Influence of Metallic Tubes on the Reliability of CNTFET SRAMs: Error Mechanisms and Countermeasures

ABSTRACT

Carbon nanotubes (CNTs) are considered as a possible successor to the CMOS technology. The adoption of these nanodevices for designing large VLSI systems, however, is limited by the unreliable manufacturing process. It is well-known that the carbon nanotube field-effect transistor (CNTFET) suffers from manufacturing defects such as metallic CNTs, misaligned/mispositioned CNTs and variations in doping, diameters and densities of tubes. In this paper, we investigate the possibility of using CNTFETs to build SRAM arrays. We analyze the error mechanisms and show how stuck-at faults and pattern sensitive faults are caused by metallic tubes in different transistors of a 6-T SRAM cell. Countermeasures based on error correcting codes are investigated. The results indicate two problems of applying ECC for building SRAM arrays using CNTFETs. First, due to the the lack of deterministic error models caused by unpredictable manufacturing variations, stronger error correcting codes than linear single-error-correcting, double-error-detecting (SEC-DED) codes used for CMOS SRAMs are required. The code should have similar encoding and decoding latency to SEC-DED codes so that the performance of CNTFET SRAMs is not compromised. Second, fault-secure decoder that can tolerate multiple faults occurring to the decoder are needed. Unless the above two problems are solved, using CNTFETs to build SRAM arrays remains impractical.

1. INTRODUCTION

SRAMs are critical elements in today's digital system designs. It was reported in [1] that more than 50% of the area of a SoC will be occupied by SRAMs in 2013. As the technologies move into the nano realm, the conventional CMOS SRAM cells encounter severe problems (e.g. the increase of the leakage current and the threshold voltage (V_T) variations [2]), which largely reduce the yield and prevent the continued improvement of their performance.

The potential usage of carbon nanotube FETs (CNTFET) for the design of SRAM cells as an alternative to CMOS technologies has been investigated in the community. In [3], a multiple-V_T 6-T CNTFET SRAM cell based on dual-chirality selection of nano tubes was presented. In [4], the authors showed that compared to the CMOS SRAM cells, the CNTFET 6T SRAM cells have 84% less standby leakage power, 1.84 times of the speed, 21% larger static noise margin and are more resistant to temperature and channel length variations.

Although promising at first glance, the above works did not thoroughly analyze the influence of manufacturing variations of CNTFETs and the results are too optimistic. In practice, the adoption of CNTFET technology has been limited by the unreliable manufacturing process. Misaligned/mispositioned CNTs, metallic CNTs [5] and variations in dopings, diameters and densities of tubes due to the imperfect manufacturing process [6] result in worse leakage power, delay, noise margin and even malfunction of transistors and logic error of the digital circuits. An initial estimation of the effects of the CNTFET imperfections described above on SRAM characteristics like the power consumption and the noise margin was presented in [7].

In this paper, we identify the possible error mechanisms caused by metallic tubes in CNTFET-based SRAM cells [3, 4] using detailed circuit level simulations and show how metallic tubes in the transistors of SRAM cells cause stuck-at faults and pattern sensitive faults. We estimate the bit error rate given the percentage of metallic tubes in a CNTFET. The results show that the bit error rate is hard to predict due to the uncontrollable manufacturing variations. As a result, error correcting codes that are stronger than single-error-correcting, double-error-detecting (SEC-DED) codes used for CMOS SRAMS should be applied.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the CNT technology and the main challenges in CNTFET manufacturing process. In Section 3, we analyze possible error mechanisms caused by metallic tubes in the SRAM cells. In Section 4, we briefly discuss the possible countermeasures based on multi-bit error correcting codes or nonlinear error-correcting codes and the associated tradeoffs.

2. CARBON NANOTUBE TECHNOLOGY

Carbon nanotube (CNT) technology has been proposed by various researchers as a potential successor to the CMOS technology due to its small dimension, large current carrying capacity and ballistic transport characteristic [10, 11]. CNTs can be used to build both transistors and wires as they can exhibit both semi-conducting and metallic behaviour.

The structure of carbon nanotube field effect transistor (CNTFET) is shown in Figure 1 [12]. Circuits using CNTFETs have less standby leakage power, faster speed, larger static noise margin and are more resistant to temperature and channel length variations compared to circuits using the conventional CMOS transistors [4].

Although promising at first glance, the large-scale adoption of CNTFET technology has been limited by its unreliable manufacturing process. The limited control over the growth of CNTs may generate misaligned/mispositioned CNTs, which can cause incorrect functioning of the logic blocks [6]. The presence of metallic CNTs (33% for a typical growth process) can result in a short between the source and drain of the transistor, which will increase the leakage power, reduce the static noise margin and may cause delay variation and logic faults. Manufacturing variations in doping, diameters and densities of tubes directly influence the delay of the

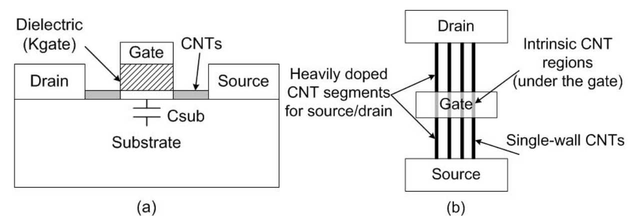

Figure 1: Carbon nanotube transistor (a) cross section view (b) top view [12]

CNTFETs and may cause degraded noise margin and malfunctions of logic gates.

