V. Leo Rideout
John J. Walker
Alice Cramer

A One-Device Memory Cell Using a Single Layer of Polysilicon and a Self-Registering Metal-to-Polysilicon Contact

The fabrication and operation of a novel one-device dynamic memory cell are described. Like the conventional double overlapping polysilicon cell, the new memory cell has a diffused bit line and a metal word line, uses five basic masking operations, and provides essentially equivalent cell area for the same lithographic feature size. Unlike the double polysilicon cell, however, the new cell uses a single layer of polysilicon to provide a more planar surface topography, and a self-registering metal-to-polysilicon contact to provide a small cell area. An essential aspect of the fabrication method of the self-registered contact cell is the use of two lithographic masking operations that define two patterns in a single polysilicon layer, the MOSFET gate electrode and the MOS capacitor electrode. The self-registering contact also facilitates a powerful polysilicon wiring technique that is applicable to the access circuits located peripherally to the array of memory cells.

Introduction

The combination of high cell packing density, moderate speed-power product, and low cost per bit of the one-device-cell dynamic random access memory (RAM) is unmatched by any other integrated circuit product. The evolutionary development of these memory cells [1] has been characterized by a series of structural innovations that have led to reduced cell area (*i.e.*, more efficient cell layouts) independent of the lithographic feature size. Two of the more important structural improvements have been the introduction of a second overlapping polysilicon layer and the sharing by two cells of etched contact holes to polysilicon or diffused silicon regions. This paper describes the fabrication and operation of a new one-device memory cell that uses a novel self-registering gate contact technique to achieve a small cell area, while retaining the attractive structural simplicity of a single polysilicon layer and a single thin gate oxide.

The operation of one-device cells has been described elsewhere [1] and will only be summarized here. A one-device cell consists of a single active device, a MOSFET switch, and an MOS storage capacitor. The FET switch is used to transfer charge from the n-type drain or bit line to the capacitor, or vice versa. Charge or its absence, representing a binary *zero* or *one*, can be stored in an inversion layer of electrons located beneath the polysilicon capacitor electrode or "plate." Many cells share one bit line, and sensing and writing circuits are located at the ends of the bit line. The n-channel or NMOS polysilicon gate FET technology is popular for dynamic RAMs by virtue of its dense cell structure and versatile interconnection wiring. Like some other polysilicon gate cells, the new cell is characterized by a diffused bit line and a metal word line that connects to the polysilicon gate of the FET switch in the cell. The metal word line must cross over and be insulated from both the capacitor plate and the n-type bit line. Unlike other conventional memory cells, the new cell has a self-registering or misregistration-tolerant contact between the metal word line and the polysilicon

Copyright 1980 by International Business Machines Corporation. Copying is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract may be used without further permission in computer-based and other information-service systems. Permission to *republish* other excerpts should be obtained from the Editor.

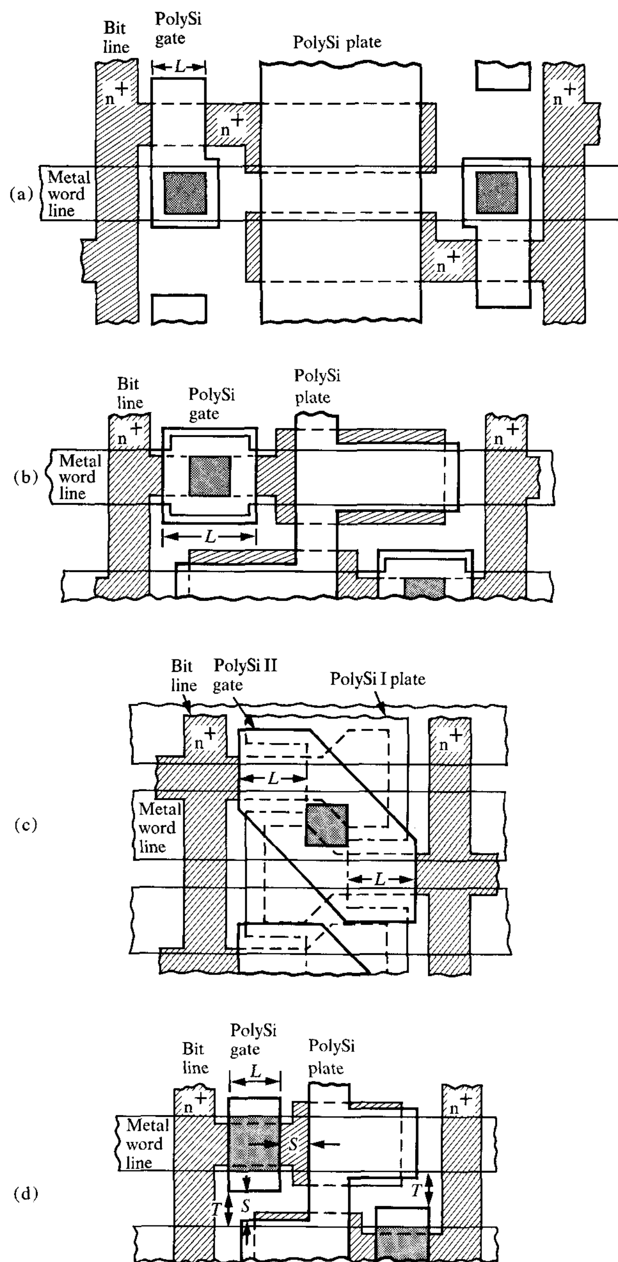


Figure 1 Dynamic one-device memory cells using diffused bit lines and metal word lines.

gate electrode. Consequently the contact area corresponds to the overall size of the polysilicon gate pattern; there are no conventional etched contact holes or vias in the array of memory cells.

