Assessing the Impact of RTN on Logic Timing Margin Using a 32nm Dual Ring Oscillator Array

Qianying Tang and Chris H. Kim

Dept. of ECE, University of Minnesota, 200 Union Street SE, Minneapolis, MN 55455, USA, Email: tangx280@umn.edu

Abstract

This paper presents a novel method for characterizing the impact of random telegraph noise (RTN) on logic timing margin under sub-0.5V supply voltages. The proposed dual ring oscillator array test structure improves the frequency measurement resolutions of the tested-and-proven beat frequency detection (BFD) technique by pairing a ROSC from one array with a ROSC from a second array having a similar frequency. Detailed circuit level RTN data was collected from a 32nm HKMG test chip, including voltage stress results. Measured data confirms that the proposed dual-array technique is effective in collecting high quality RTN statistics at sub-0.5V. The impact of RTN on logic timing margin is estimated based on the measured frequency data.

Introduction

Random telegraph noise (RTN) is becoming an increasing concern affecting critical circuit parameters such as SRAM read and write margins, analog circuit signal-to-noise ratio, and logic circuit timing margin. In particular, the logic gate delay is expected to become more sensitive to RTN effects when the circuit is operating under a Near Threshold Voltage (NTV) or Dynamic Voltage and Frequency Scaling (DVFS) environment. The temporal fluctuation of transistor threshold voltage (V_t) caused by RTN traps may result in logic errors due to fluctuation in both the clock period and logic gate delay. For a typical pipeline circuit shown in Fig.1 to operate without any logic errors, the clock period T_{clk} must be greater than $t_{clk\text{-}to\text{-}q} + t_{logic} + t_{setup} + t_{clk\_skew}$. Here, $t_{clk\text{-}to\text{-}q}$ is the flip-flop clock-to-q delay, t_{logic} is the logic path delay, t_{setup} is the flip-flop setup time, and t_{clk_skew}. is the clock skew. Fig. 1 shows three possible timing failure scenarios depending on the location of the RTN trap: (1) RTN in the clock driver increases t_{clk_skew} between two divergent clock paths therefore decreases the effective T_{clk}; (2) RTN in the logic path increases t_{logic}; and (3) RTN in the flip-flop increases $t_{clk\text{-}to\text{-}q}$ and t_{setup}.

Accurate characterization of RTN induced delay shift is of crucial importance, however, most of the reported RTN data are from individual device probing which provides only limited insight in to the circuit level RTN behavior. Deducing circuit parameters based on device I-V data is prone to error due to the complex circuit topology and fast signal switching. Recently, a ring oscillator (ROSC) based approach was used for RTN measurements wherein the beat frequency of two identical ROSCs was measured from a 65nm test chip [1]. One subtle but critical shortcoming of this method however is that the measurement resolution degrades sharply at low supply voltages due to the increased frequency variation. This makes it difficult to collect high-quality RTN statistics using the previous design. To overcome this limitation, this work proposes a dual ROSC array based test structure capable of achieving a frequency measurement resolution less than 0.01% for every single ROSC in the array, at supply voltages as low as 0.45V.

32nm Dual Ring Oscillator Array

The tested-and-proven beat frequency technique illustrated in Fig. 2 is adopted once again in the new RTN test structure. Here, a standard D-flip-flop (DFF) acts as a phase detector which continuously monitors the frequency difference between two free running ROSCs. The DFF output toggles from low to high whenever the two inputs rising edges are aligned, therefore the frequency of the DFF output signal (namely the beat signal) is equal to the frequency difference between the two ROSCs (i.e. $f_C = f_A - f_B$). The beat signal is digitalized by counting the number of ROSC cycles that fit within a single beat frequency period (i.e. $N = |f_B/(f_A - f_B)|$). The main reason this technique is effective in measuring minute RTN effects is because the measurement resolution increases exponentially when the two input frequencies f_A and f_B are closer to each other. However, in reality, the frequency difference between the selected ROSC and the reference ROSC cannot always be guaranteed to be less than ~1%, especially when measuring from a large array at low supply voltages such as 0.5V. As shown in Fig. 3 (left), the maximum frequency difference between ROSCs in the test array, and three reference ROSCs can be as high as 8% which limits the frequency measurement resolution to >0.6%. To overcome this limitation, we propose a dual-array test structure which ensures that a ROSC from main array can be paired with a ROSC from another array with a frequency difference less than 1%. This in turn guarantees a frequency measurement resolution less than 0.01% as shown in Fig. 4. As the number of reference ROSCs increases from 3 to 64, the worst case measurement resolution improves from 0.5% to 0.01%.

