PERFORMANCE MODELING OF RESONANT TUNNELING BASED RAMS

Hui Zhang, Pinaki Mazumder, Li Ding

Dept. of EECS, University of Michigan,
Ann Arbor, MI 48109,USA

Kyounghoon Yang

Dept. of EECS, KAIST,
Republic of Korea

ABSTRACT

Tunneling based random-access memories (TRAM's) have recently garnered a great amount of interests among the memory designers due to their intrinsic merits such as reduced power consumption by elimination of refreshing operation, faster read and write cycles, and improved reliability in comparison to conventional silicon DRAM's. In order to understand the precise principle of operation of TRAM's, an in-depth circuit analysis has been attempted in this paper and analytical models for memory cycle time, soft error rate, and power consumption have been derived. The analytical results are then validated by simulation experiments performed with HSPICE. These results are then compared with conventional DRAM's to establish the claim of superiority of TRAM performance to DRAM performance.

1. INTRODUCTION

Silicon dynamic random-access memories (DRAM's) are currently dominant commercial commodity in the semiconductor memories market due to their lowest cost per bit as well as gargantuan integration scale that allows DRAM manufacturers to monolithically fabricate over 256 million cells per chip. However, these mega-size DRAM chips are encountering several formidable problems due to a host of reasons, some of which are listed below.

First, DRAM's are becoming increasingly prone to soft errors. Soft error is caused by extra charge collection in the storage node of memories, generally induced by external charged particles and neutrons. The chances of loss of a stored bit depend on the amount of critical charge of the storage node. Technology scaling that achieves lowered capacitance, reduced power supply voltage, tinier transistor geometries, is generally deployed to increase the density and performance of the DRAM's; however, the scaling also concomitantly reduces the critical charge of the DRAM cell, thus increasing the Soft Error Rate (SER).

Second, DRAM's power consumption largely depends on periodic refreshing of memory cells deemed necessary due to excessive leakage currents. The continuous down scaling of the transistor threshold voltage as well as packing of memory cells more densely sharply aggravates leakage currents, thereby significantly increasing the power consumption of DRAM chips.

With the objective to solving these problems, memory manufacturers are continuously pursuing circuit and technology innovations. Tunneling based RAM's (TRAM's) proposed in [1]-[2] are of interest because of their great potentials in increasing critical charge, while reducing power consumption due to dispensing with mandatory refreshing of cells in DRAM's. A TRAM cell is composed of a conventional DRAM cell being augmented by co-integrating along the cell capacitor a pair of series connected resonant tunneling diodes (RTD), as illustrated in Fig. 1.

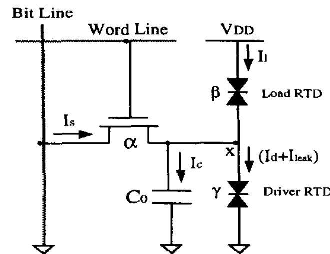

Figure 1: Schematic of 1T RTD-based RAM. α, β, and γ are sizing parameters.

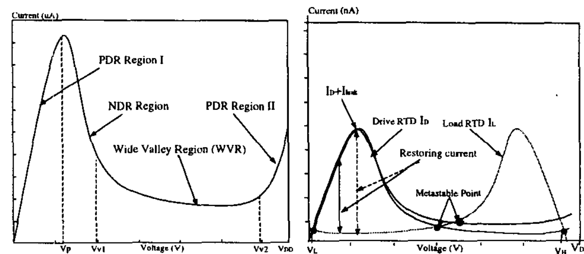

Figure 2: (a) RTD I-V characteristic. (b) Bistable property of RTD-pair.

RTD has a novel I-V characteristic as illustrated in Fig. 2(a). Instead of having a monotonic I-V characteristic, RTD has two positive differential resistance (PDR) regions interspersed by a negative differential resistance (NDR) region. This nonlinear tunneling characteristic renders RTD into a very promising device for a wide class of circuit applications, namely, multi-valued logic, high-speed and low-power circuits, and radiation-hardened reliable circuits. Two series connected RTDs have the self-latching or bistable property as is shown in Fig. 2(b). The RTD-pair can latch at either V_H or V_L, corresponding to logic '1' or logic '0', respectively. This bistable property of the RTD-pair in TRAM can be exploited to improve the soft error immunity, the standby power consumption and the speed of memories. From circuit design point of view, however, a detailed analytical study of the impact of augmentation of conventional DRAM cell by a RTD-pair is necessary.

