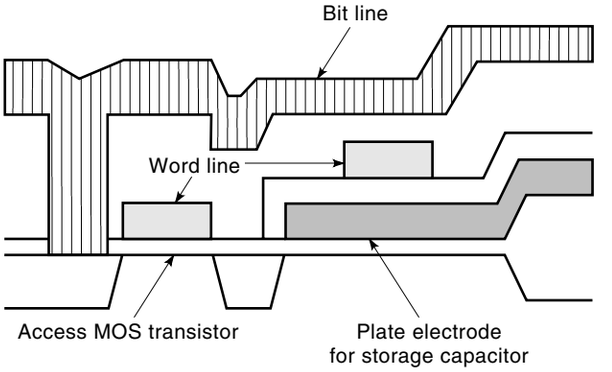

Figure 2. DRAM memory cell cross-sectional view.

DRAM CHIPS

DRAM MEMORY CELL

Dynamic random access memory (DRAM) is defined as random access memory with a refresh operation for maintaining the stored data. DRAM has always been the leading semiconductor product which requires advanced semiconductor device and fabrication technology (1–3). DRAM is the most popular memory device because of its high performance-to-cost ratio. In comparison with other kinds of memories (4,5), DRAM has a very simple memory cell. One transistor and one capacitor type is the most popular cell type in present DRAM. Figure 1 shows the schematic diagram of the basic memory cell for DRAM. The data depend on the amount of charge on the capacitor. For example, ample amount of charge stored on a capacitor is recognized as logic-1, while no charge stored on a capacitor is recognized as logic-0. This data storage capacitor is selected by the switched transistor. The storage capacitor is very important for DRAM performance. Since the capacitor leaks some charges, large capacitance is a great advantage in maintaining data integrity. Additionally large capacitance could improve speed performance. Figure 2 shows the cross section of a conventional planar DRAM cell. Since minimum area consumption is allowed in order to achieve a high-density DRAM, the structure of the cell capacitor is modified in the vertical dimension for the advanced memory cell, unlike for the conventional plane capacitor. Figures 3 and 4 show the diagrams of two present well-known fundamental three-dimensional capacitor cells (6–9). To achieve the required minimum value of capacitance, both approaches have been successful in minimizing the area consumption and fabrica-

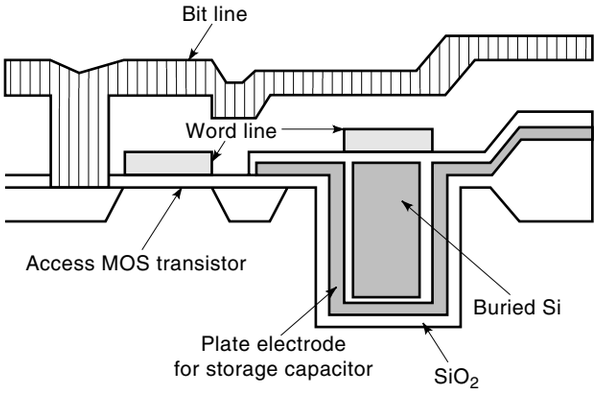

Figure 3. Cross-sectional view of trench capacitor cell.

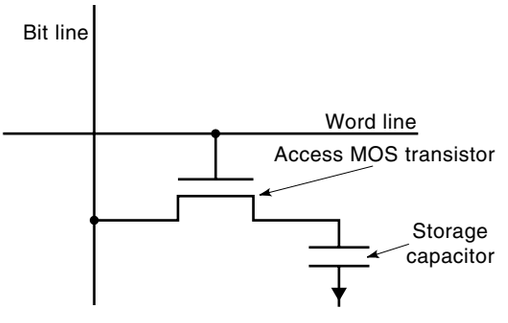

Figure 1. DRAM memory cell.

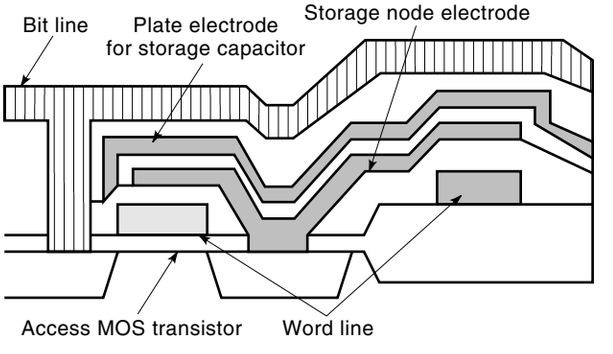

Figure 4. Cross-sectional view of stacked capacitor cell.

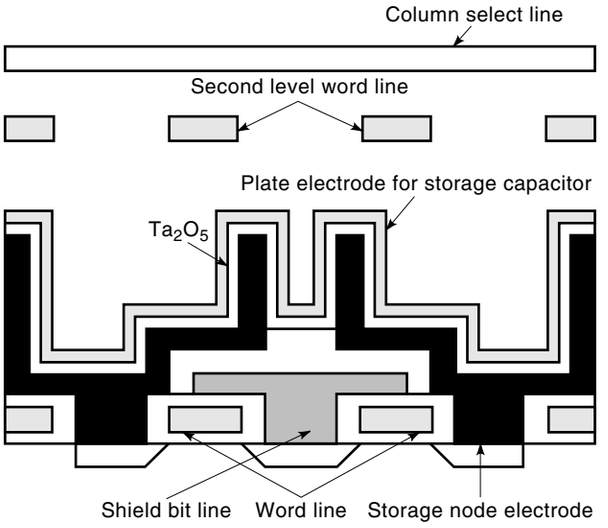

Figure 5. Cross-sectional view of advanced three-dimensional capacitor cell.

tion process complexity. Furthermore, a variety of shapes of capacitor electrodes have been investigated in order to obtain larger capacitance in a small area for higher density. Figure 5 shows an advanced three-dimensional capacitor cell. This type of capacitor will play a major role for 64 MB DRAM and higher-density versions. In addition to the above-mentioned approaches, larger capacitance with high dielectric insulator and enlarged surface area by microvillus patterning of electrode technology have been investigated for the next-generation memory cells (10,11).

BASIC DRAM SYSTEM

Figure 6 shows the block diagram of a DRAM system. In principle, the function of DRAM is very simple. It requires only writing the data into a certain address, preserving it for a certain period of time, and reading the data out. We trace this simple function using the diagram shown in Fig. 6. In the read operation, the memory chip needs to set up the address data. In DRAM, in contrast with other memory technologies, a unique address input scheme is adopted. With the address multiplex scheme, the address data are divided into two portions, defined as row address and column address, respectively. According to the enable-address signals, each address is stored in an address input buffer by individual timing. Therefore, the number of address pins is equal to one-half of the whole address digits. On the other hand, two sequential address inputs are necessary. After obtaining the address information, the stored data, which are selected by the row address, are read by sense amplifiers to amplify the signals. Then, the column address is used to transmit the amplified signals to the output buffers. In the write operation, a write amplifier allows reversal of the data on the I/O bus, if the input data are different from the data on the I/O bus.