Figure 1 compares the layouts of several polysilicon RAM cells [2]. The four cell types are drawn approximately to relative scale assuming the same lithographic feature size. All four cells use n-channel polysilicon gate technology to provide both the FET switch and the MOS inversion storage capacitor in the cell. All of the cells shown are of the diffused-bit-line/metal-word-line type.

Figure 1 illustrates that the technique used to provide the contact via between the metal word line and the polysilicon gate has a significant impact on the overall cell area. For example, the cell of Fig. 1(a) [3] is a cell layout commonly used in 4K-bit RAM designs. An area-consuming contact hole to the polysilicon gate material is located beside the channel and over the thick field oxide for greater misalignment tolerance. Four basic masking operations are required. An 8K-bit RAM design [4] shown in Fig. 1(b) placed the contact hole directly over the channel, which yielded a smaller cell but required a more precise alignment between the polysilicon gate pattern and the contact via pattern to prevent word line to bit line shorts. The cells of Figs. 1(a) and (b) use the same four-mask single polysilicon layer process.

A more recent approach is the overlapping double polysilicon gate cell used in several 16K-bit RAMs [5, 6] and shown in Fig. 1(c). Here the diffused region (*i.e.*, the FET "source") between the FET gate and the MOS capacitor storage electrode or "plate" has been eliminated, thereby shortening the cell, and a single contact via is shared between two cells, which halves the number of cell contacts. The gate and plate are fabricated using different polysilicon layers and different thin gate oxide layers. Higher packing density is achieved at the cost of an extra (fifth) masking operation and additional processing steps.

In this paper we will describe the one-device memory cell shown in Fig. 1(d), which uses a single layer of polysilicon for structural simplicity and a self-registering metal-to-polysilicon contact [7] to achieve higher density. For the same lithographic feature sizes this cell achieves an area equivalent to that provided by the overlapping double polysilicon cell shown in Fig. 1(c). Furthermore, like the double polysilicon cell, the self-registering gate contact cell requires five basic masking operations [8] and some additional processing steps to achieve a higher-density cell than the cells of Figs. 1(a) or (b).

The self-registering gate contact cell of Fig. 1(d) has one polysilicon contact per cell located directly over the FET channel region. The density improvement of the self-registering contact cell arises from a shorter channel length L (no via alignment borders are required) and a smaller spacing S between the FET gate and the capacitor storage plate in both the lateral and vertical directions. The objective here is to achieve high cell packing density by the use of this novel contacting technique, while simultaneously retaining the attractive structural simplicity of a single polysilicon layer and a single thin gate oxide.

Self-registering contact technique

The self-registering contact technique was first proposed by Kalter and Miller [7] and is fundamental to this work. As shown in Fig. 2, in its simplest implementation the technique utilizes a layer of silicon nitride to define the polysilicon gate area [Fig. 2(a)]. After the n-type source and drain regions are diffused, the entire structure is thermally oxidized. The nitride layer prevents oxidation on the top surface of the polysilicon gate [Fig. 2(b)]. Following oxidation, the nitride layer is removed by dissolving it in an etchant. Later, a metal line crosses the exposed gate and makes electrical contact to it [Fig. 2(c)]. Since there is no contact hole mask for the polysilicon gate, the alignment required between gate and metal masking patterns is relieved.

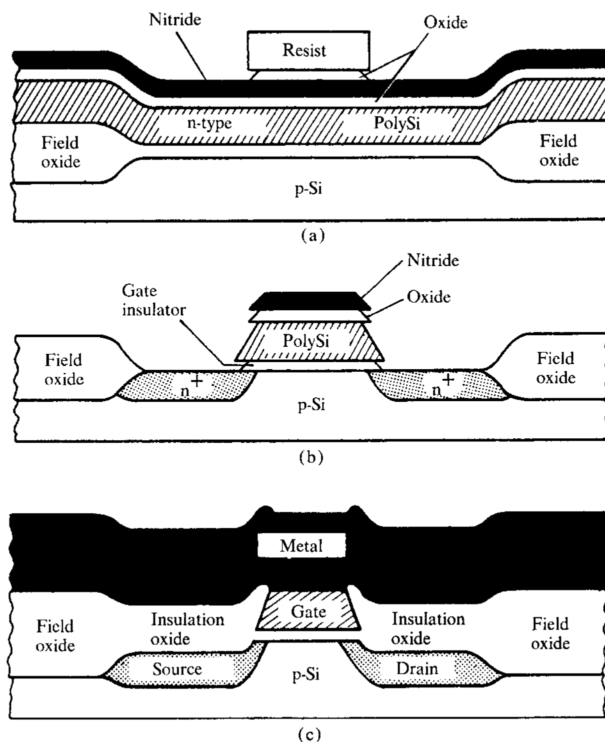


Figure 2 Self-registering contact technique.

