DRAM Module for System on LSI

Kazumasa Yanagisawa
Jun Sato

ABSTRACT: Technology for merging high density dynamic random access memory (DRAM) with LSI logic was newly developed to provide the optimum combination of memory and logic, an important part of system structure. Using LSI with embedded DRAM, various performance characteristics were improved resulting in an access time of 20 ns and a bus transfer capability increase of 20 times. The operating power for the memory section was reduced to 1/3 and the design was effective in reducing electromagnetic radiation. In addition, the design period can be reduced by moving high-speed board assembly design, previously a bottleneck in advanced designs, to the LSI. Design tools and other environmental factors are important from the standpoint of connectability in system design. The application specific integrated circuits (ASIC) design environment must be maintained in common regardless of the presence or absence of embedded DRAM. The goal of creating a merged system was achieved by combining memory cells, consisting of high density DRAM with 3-dimensional structures, and standard ASIC devices. In the DRAM module for the merged system, a micromodule method was used whereby internal micromodules are partitioned to the small level of 32 kbyte and then combined. In addition, a multiple bank structure and bank-independent access, which are special features of this method, are effectively used to greatly improve the practical application capabilities. Synchronous operation was selected as the operation mode taking into consideration the ease of ASIC development, immunity to external noise, and memory testability. Effective applications for merged DRAM include imaging in high end use and unified structure in portable systems.

INTRODUCTION
GENERALY the purpose of an information processing system is transmitting, processing, and storing the information. Among these purposes, a memory supports information processing and handles accumulation.

Problems were a few when the volume of information was small. However, information volume subsequently increasing, the transmission efficiency between memory and logic began to limit overall performance. As a countermeasure, extensive work was expended in creation of cache systems for information volume.

When building a total system using semiconductors, provision of optimum transmission capabilities is required. Optimal combination of unit memory, represented by the DRAM, logic LSI with embedded memory, and system control LSI is therefore very important. Mask ROM (read only memory) has been used for a long time as the program memory in information accumulation applications. Recently, to reduce development turnaround time (TAT), flash memory also began to be embedded.

Static random access memory (RAM) has a long history to be used as information processing applications. Capacity of SRAM is however limited because its memory cell is larger and therefore not suitable for high density. Thus technology for embedding DRAM, with their small memory cells, was developed for the purpose of achieving high density and high capacity.

Currently, practical mass production technology for embedded flash memory and DRAM products is available for customized products. The technology is now expanding into general purpose products. The DRAM module for merged integration, developed to provide large capacity memory in a system, is discussed in this report.
ADVANTAGES OF LSI WITH EMBEDDED DRAM

The advantages of embedding large-scale DRAM on logic LSIs include miniaturization, performance improvement, and reduction in development period. An example of a system structure with DRAM is shown in Fig. 1.

By embedding DRAM on an LSI, the number of components is reduced and the final product is easily miniaturized. Performance improvements attained by the merged integration are in access time, bus transmission capabilities, power consumption, and noise radiation.

Access time is very important in random access memory. As memory and logic are merged on the same chip, the transmission time between memory and logic circuits can be reduced. It can be 10–20 ns shorter than unit memory in the same product with the same design technology.

Bus transmission performance is also important in continuous information processing. The performance is determined by the product of bus width, frequency, and efficiency. With unit memory, the upper bus width is limited to about a few ten bytes due to circuit board fabrication limitations. However, with the DRAM-logic merged integration, the bus width can be increased about ten times and the bus efficiency factor about five times. Due to these improvements, compared to the conventional design, a 20 times or faster increase in performance (more than 20 Gbyte/s) is attained.

A large part of system power is consumed at the memory output buffer. This power consumption is proportional to the bus transmission performance and normally reaches 0.3 J/Gbyte. However, a large output buffer is not required with the merged DRAM-logic system and the power consumption is negligible.

