T. C. Lo
R. E. Scheuerlein
R. Tamlyn

A 64K FET Dynamic Random Access Memory: Design Considerations and Description

The major design considerations and design features of an experimental 64K-bit random access memory (RAM), implemented in a double polysilicon gate technology, are described in this paper. Design tradeoffs addressing power supply selection and chip configuration alternatives are presented. This is followed by a description of the 8K-word by 8-bit FET dynamic RAM which uses single-transistor cells with first-level-metal bit lines and second-level-metal stitched polysilicon word lines. Test results obtained on early engineering hardware chips fabricated with linear dimensions 1.2 times those of the base design are also presented.

Introduction

In comparison with the emerging 64K RAMs which use 5 V power supply and have a single-bit-per-word organization, in the first part of this paper it is shown that a configuration using 8.5 V and organized as 8K words by 8 bits has particular advantages with regard to storage charge capacity, chip size, performance, and testing efficiency without impacting reliability, error-correction-code (ECC) implementation, or system power dissipation.

In the second part of the paper, a description of an 8K-word by 8-bit FET dynamic RAM capable of 100-ns access time is given to exemplify this design philosophy. A salient design feature is the use of a single transistor cell laid out in a unique way to improve transfer ratio and to facilitate sense-amplifier and decoder interfacing.

The base design uses 2-μm-minimum-channel-length n-channel silicon gate technology with two-level metallization to improve packing density and word-line noise immunity. The experimental chips reported in this paper were fabricated with images expanded by a factor of 1.2 in linear dimensions to facilitate early parts fabrication. The chip area, before expansion, measures less than 18 mm^2.

The RAM operates with two clocks, making possible four modes of operation: *read, write, page,* and *refresh*. Power supplies required are +8.5 V, +5.0 V, and −2.2 V. All input and output signals, including the two clocks, are TTL-level compatible.

Measured results for this design are presented in the last part of this paper.

Part 1: Design approach

Two primary considerations in the design of the 64K RAM are the selection of power supply voltages and the data organization, or configuration, of the chip. These considerations are now presented.

• *Power supply selection*

For a given technology and design, the optimum applied voltage to maximize performance is the highest voltage allowed by reliability limitations. The technology under consideration is compatible with the state of the art of manufacturing tooling and has the following attributes: minimum channel length is 2 μm; oxide thickness for storage capacitors and peripheral circuits is 45 nm; lateral junction depth is 0.5 μm; and average feature size is

2.8 μm, where feature size is defined as the sum of shape width and shape spacing divided by two.

The following are the main interrelationships among charge capacity, density, performance, and supply voltage.

The total charge capacity per unit area Q_T of the storage node increases with increasing supply voltage V_{DD} according to the relationship

$$Q_T = C_G(V_{DD} - V_T), \tag{1}$$

where C_G is the storage gate capacitance per unit area and V_T is the threshold voltage of the switch device of the RAM cell.

Since leakage also increases with V_{DD}, the net effect must be examined. Charge loss per unit area due to leakage has the following three components [1]:

$$Q_{leak} = t_{ref}\left(\frac{qn_iS_0}{2} + \frac{qn_iW_D}{2\tau_0} + \frac{qn_i^2D_n^{1/2}}{N_A\tau_n^{1/2}}\right), \tag{2}$$

where t_{ref} is the refresh time, n_i is the intrinsic carrier density, q is the electronic charge, S_0 is the surface recombination velocity, W_D is the depletion depth, τ_0 is the average carrier lifetime in the depletion region, D_n is the electron diffusivity, N_A is the substrate doping concentration, and τ_n is the electron lifetime.

All three terms in Eq. (2) are temperature dependent, due mainly to n_i; only the second term is voltage dependent. Since the depletion depth of the storage element in the second term is approximately proportional to $V_{DD}^{1/2}$ rather than V_{DD}, and the third term, which is not voltage dependent, predominates at high temperature due to n_i^2, leakage charge will not increase as fast as Q_T, and therefore the net charge capacity Q_{net} $(= Q_T - Q_{leak})$ increases with increasing voltage. Calculations have shown that, going from 5 V to 8.5 V, the net charge capacity per unit storage area increases 1.9 times, thereby suggesting that the RAM cell size can be made 0.7 times smaller, assuming a sensing signal of 200 mV and 64 cells per bit-line half.

Circuit delay time T_D, to a first-order approximation, is given by

$$T_D = \frac{V_{DD}C_L}{I_{DD}},$$

$$I_{DD} \propto V_{DD}^2, \text{ and}$$

$$\therefore T_D \propto \frac{1}{V_{DD}}, \tag{3}$$

which suggest that access time is inversely proportional to V_{DD}. Here C_L is the loading capacitance of a circuit node and I_{DD} is the device drain current that charges and discharges that node.

Computer simulations, which included saturation velocity effects, have shown that for an effective channel length $\ell_{eff} = 2.8 \pm 0.8\,\mu$m, access time (both best- and worst-case) for an 8.5-V design is 0.6 times that of a 5-V design. For the same device limitation constraints, ℓ_{eff} could be made shorter by 0.4 μm for 5-V operation to enhance performance. Given a shorter ℓ_{eff} of $2.4 \pm 0.8\,\mu$m for a 5-V design, an 8.5-V design would still have 0.75 times performance improvement for maximum ℓ_{eff} and 0.93 times improvement for minimum ℓ_{eff}. In order to fully restore the performance, an oxide thickness thinner than that allowed by manufacturing considerations would be required.