Sense Amplifier

A sense amplifier is as important as the memory cell, because its performance affects the whole chip performance significantly. The function of the sense amplifier is to read the data

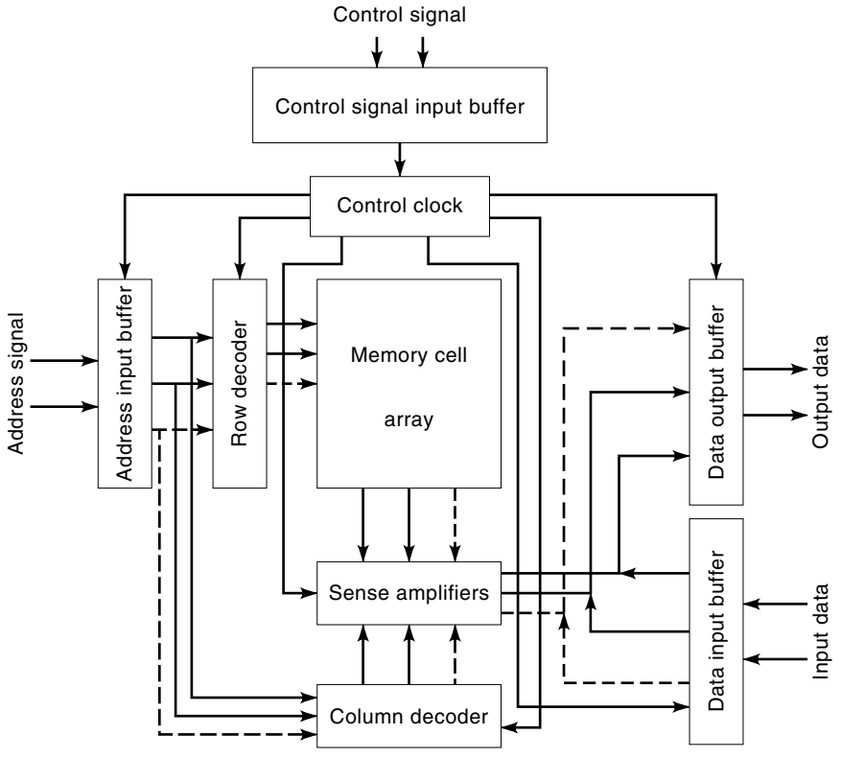

Figure 6. DRAM system block.

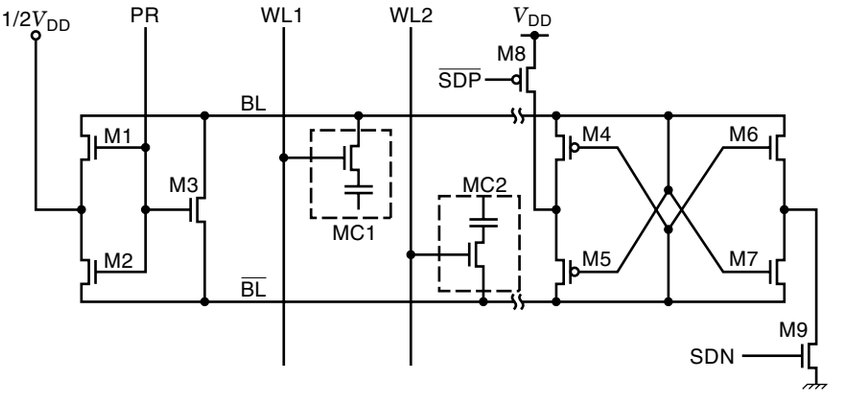

Figure 7. Conventional sense amplifier block.

from the storage cell and to magnify the signal level up to the appropriate logic level to be treated by digital logic gates. Taking a DRAM sense amplifier as an example, Fig. 7 shows the schematic diagram of a conventional sense amplifier for DRAM. This circuit block operates as follows. At the initial condition, assume that "H" level is stored in MC1. A bit line pair is precharged and equalized by M1, M2 and M3, and other MOS transistors are turned off. First, M1, M2, and M3 turn off to set up the preparation stage for sensing. Then, the selected word line (WL1) is activated by the row address. As a result, one MOS access transistor and one storage capacitor is connected to the bit line. This action causes the imbalances of the bit line pair. The voltage of the bit line changes slightly because of storage charge in the capacitor of MC1, whereas the other bit line still remains at the same level. This difference can be detected by the sensing node. Sense amplifier driver transistors M8 and M9 turn on in order to activate the amplification function. Then, a small signal value between the bit line pair becomes a large signal level by the cross-coupled MOS transistors. The basic function is simple; however, optimizing the size and timing for each clock is an essential issue for the DRAM system designer. Figure 8 shows the timing diagram for the sensing operation.

LOW-POWER DESIGN TECHNOLOGY

Low-power DRAM technology is increasingly important for general memory use (12). Since the current memory market tends to explore outdoor usage, the battery-operable feature is a crucial requirement for memories. In general, total power consumption of the DRAM chip (P), which is operated at V_{DD} and cycle time t_{RC}, can be expressed as follows:

$$\begin{aligned} P &\cong \sum C_j \cdot \left(\frac{\Delta V_j}{\Delta t}\right) \cdot V_{DD} + I_{DC} \cdot V_{DD} \\ &\cong \Delta Q_T \cdot \frac{V_{DD}}{t_{RC}} + I_{DC} \cdot V_{DD} \\ &\cong \frac{(C_{BT} \cdot \Delta V_D + C_{PT} \cdot \Delta V_P) \cdot V_{DD}}{t_{RC}} + I_{DC} \cdot V_{DD} \\ &\cong (\Delta Q_{BT} + \Delta Q_{PT}) \cdot \frac{V_{DD}}{t_{RC}} + I_{DC} \cdot V_{DD} \end{aligned} \tag{1}$$

where C_j is the capacity of node j
ΔV_j is the voltage variation at node j
ΔQ_T is the total charge of the chip during one cycle
ΔQ_{BT} is the total charge of bit lines during one cycle
ΔQ_{PT} is the total charge of peripheral blocks during one cycle
I_{DC} is the dc current component
C_{BT} is the total capacitance for bit lines
ΔV_D is the bit line charged level
C_{PT} is the total capacitance for peripheral blocks
ΔV_P is the voltage variation for peripheral blocks

There are several effective methods to reduce the power consumption. Special care on supply voltage V_{DD}, bit line capacitance charge Q_{BT}, and dc current component I_{DC} could minimize the power consumption effectively.