In this paper, an exact analysis of speed, soft errors and power consumption in a TRAM is presented. The organization of the paper is as follows. In Section II, an analytical study of speed is

0-7803-7761-3/03/$17.00 ©2003 IEEE

given and the formulas are validated by HSPICE simulation. In Section III, the critical charge, one of the most important parameters in SER analysis, is derived. A comparison between the critical charge of TRAM and that of DRAM is presented. Finally, the power consumption of TRAM is analyzed in Section IV.

2. READ AND WRITE OPERATIONS ANALYSIS

In conventional DRAM, the READ operation is destructive. The read access time is therefore increased by the decreasing drive ability. In the case of TRAM, as shown in Fig. 2(b), the restoring current generated by the RTD-pair will help to drive the bit-line capacitance. Therefore, the TRAM read access time is potentially smaller than that of the conventional DRAM. For the WRITE operation, the restoring current plays two conflicting roles. Initially, it opposes the transition between V_L and V_H. But once the voltage of the storage node reaches the meta-stable point, the restoring current begins to help the switch over. Therefore, for the WRITE operation, the RTD-pair should be sized very carefully to obtain a reasonable fast speed. In this section, we first analytically study the READ and WRITE operations. Then, the results are validated with HSPICE and compared with conventional DRAM.

2.1. Analytical study

TRAM is stable at either V_L or V_H in the standby mode. These stable values are not exactly V_{SS} or V_{DD} and are determined by the I-V characteristic of the RTD-pair. For simplicity, we use the piece-wise linear model [3] for the resonant tunneling diodes:

$$I = \begin{cases} \frac{1}{R_1}V & 0 \leq V < V_p \quad (PDR-I) \\ I_p + (V - V_p)\frac{1}{R_n} & V_p \leq V < V_{v1} \quad (NDR) \\ I_{v1} & V_{v1} \leq V < V_{v2} \quad (WVR) \\ I_q + (V - V_{DD})\frac{1}{R_2} & V_{v2} \leq V \leq V_{DD} \quad (PDR-II), \end{cases} \tag{1}$$

$$R_1 = \frac{V_p}{I_p}, \quad R_n = \frac{V_p - V_{v1}}{I_p - I_{v1}}, \quad R_2 = \frac{V_{DD} - V_{v2}}{I_q - I_{v1}},$$

where I_p and V_p represent the peak current and peak voltage; I_{v1} and V_{v1} are the valley current and valley voltage; and I_q is used to model the second peak current at $V = V_{DD}$. R_1 is the resistance of the PDR-I region; R_n is the resistance of the NDR region; and R_2 is the resistance of PDR-II region.

Let x denote the voltage of the storage node of the TRAM. We use x_0 and x_1 to respresent the voltages of logic '0' and logic '1', respectively. To balance the performance of READ '0' and READ '1', we choose β equals γ. At stable point, the driver RTD current (I_d) equals to the load RTD current (I_l). We derive x_0 and x_1 as:

$$x_0 = I_q R_{12}, \quad x_1 = V_{DD} - I_q R_{12}, \quad R_{12} = (R_1 R_2)/(R_1 + R_2).$$

Therefore, in TRAM, the voltage of the storage node is determined by RTD's characteristic instead of V_{SS} and V_{DD}.