One disadvantage associated with the use of 8.5 V is excessive power dissipation. Assuming that dynamic power (proportional to $C_LV_{DD}^2$) predominates, power dissipation will be 2.9 times higher than that of the 5-V design. But, since the actual component power dissipation remains at about the same level (300 mW) as the existing 16K RAMs, which have been proven coolable in their applications, this shortcoming is outweighed by better performance and smaller chip size.

Device limitations are next examined for adverse impact on reliability when the 8.5-V power supply is used.

Impact-ionization-induced device-sustaining voltage, which is the drain-to-source voltage above which the drain current increases regeneratively until thermal self-destruction occurs, can be improved by adding a deep implant following the conventional arsenic source/drain implant [2]. Experimental data for minimum channel devices (2.0 μm) are above 11 V. This is adequate for 8.5-V operation and overvoltage stress testing. Measured pn-junction breakdown voltage exceeds 20 V, and the gated diode breakdown voltage with grounded gate exceeds 15 V.

Hot-electron injection, which causes V_T shift, becomes worse at high voltage. There are two types of hot electrons: substrate hot electrons (SHE) and channel hot electrons (CHE). The former are normally negligible in most practical cases [3], and only the latter need be considered. CHE injection occurs predominately when the drain and the gate voltages are equal. Our experimental data have shown that the CHE-induced V_T shift after 200 hours dc stress is less than 100 mV. These data have been extrapolated to long-term V_T instability. Assuming that the device lifetime is 100 000 h and that the device is sus-

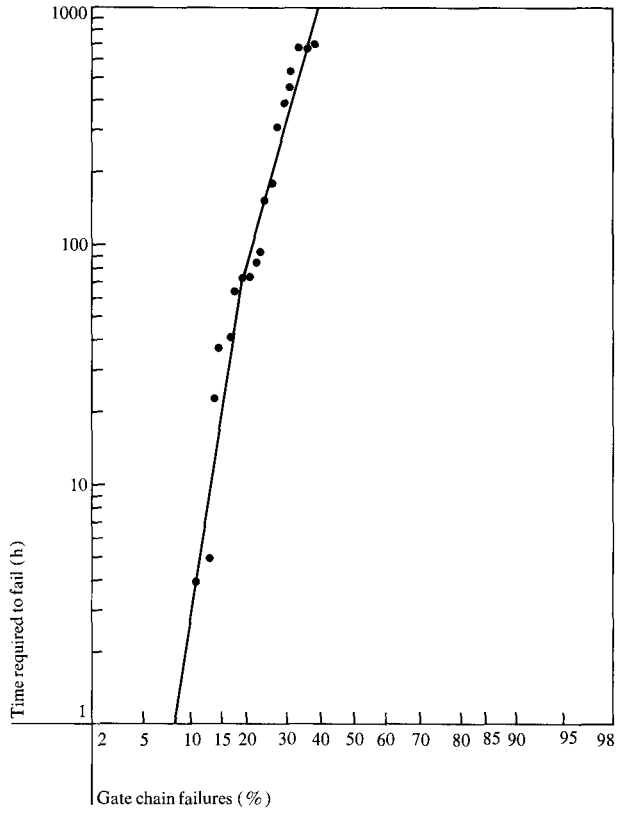

Figure 1 Failure rate (percentage of gate chain failures) as a function of stress time (time required to fail) under stress condition: $E = 4 \times 10^6$ V/cm, $T = 250$°C, and for 50 nm, 71 gate chains.

ceptible to hot electrons (*i.e.*, $V_D = V_G$) for only 5% of that lifetime, the end-of-life V_T shift doubles the stress result and amounts to 200 mV, which is acceptable from a circuit design viewpoint. It should be noted that the data were taken at $V_D = V_G = 10.4$ V, and therefore the V_T shift would be even smaller for 8.5-V operation.

Gate oxide integrity is considered next. Figure 1 shows the stress time required to produce the designated percentage of failures. A gate array with oxide thickness of 50 nm was used because, at the time this paper was written, test results for the 45-nm gate array were not available. The following equation is a model for the acceleration factor AF, which can be used to translate the stress condition into the use condition [4]:

$$AF = \frac{\text{time to fail at use condition}}{\text{time to fail at stress condition}} = \exp\left[\frac{\Delta H}{k}\left(\frac{1}{T_{use}} - \frac{1}{T_{stress}}\right) + \beta(V_{stress} - V_{use})\right], \quad (4)$$

where we have used $\Delta H = 0.5$ eV and $\beta = 61/d = 1.2$, and where

d = oxide thickness = 50 nm,
T_{use} = 358 K (85°C),
T_{stress} = 523 K (250°C),
V_{use} = 10 V (to achieve 2×10^6 V/cm),
V_{stress} = 20 V (4×10^6 V/cm), and
k = Boltzmann's constant.

Therefore, $AF = \exp(5.11 + 12) = 2.7 \times 10^7$, and

$$\text{failure rate} = \frac{40\% \text{ at } 1000 \text{ h under stress}}{AF} = 1.5 \times 10^{-6}\% \text{ per kh.}$$

Since the total gate area of the 64K RAM chip to be described is approximately 100 times larger than that of the gate chain under stress, the chip-level oxide-failure rate will be of the order of 10^{-4}% per 1000 hours, which is quite acceptable.

The foregoing discussions have shown that the use of an 8.5-V power supply is preferable to the use of 5 V from a performance and density point of view, without jeopardizing cooling and reliability.