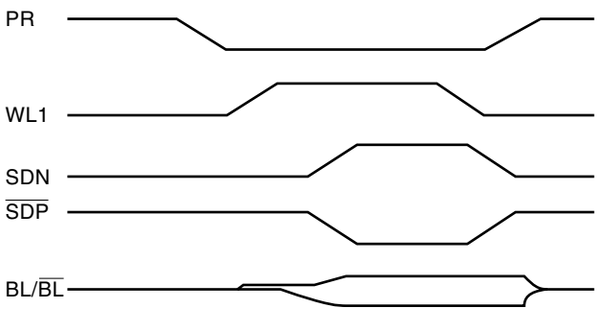

Figure 8. Timing diagram of sensing operation.

HIGH-THROUGHPUT DRAM TECHNOLOGY

Since the speed performance of the central processing unit (CPU) has been drastically increased, the DRAM is under pressure to achieve fast response. A variety of useful approaches are categorized according to three criteria: multibit data output, high-frequency control clock, and advanced interface technology. These approaches are selected with careful consideration for cost-per-bit performance, which is the most important factor for any type of memory chips. Figure 9 shows the recent market for memory, according to speed performance and capacity. Instead of common memory devices, several advanced approaches, such as extended data out (EDO) DRAM, synchronous DRAM (SDRAM), and Rambus DRAM (RDRAM), will play main roles in the future memory market. Although the

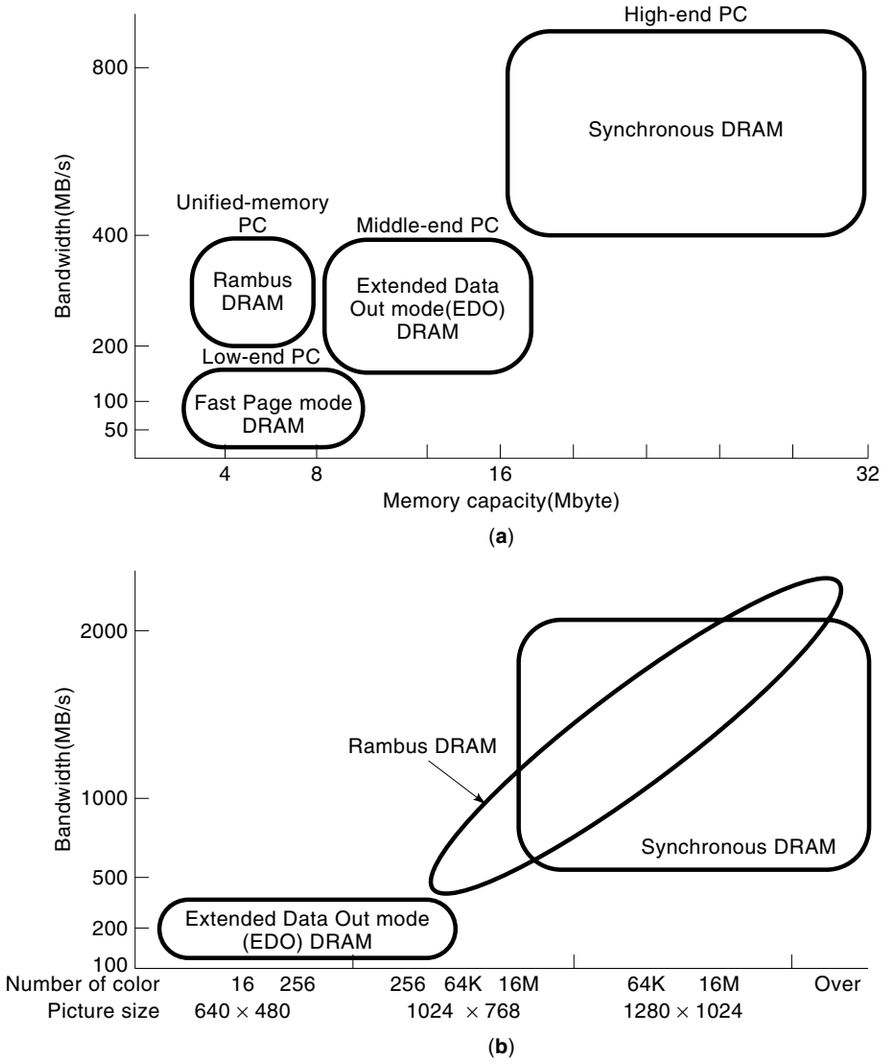

Figure 9. (a) DRAM application for PC. (b) DRAM application for graphics.

purpose of new approaches is to achieve high data throughput, each approach has unique features.

Multibit DRAM

Increasing the number of data output structures is a simple and straightforward approach to achieve high data throughput performance. This approach has already been adopted and is successful for a variety of memory chips. In fact, ×4, ×8, and ×16 bit DRAM are very popular products in the commercial market. Recently, achievement of ×32 bit 16M DRAM for mass production was reported (13). In order to realize the large number of multibit structures, development of an advanced data output buffer is indispensable. Basically, there are three inherent problems for the multibit memory: increased switching noise, increased power consumption, and increased chip area. Therefore, a data output buffer is introduced to minimize the above-mentioned effects. Switching noise is caused by parasitic inductance, which is due to the wire line connecting the external pin and the internal pad on the chip. The larger the number of data output buffers that are activated, the more the output signal is distorted by the noise. Power consumption is proportional to the number of activated data output buffers. In order to solve these problems, the following approaches are useful. Adoption of multiple power supply lines and ground lines can prevent noise interference. Minimization of the physical wire length can reduce the parasitic inductance. Lowering the power supply voltage can reduce power consumption. In practice, the combination of approaches will be able to produce stable performance of memories. Figure 10 shows a schematic diagram of a capacitor-boosted-type data output buffer (14). At the initial stage, $Q1$ and $Q3$, which are small drive-ability transistors, turn on to avoid the abrupt voltage transition that causes considerable noise. After reaching a certain output level, $Q2$ and $Q4$, which are large drive-ability transistors, turn on to increase processing speed. $T1$ and $T2$ are used to set the delay time. Since the gate voltage of $Q2$ is boosted by capacitor coupling, full V_{DD} level output signal can be achieved.