Excluding an output buffer driving external interconnects, the merged integration is also effective in reducing electromagnetic radiation. Electromagnetic radiation is proportional to wiring length and data transmission speed. Therefore, radiation in the DRAM-logic merged integration is greatly reduced since there is no external wiring for memory control.

By simply merging the memory on the LSI, the operating power for the memory section is reduced to about 1/3 of unit memory of the same generation. LSIs with embedded memory are customized products and are not constrained by standardization or other limitations. They can thus realize lower power-supply voltage and lower power consumption by improvement in control logic and application architecture.

Low power consumption is very important not only with portable equipment but also for high-performance equipment. Power consumption always increases as chip performance increases. The increased power level often exceeds the heat dissipation limits of the package. Therefore, in products seeking high speed and high
When LSI power consumption is compared, LSIs with embedded DRAM consume about the same or slightly less power than logic LSIs alone. This is because the bus power is larger than the power of the memory unit. With system structures limited by heat generation concerns, embedded memory on the LSI allows further performance to be attained than would be possible with independent LSI in the logic section.

The problem of development period is considered next. In this respect, the problem of the embedded DRAM logic is the length of time period required for prototyping. On the other hand, an advantage is the ease of assembly design. The relationship between LSI performance and overall development period is shown in Fig. 2.

In order to attain high performance, time period is required during logic design for such details as the optimization of the number of logic steps and addition of parallel operation circuits. However, this is not the only way in which high performance is attained. Balance is also very critical between the logic section and the memory section, which feeds information to the logic section.

In order to maintain the bus performance for sending information from the memory, careful attention to details such as transmission waveform and electromagnetic radiation suppression is required during circuit board interconnect design. It is well-known that design difficulty rapidly increases in proportion to frequency. With the merged DRAM, the memory bus, which is the most difficult part of design, is moved into the LSI and circuit board interconnect design is greatly simplified.

Next, increased freedom is realized in DRAM specifications. With commodity DRAM, priority is given to maintaining compatibility with standard specifications and there is no basic design freedom. Even when attempting to reduce operating power, for example, the number of sense amplifiers — which determines the operating power — is related to basic capabilities of the DRAM — such as the refresh cycle — and cannot be altered. Therefore, depending on the application, even sense amplifiers that are not used are still operated. The same is true for functions that may be used to improve performance, such as the number of banks.

Based on the commodity synchronous DRAM (SDRAM), a “direct SDRAM” specification was developed by selecting and adding only effective functions. The direct SDRAM specification is optimized for DRAM modules with ASIC and in a 1-to-1 combination with logic. Up to 1,024 sense amplifiers and 16 banks can be selected. Thus limits imposed by memory structure and internal structure can be avoided and extensive design work is not required in order to maintain performance.

Ultimately, the advantage or disadvantage is determined by design work required for performance improvement or trial production time, whichever is longer. However, when great improvements in performance operation, priority is given to reducing the power consumed by upgraded performance.

Fig. 2—Performance and Overall Development Period.
The investigation of overall development time must consider the LSI fabrication and board design time in addition to the LSI design time.
When designing a system using LSIs with embedded DRAM, custom design and ASIC development are possible. Conventionally, the performance of logic gates used in commodity DRAM are degraded by thick oxide films with high voltage breakdown. First-generation merged DRAM products, developed based on the DRAM process, used this method. Since the resulting transistor gate length is long, the performance is about 1 generation behind that of the logic gates in standard ASICs of the same time period.

On the other hand, in the case of ASIC development, continuity of environment, such as design tools, is very important. Regardless of the presence of embedded DRAM, the same environment and performance of design must be provided. Therefore, logic gate performance must be matched to advanced standard ASICs and high performance is required of the logic gate.

The inferior logic gate performance of products with conventional memory processing is caused by the processes fabricating the memory cell. The thick oxide film with high voltage breakdown, required in the memory cell, the added high-temperature processes used for capacitor fabrication, and other factors interfere with improvement of logic gate performance. Thus, memory cell selection becomes very important in realizing high performance in logic gates.

