields, hot-carrier degradation in pmos normally is no problem.

For technologies down to 0.5μm, the channel shortening dominates and screens the effect of $D_{it}$-generation.

For deep-submicron technologies, the effect of interface trap generation starts to be visible.

HOW TO DEAL WITH HC DAMAGE?

Hot carriers are always generated when a transistor operates in saturation
- Hot carrier damage cannot be avoided! It can only be limited.

HC damage is not allowed to affect circuit operation during the lifetime of the product
- Specification to guarantee this (simplified).
  - $\Delta I_{dsat} < 10\%$
  - After one year of DC operation
  - At $V_d = V_{dd} + 10\%$
  - “HC lifetime for the parameter $I_{dsat}$ must be one year under worst case operating conditions”
- This specification has to be met by every individual transistor on the chip.
- This specification is validated by many years of practice.
HOW IS THE HC LIFETIME MEASURED?

Accelerated Stress

- HC damage is proportional to the amount of hot carriers
- higher Vd = more hot carriers \(\rightarrow\) the lifetime is lower at higher Vd
  - C018 example:
    - \# hot electrons (Vd = 3.5V) \(\approx 10^6\) \# hot electrons (Vd = 2.0V)
- measurement times per transistor: 20min – 4 days, typically 3 hours

ACCELERATION MODELS

Degradation law:
\[
\Delta N_{it}(t) = C_1 \cdot \left[ \frac{I_d}{W} \cdot t \cdot \exp \left( -\frac{\phi_{t,e}}{q \lambda E_m} \right) \right]^n
\]

Multiplication factor:
\[
M = \frac{I_{sub}}{I_d} = C_2 \cdot \exp \left( -\frac{\phi}{q \lambda E_m} \right)
\]

\[
\frac{\tau I_d}{W} = C_3 \cdot \left( \frac{I_{sub}}{I_d} \right)^{-\phi_{t,e}/\phi} = C_3 \cdot M^{-m}
\]

\[
\tau = C_4 \cdot \exp \left( \frac{B}{V_d} \right)
\]

Model of Hu et al.  
Model of Takeda et al.
0.18 μm DEGRADATION

All degradation monitors degrade with power-law in time
$$\Delta I_{D_{lin}}/I_{D_{lin}} = a \cdot t^m$$, with $m=0.7$ to 0.3

---

THE HC LIFETIME AT WORK CONDITIONS IS EXTRAPOLATED FROM THE ACCELERATED STRESS Model of Hu

- Hu plot
  - X-axis: $I_b / I_d$
  - $I_b$ ~ # hot carriers
  - $I_d$ ~ # available carriers
  - Y-axis: $\frac{\tau \cdot I_d}{W}$
  - $\tau$ = lifetime (s)
  - $W$ = transistor width (μm)

- Power Law
  \[ \frac{\tau \cdot I_d}{W} = C \left( \frac{I_b}{I_d} \right)^{-\phi_{i_{t,c}}/\phi_i} \]
  - Slope $-\phi_{i_{t,c}}/\phi_i$ corresponds to the energy needed
to create an interface trap
  - $\phi_{i_{t,c}}$ = energy needed to create an interface trap
  - $\phi_i$ = energy needed to create an electron-hole pair

---

Measurements at accelerated stress

Extrapolated: $\tau = 100$ year
Specified: $\tau = 1$ year

$\tau \cdot I_d$ at worst case working conditions
LIFETIME PLOT

A more "physically" based lifetime evaluation procedure

\[ \tau \frac{I_d}{W} = C \left( \frac{I_{\text{sub}} - \phi_{i,t,e}}{I_d} \right) \]

\( \phi_i \): impact ionization energy
(\( \sim 1.3eV \))

\( \phi_{i,t,e} \): energy needed for creation of an interface trap
(\( \sim 3.5eV \))

\[ I_{\text{lin}} \text{ and } g_m \text{ are the device lifetime limiting parameters.} \]
Lifetime of more than 10 year is reached at 1.8V Vdd
HC LIFETIME OF A 90NM TECHNOLOGY

Wafers finished after ILM

Extrapolated: \( \tau = 100 \text{ year} \)
Specified: \( \tau = 1 \text{ year} \)

\( \tau / |Id / W (C/\mu m) | \)

Isub / Id at worst case working conditions

HC lifetime of this 90nm technology = ~100 years >> 1 year spec!