Fast Page Mode DRAM

The currently dominant mode of DRAM is fast page mode, although there are similar performance modes, such as static

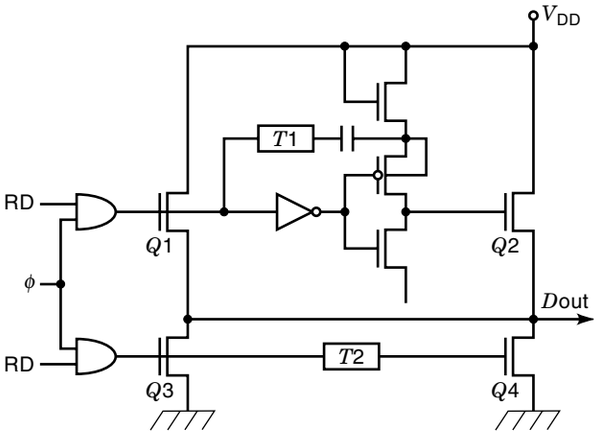

Figure 10. Schematic diagram of capacitor-boosted-type data output buffer.

column mode and nibble mode. There are three fundamental operation modes: read, write, and refresh. The user needs to set up control signals such as $\overline{\text{RAS}}$, $\overline{\text{CAS}}$, $\overline{\text{WE}}$, $\overline{\text{OE}}$ and address signals in order to activate the proper mode. $\overline{\text{RAS}}$ is the signal to strobe the row address data, whereas $\overline{\text{CAS}}$ is the signal to strobe the column address data. Since DRAM adopts multiplex addressing, the user sets both data on the same address pins at different time periods. $\overline{\text{WE}}$ is defined as the write-enable signal. Thus, $\overline{\text{WE}}$ is inactive in the read mode and is active in the write mode. $\overline{\text{OE}}$ is defined as the output-enable signal. This signal is used for $\times 4$, $\times 8$, and $\times 16$ bit operation chip. Figure 11 shows the timing diagram of standard read and write operation.

Access time is the most important specification item for DRAM. There are four kinds of access time: t_{RAC}, t_{CAC}, t_{AA} and t_{OEA}. Thus, the user has to pay attention to particular access time values. First, t_{RAC} is defined as the time from $\overline{\text{RAS}}$ low to the valid data output. This is the longest access time among them. This access time is usually used to classify the version. Second, t_{CAC} is defined as the time from $\overline{\text{CAS}}$ being low to the valid data output. For data which are stored at the same row address, DRAM can offer the minimum access time, t_{CAC}. Third, t_{AA} is defined as the time from the column address set up to the valid data ouput. The t_{AA} is a useful specification for static column mode, which can produce data by column address transition. Fourth, t_{OEA} is defined as the time from $\overline{\text{OE}}$ being low to the valid data output. Note that t_{OEA} is usually equal to t_{CAC}. The t_{OEA} is a valuable specification for $\times 4$, $\times 8$, and $\times 16$ bit DRAM. The refresh operation is a unique feature of DRAM. This unavoidable operation adds to the complexity of DRAM. The refresh operation is equivalent to row address

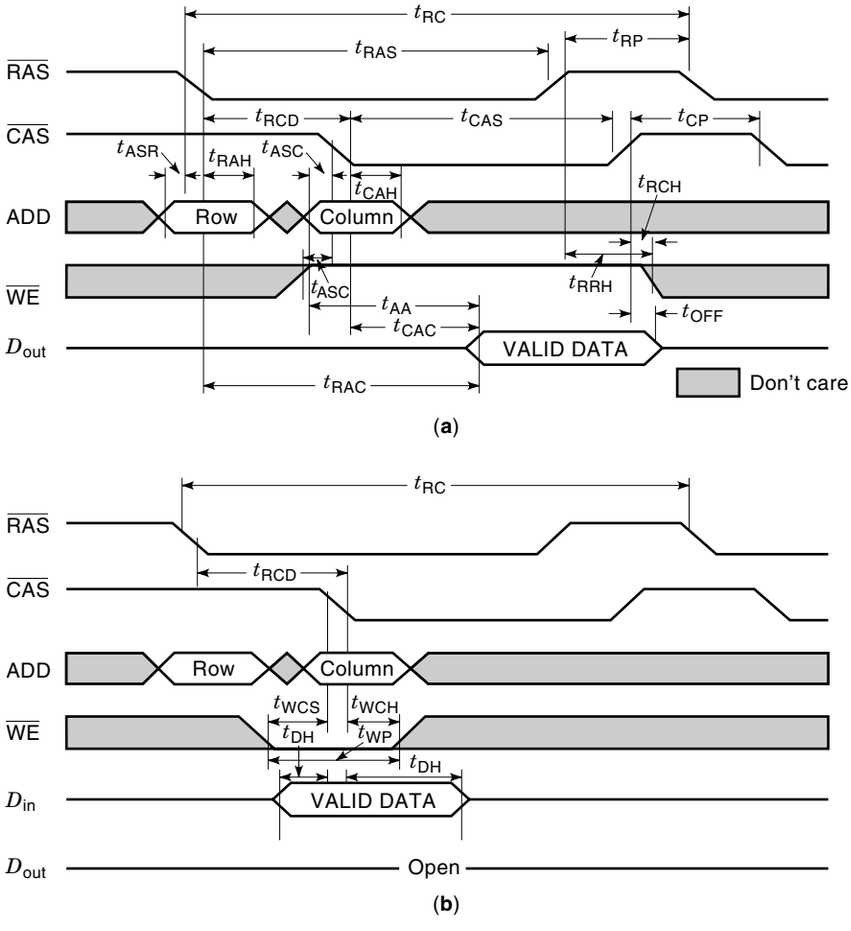

Figure 11. Fundamental DRAM operation scheme. (a) Read mode operation. (b) Write mode operation.

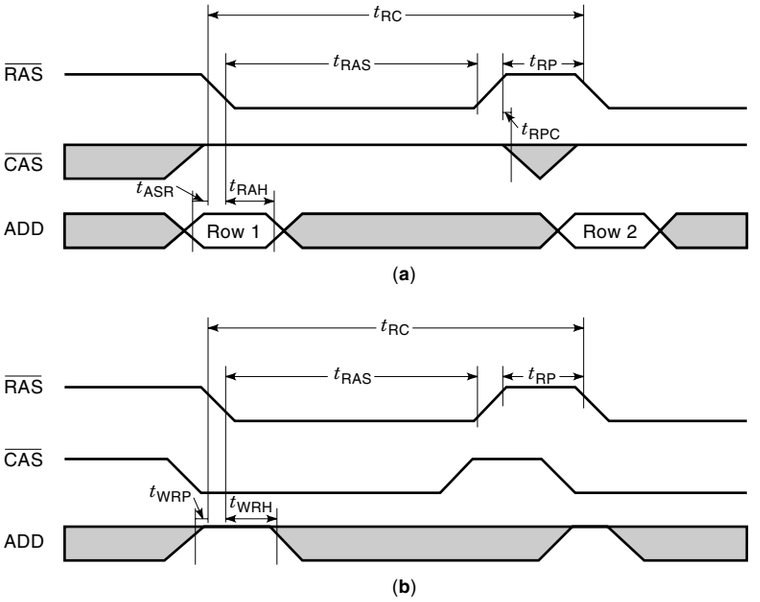

Figure 12. Refresh mode timing diagram. (a) $\overline{\text{RAS}}$-only refresh mode. (b) CBR ($\overline{\text{CAS}}$ before $\overline{\text{RAS}}$) refresh mode.

