Microelectronics Reliability 88–90 (2018) 98–102

Contents lists available at ScienceDirect

Microelectronics Reliability

journal homepage: www.elsevier.com/locate/microrel

In-situ transistor reliability measurements through nanoprobing

O. Dixon-Luinenburg, J. Fine*

Thermo Fisher Scientific, 827 Reddick Street, Santa Barbara, CA 93103, USA

ARTICLE INFO

Keywords:
Nanoprobing
Probing
NBTI
PBTI
BTI
In-situ characterization
Induced aging
Transistor aging
Transistor degradation
SRAM
AFM
HCI
TDDB
RTN
Operating lifetime
Localized variation
Intra-die comparison
Inter-die comparison

ABSTRACT

In this study we examine the feasibility of performing transistor reliability measurements with the Hyperion II nanoprobing system. Proof-of-concept bias temperature instability (BTI) measurements were run on a commercially available Intel 14 nm FinFET processor. BTI degradation was found to closely follow the expected power law over 10^3 s stress in total at 2 V with characterization done < 50 ms into recovery. Examination of 50 SRAM transistors with 30 s stress at 2 V yielded average I_{ON} reduction of 14.4% (σ = 6.6%) and 6.5% (σ = 2.5%) for pullups and pulldowns, respectively. The in-situ nature of the nanoprobing approach provides insight into transistor lifetime and performance as a function of layout as well as variations in aging between identically designed devices. This is a compelling reason to apply nanoprobing for a range of reliability measurements as a complement to the suite of established reliability testing techniques.

1. Introduction

As semiconductor fabrication moves to sub-10 nm technologies, transistor degradation mechanisms that had played relatively minor roles in the aging and reliability of past generations are proving to be formidable challenges in the design and development of leading semiconductor products [1–5]. Their prominence is largely a consequence of device scaling to dimensions roughly two orders of magnitude larger than the atoms that comprise them [6] and more stringent tolerances associated with reduced operating voltages [3]. Smaller features often lead to larger electric fields across critical transistor structures [7] and with fewer atoms comprising different structures it can be shown that random aspects in device fabrication will have a more pronounced impact on die/lot uniformity and chip aging [6]. The problem is further compounded by the introduction of new materials and device structures to keep up with Moore's law, which resolve some issues but often raise or complicate others [3, 5]. Thus, to achieve lifetimes in line with current standards when progressing to deeply scaled future product lines, a heavier burden is placed on process, design schemes, and reliability testing.

In terms of end-of-life reliability, frequently discussed aging mechanisms include Bias Temperature Instability (BTI), Hot Carrier Injection (HCI), Time Dependent Dielectric Breakdown (TDDB), and Random Telegraph Noise (RTN) [3, 7–9]. Reliability tests with large scale and scope such as high temperature operating lifetime (HTOL) tests can provide broad screening for such failures as well as a general sense of operational lifetime on a product [8], but convey little information regarding root cause in the event of individual device failure and are thus primarily used late in the development cycle. Though it has an important place within the full scope of reliability assessment there are clear disadvantages to this approach.

Dedicated test structures are widely used to assess reliability earlier in the product lifecycle but are expensive and may fail to capture the effects of layout strain or proximity to neighbouring circuitry [10, 11]. Such gaps often necessitate the use of simulations to glean useful information regarding aging of a specific functional circuit or subsystem on a given process technology [8, 9]. However, simulations require models with many assumptions and can fail to capture important behaviours, especially absent external validation. Within the suite of commonly applied reliability assessment techniques there appears to be

* Corresponding author.
E-mail address: jordan.fine@thermofisher.com (J. Fine).

https://doi.org/10.1016/j.microrel.2018.06.100
Received 30 May 2018; Received in revised form 18 June 2018; Accepted 27 June 2018

a gap in terms of providing direct reliability information while capturing complexities in behaviour stemming from the local environment of a transistor or functional subsystem.

Nanoprobing offers the ability to establish electrical contact with nodes of nanoscale devices as they exist on a product. DC biases or arbitrary waveforms within a broad bandwidth can be directly applied to the connected device [12] without the influence of large parasitics caused by interconnect layers or other front-end devices sharing access lines. The improved stability and corresponding time on contact of modern nanoprobing systems opens the possibility of running a broad suite of long-duration characterization measurements that to date have been out of reach.

Here we present results from BTI measurements on 14 nm FinFET SRAM and logic transistors using a Hyperion II Atomic Force nanoProber (AFP). BTI was chosen as a test case because of its simple measurement scheme, well documented behaviour, and significance in the development of next-generation chips.

1.1. Bias temperature instability

Negative Bias Temperature Instability (NBTI) and Positive Bias Temperature Instability (PBTI) are highlighted in the technology roadmap as key reliability concerns for PFETs and NFETs respectively [1]. BTI effects are named for their acceleration through elevated bias (B) and temperature (T) and manifest as an increase in threshold voltage (V_T) over time, thereby increasing transistor switching time and reducing ON-state current (I_{ON}) [13–15]. These effects can cause irreversible loss in functionality of subsystems that rely on specific current ratios or other precisely tuned characteristics, such as static random-access memory (SRAM) [5, 16]. While generally attributed to various forms of charge trapping in the channel or at the oxide interface, a consensus is absent in terms of a systematic theoretical framework for BTI [17]. It is, however, generally accepted that these phenomena will continue to be a problem in future CMOS technologies, and severity is expected to worsen with further device scaling [3, 7].