IMPACT OF BACK END PROCESSING

Back End Processing consists of
- Deposition of the Pre-Metal Dielectric (PMD), Inter-Metal Dielectric (IMD) and Passivation dielectrics
- metallisation of Contact, Single Damascene, Dual Damascene and Passivation levels
IMPACT OF BACK END PROCESSING

Each dielectric layer can affect the gate oxide and the Si-gate oxide interface quality (and thus the HC lifetime)

- Their deposition can create interface states
- Their presence can block efficient sintering
- Their deposition can diminish the amount of interface states
- Their presence can avoid that subsequent process steps to create interface states (they act as a barrier layer)

→ Effect of dielectric layers is very difficult to predict
→ Case by case analysis necessary!

Can metallisation steps affect the interface quality?

- There is no clear evidence gathered at IMEC
- Some reports exist in literature, e.g. on Al-metal etch

INFLUENCE OF BACKEND PROCESSING

Backend processing can have a tremendous influence on the hot carrier lifetime!

0.5µm CMOS (DLM, TLM) SOG planarization
ILD=interlevel dielectric
IMD=intermetal dielectric

Van den bosch et al., ESSDERC 1994
In digital circuits, transistors are only stressed during switching transitions: calculated as a “duty cycle”

\[
\tau = \int_{t_0}^{t_1} \frac{V_{DS}}{C_{DS} + C_L} \mathrm{d}t
\]

LIFETIME PREDICTION AC CONDITIONS

\[
\tau = C \left[ \frac{T}{T} \int_{0}^{T} I_{sub}^{m-1} dt \right]^{-1}
\]

\( I_{sub} = I_d \)
OUTLINE

1. Introduction
2. Basic reliability concepts
3. Time-dependent Dielectric Breakdown (TDDB)
4. Hot carrier degradation
5. Bias-Temperature Instability (NBTI)
   - Basics
   - Acceleration model
   - Physical models
   - Frequency dependence
   - Impact of processing

INTRODUCTION

Negative bias temperature instabilities

CMOS inverter

V_{th} shifts and reduction of I_{Dsat} occur when p-MOSFET is switched on:
Negative Bias Temperature Instabilities
CAUSE OF THRESHOLD VOLTAGE $V_{th}$ SHIFT ($\Delta$): GENERATION OF INTERFACE AND OXIDE TRAPPED CHARGE

$$\Delta V_{th} = \frac{\Delta Q_{ox} + \Delta Q_{ot}}{C_{ox}} = \frac{q\Delta N_{eff}}{C_{ox}}$$

$$\Delta Q_{ox} = q \Delta N_{o}$$

$$\Delta Q_{ot} = \frac{1}{t_{ox}} \int_{0}^{t_{ox}} (t_{ox} - z) \Delta \rho(z) dz$$

Note: sometimes used

$$\Delta N_{eff} = \frac{C_{as} \Delta V_{th}}{q}$$

$c_{ox}$ ... Oxide capacitance per area [Fcm$^{-2}$]

$t_{ox}$ ... Oxide thickness [cm]

$Q_{it}$ ... Interface charge density [Ccm$^{-2}$]

$N_{it}$ ... Interface state density [cm$^{-2}$]

$Q_{ot}$ ... Oxide trapped charge density [Ccm$^{-2}$]

$\rho(z)$ ... Oxide trapped charge density [Ccm$^{-3}$]

BTI: CHARGED OXIDE BULK AND INTERFACE STATES

<table>
<thead>
<tr>
<th>NBTI (pFET)</th>
<th>PBTI (nFET)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive oxide charge</td>
<td>Negative oxide charge</td>
</tr>
<tr>
<td>(negative $V_{th}$ shift)</td>
<td>(positive $V_{th}$ shift)</td>
</tr>
<tr>
<td>Interface states</td>
<td>X</td>
</tr>
<tr>
<td>(mobility, transconductance, subthreshold-</td>
<td>(but observed in sub 1nm EOT)</td>
</tr>
<tr>
<td>slope degradation)</td>
<td></td>
</tr>
<tr>
<td>Mitani et al., ECS 2005</td>
<td></td>
</tr>
<tr>
<td>Garros et al., ICICDT 2010</td>
<td></td>
</tr>
</tbody>
</table>