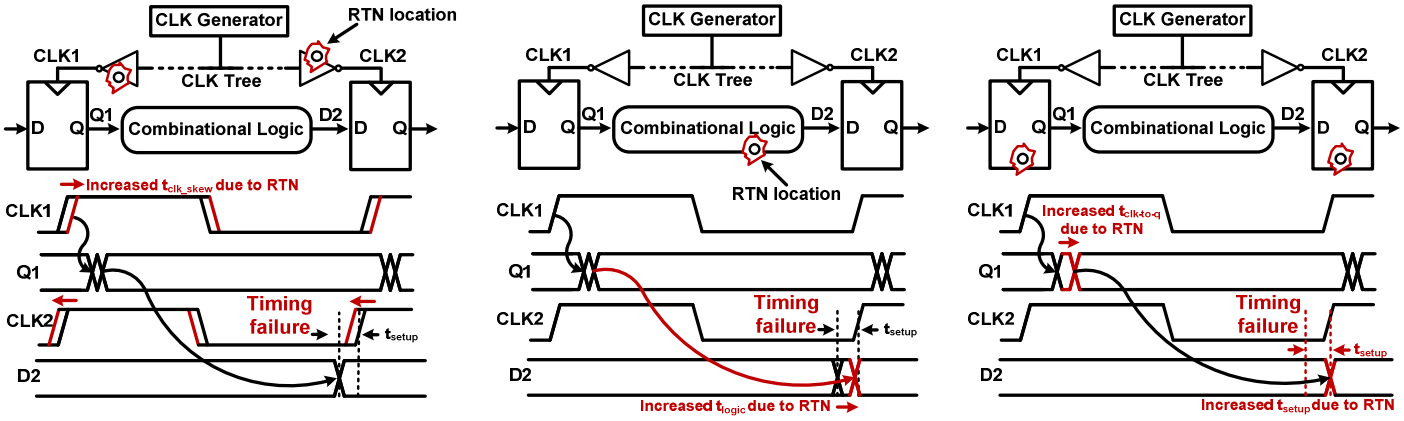


Fig. 1. Logic timing failures under different RTN locations; (left) RTN in clock tree; (middle) RTN in combinational logic; (right) RTN in flip-flop.

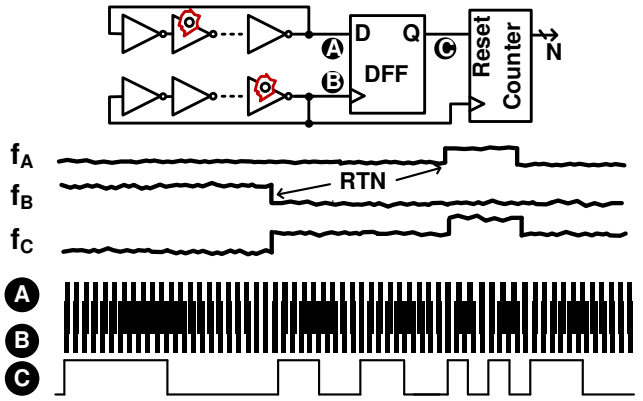


Fig. 2. Beat frequency detection circuit adopted in this work for measuring RTN induced delay shifts at sub-0.5V supply voltages with high resolution.

