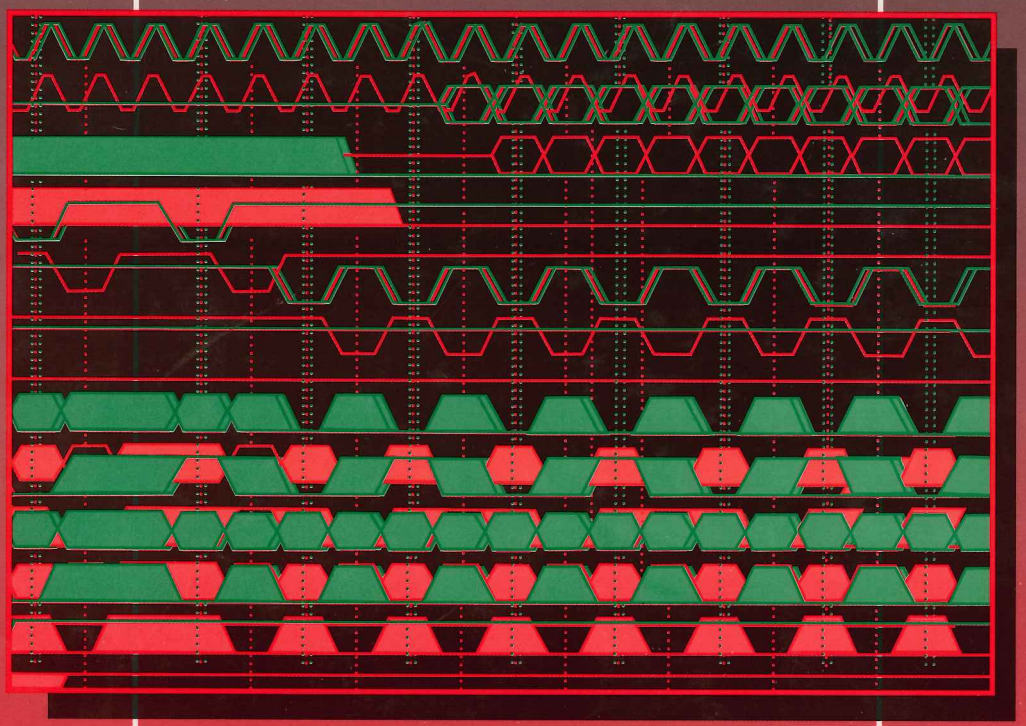
WILEY

HIGH PERFORMANCE MEMORIES

New architecture DRAMs and SRAMs evolution and function



Betty Prince

Prince

HIGH PERFORMANCE MEMORIES

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158-0012, USA

Jacaranda Wiley Ltd, 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Canada) Ltd, 22 Worcester Road,
Rexdale, Ontario M9W 1L1, Canada

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01,
Jin Xing Distripark, Singapore 0512

Library of Congress Cataloging-in-Publication Data
Prince, Betty.
High performance memories / Betty Prince.
p. cm.
Includes bibliographical references and index.
ISBN 0 471 95646 5
1. Semiconductor storage devices. 2. Very high speed integrated circuits. I. Title.
TK7895.M4P68 1996
621'.39'732–dc20

95–25742 CIP

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0 471 95646 5

Typeset in 10/12pt Palatino by Keyword Typesetting Services Ltd., Wallington, Surrey
Printed and bound in Great Britain by Bookcraft (Bath) Ltd.
This book is printed on acid-free paper responsibly manufactured from sustainable forestation,
for which at least two trees are planted for each one used for paper production.

Contents

7 Graphics DRAMs

7.1 Overview of DRAMs for Graphics Subsystems

Graphics in television systems and graphics subsystems in computers use a sufficient amount of memory to require the higher density of DRAMs. These subsystems also have special requirements for higher bandwidth than is available on the basic asynchronous DRAM.

A basic frame buffer in either the video or graphics application is required to provide at a minimum a continuous serial stream of data to refresh the display screen. A subsystem with some manipulation of bits requires, in addition, a random port for fast interface with the processor or graphics controller to provide the required data manipulations. Either these specific graphics requirements need to be met on the memory chip or a very high bandwidth memory must be available to support a graphics controller and a parallel-to-serial device which provide the required functions. Various combinations of these two approaches have been tried. These approaches are outlined in this chapter.