The present work extends the self-registering contact technique by defining not one but two patterns in a single layer of polysilicon to provide first the FET gate electrode and later the MOS capacitor electrode or plate in the cell.

Figure 3 illustrates the basic four-mask, n-channel, single polysilicon layer process used to provide the cells shown in Figs. 1(a) and (b). In this illustration a semi-recessed oxide isolation with a self-aligned p-type parasitic channel stopper [9] is shown. After the gate oxide is grown, the polysilicon is patterned and then the source and drain regions are diffused. The polysilicon regions may serve as an FET gate, a capacitor plate, an interconnection line, or a contact area. Then an insulation oxide is thermally grown and/or deposited over the structure [10]. The third masking pattern defines contact vias to source, drain, and polysilicon areas. The fourth mask defines the metal pattern. In the new process the second masking step is expanded into two masking steps.

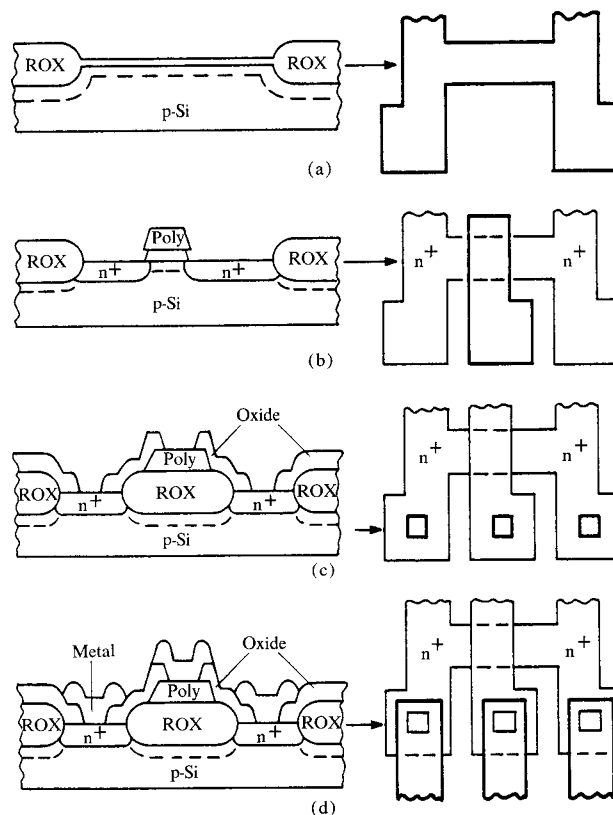


Figure 3 Basic four-mask, single polysilicon process: (a) field pattern; (b) polysilicon pattern; (c) contact via pattern; (d) metal pattern.

The fabrication procedure for the self-registering gate cell of Fig. 1(d) will now be described. With the first masking pattern the field isolation is fabricated in a conventional manner. In our experiments a semi-recessed thermal oxide with an ion-implanted channel stopper is used [9]; however, other isolation techniques can also be employed. Next the gate oxide is grown. After boron ions are implanted into the FET channel region to adjust the gate threshold voltage, a polysilicon layer is chemically vapor deposited and heavily doped n-type. Over the polysilicon, three thin layers consisting respectively of silicon dioxide, silicon nitride, and silicon dioxide are deposited and patterns delineated as follows.

Two masking patterns are used to independently define the MOS capacitor electrode and the MOSFET gate electrode as shown in Fig. 4. Two comparable methods are illustrated. Consider the method shown in Fig. 4(a). The

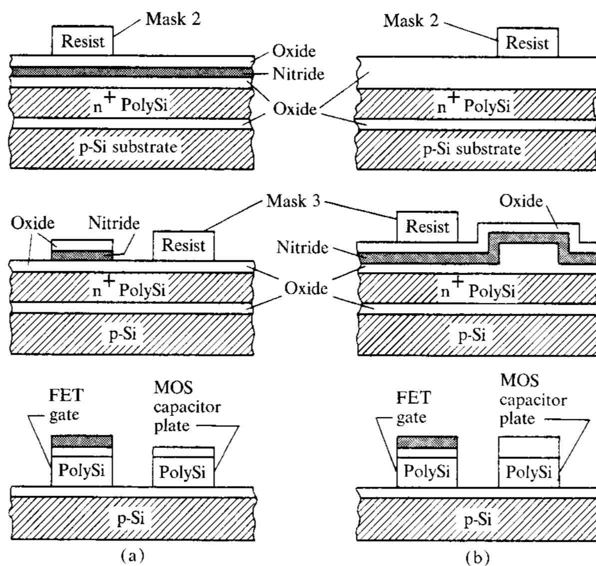


Figure 4 Polysilicon masking for self-registering contact cell.