In this paper, we concentrate our analysis on the influence of metallic tubes in CNTFETs on the reliability of SRAM cells. The possible error mechanisms caused by metallic tubes in the transistors of SRAM cells are presented in the following section.

3. INFLUENCE OF METALLIC TUBES ON THE RELIABILITY OF SRAM CELLS

To analyze the effects of metallic tubes on the reliability of CNTFET SRAM cells, we conducted simulations in HSPICE based on Stanford CNTFET models [12]. We assume that the pitch (center-to-center distance of CNTs belonging to the same transistor) is $6\,nm$ and the dielectric thickness is $3\,nm$. The supply voltage Vdd is $0.9V$. The chirality of all CNTs is selected to be $(19,0)$ resulting in a tube diameter of $1.5\,nm$ and a threshold voltage of $0.289\,V$ [7]. The gate width $Wgate$ of the CNTFETs is $pitch \times \#\,tubes$. For all the other parameters, we use the default values provided by the Stanford CNTFET models.

The conventional structure of a six-transistor (6T) SRAM cell is shown in Figure 2. Let W and L denote the width and the length of the transistor channel. The authors in [3] showed that when the chirality of the CNTs is $(19,0)$, the *cell ratio* (*CR*) (ratio of $\frac{W}{L}$ of the pull down transistors $M1/M3$ to $\frac{W}{L}$ of the access transistors $M5/M6$) should be larger than 0.5 for the purpose of avoiding the *read upset* problems. The *pull-up ratio* (*PR*) (ratio of $\frac{W}{L}$ of the pull up transistors $M2/M4$ to $\frac{W}{L}$ of the access transistors $M5/M6$) should be smaller than 1.6 to guarantee a successful write. In this work, the channel length of all CNTFETs are $32\,nm$. The number of tubes for pull up transistors, access transistors and pull down transistors are assumed to be 8, 16 and 24 respectively. The pull-up ratio is 0.5. The pull-down ratio is 1.5. Both ratios are the same as shown in [3].

According to the measurement results shown in [13], we assume that the percentage of metallic CNTs p_m satisfies a normal distribution whose probability density function is $\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/(2\sigma^2)}$, where $\mu \approx 0.102$ is the mean and $\sigma \approx 0.0257$ is the standard deviation. Since the current version of the Stanford CNTFET models does not support metallic tubes, we model each metallic tube as a resistor. The theoretical resistance of a single metallic tube is $\frac{h}{4e^2} \approx 6.5k\Omega$ when $L \leq L_0$ (Ballastic transport) and is $\frac{h}{4e^2} \cdot \frac{L}{L_0}$ when $L > L_0$ [14], where L is the length of the channel and L_0 is the mean free path. However, we note that the observed resistance of a single metallic tube typically lies in the range of $7k\Omega \sim 100k\Omega$ due to the imperfect metal-nanotube contact [15]. The effective resistance of multiple metallic tubes is derived in the same way as for parallel resistors. For the purpose of conducting a more general analysis, in this section we only consider the resistance range in which the fault

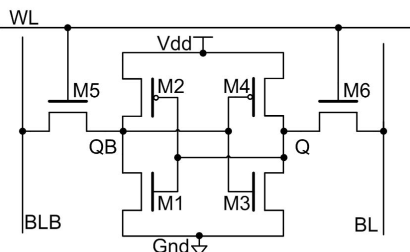

Figure 2: Six-transistor SRAM cell

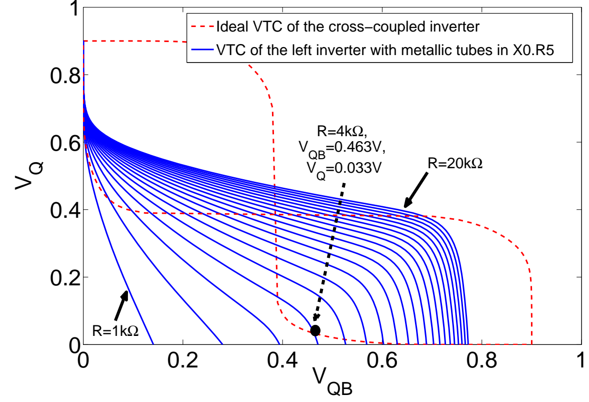

Figure 3: Influence of metallic tubes in pull down transistor on standby noise margin. The number of tubes for the pull up transistors, the access transistors and the pull down transistors are assumed to be 8,16 and 24 respectively.

will appear. The reliability of the SRAM for a given number of metallic tubes can be easily estimated once the average resistance of a single metallic tube is known.