2.1.1. Analysis of READ *operation*

In DRAM, READ operation is always destructive. In TRAM, READ operation can also be destructive if $\alpha|I_{sr}| > \beta(I_p - I_v)$, where I_{sr} is the read access current. We use κ to denote the size ratio β/α. The minimal κ value κ_{min} that makes the READ operation non-destructive is $\kappa_{min} = \alpha|I_{sr}|/(I_p - I_v)$. In READ operation, the access transistor operates within a small region. Therefore, we use a linear model for the access transistor: $I_{sr} = A - Bx$. Let the charge (discharge) current of the storage node be I_c. Using KCL at the storage node, we obtain the waveform expression for the bit-line voltage as follows:

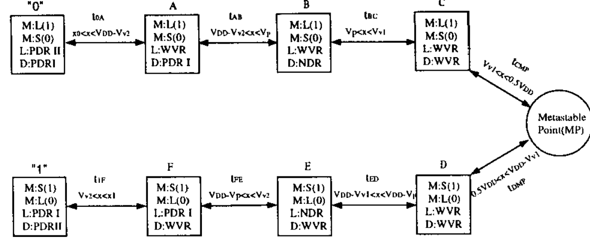

Figure 3: TRAM WRITE operation microstate diagram. $M : S(1)$, $M : S(0)$, $M : L(1)$ and $M : L(0)$ represent the saturation (S) and linear (L) operation region of the access transistor. (1) denotes WRITE '1' operation and (0) denotes WRITE '0' operation.

$$\Delta V(t) = \frac{\alpha}{C_{bit}} \left\{ \frac{B}{\lambda^*}(x_0 - \frac{A^*}{B^*})(e^{-\lambda^* t} - 1) + (A - B\frac{A^*}{B^*})t \right\}.$$

For READ '0' operation, we have $A^* = A + \Delta A_0$, $B^* = B + \Delta B_0$, and $\lambda^* = \alpha B^*/C_0$. The values of ΔA_0 and ΔB_0 are derived as:

$$\Delta A_0 = \kappa I_q, \quad \Delta B_0 = \frac{\kappa}{R_{12}} \quad (0 \leq x < V_{DD} - V_{v2}),$$
$$\Delta A_0 = \kappa I_v, \quad \Delta B_0 = \frac{\kappa}{R_1} \quad (V_{DD} - V_{v2} \leq x < V_p).$$

For READ '1' operation, $A^* = \Delta A_1 - A$, $B^* = \Delta B_1 - B$, and

$$\Delta A_1 = \kappa\frac{V_{DD}}{R_{12}} - \kappa I_q, \quad \Delta B_1 = \frac{\kappa}{R_{12}} \quad (V_{v2} \leq x < V_{DD}),$$
$$\Delta A_1 = \kappa\frac{V_{DD}}{R_1} - \kappa I_v, \quad \Delta B_1 = \frac{\kappa}{R_1} \quad (V_{DD} - V_p \leq x < V_{v2}).$$

2.1.2. Analysis of WRITE *operation*

For WRITE operation, the restoring current of the RTD-pair is against the state transition at first and then it helps to flip the state. Therefore, the write operation is only successful when $\alpha|I_{sw}| > \beta(I_p - I_v)$, where I_{sw} is the write access current. And κ should be smaller than $\kappa_{max} = \alpha|I_{sw}|/(I_p - I_v)$. The microstate diagram of WRITE operation is shown in Fig. 3.

The write access time is the time when the voltage of the storage node reaches the meta-stable point. It can be obtained according to the microstate diagram as follows:

$$T_1 = t_{0A} + t_{AB} + t_{BC} + t_{CM},$$

$$T_0 = t_{1F} + t_{FE} + t_{ED} + t_{DM},$$

where T_1 and T_0 represent the write '1' access time and write '0' access time, respectively. By using KCL at storage node, we find that for each transition region, the transition time (each term in T_1 and T_0) can be computed by using the identical expression as follows:

$$t = \frac{1}{\hat{\lambda}} \ln \frac{a^* - b^* x_{t0}}{a^* - b^* x} + t_0,$$

where, $\hat{\lambda} = \alpha b^*/C_0$. The parameters a^* and b^* are different for each transition region. They are determined by the access transistor parameters A, B and the RTD electrical parameters. Table 1 shows a^* and b^* for all different transition regions.

Table 1: Parameters for WRITE operations.