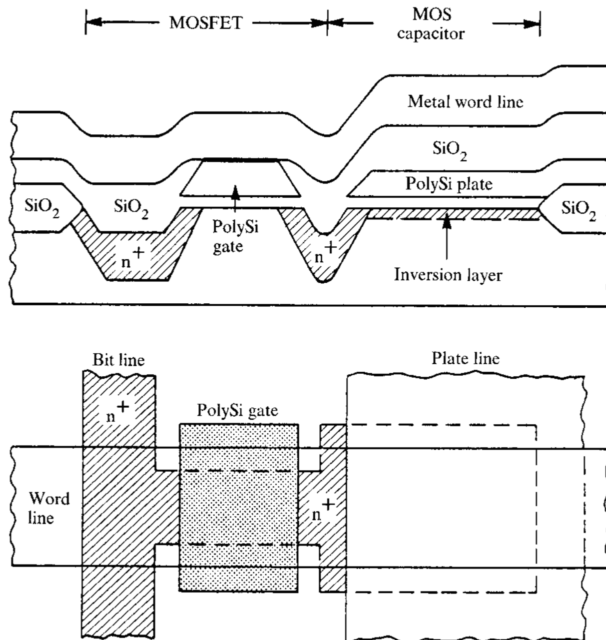


Figure 5 Self-registered contact cell: (a) cross section; (b) top view.

second masking step is used to pattern the uppermost oxide layer, which in turn is used as an etching mask for the nonoxidizing nitride layer. As a result of the second masking step, the nitride layer is retained over the future gate areas. These areas will later become self-registered gate contact regions. The third masking operation is used to retain only oxide over the future capacitor plate areas, interconnection lines, and/or insulated FET gate electrode areas. After both the gate and plate patterns have been defined, the polysilicon layer patterns are etched in a single step.

In the method shown in Fig. 4(a), the FET gate with its self-registering contact is delineated before the insulated storage capacitor plate of the cell, hence the descriptive name *gate/plate* process [8]. Figure 4(b) illustrates how the second and third masking operations can be reversed, which gives rise to the descriptive name *plate/gate* process [8]. We have exercised both approaches and favor the *gate/plate* process of Fig. 4(a) because it is somewhat simpler.

After etching the polysilicon areas, the n^+ source/drain regions are formed by diffusion or by ion implantation. Then the entire structure is thermally oxidized, which provides an insulation oxide over the capacitor plate and source/drain regions and thickens the field oxide. The self-registering contact regions do not oxidize due to the presence of the nitride layer over them [7, 8]. After oxidation the nitride layer is removed by dissolving in an etchant (*e.g.*, phosphoric acid). The fourth masking operation is used to delineate conventional etched via holes to n^+ source/drain regions in the peripheral circuits, while the fifth masking operation defines the aluminum metallization pattern.

Since the gate and plate are defined in separate masking operations, they can be spaced closer than one minimum lithographic spacing apart. This helps to reduce cell area. The gate and plate patterns are not intended to overlap in the cell. Later, however, we will discuss the advantages of overlapping the two polysilicon patterns in the peripheral circuits.

Figure 5 shows the *gate/plate* memory cell in cross section and top view. Note from Fig. 5(a) that the vertical topography is highly planar and thus vertical excursions of the metal line pattern are reduced [11], particularly in comparison with the overlapping double polysilicon RAM cell. With the *gate/plate* cell, better metal linewidth control is anticipated for optical lithography where depth of field focusing can become important.

In the *gate/plate* cell of Fig. 1(d), an FET gate to capacitor plate separation S is required and hence the cell width and length are dependent on the alignment inaccuracy between the two patterns in the single polysilicon layer. In other words, in the cell the gate and plate patterns must be separated to a required degree of precision to ensure electrical isolation. Analogously, in the double polysilicon cell of Fig. 1(c), the channel length L of the FET in the cell, and hence the cell length, are de-

pendent on the mask-to-mask alignment inaccuracy between the two patterns in the two overlapping polysilicon layers. In this case, the gate and plate patterns must overlap to a required degree of precision to ensure an operational FET.

From Fig. 5(b), the gate contact area corresponds to the area of the entire gate pattern; hence wherever the metal word line crosses the gate, an electrical connection is made. This relieves the alignment precision required between the polysilicon gate and metal line patterns. On the other hand, care must be taken to ensure that the gate does not contact an adjacent metal word line. As a result, an alignment separation T is required between the gate and the metal word line of an adjacent cell as shown in Fig. 1(d).

Both the *gate/plate* and overlapping double polysilicon RAM cells have inherent structural difficulties. For example, in the double polysilicon cell of Fig. 1(d), the "thin" gate insulator under the second polysilicon layer (the transfer gate) and the thick polysilicon layers must be provided by a common thermal oxidation step. Consequently, a compromise between gate threshold control and oxide breakdown must be obtained. Gate reliability is reported to be poorer for the first polysilicon layer [12], which typically has a thinner gate insulator under it and which experiences more high-temperature processing. Difficulties related to mask-to-mask alignment and vertical topography have been discussed earlier.

While the processing problems of the double polysilicon RAM cell center around the first polysilicon to second polysilicon overlap boundary [1], structural difficulties with the *gate/plate* RAM cell are primarily related to the self-registering metal line to polysilicon gate contact. In the cell the self-registering contact is made to the polysilicon gate directly over the channel and each cell has one such contact. The nitride layer must be applied, exposed to high-temperature thermal oxidation, and removed without degrading the polysilicon gate material or the underlying gate insulator. Reliability concerns associated with metal contacts to polysilicon gates located directly over thin insulators have been discussed elsewhere [13]. In our work, a Pd_2Si layer [14] was used between the polysilicon gate contact and the metal interconnection line to improve contact reliability [15]. Finally, the metal word line does not completely cover the exposed polysilicon gate area; hence one must rely on additional passivating layers.