3.1 Stuck-at Faults

Metallic tubes in SRAM cells will introduce a short between its internal nodes and the bit line, the power supply or the ground depending on their locations. These shorts can cause stuck-at faults which result in read upsets or write failures.

3.1.1 Metallic tubes in the cross-coupled inverters

Let us assume there is a short between QB and Gnd introduced by metallic tubes in $M1$ (see Figure 2). As shown in Figure 3, the VTC of the inverter composed of $M1$ and $M2$ will move bottom left and the standby noise margin of the inverter will decrease as the effective resistance R of metallic tubes changes from $20k\Omega$ to $1k\Omega$. Because of the short between QB and Gnd, QB cannot be pulled up to Vdd when a 0 is stored in the cell (QB is high when a zero is stored). The stable state of V_{QB} can be derived from Figure 3 and is determined by the value of R. The existence of metallic tubes in $M1$ may cause read upsets. Assume a logic 0 is stored in the cell (QB is high). When $R = 4k\Omega$, V_{QB} is only $0.463V$ (Figure 3) and a logic 1 will be mistakenly written in the cell after the word line WL is enabled during a read operation. When $R \leq 3k\Omega$, logic 0 cannot be written into the cell. Both of the cases will result in a stuck-at-1 fault ($V_Q = Vdd, V_{QB} = 0$).

Remark 3.1 *The above simulation ignores the influence of the remaining semi-conducting tubes in $M1$ and the resulting transistor $M1'$ with a shrinked size. However, the effect of $M1'$ only shows up when V_Q is large enough to turn $M1'$ ON and will not improve the noise margin when a logic 0 is stored. The analysis of possible stuck-at-1 faults will still be valid if $M1'$ is considered. Similar analysis can be conducted for metallic tubes in $M2, M3$ and $M4$.*

3.1.2 Metallic tubes in the access transistors

Suppose a logic 0 is stored in the cell and there are metallic tubes in $M5$. Let R be the effective resistance of metallic tubes. Let $R_{M5'}$ be the ON resistance of $M5'$ composed of the remaining

Table 1: Stuck-at faults in SRAM cells caused by shorts introduced by metallic tubes

Position of metallic tubes	Initial state of QB	Final state of QB	Defects	Conditions
$M1$	-	-	Degraded noise margin	$R \geq 5k\Omega$
$M1$	1	0	Read upset (Stuck-at-1)	$R \approx 4k\Omega$
$M1$	0	0	Write failure (Stuck-at-1)	$R \leq 3k\Omega$
$M2$	-	-	Degraded noise margin	$R \geq 3k\Omega$
$M2$	1	1	Write failure (Stuck-at-0)	$R \leq 2k\Omega$
$M5$	1	1	Write failure (Stuck-at-0)	$R_A \geq 3k\Omega$

R is the effective resistance of metallic tubes is $M5$. $R_A = \frac{R_{M5'}R}{R_{M5'}+R}$ where $R_{M5'}$ is the ON resistance of $M5'$. Stuck-at faults caused by metallic tubes in $M3, M4, M6$ can be analyzed using similar methods presented in this section.

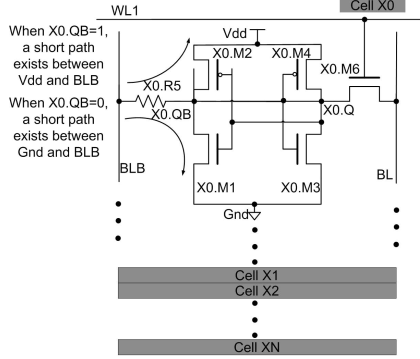

Figure 4: Multiple cells sharing the same bit lines in the SRAM array. Metallic tubes in the access transistors can cause pattern-sensitive faults.

semi-conducting tubes in $M5$. Let R_A be the resistance between BLB and QB when the wordline is enabled, then $R_A = \frac{R_{M5'}R}{R_{M5'}+R} < R$. Assume a logic 1 is being written in the cell by setting $BLB = 0$ and $BL = Vdd$. Ideally, a reliable write is guaranteed by pulling QB below the threshold voltage of the transistor $M4$. However, when R_A is large, the effective pull-up ratio of the cell will increase. Simulation shows that when $R_A \geq 3k\Omega$, V_{QB} cannot be pulled low enough to ensure the writing of 1 in the cell, which results in a stuck-at-0 fault.

Possible stuck-at faults caused by metallic tubes in SRAM cells are summarized in Table 1. These errors can be detected by off-line testing methods, e.g. first write data into the memory and then verify it through read operations. For the case where more than one CNTFET contains metallic tubes, the overlay effect depends on the position, the number and the distribution of metallic tubes and can be derived using similar analysis.