• *Memory chip configuration selection*

The key system design considerations which must be taken into account in the chip configuration selection are the following.

The typical CPU-to-memory-bus bandwidth requirement is many times larger than that available from individual memory chips. This mismatch forces the simultaneous selection of many memory chips at a higher level of assembly. Therefore, memory chips organized as multiple-bit words offer advantages over a by-one-bit organization from a system-power-dissipation point of view because fewer chips are selected to satisfy the fundamental bandwidth requirement of the memory bus. With the use of the by-eight organization described in this paper, the memory system power dissipation of an 8.5-V system is actually lower than that of a 5-V system using a by-one-chip organization.

Also, the multiple-bit-word chip organization gives rise to finer granularity for storage expansion. This is particularly important to medium- and low-end users.

The apparent inconvenience with multiple-bit-word organization is ECC implementation. The ECC feature is more effective when each bit in the accessed word comes from a different chip, so that the failure of a chip merely results in the loss of a single bit in a word, which can easily be corrected. The by-one configuration inherently fulfills this requirement. For systems using by-eight

chips, the ECC criterion can be met by using a parallel-bit-serial-word (PBSW) memory architecture. In this scheme, a large number of bits called a page is first loaded in parallel into a fast memory called a fetch register, and a page is subdivided into several groups of bits called words so that each bit of a word comes from a different chip. The words are then read out one at a time from the register. A store operation is performed in a similar fashion.

The particular chip configuration chosen also has implications for testability. A by-eight chip requires shorter test time because each channel, which amounts to 1/8 of the total number of bits, can be tested in parallel. This parallel configuration significantly increases testing throughput.

All in all, a by-eight chip configuration outstrips the by-one configuration in terms of system power dissipation and granularity without penalizing ECC implementation. Using an 8M-bit capacity as a design example, the tradeoffs showing the advantages of the by-eight chip configuration are given in Table 1.

Part 2: Description of a 64K RAM design

Key design features are highlighted in this part of the paper. Among these are novel cell layout, optimum chip organization for maximum density, and self-gated clocking logic to compensate for process and environmental parameter variations.

- *Technology*

The developmental 64K RAM is designed in an n-channel silicon-gate technology with two layers of polysilicon to reduce cell size, and with two layers of metallization to improve peripheral circuit layout efficiency and word-line noise suppression. No buried contacts or depletion load devices are used in this design, since the area savings were found to be insufficient to justify additional processing complexity. Other processing attributes were discussed in Part 1.

To alleviate the photolithography tooling limitations for early fabrication, the RAM was initially fabricated with images expanded by a factor of 1.2 over the base design. The gate oxide thickness of the poly-1 (first-level polysilicon) devices, used in all peripheral circuits and storage capacitors, is 45 nm, and the gate oxide thickness of the poly-2 (second-level polysilicon) devices, used only in the transfer switches of all RAM cells, is 65 nm. The thicker poly-2 gate oxide improves poly-1 to poly-2 inter-oxide integrity as well as RAM cell refresh stability (discussed subsequently in the sense amplifier section).

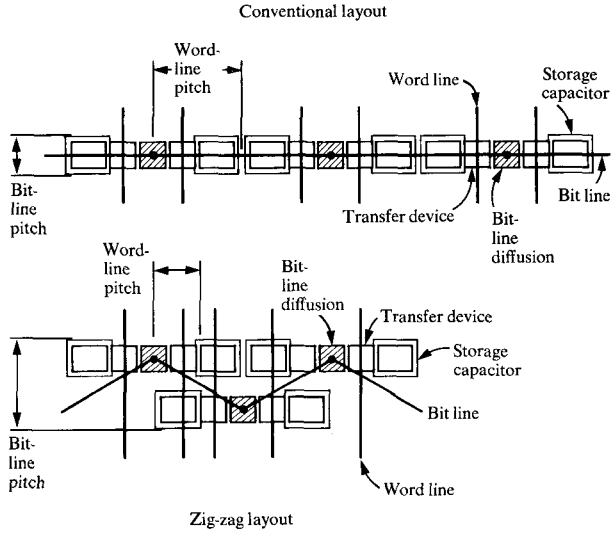

Figure 2 Zig-zag concept for metal-bit-line cell.

Table 1 Memory component architecture comparison for an 8M-bit capacity example using 64K RAMs.

	8 bits out	*1 bit out*
Relative array power and current surge	8 chips on 120 chips standby	64 chips on 64 chips standby
ECC implementation	Yes with PBSW architecture	Yes
Granularity	512K bits	4096K bits
Chip design complexity	8 I/O buffers required	One I/O buffer required
Memory testing*	$\left(\frac{N}{8}\right)^a$ parallel testing assumed	N^a

*N = total number of bits on chip. a = test pattern coefficient ($1 < a < 2$).

- *Cell and array structure*

The one-device RAM cell used in this design has a metal bit line and a polysilicon word line with bit-line contacts shared with adjacent cells, resulting in half of a contact per bit. A novel cell arrangement with a zig-zag bit line is used to achieve better storage efficiency and easier peripheral circuit interfacing.

This zig-zag RAM cell used in the present design is different from the conventional metal-bit-line cell discussed in previous work [5]. Figure 2 illustrates the key advantages of this zig-zag cell over the conventional metal-bit-line cell. Both cell groups are drawn in a symbolic form.