Results

The mask set developed to test the *gate/plate* process is shown in Fig. 6. The test chip measures 2.54 by 2.54 mm^2.

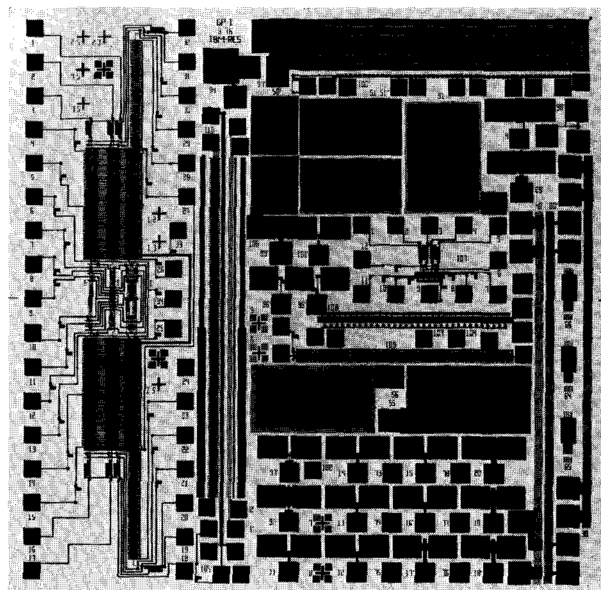


Figure 6 *Gate/plate* mask set.

On the left side of the chip are four rows of one-device cells located in a balanced fashion on either side of four sense amplifiers. The three shorter lines have 64 cells on either side of the sense amplifier, and the one longer line has 128 cells on either side of the same amplifier. In the center of the chip are various inactive line structures which can be used to examine the vertical topography. The lower right-hand quadrant of the chip contains single FETs of varying channel length, gated diodes for leakage current tests, and four-point probes for sheet resistance measurements. The upper right-hand quadrant contains capacitors to monitor oxide thickness, breakdown voltage, and retention time.

Figure 7 shows part of the four sense amplifiers and the lines of one-device cells. There are two cell areas, one of 189 μm^2 (7.5 × 24.5 μm^2) and another with a larger storage capacitor which has a cell area of 271 μm^2. Both sizes of cells have been successfully operated. The chips were fabricated using optical proximity printing with a minimum resist feature size of about 2.5 μm. Figure 8 is a scanning electron microscope (SEM) photograph of the smaller one-device cell (189 μm^2). The metal word line is misaligned with respect to the gate pattern but experiences no difficulty in making a self-registering contact to the gate. There are no conventional etched contact holes in the array of cells.

The operation of the balanced sense amplifier with the 271-μm^2 cell is shown in Fig. 9. A binary *zero* or *one* is

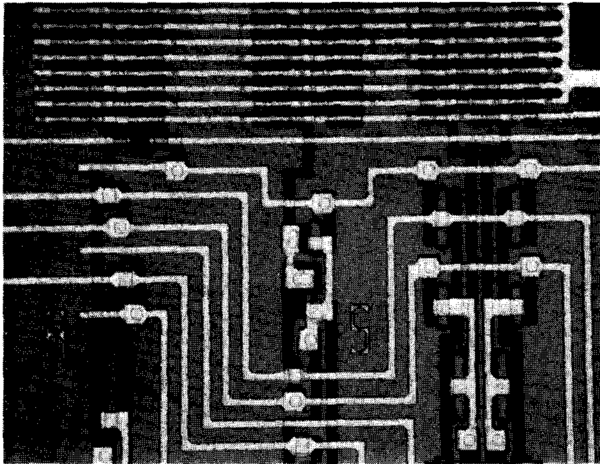


Figure 7 Part of the four sense amplifiers and the lines of one-device cells.

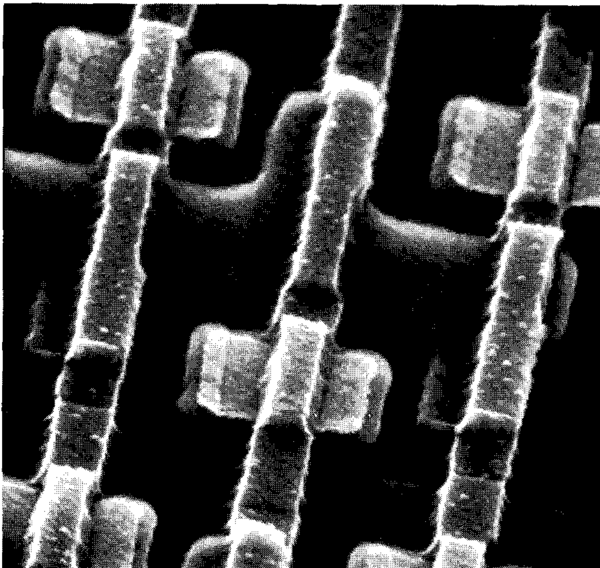


Figure 8 SEM photograph of smallest one-device cell (cell area 189 μm^2).

written into a selected cell. Upon reading, the MOS storage capacitor is discharged onto the bit line. In the case of a *one*, this causes a small signal voltage V to be impressed onto the bit line. This voltage is amplified and compared to an amplified reference voltage ($V/2$) generated by a dummy cell located on the other side of the sense amplifier.