3.2 Pattern-Sensitive Faults

SRAM cells belonging to the same column of the SRAM array share the same bit lines. Metallic tubes in the access transistors will introduce extra charging and discharging paths between the bitlines and the internal nodes of the SRAM cells and may affect the reading and writing operation on other cells in the same column (Figure 4).

3.2.1 Errors During Read Operations

Assume a logic 1 is stored in $X0$ ($X0.Q = Vdd, X0.QB = 0$) which contains metallic tubes (modeled as $X0.R5$) in the access transistor connecting BLB and $X0.QB$ and we try to read from another cell $X1$ in the same column storing a logic 0. To complete the read operation, BL and BLB will be precharged to Vdd and then the wordline of $X1$ will be enabled. Ideally, during the read operation BL will be discharged to 0 through $X1.M6$ and $X1.M3$ and BLB will maintain a voltage level close to Vdd given the fact that $X1.QB = Vdd$. Assume a basic differential sense amplifier [16] is used to generate the final output of the SRAM cell (Figure 5 (a)). When SAE is enabled, $SA.A$ will start discharging through $SA.M2$ and $SA.M4$.

Due to the short introduced by metallic tubes between BLB and $X0.QB$, BLB will be simultaneously charged through $X1.M2$ and $X1.M5$ and discharged through $X0.R5$ and $X0.M1$. As the resistance of $X0.R5$ becomes smaller, the stable voltage of BLB decreases and the ON resistance of $SA.M2$ becomes larger. The current I_D flowing through $SA.M2$ may be too small to discharge $SA.A$ fast enough to flip $SA.B$ and the output signal in the available time.

Assume the number of CNTs in each transistor of the sense amplifier is 16 and the clock period is 500 ps. The wordline of $X1$ and SAE are enabled for 300 ps after BL and BLB are precharged to Vdd. The voltage curve of $SA.B$ for different values of $X0.R5$ is plotted in Figure 6. Simulation results show that when $X0.R5 \leq 0.8k\Omega$, the voltage of $SA.B$ cannot be charged high enough to pull down the output signal and a logic 1 is mistakenly read out from the cell.

Even worse, it is possible that the read operation on $X1$ will result in an un-wanted write of 1 in the same cell. Simulation results show that when $X0.R5 \leq 0.3k\Omega$, the voltage of $X1.QB$ can be pulled low enough through $X0.R5$ to flip the bit stored in the cross-coupled inverter. The error will stay until new contents are written in the cell.

Similar problems also exist for other types of sense amplifiers whose output converging speed depends on the absolute voltage difference between the two input signals. If more than one cell storing logic 1 contain metallic tubes in the access transistors connecting BLB and QB, as long as the effective resistance of all resistors introduced by metallic tubes in particular access transistors is less than a certain value (e.g. $0.8k\Omega$ in our simulation), the failures described above will occur.

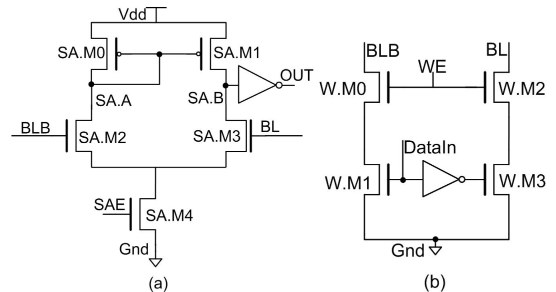

Figure 5: (a) Differential sense amplifier (b) Write driver

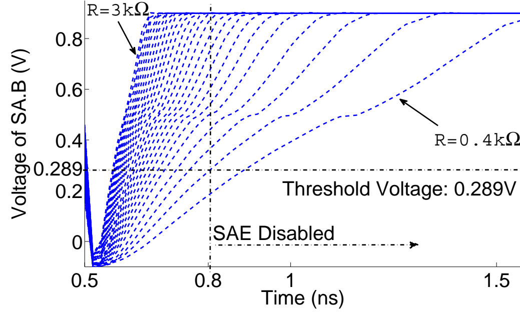

Figure 6: Voltage curve of $SA.B$ under the situations of different resistance of $X0.R5$. Here we are reading a logic 0 in the cell $X1$, which is in the same column as $X0$.

3.2.2 Errors During Write Operations

Write 0 failure: Suppose $X0$ has metallic tubes in $M5$ and a logic 1 is stored in both $X0$ and $X1$ ($X0.QB = X1.QB = 0$). A 0 is being written in $X1$ by setting BLB to Vdd and BL to 0. Write failures are observed when $X0.R5 \leq 0.4k\Omega$, assuming a basic write driver is used [16] and each transistor in the write driver contains 48 CNTs(Figure 5 (b)). Ideally, the write operation is completed by pulling down $X1.Q$ to a voltage level lower than the threshold of $X1.M2$. However, due to the short introduced by the metallic tubes in $X1.M5$, $X1.QB$ will also be pulled down through the path $X1.M5 \rightarrow X0.R5 \rightarrow X0.M1$ when the wordline of $X1$ is enabled. As a result, $X1.Q$ cannot be pulled low enough to complete the write operation.