block activation. In other words, data are read out on the bit line and stored back to the same memory cell. For this operation, column address block is not activated. If the user sends the specific row address, the chip can be refreshed by $\overline{\text{RAS}}$-only refresh mode. If the user sets up $\overline{\text{CAS}}$ low before $\overline{\text{RAS}}$ low (CBR), the internal counter provides the specific row address instead of the external row address and starts the refresh operation. Figure 12 shows the refresh operation timing diagram. In addition to the above-mentioned operation modes, there are additional operation modes, such as read-modify-write mode, hidden-refresh mode, and so on. Data books provide detailed information on various operation modes (15).

In order to achieve high throughput rate, the fast page mode is useful. In the fast page mode, access time is just the $\overline{\text{CAS}}$ access time (t_{CAC}), although random access is limited to the selected row address. This is based on the characteristics of parallel structure of the DRAM memory cell array. In this mode, every data cell which is connected to the selected word line is amplified and sent to the bit line. The external column address data is used to select the data on the bit line. Theoretically, it is possible to fetch all the data which are connected to the selected word line successively. This operation does not require repeating the word line boosted operation and sensing operation. This access time is the same as $\overline{\text{CAS}}$ access time (t_{CAC}), which is defined as the column address access time. Figure 13 shows the fast page read cycle mode.

Extended Data Out Mode (EDO)

Recent progress has improved throughput rate efficiently. The technique is called extended data out (EDO) mode, or hyper page mode. With simple modification on data output buffers, it can increase the data throughput performance. Figure 14 shows the read cycle timing chart in the EDO mode. The difference between conventional fast page mode and EDO mode is the data output control. In fast page mode, output data are reset by rising $\overline{\text{CAS}}$; however, output data can be maintained until next $\overline{\text{CAS}}$ falling edge in EDO mode. This

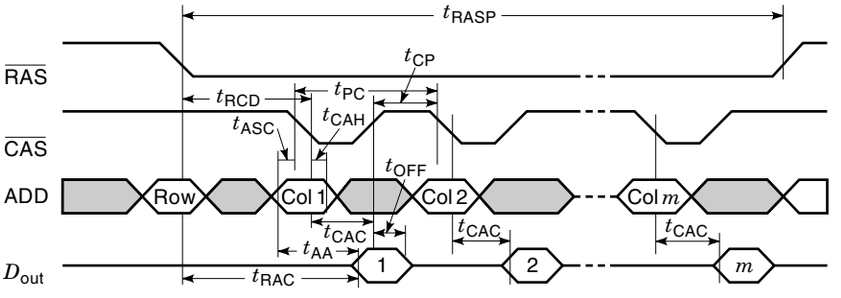

Figure 13. Fast page read cycle mode.

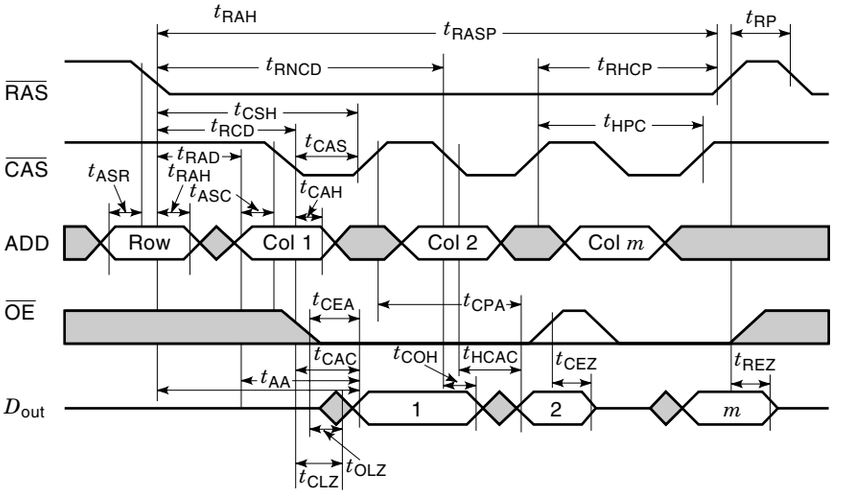

Figure 14. Extended data out DRAM read cycle mode.

unique feature enables one to use DRAM efficiently in terms of data throughput. Figure 15 shows the comparison of conventional fast page mode DRAM with word interleave technique and EDO mode. In fast page mode, the minimum value of the pulse duration of $\overline{\text{CAS}}$ high cannot be used, because it requires some period to reset the data after the $\overline{\text{CAS}}$ rising edge. Although the word interleave technique is used, it cannot reach the same performance as the EDO mode. This modification considerably improves the DRAM speed performance in terms of data throughput rate. Since the trend of the control DRAM is toward synchronous operation, this synchronization capability is a great advantage of the EDO. In terms of productivity, EDO has also an advantage because it can be manufactured based on the current fast page mode DRAM with small modifications rather than major changes. In addition, full compatibility with conventional fast page mode