The oscilloscope traces in Fig. 9 show the waveforms of the bit-line voltage as seen through a source-follower circuit on the end of the line. After about 70 ns the sense amplifier latches into the *one* or the *zero* state. The total bit access time will be somewhat greater than 70 ns because there are no bit-line or word-line decoders and no input multiplexers on the test chip.

Polysilicon devices were made with and without the self-registering contact on the same chip, and the substrate sensitivity or "back-gate" bias relationships [16] were measured. As expected, no difference was observed between the characteristics of the two devices, which demonstrates that the self-registering contact does not affect the FET device characteristics, even when the metal interconnection line contacts the polysilicon gate material over the channel. The subthreshold conduction characteristic [16] of an FET in a cell with a self-registering contact was also measured. The memory circuits were designed to operate with a substrate bias of 1 V. The FET turns off correctly to below 10^{-11} A, which is satisfactory source-to-drain leakage current for dynamic RAM operation. Again, this is typical of a device without a self-registering contact.

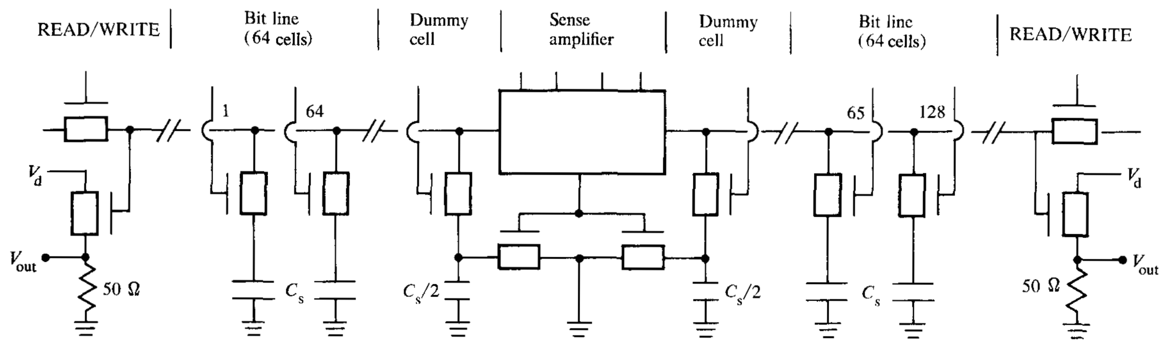
In the *gate/plate* process the entire structure is thermally oxidized, except for the gates. This provides a thick oxide over the source, drain, and capacitor plate. It also thickens the field oxide and causes additional boron depletion [17] in the field oxide regions between cells.

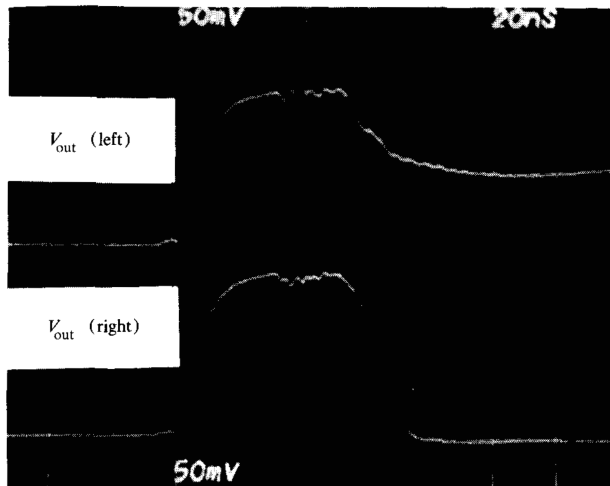
Unlike conventional polysilicon-gate processes, chemically vapor-deposited SiO_2 or phosphorus-doped glass layers were not used in the self-registered gate process. No metal step coverage problems [11] were encountered.

An initial concern was that the oxidation step would excessively lower the threshold voltage in the field regions. That this is not the case is confirmed by Fig. 10, which shows the substrate sensitivity of the thick field oxide under the polysilicon plate and under a metal interconnection line. The field oxide is comparable in thickness for the two cases, and the metal line has a lower threshold due to the additional boron depletion; however, the threshold is not so low as to present a problem. The memory circuits were designed to operate with 1 V substrate bias, and for that bias the parasitic threshold voltage under the metal line is an acceptable 20 V.

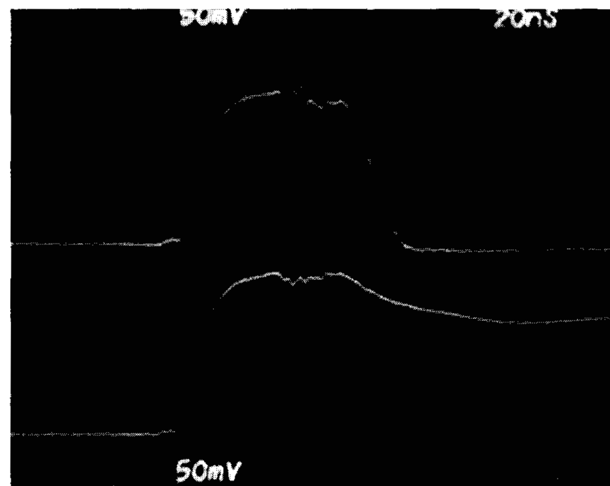
Extensions

Although the primary emphasis of this work is on the fabrication of a novel one-device cell, it is of interest to consider other applications of the self-registering contact technique. Thus far, we have only considered cases in which the two polysilicon masking patterns do not interact, *i.e.*, they do not overlap. As shown in Fig. 11, however, the patterns can be overlapped [18] to connect a