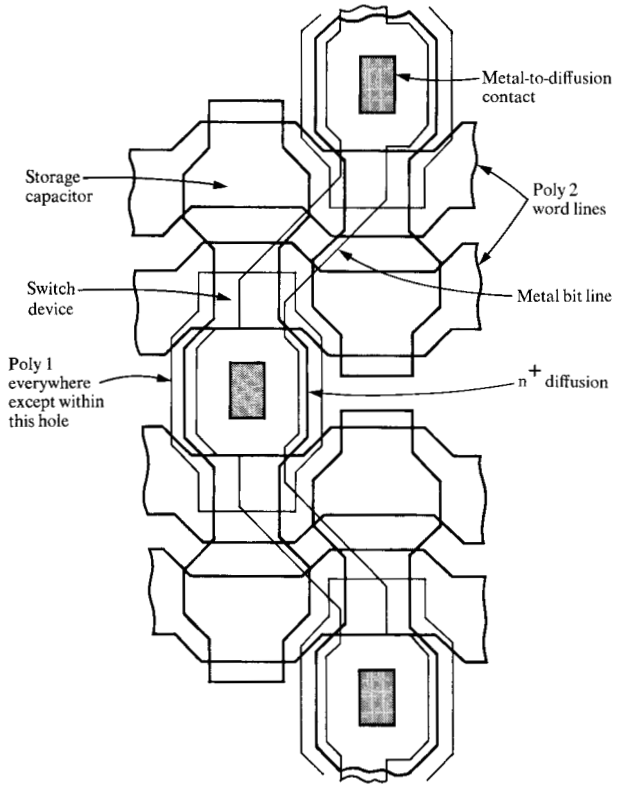

Figure 3 RAM cell with zig-zag metal bit line and polysilicon word line. Second-level-metal word-line strap is not shown.

One advantage is that the pitch between two adjacent bit lines is approximately doubled in the zig-zag metal-bit-line cell, which facilitates the sense amplifier layout. Although this can also be realized with a folded-bit-line approach, the folded-bit-line technique has inherently higher bit-line capacitance because the number of word-line crossings is doubled. The second advantage is that the length of the bit line per cell for the zig-zag cell is approximately half that of the conventional arrangement, resulting in lower bit-line capacitance and therefore higher transfer ratio. An additional advantage of the zig-zag bit-line cell arrangement is that a first-metal wiring path can be used to connect the sense amplifier and the column decoders. The latter can now be placed outside the array, thereby reducing the bit-line overhead capacitance since the bit line does not have to pass through the column decoders, which otherwise are located at the sense amplifier region. In the 2-μm technology the minimum layout limited cell area is 108 μm^2. In this specific design a cell area of 135 μm^2 before expansion was chosen to provide sufficient signal to operate a fast cross-coupled symmetrical sense amplifier (described in a subsequent section). With 64 bits per bit-line half, the ratio of storage capacitance to the capacitance of a bit-line half (defined as transfer ratio) is about 0.07 in the worst-case condition.

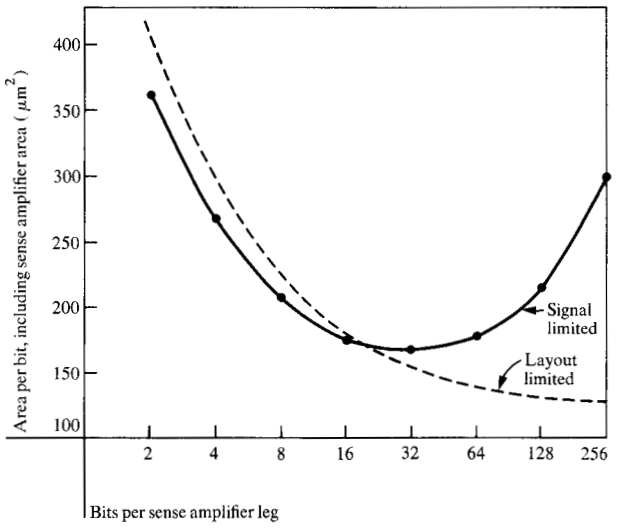

Figure 4 Optimization of number of bits per bit-line half. Assumptions: (1) average feature size = 2.8 μm; (2) minimum signal to sense amplifier = 150 mV.

The key disadvantage of metal-bit-line cell structures is the time constant in the polysilicon word line, which not only degrades word-line rise time but also, more importantly, limits the effectiveness of word-line bounce-suppression circuits used to clamp the unselected word lines, thus resulting in poor word-line noise immunity. This can cause a pattern-dependent degradation of stored data [1]. This concern is removed in this design by using parallel second-level metal which makes contact periodically with the polysilicon word line, with contacts 64 cells apart. Thus the word-line time constant can be reduced from 132 ns to 2 ns. The RAM cell layout used in this design is shown in Fig. 3. In addition, second-level metal also improves the peripheral circuit packaging density, notably in the row decoder region and power bussing.

• *Chip organization*

The number of bits per bit-line half can be optimized for minimum chip area. Having a smaller number of bits per bit-line half requires more sense amplifiers; conversely, having a larger number of bits per bit-line half suggests a larger cell area to satisfy the sensing signal requirements. Figure 4 shows the area per bit, including the sense-amplifier effect. It shows that the optimum number of bits per sense-amplifier leg is 32. Since the optimum point is rather broadly tuned, going from 32 bits to 64 bits makes a negligible difference. In this design 64 bits per sense-amplifier leg is chosen to reduce the peripheral circuit complexity.