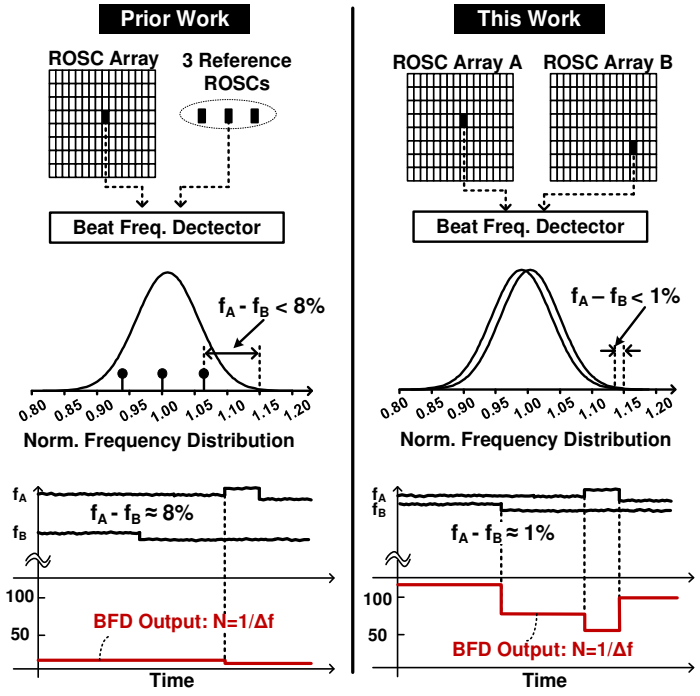


Fig.3. Comparison between prior work (3 reference ROSCs) and this work (utilizing an array of reference ROSCs). By pairing a ROSC from array A with a ROSC from array B having a similar frequency (e.g. Δf <1%), a frequency measurement resolution of 0.01% can be achieved for a wide frequency range.

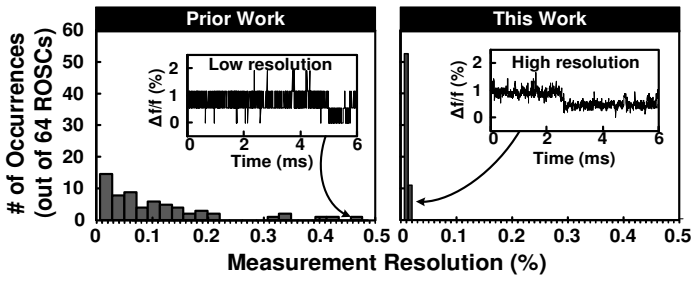


Fig. 4. Measurement resolution when pairing a 64 ROSC array with 3 (left figure) and 64 (right figure) reference ROSCs. A more precise waveform can be reconstructed using 64 reference ROSCs which is critical for collecting high quality RTN statistics at low supply voltages such as 0.5V.

The 32nm test chip consists of two ROSC arrays each comprised of 64 ROSCs, along with two separate beat frequency detectors to determine which of the two input frequencies is higher (Fig. 5). The beat frequency detector (BFD) operates at a nominal supply voltage to prevent any meta-stability issues induced by RTN. In the previous design [1], the three reference ROSC frequencies had to be manually trimmed before taking each measurement in order to achieve the desired resolution. In the proposed dual ROSC array, a ROSC in the main array is sequentially compared with each ROSC in the second array until the BFD count falls within the desired range (e.g. >100). The ROSCs are designed with a programmable number of stages ranging from 9 to 15, to study the impact of logic depth on RTN induced frequency shift. The new test structure is also amenable to BTI stress experiments since tri-state gates can be disabled simultaneously allowing the ROSC to be configured into an open-loop inverter chain. A detailed comparison is provided in Fig. 6.