Sensing a binary *one*



Sensing a binary *zero*

Figure 9 Operation of *gate/plate* cell and sense amplifier.

revealed polysilicon contact area to an insulated polysilicon gate, plate, or interconnection line. Three masking sequences are feasible and they yield slightly different contact boundary shapes where the mask patterns overlap. We have successfully practiced the interconnection approach shown in Fig. 11(a), but generally prefer the somewhat simpler approach of Fig. 11(c). Figure 12 illustrates the various polysilicon interconnection combinations that can be achieved using the overlapping masking approach of Fig. 11(a). A very powerful interconnection wiring technique can be developed in this manner. This is particularly attractive for the peripheral circuits (decoders, multiplexers, sense amplifiers, etc.) of a dynamic memory chip, and for other MOSFET integrated circuits such as ROMs, PLAs, and microprocessors. The wiring approach can be made even more effective by directly connecting the polysilicon patterns to the n^+ diffused source and drain regions [19] by use of a "buried contact" step.

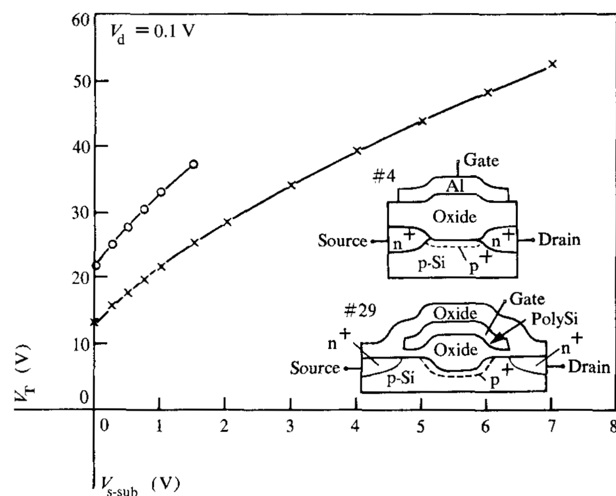


Figure 10 Parasitic threshold voltages for thick oxide field isolation regions: ○, Device #29—thick oxide FET with polysilicon gate; ×, Device #4—thick oxide FET with aluminum gate.

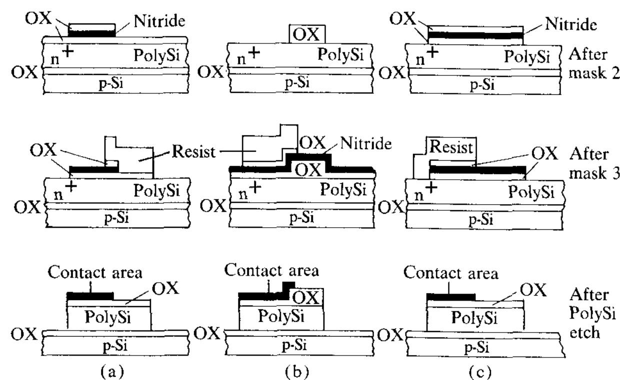


Figure 11 Result of overlapping the two polysilicon masking patterns.

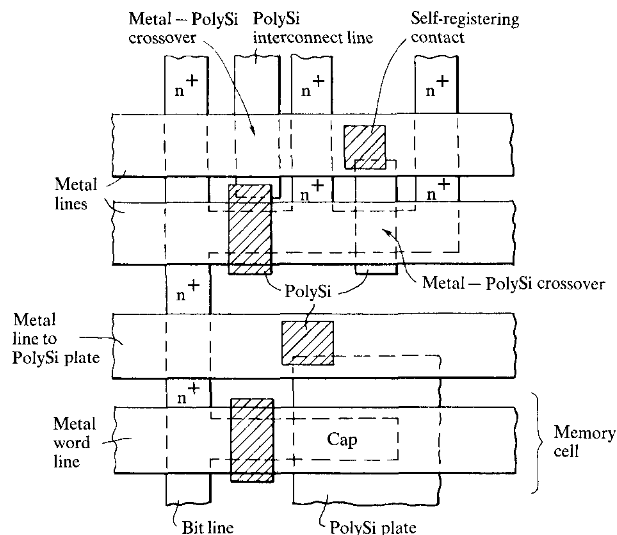


Figure 12 Polysilicon interconnections using self-registering contacts.

A more general extension of the self-registering metal-to-polysilicon contact is the analogous self-registering metal-to-diffusion contact which provides a single polysilicon metal bit-line cell [20]. In this case, part of the nitride used initially to form the semi-recessed oxide isolation regions is retained to later form the metal-to-diffusion contact areas. Finally, one can combine both techniques into a five-mask process that yields self-registering contacts to both polysilicon and diffused regions [20], thus eliminating etched contact holes on the chip.