	State Transition (WRITE '1')			
	'0'→A	A→B	B→C	C→M
a^*	$A+\kappa I_q$	$A+\kappa I_v$	$A+\kappa(I_v-I_p+\frac{V_p}{R_n})$	A
b^*	$B+\frac{\kappa}{R_{12}}$	$B+\frac{\kappa}{R_1}$	$B+\frac{\kappa}{R_n}$	B
	State Transition (WRITE '0')			
	'1'→F	F→E	E→D	D→M
a^*	$-A-\kappa Iq$ $+\kappa\frac{V_{DD}}{R_{12}}$	$-A-\kappa I_v$ $+\kappa\frac{V_{DD}}{R_1}$	$-\kappa(I_v-I_p+\frac{V_p}{R_n})$ $-A+\frac{\kappa V_{DD}}{R_n}$	$-A$
b^*	$-B+\frac{\kappa}{R_{12}}$	$-B+\frac{\kappa}{R_1}$	$-B+\frac{\kappa}{R_n}$	$-B$

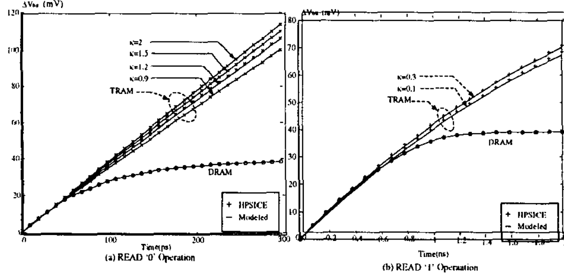

Figure 4: Bit-Line waveforms comparison during READ operation.

2.2. Validation and comparison

Figure 4 shows the bit-line waveforms during READ operation for TRAM and DRAM. We have used $V_{DD} = 1.6V$, $C_0 = 20fF$, $C_{bit} = 400fF$ and $\alpha = 1$. The κ_{min} is 0.8 for READ '0' operation and a much smaller value 0.1 for READ '1' due to small I_{sr} caused by the body effect of the access transistor. Since increasing κ does not help the READ '1' access speed which is limited by the access transistor current, we have used κ to be 0.1 and 0.3 to compare with conventional DRAM as shown in Fig. 4(b). We finally choose the κ value of 0.9 in the effort to optimize both READ '0' and READ '1' operations in the comparison. The speed improvements of READ '0' and READ '1' operations compared with DRAM are 27.8% and 9.2% at $\Delta V_{bit} = 30mV$, respectively. As is shown, the derived result ΔV_t matches the result of HSPICE simulation very well in a large ΔV_{bit} region and the relative value keeps increasing when ΔV_{bit} of DRAM saturates. Therefore, TRAM does not require stringent sensitivity of sense amplifier and can get by using simple sense amplifier.

Further research shows that the derived WRITE results also agree with the experimental results obtained by HSPICE simulation. For the WRITE operation, with $\alpha = 1$, $V_{DD} = 1.6V$, the κ_{max} in our case is 1.4. The WRITE '0' speed improvement compared with DRAM is 37.8% at $\kappa = 0.9$. For WRITE '1' operation, due to the body effect and V_T drop, the write time is determined by the access transistor. However, TRAM also shows comparable speed as conventional DRAM.

3. CRITICAL CHARGE ANALYSIS

Soft error in memories is becoming a critical issue as technology continues to shrink. Soft error occurs at the storage node of memory cells when the induced external charge is larger than the critical charge (Q_c). Therefore, the critical charge is one of the most important parameters for estimating the soft error rates in memories. In this section, we analytically study the critical charge in TRAM and compare it with conventional DRAM technology.

In both DRAM and TRAM, the worst case for charge collection happens during the READ operation. The critical charge for DRAM cell can be expressed as [4]:

$$Q_c = \frac{1}{2}C_0V_{node} - (C_{bit}+C_s)\Delta V_{sen},$$

where C_0, C_{bit}, V_{node} and ΔV_{sen} are the storage capacitance, bitline capacitance, storage node voltage and sense margin voltage, respectively.