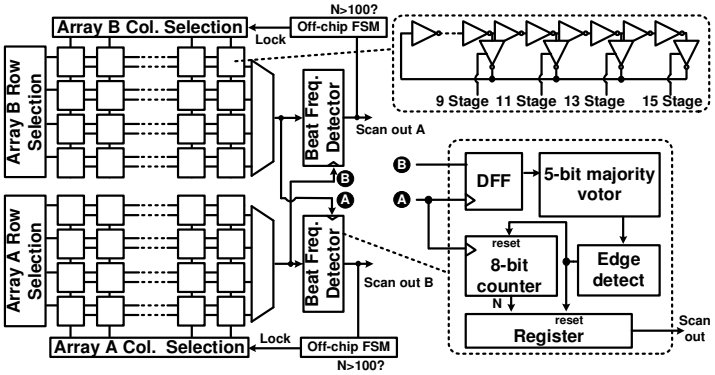


Fig. 5. Proposed dual ROSC array based RTN characterization vehicle with programmable number of ROSC stages. The beat frequency detection circuit measures the delay shift of each ROSC with a frequency resolution less than 0.01%.

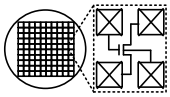
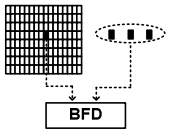
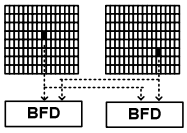
	Traditional Probing	Prior Work [5]	This Work
Schematic			
Description	Individual device probing	Single ROSC array	Two ROSC arrays
Technology	-	65nm	32nm HKMG
Automated measurement	No	Semi-automatic (requires freq. trimming)	Fully-automatic
Silicon area	Large	Small	Small
Min. freq. meas. resol.	-	0.05% (after trimming)	0.01% (no trimming)
# of stages	-	11	9, 11, 13, 15
Stress results	Yes	No	Yes (@ 2xVDD)
Meas. temp.	-	25°C	27, 55, 76°C
Meas. voltage	-	Poor resolution below 0.8V	Good resolution down to 0.45V

Fig. 6. Comparison with prior art.

RTN Induced Frequency Shift Measurements

Figs. 7-9 show frequency shift traces for a ROSC containing a single RTN trap measured under different test conditions. The RTN induced frequency shift decreases from 0.38% to 0.15% as the supply voltage is increased from 0.45V to 0.6V. As the number of stages increases from 9 to 15, the frequency fluctuation reduces from 0.38% to 0.24% for the same RTN trap which indicates that the

same trap has a weaker influence on the overall circuit delay for longer ROSCs due to the circuit averaging effect. RTN time constants show strong dependence on temperature. The frequency shift measured from 6 different ROSCs are displayed in Fig. 10, all showing a monotonic decrease in magnitude as the supply voltage is increased. This suggests that RTN will be negligible at the full nominal supply. Fig. 11 shows that the RTN trapping and de-trapping time constants become shorter at higher temperatures which is in line with previously reported data [2].

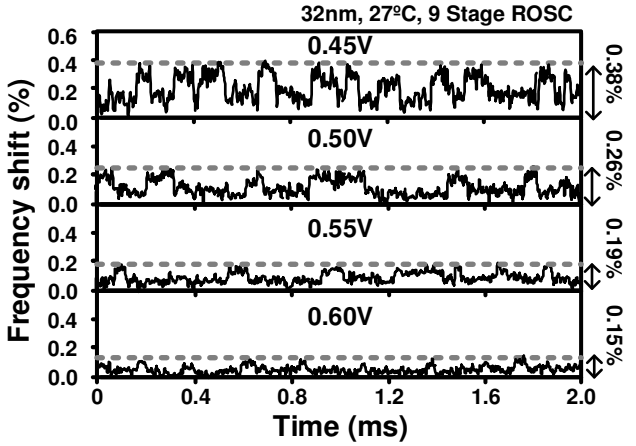


Fig. 7. RTN induced frequency shift of a ROSC measured at different supply voltages.

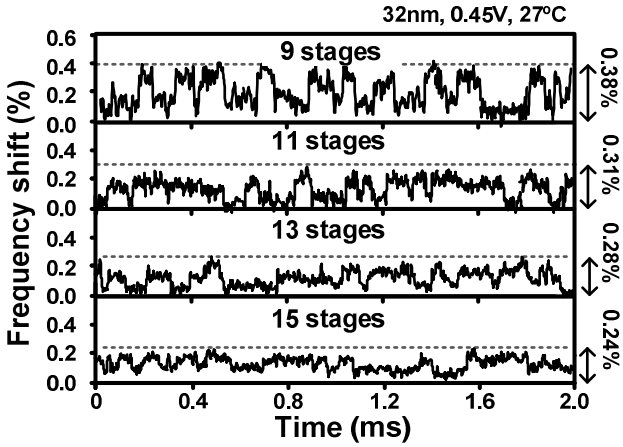


Fig. 8. RTN induced frequency shift measured for different number of stages.