Summary

A dynamic one-device RAM cell has been described which uses a single layer of polysilicon and a self-registering metal-to-polysilicon gate contact. Five basic masking operations are required. Attractive features of the *gate/plate* process include a highly planar surface topography, high packing density, and increased mask-to-mask misalignment tolerance. MOSFETs with self-registering gate contacts were fabricated which exhibited satisfactory subthreshold turn-on and substrate sensitivity characteristics. Using optical lithography, memory cells 7.5 by 24.5 μm in area (189 μm^2) were fabricated and operated successfully. The *gate/plate* process is, of course, applicable to circuits other than dynamic RAMs.

Acknowledgments

The authors wish to acknowledge the processing support of the silicon technology group at the T. J. Watson Research Center and helpful discussions with C. M. Osburn and R. H. Dennard.

References

1. V. L. Rideout, "One Device Memory Cells for Dynamic Random Access Memories: A Tutorial," *IEEE Trans. Electron Devices* **ED-6,** 839–852 (1979).
2. V. Leo Rideout, John J. Walker, and Alice Cramer, "A One-Device Memory Cell Using a Single Layer of Polysilicon and a Self-Registering Metal-to-Polysilicon Contact," *IEDM Tech. Digest*, 258–261 (1977).
3. K. U. Stein, A. Sihling, and E. Doering, "Storage Array and Sense/Refresh Circuit for Single-Transistor Memory Cells," *IEEE J. Solid-State Circuits* **SC-7,** 336–340 (1972).
4. W. K. Hoffman and H. L. Kalter, "An 8K-Bit RAM Chip using the One-Device FET Cell," *IEEE J. Solid-State Circuits* **SC-8,** 298–304 (1973).
5. C. N. Ahlquist, J. R. Breivogel, J. T. Koo, J. L. McCollum, W. G. Oldham, and A. L. Renninger, "A 16,384-Bit Dynamic RAM," *IEEE J. Solid-State Circuits* **SC-11,** 570–574 (1976).
6. P. R. Schroeder and R. J. Proebsting, "16K × 1 Bit Dynamic RAM," *1977 IEEE ISSCC Digest of Technical Papers* **20,** 12–13 (1977).
7. H. L. Kalter and D. A. Miller, "Self-Aligning Metal to Polysilicon Contacts," *IBM Tech. Disclosure Bull.* **4,** 3176 (1972).
8. V. L. Rideout, "Masking for One-Device Cell Memories using Self-Registering Metal-to-Polysilicon Contacts," *IBM Tech. Disclosure Bull.* **17,** 2802–2804 (1975).
9. P. Richman, *MOS Field-Effect Transistors and Integrated Circuits*, John Wiley & Sons, Inc., New York, 1973, p. 207.
10. J. T. Clemens, R. H. Doklan, and J. J. Nolen, "An N-Channel Si-Gate Integrated Circuit Technology," *IEDM Tech. Digest*, 299–302 (1975).
11. V. J. Silvestri, V. L. Rideout, and V. Maniscalco, "Al Coverage of Surface Steps at SiO_2 Insulated Polycrystalline Si Boundaries: Al Evaporation in Vacuum and Low Pressure Ar," *J. Electrochem. Soc.* **126,** 1335–1338 (1979).
12. C. R. Barrett and R. C. Smith, "Failure Modes and Reliability of Dynamic RAMs," *IEDM Tech. Digest*, 319–322 (1976).
13. M. Revitz and J. F. Shepard, "A Failure Mechanism in Silicon-Gate Structures," *Electrochemical Society Extended Abstracts*, Los Angeles meeting, October 1979, Abs. No. 569.

14. C. J. Kircher, "Contact Metallurgy for Shallow Junction Si Devices," *J. Appl. Phys.* **47**, 5394–5399 (1976).
15. W. R. Hunter, L. M. Ephrath, W. Grobman, C. M. Osburn, B. L. Crowder, A. Cramer, and H. E. Luhn, "One-Micron Electron-Beam Lithography FET Technology," *IEDM Tech. Digest*, 54–57 (1978).
16. V. L. Rideout, F. H. Gaensslen, and A. LeBlanc, "Device Design Considerations for Ion Implanted n-Channel MOSFETs," *IBM J. Res. Develop.* **19**, 50–59 (1975).
17. B. E. Deal, A. S. Grove, E. H. Snow, and C. T. Sah, "Observation of Impurity Redistribution During Thermal Oxidation of Silicon Using the MOS Structure," *J. Electrochem. Soc.* **112**, 308 (1965).
18. V. L. Rideout, "Self-Registering Metal-to-Polysilicon Contacting Technique," *IBM Tech. Disclosure Bull.* **21**, 3818–3821 (1979).
19. S. Muramoto, T. Hosoya, and S. Matsuo, "A New Self-Aligning Contact Process for MOS LSI," *IEDM Tech. Digest*, 185–188 (1977).
20. V. L. Rideout, "Polysilicon-Gate Field-Effect Transistors with Self-Registering Metal Contacts to Both Polysilicon and Diffused Silicon Regions," *IBM Tech. Disclosure Bull.* **21**, 3833–3835 (1979).

Received June 22, 1979; revised November 19, 1979

V. Leo Rideout and Alice Cramer are located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598. John J. Walker is located at the IBM Data Systems Division laboratory, East Fishkill (Hopewell Junction), New York 12533.