DRAM memory cells are primarily fabricated by a planar method, used up to the 1-Mbit generation, and by a 3-dimensional method, used in products since the 4-Mbit generation. In addition, the 3-dimensional method can be divided into the trench method, in which holes are opened in the substrate and the capacitors formed internally, and the stacked method in which a capacitor is formed as a stack of layers on the substrate.

Of the above methods, the planar method is the most compatible with ASIC processes and eases the task of fabricating high-performance devices. However, with this method the memory cell density is about the same as with SRAM and high density cannot be realized. On the other hand, with the 3-dimensional methods, high density is easier to achieve because only the selection transistors of the memory cells occupy substrate real estate and the capacitors are stacked above the silicon surface.

The trench and stacked methods were evaluated. When fabricating memory cells by a lengthy thermal process, the trench method for forming logic gates has conventionally been advantageous. However, as processing has advanced toward lower temperatures,
both methods are expected to attain almost the same target performance from the 0.35-µm generation onward.

Technological details, such as cell density and process complexity, are the same but the stacked method was selected due to Hitachi’s extensive mass-production experience with this method. The relationship between memory density and logic gate performance is shown in Fig. 3.

**SPECIFICATIONS OF DRAM FOR EMBEDDED LSI**

**Micromodule Structure**

The goal of merged DRAM is fabrication of DRAM with optimal structures. Therefore automated creation, similar to the SRAM of the ASIC environment, is desired. However current DRAM memory cells, with their advanced 3-dimensional structure, cannot easily meet this goal. As with commodity DRAM, when manual work is added, the product development time cannot be satisfied. With this in mind, the memory was divided into small micromodules with units of 32 kbyte. A micromodule design, wherein only the required number of units are incorporated, is used.

The DRAM module consists of 3 types of micromodules; input-output (I/O), power, and bank. The I/O section includes the input and output amplifiers, related control circuits, and overall control circuits. The power section generates and provides potentials at voltages other than the power used within the DRAM. The bank section consists of direct control circuits for memory related to the memory cell, sense amplifier, and other units. The section also contains control circuits. The bank section contains all functions required to independently control the memory cell and can operate completely independently.

Since this method makes banks independent (surrounding regions are free), surface area is increased by up to 10% due to interference region and operating control circuits within each bank. However, this is equivalent to about 4 months of process advances (about 2.2% of surface area loss per month) and is more advantageous than customized design.

The micromodule structure is shown in Fig. 4.

**Operation Mode**

In addition to the synchronous operation mode SDRAM, DRAM operation modes also include the semi-synchronous mode extended data output (EDO) DRAM.

In the semi-synchronous mode, the amplifier is controlled by an address transition detection circuit. Since analog-type constant adjustment is required in the address transition detection circuit, the “prototype = mass production” required in ASIC development is not possible. This problem is not present with the synchronous mode using all clocks.

The next important item is inspection of the memory
section. Memory inspection items include normal relationships between signals, represented by “setup”, and delay components, represented by access time. The most important detail is measurement of delay time.

Delay time is not confined to a single cycle and extends over multiple cycles. This is a basic problem with merged DRAM, with random operation speeds at the 10-MHz level, and logic sections, which can operate at a few hundred MHz level. For the synchronous mode the length of the delay can be expressed as the product of easy-to-detect latency and frequency. For the semi-synchronous mode, the delay of the memory data bus changes with the number of modules on the chip and cannot be measured easily.

In addition, immunity to noise is a problem. With semi-synchronous products, immunity to noise generated by external circuit operation is required. Even in design alone, inspection of all numbers is not a simple matter. Therefore the synchronous mode is thought to be the most generally advantageous mode and has been adopted.

Multiple Bank Structure

Often, when attempting to increase system performance, a memory bottleneck occurs. SRAM has been used to eliminate this problem in general. Excluding memory density, SRAM is superior to DRAM from the standpoint of performance.