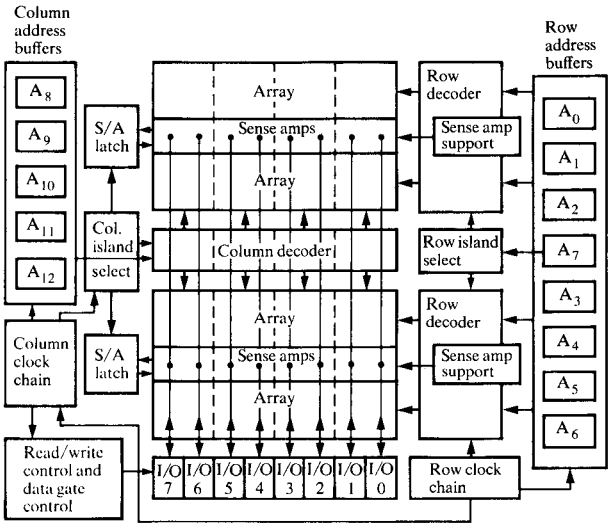

Figure 5 Chip organization of the 64K RAM.

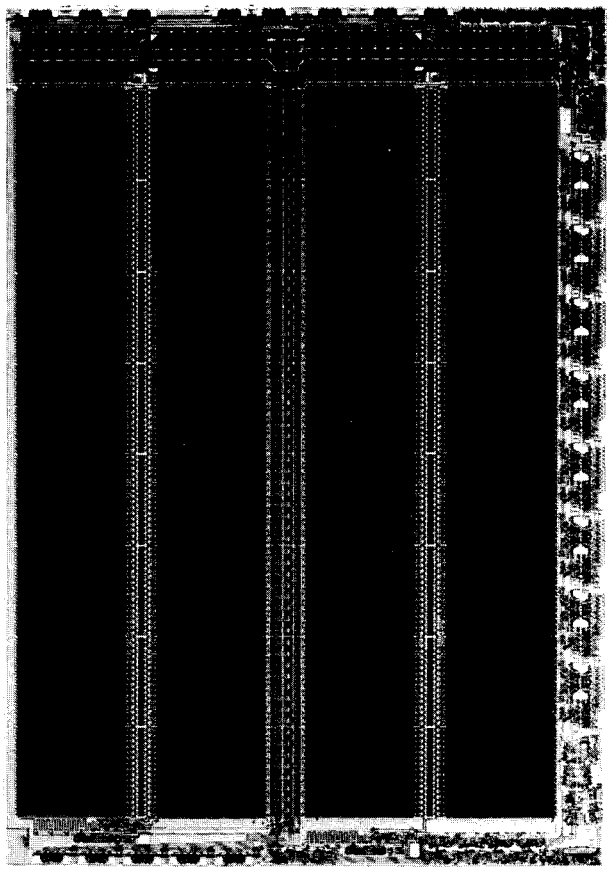

Figure 6 Photomicrograph of the 64K RAM chip. (Note that the chip-to-package interconnections are not shown, for clarity.)

Topologically, this design consists of two islands configured in the manner shown in the chip organization diagram of Fig. 5. A photomicrograph of the chip is shown in Fig. 6. Each island has 256 sense amplifiers and 128 word lines. The column decoder is located in the middle and is shared by both islands. The row decoders are located at the side in order to facilitate peripheral circuit interfacing. In any memory cycle only one island will be selected. The unselected island remains quiescent, thus dissipating less transient power than a single-island approach, because only half the bit lines are precharged per cycle. This also reduces, by a factor of two, the peak current density in the bit lines, sense-amplifier grounding busses, and bit-line precharge busses. These busses would have to be doubled in width if a single-island design were used.

Data masking can be implemented on-chip to provide flexible system applications. The eight I/O buffers shown in Fig. 5 have their own data-gate control which determines which I/O channel should be examined or inhibited. Data-gate control allows the device to function in a flexible manner. Different chip organizations, from 64K by 1 to 8K by 8, can be achieved by manipulating the state of the data gates. This also facilitates the use of chips with one or more nonfunctional octants in the bit dimension.

The 8K by 8 array is divided into four portions, as shown in the dotted lines of Fig. 5. Each portion consists of two 8K by 1 arrays interwoven with each other sharing the corresponding column decoders. The four 8K by 2 arrays are separate except for common word lines. Sense amplifiers, column decoders, and I/O logic circuits are identical for the four different portions. Since the cell plates are also divided into four groups, a defective plate will not cause the whole chip to fail; therefore the yield of partially good chips is enhanced.

• *Principles of chip logic operation*

The operation of the chip can best be described by the flowchart shown in Fig. 7. This design requires two TTL clocks to operate: the row address select clock $\overline{RAS}$, and the column address select clock $\overline{CAS}$. When the $\overline{RAS}$ clock goes low, a chain of circuit activities occurs, starting with the selection of the word line and ending with the sensing and latching of the stored data at the sense amplifiers. Likewise, when the $\overline{CAS}$ clock goes low, another chain of circuit activities occurs to transfer the contents of the eight selected sense amplifiers to the respective output buffers in read mode, or to alter their contents and hence the contents of the accessed cells in write mode via the respective input buffers. The mode of operation is determined by an external TTL signal $\bar{W}$. For a write cycle $\bar{W}$ is low, and for a read cycle $\bar{W}$ is high.