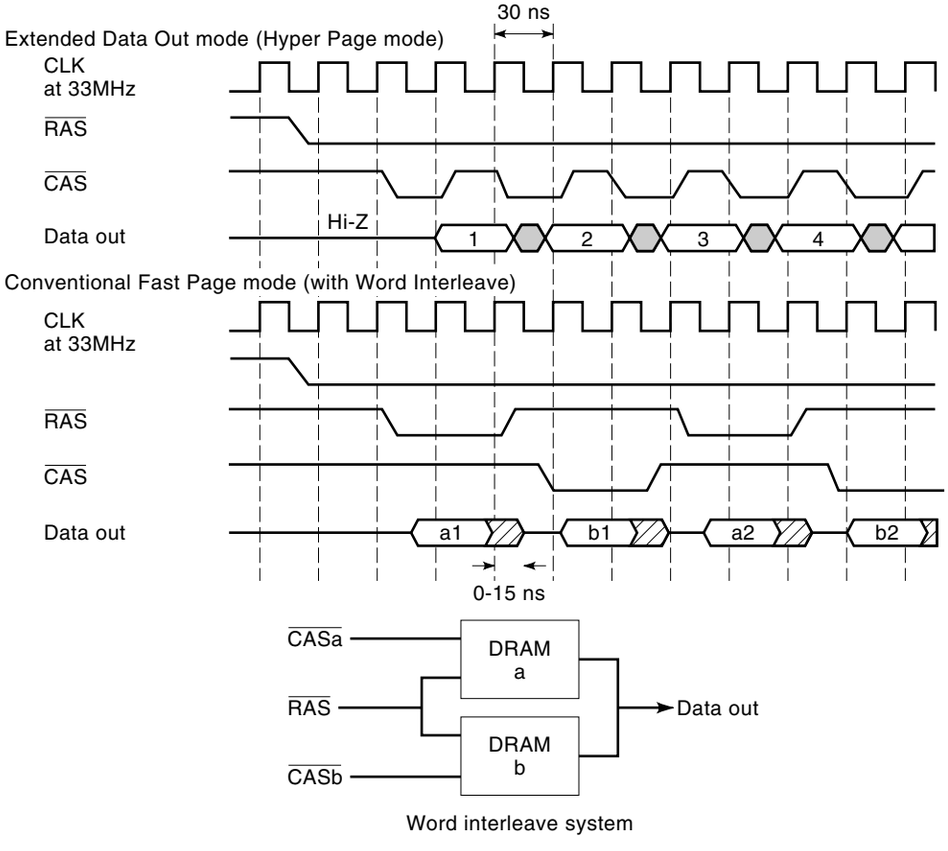

Figure 15. Comparison of EDO and fast page mode by word interleave operation.

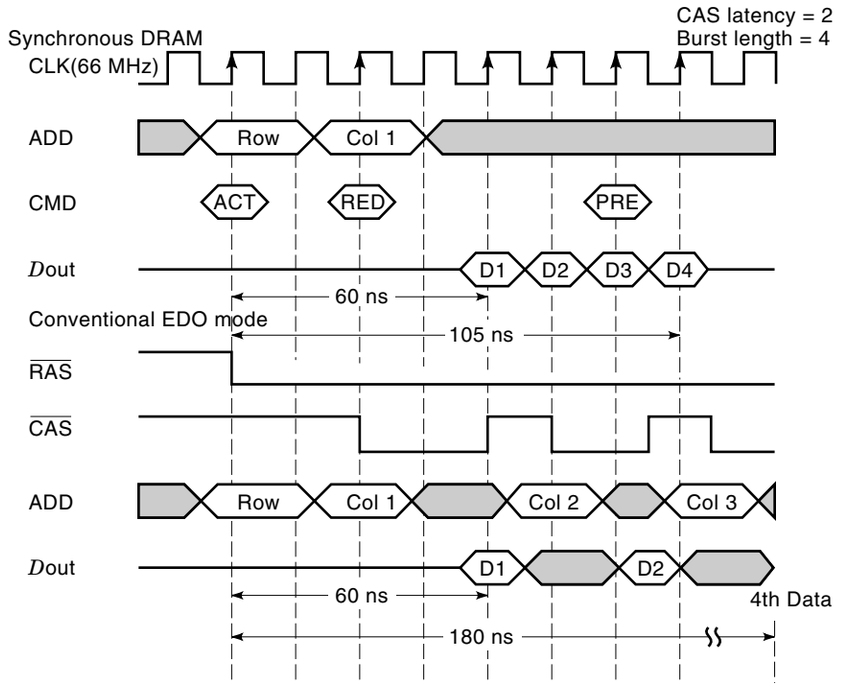

Figure 16. Comparison of operation methods of SDRAM and EDO mode DRAM.

makes EDO an acceptable product in the commercial market. Therefore, between the fast page mode and synchronous DRAM, the EDO scheme could be recognized as a suitable bridge in practice.

Synchronous DRAM

Although the EDO mode approach has improved speed performance, it cannot reach the level that memory application designers have expected. As one of the future memory technologies, synchronous DRAM (SDRAM) has been proposed to improve the throughput performance by introducing synchronous operation to DRAM (16). Figure 16 shows a comparison of data throughput performance between SDRAM and EDO DRAM.

In SDRAM operation, read or write operation is set by command, which is combination of levels of input pins at the rising edge of the clock. The command is decoded by the command decoder, and it starts to operate according to the decoded signal. Figure 17 shows how to set up commands, such as active, read/write, and precharge, by the input pin level. Owing to the adoption of command input scheme, the designer only has to pay attention to the set-up time and hold time of each input signal related to the clock rising edge. It significantly contributes to high-performance system design. SDRAM requires initial conditions to set several important characteristics of the chip, such as CAS latency, the length of burst data, and address generation type. CAS latency is the number of required clocks after receiving the read command to produce the first data. The length of burst data is defined

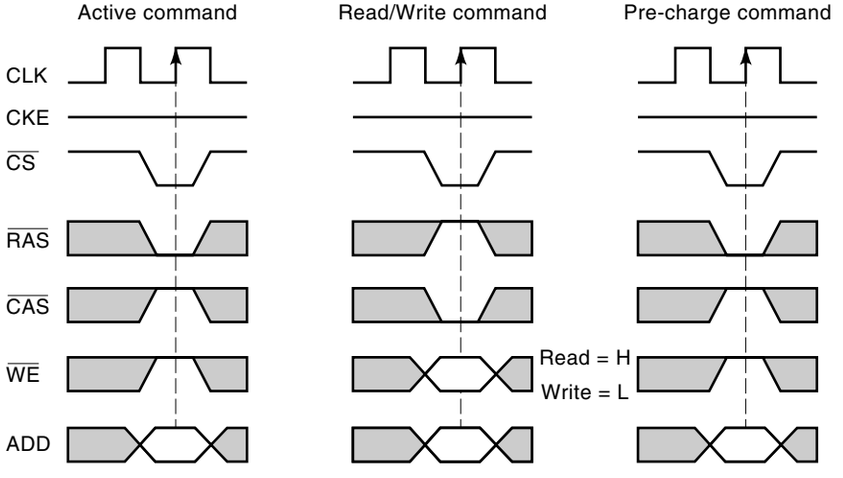

Figure 17. Command input SDRAM.