Generally at substrate interface

Zafar et al., VLSI 2006
Kerber et al., TED 55, 3175 (2008)
Aoulaiche et al., IRPS 2009
NEGATIVE BIAS TEMPERATURE INSTABILITY

Occurs at relatively low voltage/fields and high temperatures
Mostly pMOS problem (negative \(V_g\))
Problem exists from early days of MOSFET technology!
Damage not related to charge transport through dielectric
Interpretation: fixed positive charge and interface trap generation (acceptor in upper, donor in lower half of bandgap)

Jeppson, Svensson,
JAP 1977

ORIGIN OF NBTI SHIFTS

Other interpretation:
- no positive charge, no \(V_{fb}\) shift
- donor type interface traps

Reddy et al, IRPS 2002
INTERFACE STATE DENSITY GENERATION OBSERVED BY CHARGE PUMPING

\[ \Delta N_a(t) = \frac{\Delta I_{CP,\text{max}}}{qA_f} \]

e.g. M. Aoulaiche et al., Microelectron. Eng. 80, 134-137 (2005).

Note: also gives \( V_{th} \) and \( V_{FB} \) shifts

ORIGIN OF NBTI SHIFTS

M. Alam et al., IEDM 2003

1:1 Correlation between threshold voltage shift and interface trap generation indicates that \( D_{it} \) generation is dominant!

V. Huard et al., IRPS 2003

But \( V_t \)-shift and \( D_{it} \) increase exhibit different activation energies: no consensus!
TIME DEPENDENCE

Threshold voltage shift and Dit-generation show power law dependence with slope $n=0.2-0.3$

$\Delta V_T = A t^n$

OXIDE THICKNESS DEPENDENCE

Normalized $V_t$ shift as a function of $E$ shows universal behaviour for all oxide thicknesses.

Data show power law field dependence: $\sim E^b$
ACCELERATION MODELS

\[ \Delta V_T = C \cdot \exp\left( -\frac{E_a}{kT} \right) \cdot \exp(\beta|V_G|) \cdot t^n \]

Voltage/field acceleration: no consensus
- linear (\( \sim E \))
- power law (\( \sim E^n, n=1.5 \))
- exponential (\( \sim \exp(E) \))
- \( 1/Vg \)

Temperature acceleration: consensus
- Arrhenius

EXAMPLE OF BTI LIFETIME EXTRAPOLATION

- Failure Criterion: \( \Delta V_{th} = 30 \text{mV} \) at \( T = 125 \text{ C} \)
- Minimized relaxation: \( t_{\text{sense, delay}} = 2 \text{ms} \)
- Stress experiment at different voltages (\( V_{\text{Gstress}} > V_{DD} \)) \( \rightarrow \)
  Estimate lifetime \( \rightarrow \) Power Law extrapolation at 10 Years
ACCELERATION MODEL: TEMPERATURE

Activation energies are usually in the order of 0.12-0.15 eV

NBTI DEGRADATION—SEEMINGLY STRAIGHTFORWARD MEASUREMENT

\[ \Delta V_{th} = A t^{b} \]

\( T = 125 \, ^{\circ}C \)

\( V_{G,\text{stress}} = -2 \, V \)

\( V_{th,\text{init}} \)

\( V_{th} \)

\( f_{stres} \)

\( \Delta V_{th} = V_{th} - V_{th,\text{init}} \)

\( I_{D} \)

\( V_{G} \)

\( V_{G} = -2V \)

10 s

20 s
“MINOR” COMPLICATION: THRESHOLD VOLTAGE SHIFT STARTS RELAXING IMMEDIATELY AFTER STRESS REMOVED

Crucial issue for all BTI measurements and interpretation!
Observed in p/nFETs, P/NBTI, DC and AC stress...