The TRAM will flip when the meta-stable point is met. Let x_m represent the meta-stable point voltage. When the memory cell is in the standby mode, the critical charge for logic '0' and logic '1' are $Q_{sc0} = C_0(x_m - x_0)$ and $Q_{sc1} = C_0(x_1 - x_m)$, respectively. Since the two RTDs are identical, in the standby mode, we have

$$Q_{sc0} = Q_{sc1} = C_0(\frac{1}{2}V_{DD} - I_qR_{12}).$$

In the worst case scenario, the critical charge Q_{r1} for READ '1' operation is obtained as:

$$Q_{r1} = Q_{sc1} - C_0(x_1 - x), \qquad (2)$$

$$x = \begin{cases} (\kappa(\frac{V_{DD}}{R_1} - I_v) - A)/(\frac{\kappa}{R_1} - B) & (V_{DD} - V_p \le x < V_{v2}) \\ (\kappa(\frac{V_{DD}}{R_{12}} - I_q) - A)/(\frac{\kappa}{R_{12}} - B) & (V_{v2} \le x < x_1). \end{cases}$$

Similarly, the critical charge Q_{r0} for READ '0' operation is obtained as:

$$Q_{r0} = Q_{sc1} - C_0(x - x_0), \qquad (3)$$

$$x = \begin{cases} (\kappa I_v + A)/(\frac{\kappa}{R_1} + B) & (V_{DD} - V_{v2} \le x < V_{vp}) \\ (\kappa I_q + A)/(\frac{\kappa}{R_{12}} + B) & (x_0 \le x < V_{DD} - V_{v2}). \end{cases}$$

Equation (2) and (3) show that critical charge is determined by the electrical characteristic of the RTD device, the supply voltage, as well as the size ratio κ. As shown in Fig. 5. Q_c initially increases very quickly and then tends to be saturated with the increase of κ. In order to get large critical charge for both READ '1' and READ '0' operations with small area penalty, κ should be choose properly. Table 2 shows the comparison of the critical charge in TRAM and DRAM. We have used $V_{DD} = 1.6V$, $\kappa = 0.9$. A and B are extracted from a 0.18 micron process technology by curve fitting the I-V characteristic generated by the HSPICE Level-49 model. The result shows that even in the worst case, TRAM still has considerably larger critical charge than DRAM.

4. POWER CONSUMPTION ANALYSIS

TRAM has lower power consumption than conventional DRAM. First of all, the bistable property eliminates the requirement of refreshing operation. Leakage currents are replenished by restoring current of the RTD-pair. Second, in DRAM, due to the cell leakage current variation, the worst-case leakage current has to be accounted for at each cell when performing the refreshing of cells which means that power consumption for refreshing operation is

Table 2: Critical charge comparison of TRAM and conventional DRAM

	DRAM Q_C (fC)						TRAM		
$C_{bit}(fF)$	250		300		350		Q_C (fC)		
C_0 (fF)	$\Delta V_{sen}(mV)$ 40	$\Delta V_{sen}(mV)$ 50	$\Delta V_{sen}(mV)$ 40	$\Delta V_{sen}(mV)$ 50	$\Delta V_{sen}(mV)$ 40	$\Delta V_{sen}(mV)$ 50	Q_{sc} (fC)	Q_{r1} (fC)	Q_{r0} (fC)
40	20.4	17.5	18.4	15.0	16.4	12.5	28.0	27.2	22.1
35	16.6	13.8	14.6	11.3	12.6	8.80	24.5	23.8	19.4
30	12.8	10.0	10.8	7.50	8.80	5.00	21.0	20.4	16.6
25	9.00	6.25	7.00	3.80	5.00	1.25	17.5	17.0	13.8

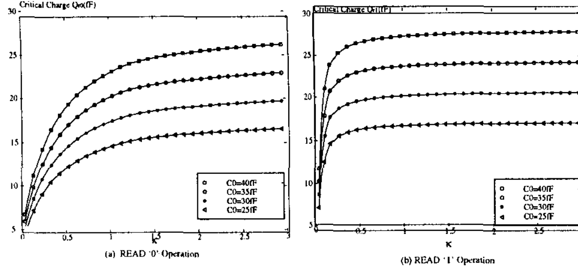

Figure 5: Critical charge of READ operation w.r.t. κ.

much more than what is actually required. In case of TRAM, the restoring current always equals to the actual leakage current of each single cell. Therefore, the power consumption is reduced further. In this section, we derive the power consumption of TRAM and then compare it with conventional DRAM.