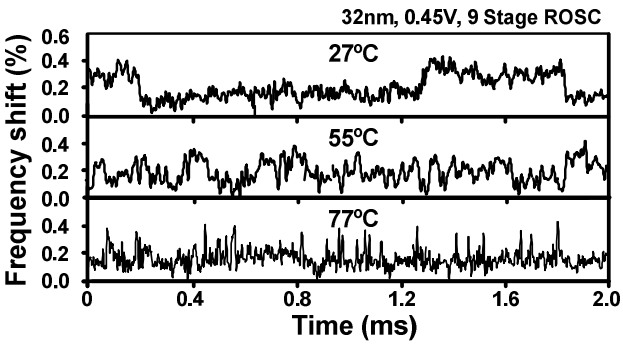


Fig. 9. RTN induced freq. shift of a ROSC measured at different temperatures.

Statistical Measurements and Stress Data

Fig. 12 shows that the number of detectable RTN traps slightly decreases at higher temperatures. This can be attributed to increased device noise (e.g. jitter) which makes it difficult to clearly distinguish between RTN and other noise sources. Fig. 13 shows the occurrences and locations of RTN traps across a single test chip from 0.45V to 0.6V. A given RTN trap can either disappear or reappear depending on the supply voltage which can be attributed to the Fermi level change. The number of ROSCs affected by RTN stays relatively constant under different supply voltages. RTN is believed to have major implications on device reliability issues [3]. For instance, both RTN and Bias Temperature Instability (BTI) have been reported to originate from the same defect source [4]. To understand the interplay between RTN and BTI better, we measured the location and magnitude of RTN while applying a stress voltage that is two times the nominal supply. To minimize the effect of fast BTI recovery, we turned off the local supply of the ROSC for a short period of time before turning it back on for measurements. This extra relaxation period ensures that devices are sufficiently recovered before they are characterized. The ROSC frequencies were sampled periodically while the test chip was under a 1.8V voltage stress for 14 hours. Stress results in Fig. 16 show several new RTN traps generated as a result of the voltage stress as well as few that have disappeared.

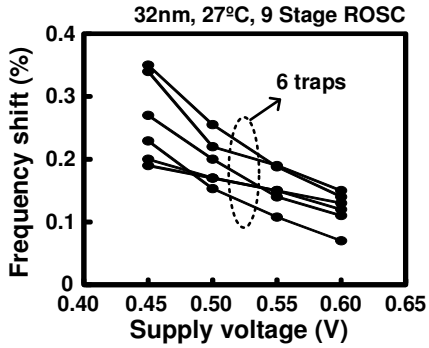


Fig. 10. RTN induced frequency shift for different traps measured at 0.45-0.6V.

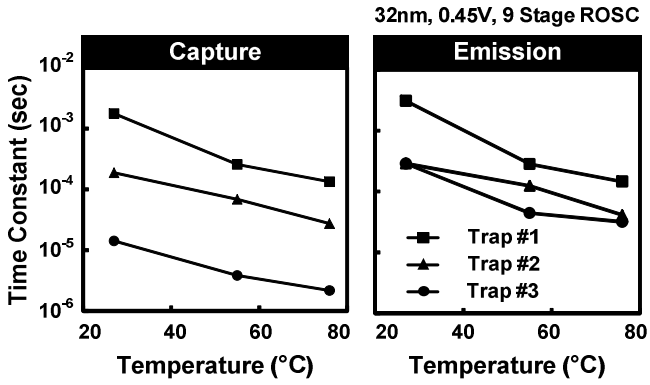


Fig. 11. Capture and emission time constants versus temperature. Both time constants decrease at higher temperatures.