Write 1 failure: If both $X0$ and $X1$ store a logic 0 ($X0.QB = X1.QB = 1$) and a logic 1 is written in $X1$, $X1.QB$ will be pulled down through $W.M0$ and $W.M1$ in the write driver. Since $X0.QB = 1$, $X0.QB$ will pull up BLB as well as $X1.QB$ through $X0.R5$ introduced by the metallic tubes. When $1.4k\Omega \leq X0.R5 \leq 12k\Omega$, $X1.QB$ cannot be pulled low enough to ensure a writing of 1 in the cell. When $X0.R5 \leq 1.7k\Omega$, a 1 will be mistakenly written in $X0$.

3.2.3 Further Discussions

Errors described in Section 3.2.1 and 3.2.2 are summarized in Table 2. These errors are pattern-sensitive in a sense that they only occur when all the involved cells containing metallic tubes in the access transistors store a particular logic value. These errors cannot be efficiently detected by off-line testing methods.

The chance of un-wanted write of 1 in $X0$ (Section 3.2.2) is small since all metallic tubes must be in one single CNTFET to create a short path whose resistance is low enough for the failure to occur. On contrary, read upsets and un-wanted write in $X1$ are more probable because they can also occur when metallic tubes spread over multiple cells as long as the effective resistance between BLB and Gnd is less than a threshold value.

The most probable error is the write 1 failure in $X1$. Assuming each metallic tube is modeled as a resistor of $70k\Omega$ (including the contact resistance), 6 metallic tubes are enough to introduce a resistor of resistance lower than $12k\Omega$ so that the write failure may occur. Given the assumption that around 10% of tubes are metallic [13], the probability of write failures will be very high (see Section 3.3). The problem can be mitigated by increasing the pull down strength of the write driver. The most straightforward way is to increase the size of the pull down transistors of the write driver. Stronger pull down networks can reduce the stable voltage level of $X1.QB$ for a given value of $X0.R5$. As a result, a smaller value of $X0.R5$ is required to pull up $X1.QB$ high enough to incur the write failure. Simulation results show that when the size of the transistors in the write driver is doubled, the effective resistance of metallic tubes required for the write failure to occur is reduced to approximately $1k\Omega$ and the write failure rate will also be reduced because more than 6 metallic tubes would be required to incur the failure. Similarly, the mis-read problem (Section 3.2.1) can also be solved by increasing the size of the transistors in the sense amplifier (or slowing down the circuit).

The method based on re-sizing transistors, however, cannot solve the problem of the un-wanted write in $X1$ during read operations (Section 3.2.1). Moreover, as the manufacturing techniques improve and the contact resistance is reduced, errors due to metallic tubes will become more probable because less metallic tubes are required to introduce a resistance low enough to incur the errors.

If other transistors in $X0$ also contain metallic tubes, the probability of pattern-sensitive errors may either increase or decrease depending on the number and the position of metallic tubes. For example, suppose a 1 is stored in $X0$ and a 0 is stored in $X1$. If both $X0.M5$ and $X0.M2$ contain metallic tubes, the short between $X0.QB$ and Vdd will reduce the pull down effect of $X0.R5$ on BLB and $X1.QB$. As a result, the value of $X0.R5$ required to incur a mis-read failure (Section 3.2.1) will be smaller and the error rate will decrease. On contrary, if $X0.M5$ and $X0.M6$ contain metallic tubes, the short between BL and $X0.Q$ will pull up BL so that a smaller $X0.R5$ will be enough to cause the mis-read failure and the error rate will increase.

Metallic tubes in sense amplifiers and write drivers (Figure 5) may also cause stuck-at faults. For example, suppose $SA.M2$ contains metallic tubes and a logic 1 is read out from the cell. Ideally, $SA.M3$ is ON and $SA.B$ is discharged to 0. $SA.M2$ is OFF and $SA.A$ should maintain a voltage level close to Vdd. However, due to the metallic tubes in $SA.M2$, $SA.A$ will also be discharge when SAE is enabled. The equivalent voltage difference between BLB and BL is reduced and it is possible that $SA.B$ cannot be pulled down fast enough to flip the output in the given time resulting in a stuck-at 0 fault. The probability of stuck-at faults caused by metallic tubes in the sense amplifiers and the write drivers is small given the fact that the number of tubes in these two parts are really small compared to the number of tubes in the SRAM array.

3.3 Estimation of the Probability of Errors in Columns

Stuck-at errors can be detected by off-line testing methods [17]. The most straightforward way is to first write data into the memory and then verify it through read operations.