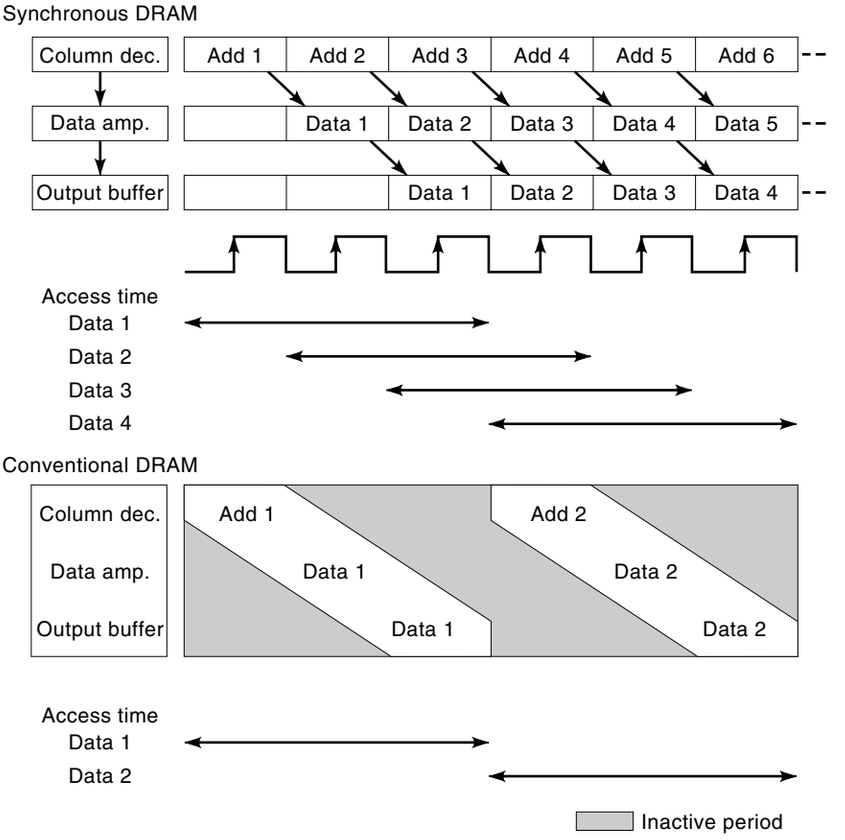

Figure 18. Comparison of pipeline process of SDRAM and DRAM internal process.

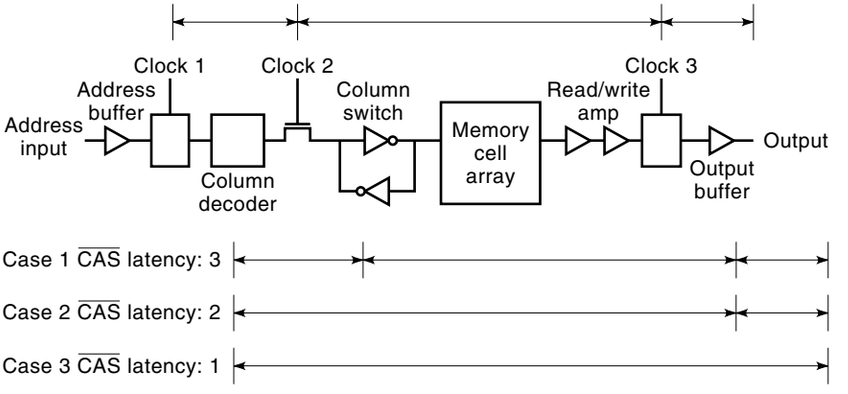

Figure 19. Internal pipeline structure of SDRAM.

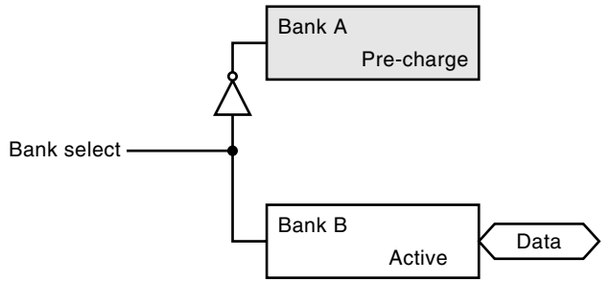

Figure 20. Multibank structure of SDRAM.

as the consecutive number of data. Address generation can be classified into two types, one being sequential and the other interleave.

SDRAM has great advantage in data throughput performance. Because of the synchronous operation, SDRAM has several crucial advantages to enhance its performance. First, all input signals are latched by the clock. System designers do not have to worry about complicated timing problems, such as signal skew. In addition, pipeline operation, which is defined as the parallel operation for sequentially separated blocks, can be more efficient than asynchronous operation. Figure 18 shows the timing chart of pipeline operation, and Fig. 19 shows the diagram for internal circuit structure of pipeline operation. According to the control clock, each seg-

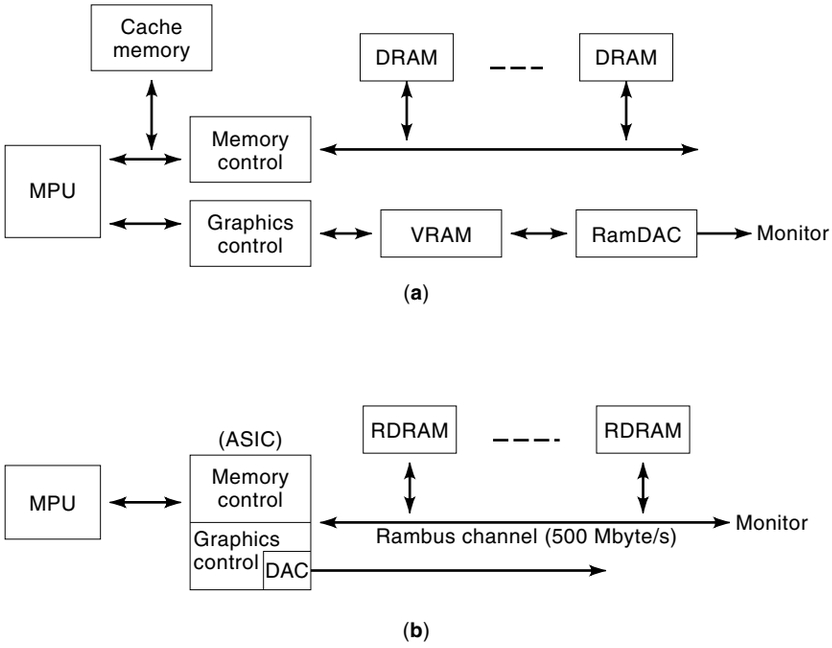

Figure 21. Comparison of (a) conventional memory system and (b) RDRAM system.

ment of column block is operated in parallel. Although access time for the first data remains the same, this approach can improve access time from the second piece of data to the last one. Furthermore, synchronous operation allows the implementation of multibank configuration on the chip, instead of using several discrete chips. In fact, the multibank structure can provide a nonprecharged period for the user, because some banks could be accessed while others are in the precharged period. Figure 20 shows the block diagram of a two-bank scheme. If Bank B is active, data are produced from only bank B memory cell array. During this period, Bank A is prepared for the next active period through precharging. This hidden multibank technique allows the designer to use memory in a very efficient manner. The SDRAM could play a major role in high-performance memory, especially for high-speed applications.