The difference between SRAM and DRAM performance is primarily due to differences in access time — column and row access time differs greatly. In many applications high-speed column access is used to realize higher performance. However, freedom is only 1-dimensional in the column direction and therefore this design cannot handle many degrees of freedom. In addition, making row and column access time equivalent is not possible in principle with DRAM since the basic operation of rows is destructive addressing.

Thus a multiple bank structure, an advantage of the micromodule method, was applied and a new high-speed axis of bank access was added to handle the problem (see Fig. 5).

**APPLICATION EXAMPLES**

**Image Application (Frame Buffer)**

Picture elements of images are arranged 2-dimensionally relative to the screen. Each element contains color information, such as RGB (red, green, and blue), and processing support information for the element, such as a Z buffer for accumulating the depth information of the element. In image processing, high speed random access is required for neighboring elements. To handle these requirements, elements are often gathered into tiles and stored in memory in a page mode. For this type of organization, the multiple bank structure of the DRAM module is especially advantageous.

The horizontal direction is assigned to columns and the vertical direction is assigned to banks to create the tile structure. As before, rows, with their low operating
speeds, are assigned to tile arrangements. Since the required bank only operates on a single tile, low power consumption is realized and large tiles can be organized. In addition, since interleave operation is possible for each bank, the access delay (row operation) accompanying transition between tiles can be concealed, making seamless access possible. The operating frequency relative to neighboring elements is fast and on the level of SRAM. An example of memory assignment in the imaging field is shown in Fig. 6.

Unified Memory System

As with the relationship between image memory and main memory, conventional memory systems distribute memory according to the application function. This is done to eliminate the limitations of the memory bus (bus neck) and makes it possible to divide and use a single memory for multiple application functions (unified structure).

The advantage of the unified structure is the ability to provide optimum memory allocation for the application. For example, assume an LSI contains 2 Mbyte (64 banks) of internal memory. If imaging uses the 16 color scale of video graphics array (VGA), 160 kbyte (5 banks) can be allocated to the image memory and the remaining 1,888 kbyte (59 banks) can be allocated to the main memory.

In order to make the same hardware handle the 64 colors of extended graphics array (XGA), 1,536 kbyte (48 banks) are allocated to the image memory and the remaining 512 kbyte (16 banks) are allocated to the main memory. The same hardware is therefore compatible with different brightness scales. The DRAM modules installed in ASIC have many banks and better bus performance than commodity SDRAM. Therefore, even more effective unified structures can be realized, as shown in Fig. 7.

With the unified structure, demands from the various application functions on the memory are mingled and expectations are not good for long-page mode applications that monopolize the bus. This means that control commands are frequently issued to the memory control bus and control bus conflicts occur with the address multiplexing scheme used in commodity DRAM. To overcome this problem, the control bus is
CONCLUSIONS
The DRAM module developed to place large capacity memory on the LSI was discussed here. Previously, when the volume of processed information was small, memory and logic were arranged separately. However, as the volume of information has increased, the connections between memory and logic have become a bottleneck resulting in the demand for diverse memory layers. The history of computer development tells the story.

As information volume increases, demands for a larger number of memory levels will impact all systems and requirements for logic LSI containing large capacity memory will increase. In the future, various types of memory, including large capacity SRAM and DRAM modules for functional improvements, are expected to be embedded to advance system integration. Hitachi will make every effort to meet these demands.

REFERENCE

ABOUT THE AUTHORS
Kazumasa Yanagisawa
Joined Hitachi, Ltd. in 1979. Belongs to the System LSI Business Operation in the Semiconductor & Integrated Circuits Division. Currently engaged in development of DRAM modules for merged LSI. Member of Electronic Communication Association. E-mail: yagani@cm.musashi.hitachi.co.jp

Jun Sato
Joined Hitachi, Ltd. in 1984. Belongs to the System LSI Business Operation in the Semiconductor & Integrated Circuits Division. Currently engaged in development of logic products with emerged DRAM. E-mail: satojun@cm.musashi.hitachi.co.jp