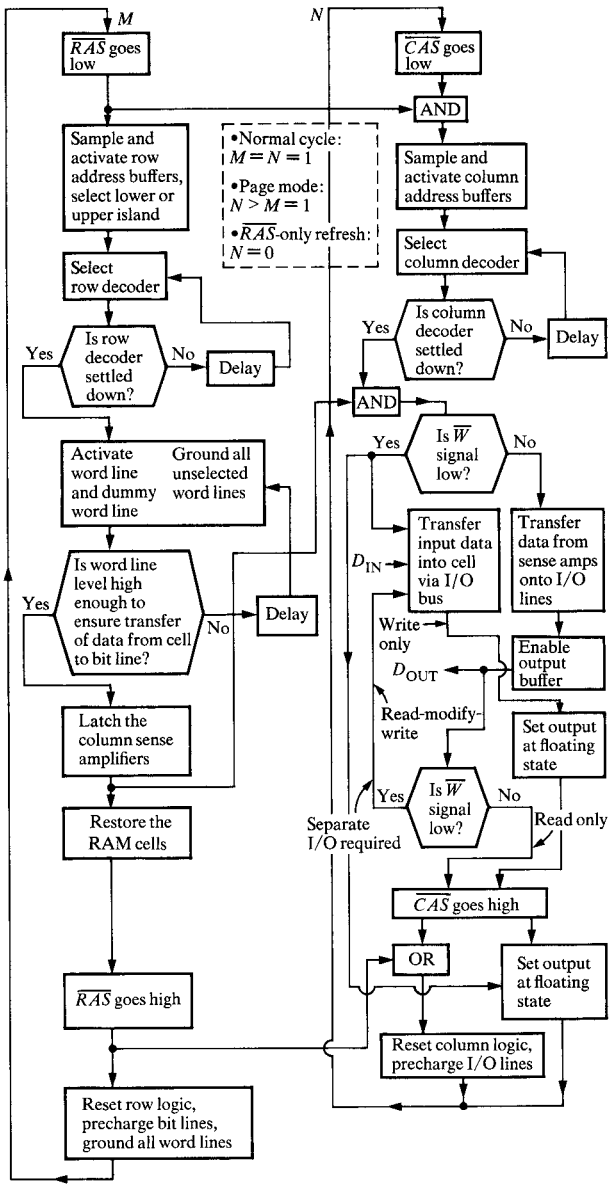

Figure 7 Principles of chip logic operation.
Normal cycle: $\overline{RAS}$ goes low, $\overline{CAS}$ goes low.
Page mode: $\overline{RAS}$ goes low, $\overline{CAS}$ goes low several times.
$\overline{RAS}$-only refresh: $\overline{RAS}$ goes low, $\overline{CAS}$ remains high.

As shown in Fig. 7, both $\overline{RAS}$- and $\overline{CAS}$-initiated activities are self-gated; for example, to avoid multiple selection, the selected word line will not go high until all decoders have been stabilized. Furthermore, the column activities are gated by the row activities; for instance, transfer of the sense-amplifier data initiated by $\overline{CAS}$ will not occur until all sense amplifiers have been made stable

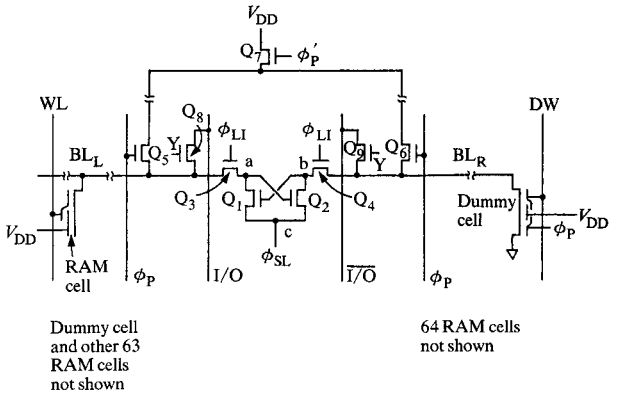

Figure 8 Sense amplifier.

by $\overline{RAS}$ to avoid a premature disturbance of the sensing process. This automatic timing control removes clock delay uncertainties due to process and environmental parameter variations.

In each row activation, data from all RAM cells along the accessed word line are latched and stored in the sense amplifiers.

To read or write the sense amplifiers, one simply activates the $\overline{CAS}$ clock alone; this is known as page-mode operation. Page-mode operation offers shorter access time and faster data rate.

Since row activation involves the latching of sense amplifiers, which in turn regenerate the stored data, memory cells can be refreshed by cycling the $\overline{RAS}$ clock alone, thereby saving the power required to activate the $\overline{CAS}$-related circuits. This is known as $\overline{RAS}$-only refresh.

• *Sense amplifier*

A symmetrical, fully dynamic sense amplifier is used for low-power high-sensitivity signal detection. The basic sense amplifier is shown in Fig. 8 and consists of two devices Q_1 and Q_2, forming a symmetrical cross-coupled flip-flop with 64 cells plus one dummy cell on each side. During the quiescent state, the bit-line halves BL_L and BL_R are precharged to full supply level by means of Q_5, Q_6, and Q_7, whose gates, and also the gates of the loading isolation devices Q_3 and Q_4, are higher than the supply voltage at that time. Prior to access of the RAM cell, device Q_7 is turned off, allowing the potentials on BL_L and BL_R to equalize. Then Q_5 and Q_6 are turned off, causing the pair of bit-line halves to float at precisely the same voltage. The memory cells are interrogated by bringing up