RAMBUS DRAM

Besides the improvement of DRAM chip performance itself, development of a high-performance interface is the future alternative approach. Rambus DRAM (RDRAM) was proposed to provide the optimized interface solution for data transfer between CPU and memory (17). In order to achieve this goal, RDRAM adopts a new scheme for memory architecture. Figure 21 shows the comparison of a conventional memory hierarchy and the Rambus system. The system-on-a-chip approach is the target goal for RDRAM. RDRAM adopts a 9-bit data bus. Since there is no specified address bus, a request packet which includes the command to set the bus for address bus should be sent to the chip first through the control bus when a chip is accessed. After the request packet, next come the acknowledge packet and the data packet. Because the initial condition has to be set, it will be late for the first data access. However, once a transfer condition is set, data access is achieved at considerably high speed, such as at 500 Mbyte/s. For the purpose of stable data input and output, a phase-locked loop (PLL) circuit is located on the chip to synchronize operation between the chip and the external clock. Sense amplifiers in the memory array are used for cache memory to realize fast response. Figure 22 shows the block diagram of the read cycle operation for RDRAM.

ACKNOWLEDGMENTS

Yoichi Oshima would like to thank Kunio Matsumoto for his continual encouragement, and appreciates the helpful support of Akira Shoyama, Jun Ihara, and Masaaki Awano, Semiconductor Memory Technical Group, Semiconductor Device Section, Japanese Patent Office. Dr. Bing Sheu and Mr. Steve Jen would like to thank Integrated Media Systems Cen-

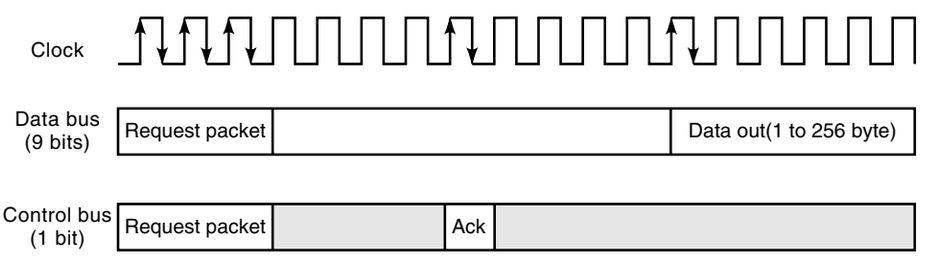

Figure 22. RDRAM operation scheme.

ter, an NSF Engineering Research Center in the multimedia field.

BIBLIOGRAPHY

1. T. Masuhara et al., VLSI memories: Present status and future prospect, *J. Inst. Electron. Inf. Commun. Eng.,* **E74** (1): 130–141, 1991.
2. H. Komiya, Future technological and economical prospects for VLSI, *IEEE Int. Solid-State Circuits Conf.,* San Francisco, pp. 16–19, 1993.
3. H. Sasaki, Multimedia complex on a chip, *IEEE Int. Solid-State Circuits Conf.,* San Francisco, pp. 16–19, 1996.
4. K. Ishibashi et al., A 6-ns 4-Mb CMOS SRAM with offset-voltage insensitive current sense amplifiers, *IEEE J. Solid-State Circuits,* **30** (4): 480–485, 1995.
5. K. Kashihara, T. Okudaira, and H. Itoh, A novel metal-ferroelectric insulator-semiconductor (MFS) capacitor using PZT/$SrTiO_3$ layered insulator, *IEEE Symp. VLSI Tech. Dig.,* Kyoto, Japan, pp. 49–50, 1993.
6. T. V. Rajeevakumar and G. B. Bronner, A novel trench capacitor structure for ULSI DRAMs, *IEEE Symp. VLSI Tech. Dig.,* Oiso, Kanagawa, Japan, pp. 7–8, 1991.
7. D. Kenney et al., A buried-plate trench cell for 64-Mb DRAM, *IEEE Symp. VLSI Tech. Dig.,* Oiso, Kanagawa, Japan, pp. 14–15, 1991.
8. T. Kaga et al., Crown-shaped stacked-capacitor cell for 1.5-V operation 64-Mb DRAM's, *IEEE Trans. Electron Devices,* **38**: 255–261, 1991.
9. H. Watanabe et al., A novel stacked capacitor with porous-Si electrodes for high density DRAMs, *IEEE Symp. VLSI Tech. Dig.,* Honolulu, HI, pp. 19–20, 1994.
10. K. W. Kwon et al., Ta_2O_5/TiO_2 composite films for high density DRAM capacitors, *IEEE Symp. VLSI Tech. Dig.,* Kyoto, Japan, pp. 45–46, 1993.
11. J. H. Ahn et al., Microvillus patterning technology for 256 Mb DRAM stack cell, *IEEE Symp. VLSI Tech. Dig.,* Seattle, WA, pp. 12–13, 1992.
12. K. Itoh, K. Sasaki, and Y. Nakagome, Trends in low-power RAM circuit technologies, *IEEE Symp. Low Power Electron.,* San Diego, CA, pp. 84–87, 1994.
13. H. Koinumai and T. Kimura, ×32 bit 16 M DRAM structure, *Nikkei Microdevices,* June, pp. 84–85, 1995.
14. M. Aoki et al., New DRAM noise generation under half-V_{cc} precharge and its reduction using a transposed amplifier, *IEEE J. Solid-State Circuits,* **24** (4): 889–894, 1989.
15. Texas Instruments, *MOS Memory Data Book,* June 1995.
16. Y. Takai et al., 250 M Byte/s sychronous DRAM using a 3-stage-pipelined architecture, *IEEE J. Solid-State Circuits,* **29** (4): 426–431, 1994.
17. N. Kushiyama et al., 500 M Byte/sec data-rate 512 K bits ×9 DRAM using a novel I/O interface, *IEEE Symp. VLSI Circuits Dig.,* Seattle, WA, pp. 66–67, 1992.

YOICHI OSHIMA
Japanese Patent Office

BING J. SHEU
STEVE H. JEN
University of Southern California

DRIVES, DC. See MOTOR DRIVES, DC.

DRY TYPE TRANSFORMER. See TRANSFORMERS, DRY TYPE.

DTL. See DIODE-TRANSISTOR LOGIC.

DTV TRANSMITTERS. See TRANSMITTERS FOR DIGITAL TELEVISION.