In this section we estimate the probability for a column to have pattern-sensitive errors due to metallic tubes in the access transistors. To simplify the analysis, we only consider metallic tubes in the access transistors connecting BLB and QB. (A similar analysis can be performed for the case when there are also metallic tubes in access transistors connecting BL and Q.) We assume that the sense amplifier and the write driver are designed with enough driving capability and the effective resistance of metallic tubes in the access transistors required for a mis-read or a write failure is around $0.3k\Omega$. (Note that the un-wanted write in $X1$ during a read operation happens when $X0.R5 \leq 0.3k\Omega$ (see Table 2) and cannot be solved by re-sizing the sense amplifier.)

Table 2: Pattern sensitive errors due to metallic tubes in $M5$ of $X0$

Operation*	$X0.QB$ (Initial)	$X1.QB$ (Initial)	$X0.QB$ (Final)	$X1.QB$ (Final)	Failures	Conditions
Read	0	1	0	1	Mis-read	$0.3k\Omega < X0.R5 \leq 0.8k\Omega$
Read	0	1	0	0	Mis-read, Un-wanted write in $X1$	$X0.R5 \leq 0.3k\Omega$
Write 0	0	0	0	0	Write Failure	$X0.R5 \leq 0.4k\Omega$
Write 1	1	1	1	1	Write Failure	$1.7k\Omega \leq X0.R5 \leq 12k\Omega$
Write 1	1	1	0	1	Write Failure, Un-wanted write in $X0$	$1.4k\Omega \leq X0.R5 \leq 1.7k\Omega$
Write 1	1	1	0	0	Un-wanted write in $X0$	$X0.R5 \leq 1.4k\Omega$

$*$: All the operations are on $X1$.

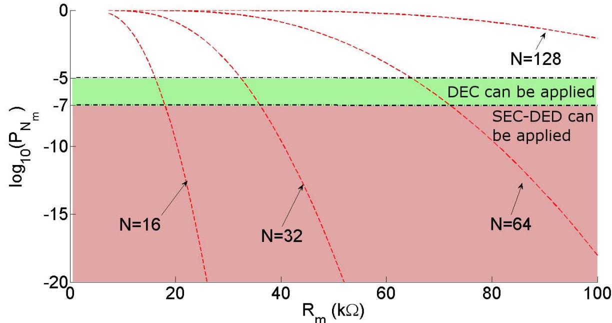

Figure 7: Column Failure Rate of SRAM Cells Due to Metallic Tubes in the Access Transistors. N is the number of rows in the SRAM array. R_m is the resistance per metallic tube. Here we assume each access transistor contains 16 tubes and the percentage of metallic tubes is approximately 10%.

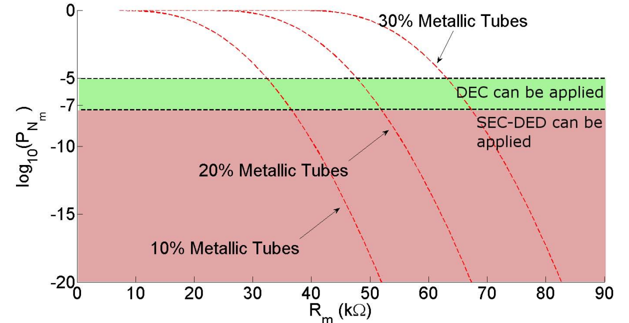

Figure 8: Column Failure Rate of SRAM Cells Due to Metallic Tubes in the Access Transistors. R_m is the resistance per metallic tube. Here we assume each access transistor contains 16 tubes and the number of rows N in the SRAM array is 32.

Let R_m be the resistance per metallic tube. To have a resistance no larger than $0.3k\Omega$, at least $N_m = \frac{R_m}{0.3k}$ metallic tubes are required. Let N be the number of rows in the SRAM array. The total number of tubes in the access transistors connecting BLB and QB is $16N$ (each access transistor contains 16 tubes). Assume the percentage of metallic tubes p_m satisfies a normal distribution with $\mu \approx 0.102$ and $\sigma \approx 0.0255$ [13] (Section 3). Let F_C be the complementary cumulative distribution function for the normal distribution. We have $F_C(x) = 0.5(1 - erf(\frac{x-\mu}{\sqrt{2}\sigma}))$, where $erf(x)$ is the *error function*. The column may have pattern-sensitive errors if and only if the number of metallic tubes in all the access transistors connecting BLB and QB is at least N_m, which can be computed as $P_{N_m} = F_C(\frac{N_m}{16N}) = F_C(\frac{R_m}{N \cdot 4.8k})$. The probability that M columns have potential pattern-sensitive errors is $P_{N_m}^M$.