Kerber et al., TED 55, 3175 (2008); Kaczer et al., IRPS 2008

A AND B DEPEND ON $T_{MEAS}$

$\Delta V_{th} = A t_{stress}^b$

$A$ decreases while $b$ increases with increasing $t_{meas}$
MEASURING BTI

• Several methods developed to compensate relaxation, based on
  ▶ Measure-Stress-Measure
  ▶ On-the-fly
• No single “silver bullet” approach to measuring BTI
• Relaxation time needs to be considered otherwise $\Delta V_{th}$ measurement meaningless
• JEDEC standards give some guidance
• Rule of thumb: when comparing stacks/processing, use the same measurement setup (fields, timing, etc)

FAST MEASURE-STRESS-MEASURE (MSM) TECHNIQUE

...devised to speed up BTI threshold voltage shift measurements

Measurement speed increased to ~ 0.1 s
**MOST POPULARIZED EXPLANATION FOR NBTI:**

**REACTION-DIFFUSION MODEL:** DIFFUSION OF HYDROGEN IS THE RATE-LIMITING STEP

- All effects explained by interface states generation ($\Delta N_i$):
  - Channel holes break hydrogen bonds passivating interface state
  - Hydrogen released, diffuses through $\text{SiO}_2$
  - Repassivation of interface states during relaxation

M. A. Alam, IEDM 2003

**REACTION-DIFFUSION MODEL**

= Most popular models for NBTI

$$Si_3 \equiv SiH + X \xrightarrow{k_i} Si_3 \equiv Si^\bullet + Y$$

$Y_{\text{interface}} \rightarrow Y_{\text{bulk}}$

Model orginally proposed by Jeppson and Svensson, 1977, further elaborated by Ogawa et al, 1995 and recently by many other groups
- Si-H bonds interact with holes (species X) and are broken (reaction)
- Result: Si dangling bond (interface trap)
- H-related species Y diffuses away from the interface (diffusion): $H_2$, H, $H^\bullet$, OH
- H-species diffusion is rate limiting step: explains power law slope $n=0.25$

S. Ogawa et al., JAP **77**, 1137 (1995)
REACTION-DIFFUSION MODEL

Generation Annealing

\[
\frac{\partial N_{H}}{\partial t} = k_{1}(N_{0} - N_{H}) \cdot p - k_{2}N_{H}N_{H}(x, t)
\]

Reaction

\[
N_{H}(t) = \int_{0}^{\infty} N_{H}(x, t)dx
\]

Diffusion

\[
N_{0} = \text{Initial SiH bonds}
\]

Early stress phase:
Reaction rate \(k_{1}\) is dominant

Later stress phase:
Diffusion rate is dominant

M. Alam et al, IEDM 2003

---

Using the reaction-diffusion model, Alam proposed that partial \(V_{th}\) recovery is due to the passivation of \(P_{B0}\) centers by \(H_{2}\) (or H) that diffuses back towards the Si/SiO\(_{2}\) interface when the gate bias is switched or set to zero.

M. Alam, IEDM 2003, 345
REACTION-DIFFUSION MODEL

M. Alam et al, IEDM 2003

R-D model allows good simulation of stress/relaxation cycles

PREDICTIONS OF THE CLASSIC R-D MODEL

Stress (kf) Relax (kf=0)

$\Delta = f \left( G_t, R \right)$

stress

relaxation

$\Delta = f \left( 0, R \right)$

$\Delta V_{th} = A t^{\frac{1}{3}}$

$\Delta V_{th}$

$\Delta V_{th}$

Intricate balance between generation (relaxation)
PROPERTY OF BTI RELAXATION: RELAXATION TRANSIENTS CAN BE RESCALED TO THE SAME UNIVERSAL RELAXATION CURVE

- $\xi = \frac{t_{\text{relax}}}{t_{\text{stress}}}$
- $R(t_{\text{relax}}) / S(t_{\text{stress}})$

Grasser et al., IEDM 2007

DISPERNSIVE NBTI RECOVERY: LONG LOGARITHMICAL RELAXATION TAIL WELL DESCRIBED

- $\Delta V_{th} = \frac{1}{1 + \frac{\alpha}{2}}$

Kaczer et al., IRPS 2005

$\xi = \frac{t_{\text{relax}}}{t_{\text{stress}}}$ is the dimensionless relaxation time
NEW, “EXTENDED” MSM TECHNIQUE: SEVERAL DECADES OF RELAXATION TREND RECORDED AFTER EVERY STRESS