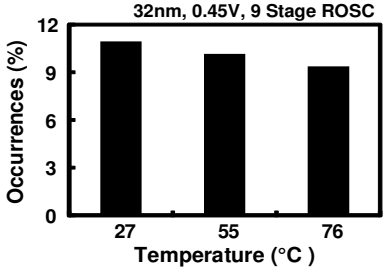


Fig. 12 Occurrence of RTN at 3 different temperatures. The % of ROSCs exhibiting RTN behavior decreases slightly at higher temperatures.

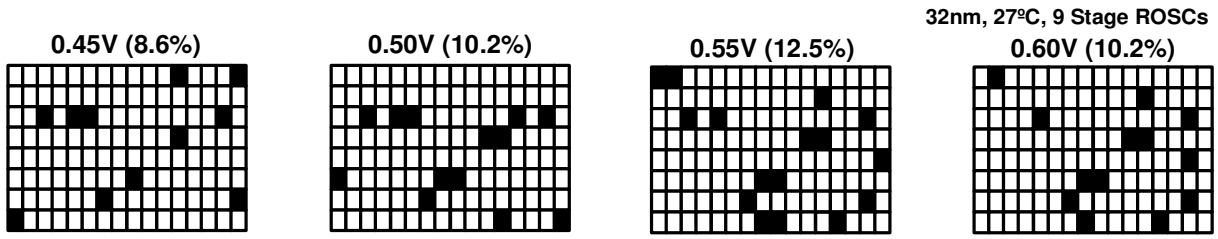


Fig. 13. RTN trap location in ROSC array at different supply voltages. Nominal supply voltage is 0.9V.

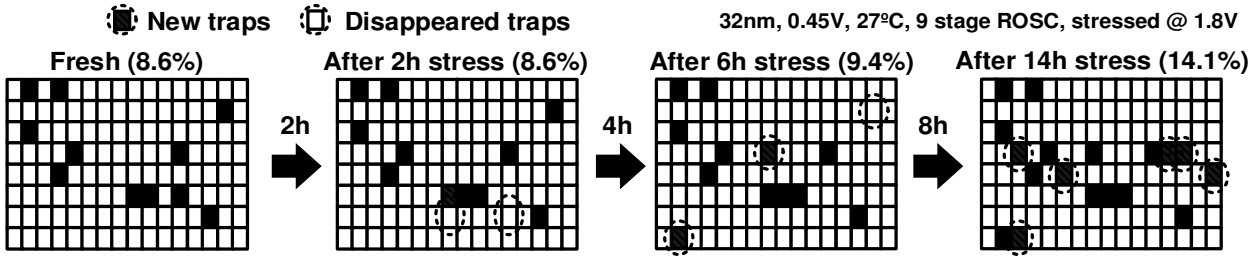
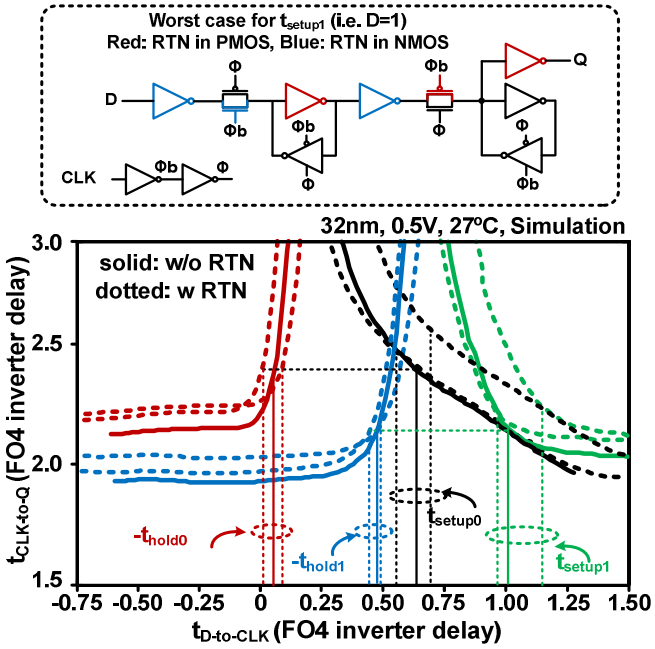