Given the number of tubes per transistor, P_{N_m} is a function of the number of rows N and the resistance per metallic tube R_m. Generally speaking, larger N will increase the chance of having more than N_m metallic tubes in the access transistors. Smaller R_m will decrease the number of metallic tubes required to incur the mis-read or the write failure. Both cases will result in a worse P_{N_m} (Figure 7). Thereby, to increase the reliability of the SRAM, small N and large R_m should be targeted.

Larger percentage of metallic tubes will result in worse column failure rate of SRAM cells. As shown in Figure 8, the curve of the column failure rate moves right as the percentage of metallic tubes increases. When $N = 32$ and there are approximately 30% metallic tubes, R_m needs to be almost twice larger compared to the case when only 10% tubes are metallic.

Metallic tubes in the access transistors of one cell will have similar influence on the read and write operations of all the other cells in the same column. Thereby, the same error mechanism tends to repeat for the data stored in cells belonging to the same column. Moreover, if more than one columns have potential pattern-sensitive errors, since the locations of these columns are determined by manufacturing process and will not change when the device is in use, it is highly probable that the multi-bit error patterns will repeat for several consecutive clock cycles.

4. IMPROVE THE RELIABILITY OF CNTFET SRAMS USING ECC

Error correcting codes are widely used to correct errors in memory array. The bit error rate for CNTFET SRAMs is difficult to predict due to the lack of a deterministic manufacturing process. For an initial investigation of the error correcting codes needed to ensure reliable operation of CNTFET SRAM arrays with metallic tubes, we analyze the probability of errors caused by the metallic tubes in the access transistors of the 6-T SRAM cells. We consider two cases: first where we keep the percentage of tubes metallic tubes per transistor fixed (10%) and determine the column failure rate for different row count (see Figure 7), and the second where we keep the number of rows in the memory array fixed to 32 and vary the percentage of metallic tubes per transistor (see Figure 8).

Figure 7 and 8 shows the regions where ECC with different error correcting capabilities are enough to provide a satisfactory reliability for CNTFET SRAM arrays. The well-known single-error correcting, double-error detecting (SEC-DED) linear extended Hamming codes can correct all single bit errors and detect all double bit errors. When the bit error rate is smaller than 10^{-7}, the chance that the error is uncorrectable is less than 10^{-14} for systems protected by SEC-DED codes. If the bit error rate is larger than 10^{-7}, codes with higher error correcting capabilities are required. When the bit error rate is around 10^{-5}, the probability that an error is uncorrectable by SEC-DED codes is increased to 10^{-10}. In this case, double error correcting (DEC) codes are needed to reduce this probability to 10^{-15} so that the reliability of the system can be guaranteed.

The area overhead of using ECC to protect memories is mainly determined by the number of extra cells needed to store the redundant bits of the ECC. As an example, we compare the number of redundant bits and the corresponding overhead area of a SEC-DED code, a DEC code and a triple error correcting code (TEC) when

protecting a 64-bit SRAM in Table 3. The area overhead is computed by dividing the number of redundant bits of the codes by 64. For systems protected by DEC codes, the area overhead is nearly doubled compared to systems protected by SEC-DED codes. The overhead increases as the error correcting capability of the code becomes higher.

Table 3: Area overhead when protecting a 64-bit SRAM using codes with different error correcting capabilities

Code	Number of Redundant bits	Percentage Overhead
SEC-DED	8	12.5%
DEC	14	21.9%
TEC	21	32.8%

One challenge of using linear codes with higher error correcting capabilities is in the design of the encoder and decoder. Compared to linear SEC-DED codes, the encoder and the decoder for linear multi-bit error correcting codes tend to have longer critical path because of the more complex encoding and decoding logic. Moreover, the standard decoder for linear multi-bit error correcting codes, e.g. BCH codes, may require several clock cycles to complete the decoding procedure, which may be unacceptable for SRAM CNTFET-based memory caches.

Furthermore, since the decoder for error correcting codes protecting CNTFET SRAMs is also built using unreliable technology making it inherently unreliable. Hence there is a need for *fault secure* (FS) decodeer. A circuit is fault secure for a fault set ε if every fault in ε can be detected as long as it manifests at the output of the circuit. Fault secure decoders based on odd parity code and multi-input two rail code checker [18] are well studied in the community. However, most of the existing methods can only tolerate single errors in the decoder. Ideally, the fault secure decoder for codes used to protect CNTFET SRAMs should also be able to tolerate multi-bit errors. Since faults in the encoder will result in bit flips in the memory array and can be detected by the code, no fault secure encoder is required.

The problems of applying error correcting codes for the protection of nano devices exist not only for CNTs but also for other technologies such as graphene, quantum molecular conductors, etc. As IC designs move into the nano realm, it is highly probable that the future digital systems will have to be built using *unreliable* devices with much higher error rates or error models that are more difficult to predict than the current CMOS technology. Hence, ECC techniques that can handle higher multiplicity errors with low area and power overhead need to be developed. An approach that can jointly optimize all parameters of the system, including reliability, area, latency and power consumption will be indispensable for the development of the nanotechnology-based systems.