The market volumes of the systems involved are historically high enough to have generated a number of applications-specific DRAMs to serve the special requirements of the graphics subsystems. These included through the 1980s a variety of simple serial frame memories used in television applications and the more standard dual ported video DRAMs which have one serial and one random port which have been used in computer applications requiring more graphics manipulations.

In the 1990s the datarate of the single port DRAM has been increased significantly with the introduction of the EDO (Hyperpage) mode and synchronous DRAMs, with very wide interfaces. This has led to increased use of fast, wide single port DRAMs in graphics subsystems in PCs coupled with the use of standard graphics controllers which provide the multi port interface to the processor and the RAMDAC. Many new single port DRAM variants are also being developed such as the synchronous graphics DRAMs and the Multibank DRAM.

Multi-port VRAMs continue to be used in high end applications and appearing on the horizon are chips which gain performance and save space by integrating larger parts of the graphics subsystem, either the logic and DRAM separately or the entire subsystem, into the DRAM. These integrated chips are being initially developed for portable applications such as notebook computers. They are called "multi-ported" because the two ports are dissimilar.

Meanwhile for low end PC applications there is a trend toward unified memory which returns the frame buffer function to the main memory of the computer. Unified memory is made possible by advances both in processor and DRAM bandwidth and in software development.

This chapter considers first simple serial video memories for television applications, then the high bandwidth single port graphics memories, the multi-port memories, and finally the new phenomenon of the integrated graphics memory.

7.2 Frame Memories for Television Applications

7.2.1 *Simple Serial Field Memory for Temporary Frame Storage*

In a basic video subsystem such as is used in television sets to reduce visible lines or visible flicker on the screen, only a simple serial access storage device with four-bit input and output is required to store a frame and recycle it to the screen. A random access port is not required since there is no graphics manipulation involved.

Such devices are called field memories and are frequently made from DRAMs configured with serial input and output ports instead of a single parallel port. The DRAM core array is not changed. Read and write frequencies range from 33 to 50 MHz for a line of serial data.

An example is a 256K×4 field memory from Texas Instruments. This part has a 5 V power supply, two four-bit wide ports for fast FIFO (first-in-first-out) operation, and asynchronous read and write at 33 MHz providing a bandwidth of 16MB/sec. It has cascade connection capability so two or more parts can be connected to increase storage size.

Cost is a pressing issue in consumer systems such as televisions and a smaller package with fewer leads helps reduce the cost of the part. There are no external address pins in a serial access RAM, which eliminates nine pins. The inputs and outputs are demultiplexed to improve control, which adds four pins, plus the write enable (WE\) becomes a separate Read (R) and Write (W) control. The Output Enable (OE\) is replaced by a Reset Read (RSTR) and Reset Write (RSTW) pin, resulting in the addition of one control pin, giving a total of 16 pins, so the field memory fits in a 16-pin package rather than the 20 or 20/26-pin package used by the random access 1M DRAM.

The 20/26 package is a 26-pin package with the six middle pins removed to accommodate the large early generation 1M DRAM chips. A comparison between the pinouts of the serial field memory and the standard DRAM is shown in Figure 7.1.

Another change is that the RAS\ and CAS\ control signals are replaced with serial clocks.

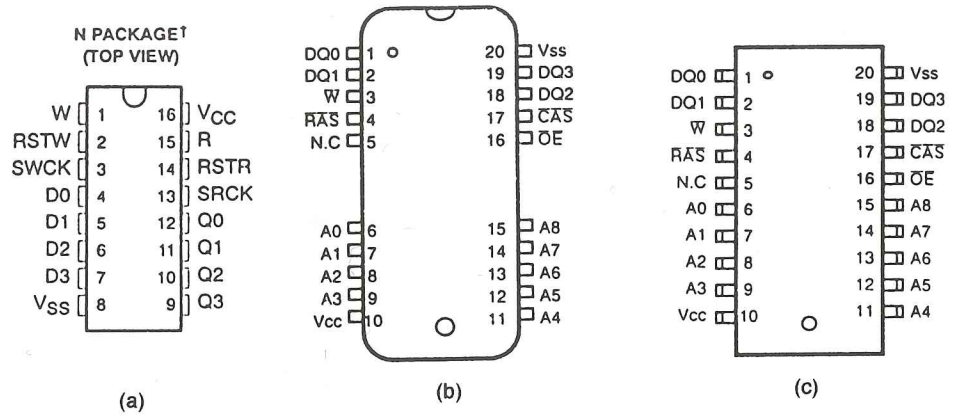


Figure 7.1 Pinout comparison of 1M serial frame memory and 1M DRAM: (a) 16-pin serial field 1M 4-bit DRAM package; (b) early 20/26-pin 1M ×4 DRAM package; (c) late generation 20-pin 1M ×4 DRAM package