the selected word line WL on one side of the sense amplifier and the dummy word line DW on the opposite side. The dummy cell stores a *zero* and is approximately half the RAM cell size. If the RAM cell being accessed stores a *zero*, it will be read out destructively, causing the corresponding bit-line-half potential to drop below the reference level set by the dummy cell. If the RAM cell stores a *one*, it will be read out nondestructively, causing its bit-line-half voltage to either drop slightly if a "poor" *one* is stored or remain intact if a "good" *one* is stored. The voltage difference between two bit-line halves is presented to the sensing nodes a and b via devices Q_3 and Q_4, which are in a low impedance state at that time. The common source node c, which has been precharged high to ensure the turn-off of Q_1 and Q_2, starts to pull down slowly, giving the cross-coupled devices the opportunity to "feel" the voltage difference on their drains. As the conductance difference increases regeneratively, node c drops at an increasing rate to discharge the intended low side. This two-step sensing optimizes the speed-sensitivity performance of the sense amplifier [6]. During this latching process, isolation devices Q_3 and Q_4 are switched to a high impedance state to preserve the charge on the intended high side. The bit switches Q_8 and Q_9 transfer the selected data to the output buffer via the I/O lines.

After a read operation the contents of the accessed RAM cells are to be restandardized. If a *zero* has been read, the bit-line half that has been pulled down will confirm a good *zero* in the RAM cell again. If a *one* has been read, the precharged high bit-line half will share charge with the RAM cell. Note that an instability can occur when the *one* level written back into the storage capacitor becomes a function of the voltage of the bit-line half from which the RAM cells replenish charge: a "poor" bit-line-half voltage rewrites a "poor" *one* into the RAM cell by way of charge sharing, which in turn translates to an even worse bit-line-half voltage after a subsequent read. The stored *one*, therefore, degenerates in successive read cycles.

To avoid this multiple-read instability, it is important to make sure that the RAM cell is refreshed to a good *one* level which is not a function of the final bit-line-half voltage. One convenient way to restore the RAM cell is to write back on the storage capacitor a saturated voltage which is a constant equal to a threshold drop below the word-line voltage. This is accomplished with the following measures: a) precharge the bit-line halves to the highest possible potential to increase the amount of charges available to share; b) use the loading isolation devices Q_3 and Q_4 to avoid charge lost on the intended high side during the latching process; and c) use a slightly higher threshold device in the switch element of the RAM cell to facilitate the writing of a maximum *one*. Such a slightly higher threshold device can be implemented by using a slightly thicker gate oxide for the switch device of the RAM cell. This also improves interpoly oxide integrity in the double polysilicon gate technology and reduces the bit-line capacitances. Since too high a threshold for the transfer device will degrade the charge capacity of the RAM cell, the desirable poly-2 V_T is set about 300 mV higher than the poly-1 V_T.

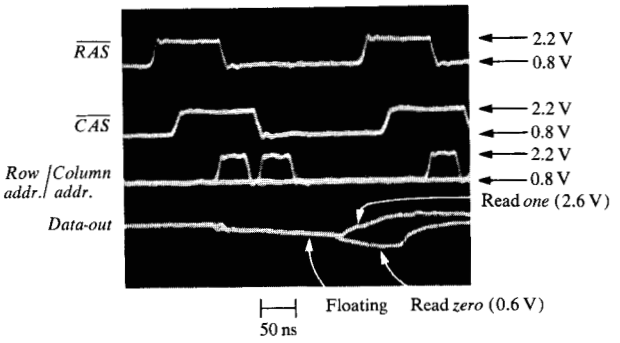

Figure 9 Read-cycle waveforms for scaled-up chip. Notes: (1) $\overline{W}$ set at high; (2) addresses are in multiplexing mode; (3) cable delays have added 24 ns to the actual $\overline{CAS}$ access time.

To write the RAM cell, the bit switches Q_8 and Q_9 are turned on, passing the input data onto the bit-line halves via the complementary I/O lines. The storage capacitor of the RAM cell thus takes the bit-line voltage.

Part 3: Measured performance

A 64K dynamic RAM has been successfully fabricated with a 2.4-μm technology, which is a scaled-up version of our developmental 2-μm technology. Figure 9 illustrates the read-cycle operating waveforms with a row access time of 150 ns and a column access time of 100 ns under worst-case voltage conditions of V_{DD} = (8.5 V − 10%), V_{BB} = − (2.2 V + 10%), and still air ambient, using "march" pattern [1]. A diagnostic study of the internal chip clock nodes has indicated that performance improvements for this particular chip can be made to achieve 98 ns row access time and 68 ns column access time under the same conditions.

On the basis of these measurements, a 2σ worst-case chip built with the 2-μm nonexpanded dimensions is projected to have 100 ns row access time and 70 ns column access time at 75°C junction temperature and worst-case voltage conditions.

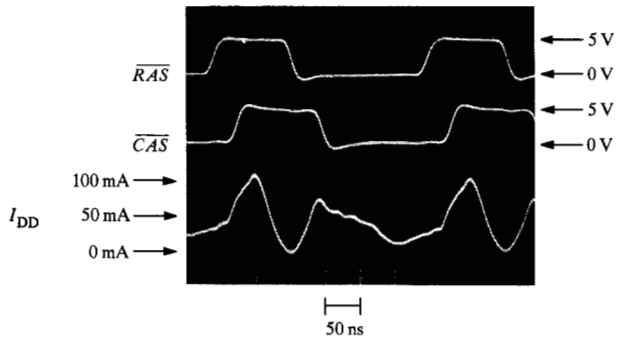

Figure 10 I_{DD} current at 300 ns cycle time.