Static power consumption of TRAM is determined by the leakage current and the dc current of the RTD-pair [5]. Assuming the average leakage current of the cell is I_{leak}, in order to guarantee the bistability of the TRAM, the following condition should be satisfied: $I_P > I_{leak} + I_v$. Let δ be the ratio of the maximum leakage current I_{leak}^{max} and the average leakage current I_{leak} in a DRAM chip. For all cells, the valley current I_v should be larger than a minimum valley current I_v^{min}:

$$I_v^{min} = \delta \frac{I_{leak}}{PVCR - 1}.$$

Therefore, for the worst case that the maximal leakage current is assumed for all the cells, the standby power is obtained as:

$$P_{TRAM} = I_{leak} V_{DD} \left(1 + \frac{\delta}{PVCR - 1}\right).$$

Because of the dynamic compensation of the leakage current, the standby power is then given by

$$P_{TRAM}^{Avg} = I_{leak} V_{DD} \left(\frac{1}{\delta} + \frac{1}{PVCR - 1}\right).$$

The power consumption of the DRAM due to the refresh operation is given by [5]:

$$P_{DRAM} = \delta I_{leak} V_{DD} \left(1 + \frac{C_{bit}}{C_0}\right) \frac{1}{1 - \frac{C_{bit}}{C_0} \frac{1}{\frac{V_{DD}}{2V_r} - 1}},$$

where V_r is the sense margin of the sense amplifier. Therefore, the power consumption of TRAM versus that of DRAM is derived as:

$$\eta = \frac{P_{TRAM}^{Avg}}{P_{DRAM}} = \frac{\left(\frac{1}{\delta} + \frac{1}{PVCR-1}\right)}{\delta\left(1 + \frac{C_{bit}}{C_0}\right) \frac{1}{1 - \frac{C_{bit}}{C_0} \frac{2V_r}{V_{DD} - 2V_r}}}$$

Using the typical value $\delta = 50$, $PVCR = 10$, $C_{bit}/C_0 = 10$, the power consumption ratio is estimated to be $\eta \simeq 10^{-2}$ for an ideal sense amplifier with $V_r = 0$. With increase of the DRAM density, the power consumption ratio η will continue to decrease.

5. CONCLUSIONS

In this paper, we analytically model the READ and WRITE operations, critical charge for soft errors, and power consumption of single-transistor tunneling based RAM. The results are validated by HSPICE simulation. The size ratio of the RTD pair to the access transistor plays an important role in determining the access speed and the critical charge of TRAM. We show that critical charge is not as sensitive to ΔV_{sen} as in DRAM due to its self-latching property. The analytical study of power consumption shows that the dynamic compensation to the leakage current by the restoring current of RTD-pair will reduce the power consumption by one or two orders of magnitude. In a nut shell, TRAM has a great potential in future high-density, low-power, fast and highly reliable memory design. This paper analytically establishes the above claim.

6. REFERENCES

[1] J. P. A. van der Wagt, "Tunneling-Based SRAM," *Proceedings of the IEEE*, vol. 87, no. 4, pp. 571-595, Apr. 1999.

[2] E. Goto, *et al.*, "Esaki Diode High-speed logical circuits," *IRE Trans. Electron. Comput.*, vol. 9, pp. 25-29, Mar. 1960.

[3] S. Mohan, J. P. Sun, P. Mazumder, and G. I. Haddad, "Device and Circuit Simulation of Quantum Electronic Devices," *IEEE Trans. CAD*, vol. 14, no. 6, pp. 653-662, June 1995.

[4] S. Hyungsoon, "Modeling of Alpha-Particle-Induced Soft Error Rate in DRAM," *IEEE Trans. Elec. Dev.*, vol. 46, no. 9, pp. 1850-1857, Sept. 1999.

[5] T. Uemura and P. Mazumder, "Design and Analysis of Resonant Tunneling Diode (RTD) Based High Performance Memory System," *IEICE Trans. on Electrons*, vol. E82-C, no. 9, pp. 1630-1637, Sept. 1999.