Fig. 14. RTN trap location in ROSC array after 0, 2, 6 and 14 hours of 1.8V stress. Nominal supply voltage is 0.9V.



	t_{hold0}	t_{hold1}	t_{setup0}	t_{setup1}
w/o RTN	-0.05	-0.47	0.64	1.01
w/ RTN (best case)	-0.08	-0.48	0.56	0.97
w/ RTN (worst case)	-0.01	-0.45	0.69	1.15

**Delay values normalized to FO4 inverter delay at 0.5V*

Fig 15. RTN impact on D-flip-flop setup and hold times based on test chip data.

Logic Timing Margin Analysis

To better understanding RTN's impact on circuit timing, we simulated the setup/hold times of a standard DFF circuit using the measured RTN data. As shown in Fig. 15, RTN either degrades or improves the setup/hold times depending on the location of RTN trap in the DFF. For example, the worst case setup time for data '1' occurs when traps appear in alternating PMOS and NMOS devices on the path from D to Q. Simulation result shows that the fluctuation in setup and hold times ranges from 0.03 to 0.18 FO4 inverter delays. In sequential circuit consisting of logic and flip-flips, the DFF setup time and hold time limit the max-delay and min-delay on the data path, respectively. The impact of RTN on the min-delay constraint is negligible because the hold time is generally negative. Fig. 16 compares the max-delay time under different RTN scenarios. In the absolute worst case, traps may be present in the input and output DFFs as well as the clock tree and logic path. The max-delay time allowed for correct operation is reduced by up to 0.21 FO4 inverter delays under this worst case condition.

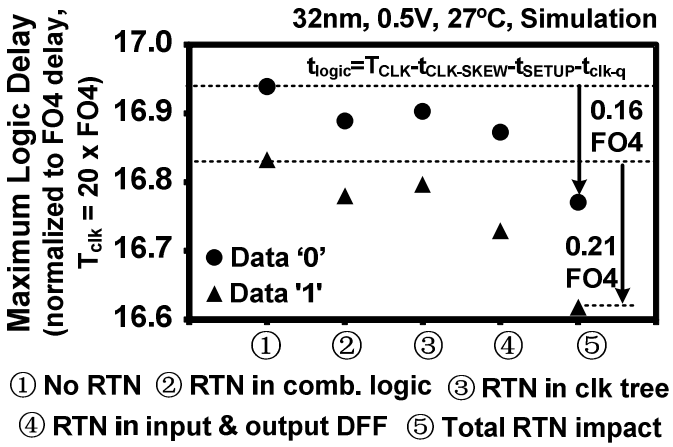


Fig 16. RTN impact on logic path max delay time estimated using test chip data (assumes T_{clk} = 20 FO4 inverter delay, and single RTN trap in logic, clk tree, input DFF, and output DFF).

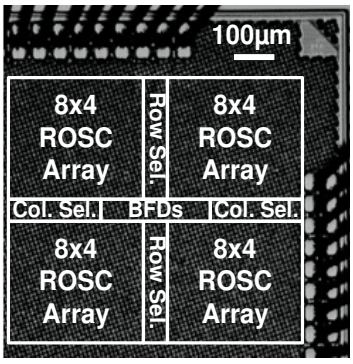


Fig. 17. 32nm chip die photo.

Acknowledgements: This work was supported in part by the Failure-Resistant Systems (FRS) program, a joint initiative between the National Science Foundation (NSF) and the Semiconductor Research Corporation (SRC).

References: [1] T. Kim, et al., JSSC, Apr. 2008. [2] H. Miki, et al., VLSI Tech. Symp., 2012. [3] H. Miki et al., VLSI Tech. Symp., 2011. [4] T. Grasser, et al., IRPS, 2014. [5] Q. Tang, et al., VLSI Tech. Symp., 2013.