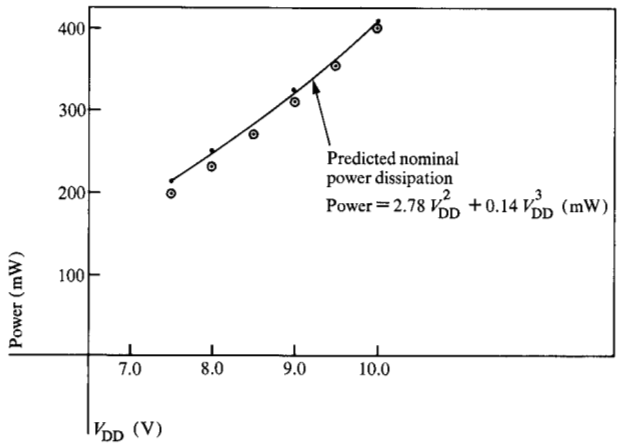

Figure 11 Selected power dissipation *vs.* V_{DD} at 300 ns cycle time and 25°C. ⊙—hardware measurement.

Instantaneous current from the V_{DD} supply as a function of time is shown in Fig. 10. Current surges have been minimized in the design by staggering precharge and discharge activities as much as possible. The current measurements were made at room temperature, (8.5 V + 10%) V_{DD}, and −(2.2 V − 10%) V_{BB}. The V_{DD} voltage dependence of power dissipation for the 64K RAM design is plotted in Fig. 11. The measurements agree with the predicted results and are predominantly dependent on V^2_{DD} due to the capacitive charging and discharging power.

The RAM has shown wide supply voltage margins, as illustrated in Fig. 12.

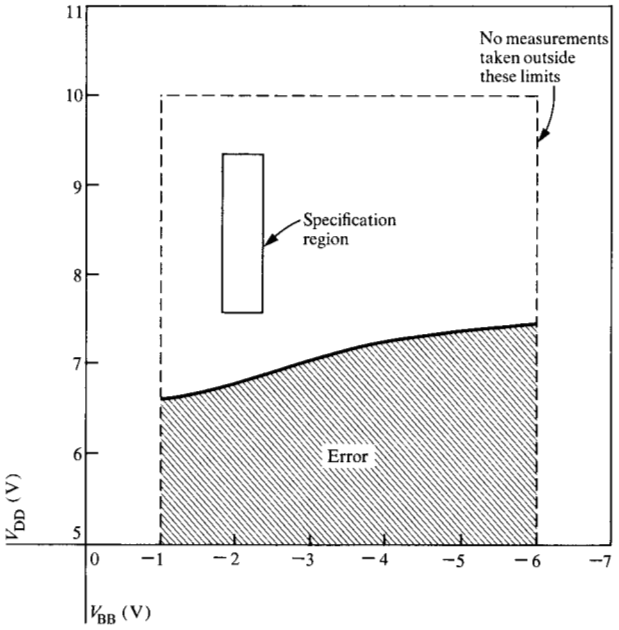

Figure 12 Power supply margin plot at room temperature using "march" test pattern.

Summary

The key design considerations for a 64K RAM have been presented. These considerations show that an 8K by 8 RAM design using an 8.5-V supply will exhibit major advantages over a 64K by 1 RAM design using a 5-V supply in the areas of charge capacity, chip size, performance, testing efficiency, and system power dissipation, without impacting either reliability or the capability for ECC implementation.

These salient advantages were exemplified by a design using state-of-the-art manufacturable n-channel FET technology with two layers of polysilicon and two layers of metallization. The unique cell structure, the chip organization, and the sensing circuitry were described, and measurement results were presented demonstrating that key design objectives were met.

Acknowledgments

The authors thank J. A. Palmieri and G. Sonoda for management support and technical discussions.

References

1. T. C. Lo and M. R. Guidry, "An Integrated Test Concept for Switched-Capacitor Dynamic MOS RAMs," *IEEE J. Solid-State Circuits* **SC-12,** 693 (1977).
2. S. A. Abbas and E. E. Davidson, "Reliability Implications of Hot Electron Generation and Parasitic Bipolar Action in an IGFET Device," *14th Annual Proceedings, Reliability Physics Symposium*, 1976, p. 18.

3. T. Ning, P. W. Cook, R. H. Dennard, C. M. Osburn, S. E. Schuster, and H. N. Yu, "1 μm MOSFET VLSI Technology—Part IV—Hot-Electron Design Constraints," *IEEE Trans. Electron. Devices* **ED-26,** 346 (1979).
4. E. S. Anolick and G. R. Nelson, "Low Field Time Dependent Dielectric Integrity," *17th Annual Proceedings, Reliability Physics Symposium,* 1979, p. 8.
5. K. Itoh, K. Shimohigashi, K. Chiba, K. Taniguchi, and H. Kawamoto, "A High-Speed 16K-Bit n-MOS Random-Access Memory," *IEEE J. Solid-State Circuits* **SC-11,** 585 (1976).
6. W. T. Lynch and H. J. Boll, "Optimization of the Latching Pulse for Dynamic Flip-Flop Sensors," *IEEE J. Solid-State Circuits* **SC-9,** 49 (1974).

Received May 17, 1979; revised December 14, 1979

The authors are located at the IBM Data Systems Division laboratory, East Fishkill (Hopewell Junction), New York 12533.