Advantage: can be done with standard off-the-shelf equipment

MEASURING $V_{TH}$ SHIFT (2)

Flexible: 1) fast measurement, or 2) the effect of recovery can be studied
BTI DEGRADATION RELAXES “DISPERSIVELY” OVER MANY TIME SCALES: MICROSECONDS OR LESS TO DAYS OR MORE

Improvement of a few decades negligible on the logarithmic scale

TYPICAL EXTENDED MSM MEASUREMENT: RELAXATION TRENDS RECORDED INSTEAD OF SINGLE POINTS

The effect of techniques measuring at different speeds (Fast MSM, DC-I-V) can be looked up on this graph (allows “fallback”)

HfSiO pFET

\[ T = 125 \, ^\circ\text{C} \]

\[ V_{stress} = V_{meas} - 1.8\text{V} \]
RELAXATION NOT COMPLETE: PERMANENT DEGRADATION

\[ \Delta V_{th} = R(t_{stress}, t_{relax}) + P(t_{stress}) \]

BTI RELAXATION IS PRESENT IN MOST SYSTEMS

SiON PFET PBTI
\[ V_{stress} = +2V \quad T = 125^\circ C \]

SiON NFET PBTI + SHCI
\[ V_{stress} = +2V \quad T = 25^\circ C \]

SiON pFET AC NBTI
\[ V_{stress} = -2V \quad T = 125^\circ C \]

High-k PFET NBTI
\[ V_{stress} = -2.3V \quad T = 125^\circ C \]
SCHEMATIC SUMMARY: PREVIOUS VS. NEW ANALYSIS

Recorded relaxation transients carry additional information, used to separate recoverable and permanent components.

TYPICAL RESULT OF EXTENDED MSM ANALYSIS

Recall: total $\Delta V_{th} = R(t_{stress}, t_{relax}) + P(t_{stress})$
PRESENT UNDERSTANDING OF THE NBTI COMPONENTS

Recoverable component “R”:
- Interface states ($N_{it}$)
  - Cannot be due to discharging of interface states (too fast)
  - Cannot be due to R-D passivation of interface states (long relaxation tails)
  - Cannot be due to R-dispersive Diffusion passivation of interface states
- Hole trapping ($N_{ot}$)
  - Elastic tunneling too fast in thin oxides
  - Most likely after inclusion of inelastic phonon component

Permanent component “P”:
- Hole trapping ($N_{ot}$)
  - Possibly not fully permanent but really long relaxation
- Interface states ($N_{it}$)
  - Most likely component

PFET AC NBTI: (FAST) SEQUENCE OF STRESS AND RELAX PHASES

- $V_{select} = 2$ V
- $V_{select} = 0$ V
- $V_G = 0$ V
- $V_{select} = 0$ V
- $V_{select} = 2$ V

- $T = 125$ C
- $t_{stress} = 1000$ s
- $V_S = V_D = V_B = V_{cc} = 2$ V
- $V_{select} = 2$ V
- $V_{select} = 0$ V
- $V_G = 0$ V

- AC degradation ~ DC degradation / $2$
- Frequency (Hz)

- AC BTI degradation lower than DC
- $f$-independent up to GHz ⇒ no “cut-off” time scale down to ~1 ns
(110) AND (100) SURFACE ORIENTATIONS

- Higher electron mobility
- Higher hole mobility
- Higher density of Si-H bonds


(110) SURFACE ORIENTATION DEGRADED MORE BY NBTI

M. Aoulaiche et al., SISC, 2006
INFLUENCE OF SIDEWALL ORIENTATION

NBTI was reported to be slightly worse for MuGFET devices due to (110) sidewall orientation.

IMPROVEMENT DUE TO DEUTERIUM AND FLUORINE

\[ Si_D = Si - D \approx Si_H = Si - H \]

Difference in the dissociation dynamics multiple-vibrational excitation [Van de Walle et al.]

\[ Si - H \approx 2.2eV \]

\[ Si - F \approx 5.8eV \]

M. Houssa et al., Electrochemical and Solid-State Letters, 2006
INCREASED NBTI DEGRADATION DUE TO NITROGEN INCORPORATION

NBTI has become important again with the introduction of nitrided oxides: $V_{th}$ shifts and $I_{dsat}$ reduction are much enhanced in oxynitride layers, the more so when N content at the Si/dielectric interface increases.