BTI effects on threshold voltage can be simply modelled by combining Arrhenius' law with power laws for electric field and stress time shown by Eq. (1) [15], where E_A is an activation energy, k_B is the Boltzmann constant, T is the stress temperature, t is stress time, γ and n are constants on the order of 2.5–3 and 0.1–0.25 respectively, and electric field has been expressed as the overdrive voltage $V_{OV} = |V_G - V_{T0}|$ divided by oxide thickness t_{ox}. A typical failure criterion for ΔV_T ranges from 30 mV [15] to 50 mV [14].

$$\Delta V_T \propto \exp\left(-\frac{E_A}{k_B T}\right)\left(\frac{V_{OV}}{t_{ox}}\right)^{\gamma} t_{stress}^{n} \quad (1)$$

In addition to the simple relationships presented in Eq. (1), it is well reported that BTI-stressed transistors exhibit partial recovery in V_T and I_{ON} with the recovery occurring on multiple timescales across several orders of magnitude and with 50% of recovery occurring in the first few seconds [14, 18]. Using the proportionalities shown in Eq. (1) and elevated-voltage stress one can extrapolate long-term aging under normal device operation and thus understand end-of-life characteristics. However, recovery must be considered, generally by measuring as soon as possible after stress.

2. Measurements

2.1. Sample and preparation

NBTI and PBTI measurements were run on a commercially available Intel Pentium G4400 Skylake Dual-Core CPU. Higher metal layers were removed in order to access electrical contacts of individual transistors at the local interconnect layer. Surface topography at the target layer was on the order of 10 nm to accommodate AFM topography imaging.

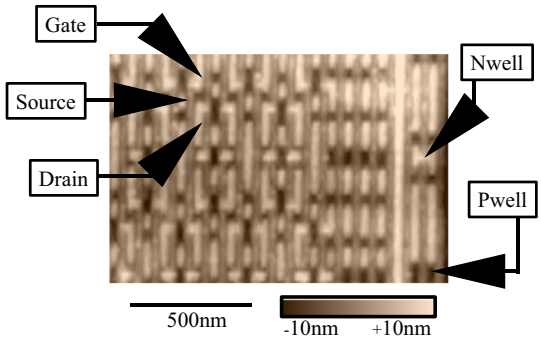

Fig. 1. AFM topography image of the edge of a 6 T SRAM array. Positioning of probes to measure a pullup PFET transistor with biasing for Source, Gate, Drain, Nwell, and Pwell nodes are shown with black triangles. The conductive chuck on which the sample sits when landing probes also serves as a node for measurements.

2.2. Measurement system and parameters

A Hyperion II AFP was used to establish electrical connection. Measurements were performed in a region near the corner of an SRAM array due to the authors' familiarity with the structure and transistor layout. A tapping-mode atomic force microscopy (AFM) image of the edge of the SRAM array is seen in Fig. 1. Images were captured with each of 8 electrically conductive AFM tips which were subsequently positioned onto Source, Gate, Drain, and Well contacts.

A Keysight B1500A Semiconductor Device Analyzer, hereafter referred to simply as the B1500, was used to apply DC biases and sweep/step sequences to the transistors. Nominal operating range was taken as 0–700 mV based on Intel publications regarding their 14 nm Tri-Gate transistor technology [19] and 6 different tests were run with duration and node voltages as described in Table 1. Current was sampled with 10 ms integration for I_{ON} measurements including that for recovery. Auto sampling at 101 points was used for current measurements in I_DV_G curves.

Stress times, voltages, and characterization methods were chosen to work within the capabilities of the nanoprobing platform, namely contact times on the order of 10^3 s, voltages within the range of ± 40 V, currents < 1 mA, and frequencies from DC to ~1 GHz.

With −2.4 V the permanent degradation from NBTI in a PFET is evident within a few minutes. The I_DV_G approach to characterizing transistor performance was used to provide a full picture of changes in DC behaviour whereas I_{ON} sampling was used to quickly characterize behaviour early in recovery revealing transient BTI effects. For these transient effects, ± 2 V stress was adequate to show clear trends from 10^0 to 10^3 s.

Table 1
Parameters for stress conditions and assessment measurements including measurement durations and biasing for Source, Gate, Drain, Nwell, Pwell, and Chuck nodes, with values listed in that order.

	Duration	Bias (V) S/G/D/N/P/C
PFET 2.4 V stress	10 to 500 s	0/−2.4/0/0/−2.4/float
PFET 2 V stress	400 ms to 1000 s	0/−2.0/0/0/−2.0/float
PFET I_{ON}	< 50 ms	0/−0.7/−0.7/0/−0.7/float
PFET recovery	200 s	0/−0.7/−0.7/0/−0.7/float
PFET I_DV_G	~30s	0/sweep 0 to −0.7/step −0.05, −0.7/0/−0.7/float
NFET 2 V stress	400 ms to 1000 s	0/2.0/0/2.0/0/float
NFET I_{ON}	< 50 ms	0/0.7/0.7/0.7/0/float