5. CONCLUSIONS

We analyzed possible error mechanisms due to metallic tubes in the CNTFET SRAM cells. CNTFET SRAM array can have stuck-at faults or pattern-sensitive faults depending on the count and the positions of metallic tubes. Stuck-at faults can be detected using off-line testing methods. The errors from patten-sensitive faults are transient and their error rate is hard to predict due to the manufacturing variations (e.g. variations of the number of metallic tubes and the resistance per metallic tube). Moreover, the same error pattern tends to repeat since the locations of columns having potential patten-sensitive faults will not change.

Error correcting codes can be used to improve the reliability of CNTFET SRAMs. Due to the unpredictable error models and possibly higher bit error rate, multi-bit error correcting codes should be considered when designing reliable CNTFET SRAMs. The primary challenges of applying multi-bit error correcting codes for the protection of memories are twofold: first, the latency and the area overhead of multi-bit error correcting codes should be reduced to a reasonable level and second, the decoder should be able to tolerate multi-bit errors due to the manufacturing variations. These two challenges for using ECC are not only for the design of CNTFET SRAMs but also for the design of nearly all digital systems based on the unreliable nano-scale technologies in the future.

REFERENCES

[1] "International technology roadmap for semiconductors," 2008.

[2] C.-T. Chuang, S. Mukhopadhyay, J.-J. Kim, K. Kim, and R. Rao, "High-performance sram in nanoscale cmos: Design challenges and techniques," in *MTDT '07: Proceedings of the 2007 IEEE International Workshop on Memory Technology, Design and Testing*, 2007, pp. 4–12.

[3] S. Lin, Y.-B. Kim, and F. Lombardi, "Design of a cntfet-based sram cell by dual-chirality selection," *Nanotechnology, IEEE Transactions on*, vol. 9, no. 1, pp. 30 –37, jan. 2010.

[4] A. Kureshi and M. Hasan, "Performance comparison of cnfet- and cmos-based 6t sram cell in deep submicron," *Microelectronics Journal*, vol. 40, no. 6, pp. 979 – 982, 2009.

[5] J. Zhang, N. P. Patil, and S. Mitra, "Design guidelines for metallic-carbon-nanotube-tolerant digital logic circuits," in *DATE '08: Proceedings of the conference on Design, automation and test in Europe*, 2008, pp. 1009–1014.

[6] N. Patil, J. Deng, A. Lin, H.-S. Wong, and S. Mitra, "Design methods for misaligned and mispositioned carbon-nanotube immune circuits," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 27, no. 10, pp. 1725–1736, Oct. 2008.

[7] B. Ebrahimi and A. Afzali-Kusha, "Realistic cnfet based sram cell design for better write stability," july 2009, pp. 14 –18.

[8] A. Sil, S. Ghosh, and M. Bayoumi, "A novel 90nm 8t sram cell with enhanced stability," in *Integrated Circuit Design and Technology, 2007. ICICDT '07. IEEE International Conference on*, 30 2007.

[9] R. Aitken, N. Dogra, D. Gandhi, and S. Becker, "Redundancy, repair, and test features of a 90nm embedded sram generator," *Defect and Fault-Tolerance in VLSI Systems, IEEE International Symposium on*, vol. 0, p. 467, 2003.

[10] N. Patil, J. Deng, S. Mitra, and H.-S. Wong, "Circuit-level performance benchmarking and scalability analysis of carbon nanotube transistor circuits," *Nanotechnology, IEEE Transactions on*, vol. 8, no. 1, pp. 37 –45, jan. 2009.

[11] A. H. Johnston, "Scaling and technology issues for soft error rates," ser. 4th Annual Research Conference on Reliability, 2000.

[12] "Stanford cntfet model," http://nano.stanford.edu/models.php.

[13] Y. Li and etc, "Preferential growth of semiconducting single-walled carbon nanotubes by a plasma enhanced cvd method," *Nano Letters 2004 4 (2), 317-321*.

[14] A. Raychowdhury and K. Roy, "Modeling of metallic carbon-nanotube interconnects for circuit simulations and a comparison with cu interconnects for scaled technologies," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 25, no. 1, pp. 58 – 65, jan. 2006.

[15] N. Srivastava and K. Banerjee, "Performance analysis of carbon nanotube interconnects for vlsi applications," in *ICCAD '05: Proceedings of the 2005 IEEE/ACM International conference on Computer-aided design*, 2005, pp. 383–390.

[16] *CMOS SRAM Circuit Design and Parametric Test in Nano-Scaled Technologies*, 2008, ch. SRAM Circuit Design and Operation.

[17] L.-T. Wang, C.-W. Wu, C.-W. Wu, and X. Wen, *VLSI test principles and architectures: design for testability*, 2006.

[18] T. R. N. Rao and E. Fujiwara, *Error-control coding for computer systems*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.