M. Aoulaiche et al., JVST B, 2009

NBTI IN OXYNITRIDE LAYERS

Consensus more or less reached on the cause of NBTI acceleration in nitrided oxides:

presence of nitrogen-induced defects can trap holes and/or protons, leading to the generation of positive charges in the oxide:

$$Si_2 = N - Si + h^+ \rightarrow (Si_2 = N - Si)^+$$

$$Si_2 = N - Si + H^+ \rightarrow (Si_2 = NH - Si)^+$$

Consequence: Larger $V_{th}$ shift when [N] increases

N. Kimizuka et al., 2000 VLSI Symp., p.92.
K. Kushida-Abdelghafar et al., APL 81, 4362 (2002)
S.S. Tan et al., APL 82, 1881 (2003)
S. Fujieda et al., APL 82, 3677 (2003)
M. Houssa et al., JAP 95, 2786 (2004)
NITROGEN INDUCES ADDITIONAL HOLE BULK TRAPS

CP

SILC

Nitrogen related defects

IMPACT OF NITROGEN—RECOVERABLE COMPONENT ONLY

Same oxide electric field

log stress time

log degradation (ΔVth)

M. Aouaiche et al., JVSTB, 2009
IMPACT OF HfO$_2$ THICKNESS: MECHANISM CONTROLLED BY ELECTRIC FIELD AT SI/SiO$_2$ INTERFACE

NBTI lifetime $\Delta N_{\text{eff}}$ $\Delta V_{\text{th}}$ from IV-pulse

NBTI scales with $E_{\text{ox}}$ at Si/SiO$_2$ interface
$\rightarrow$ NBTI increases for thinner HfO$_2$ due to higher $E_{\text{ox}}$ at the Si/SiO$_2$ interface

BTI IS A SERIOUS RELIABILITY PROBLEM

Example: Tolerable NBTI overdrive voltage for 10 years operation drops to zero in sub-1-nm EOT region
Lower activation energy for 0.58nm EOT devices: more direct tunneling of holes from the channel into the oxide

Minimum interfacial layer thickness is required to maintain reliability!

Minimum of 0.4nm interfacial layer needed to prevent sudden increase of the NBTI degradation consistent with ab-initio modeling by Monaghan et al. (Phys. Rev. B, 2007)
NBTI IMPROVEMENT ON Ge-SUBSTRATE pFET

J. Franco et al., accepted to IRPS 2010

Si CAP THICKNESS
THINNER CAP IMPROVES NBTI RELIABILITY

- Si cap thickness has a strong impact on NBTI
- **Thinner** Si cap → Reduced NBTI
THIN SI CAP IMPROVES NBTI LIFETIME WHILE SCALING TINV

Franco et al., IEDM 2010

Only solution so far: switch to SiGe substrates!

SiGe→LOWER $\Delta N_{OT}$
MODEL: REDUCED INTERACTION WITH OXIDE DEFECTS

Thin Si cap: channel carriers ($E_F$) ‘miss’ part of the oxide defects

- Less defects, located on the gate side, are energetically favorable for holes
- Stronger field acceleration due to ‘unaccessible’ traps becoming ‘accessible’
**CONCLUSIONS**

1) Reliability becomes and more and more critical when downscaling technologies and needs to be taken into account in technology and material choices

2) TDDB, BTI and HCD are the three most critical device related failure mechanisms, lot of progress in the understanding of these failure mechanisms has been obtained in the past 3 decades

3) TDDB and HCD is under control provided the necessary technology measures are taken (e.g. corner rounding for FinFET’s, drain engineering techniques)

4) NBTI reliability becomes a showstopper for Si based pMOS devices with sub-1nm EOT

5) (Si)Ge/SiO$_2$/HfO$_2$ pMOSFETs offer superior reliability w.r.t. Si for sub-1nm EOT nanoscaled pMOSFETs, optimized stack has been identified