The addressing is controlled by write and read address pointers which clock the data read out or written in. These must be reset to zero before a new memory access begins. The chip provides self refresh and arbitration logic to prevent conflict between memory refresh requests and data input and output cycles.

A functional block diagram is shown in Figure 7.2 which shows the DATA input and output buffers, the data cache ("A" and "B" line) buffers, the read and write counters along with the ring oscillator, the address pointers, the serial read and write timing controllers, and the read and write reset controllers.

The first 120 words of data input are stored in the "A" line cache buffer for fast access without having to access the main memory. The next data starting with word 120 goes into the 256-word write line buffer and thence into the main array. While the upper write line buffer is transferred to memory, the bottom write line buffer will be filling.

The simple frame memory storing up to 4M bits running at 33 MHz is useful for storing one frame in conventional television. High definition TV (HDTV) requires a sampling frequency of 70–80 MHz and 16Mbit of serial RAM for field memory. Two interleaved 8Mbit serial DRAMs running at 50 MHz or one 16Mbit serial DRAM running at 100 MHz can serve this application.

7.2.2 *Serial DRAM for High Definition TV Frame Storage*

The block diagram of an 8Mbit 50 MHz serial DRAM designed by Matsushita for HDTV applications is shown in Figure 7.3. The part has eight serial inputs and eight serial outputs. Each of the internal 128K × 8 subarrays has a serial-to-parallel and parallel-to-serial converter [10]. The bandwidth is 50MB/sec and for two of these parts interleaved is 100MB/sec.

A higher density DRAM can permit storage of multiple fields of video data. For example, a 256Mbit serial access DRAM, also from Matsushita, can store two seconds

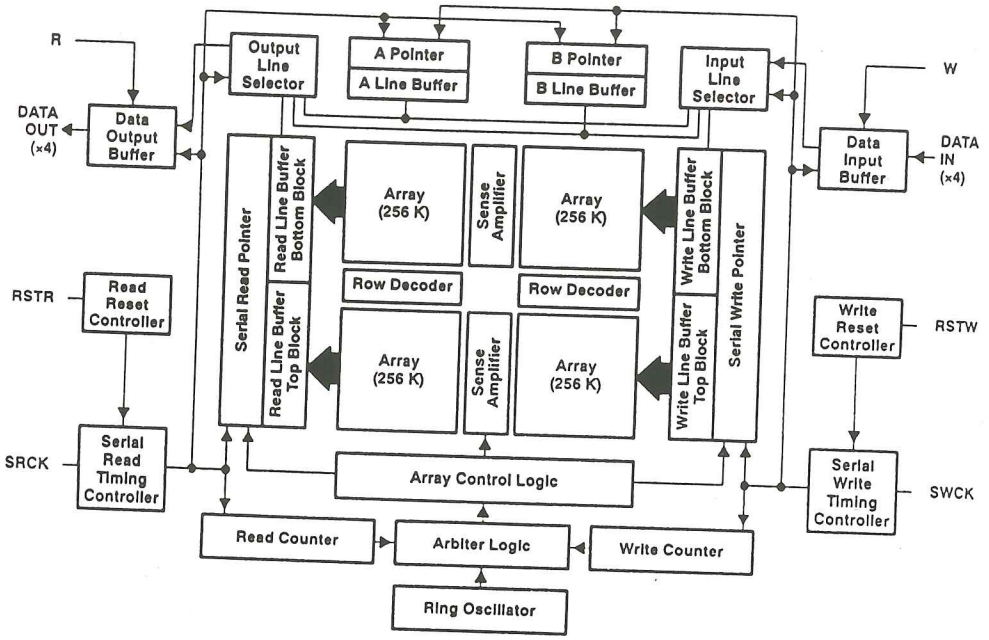


Figure 7.2 Functional block diagram of 1M field memory for TV applications (source: Texas Instruments [4])

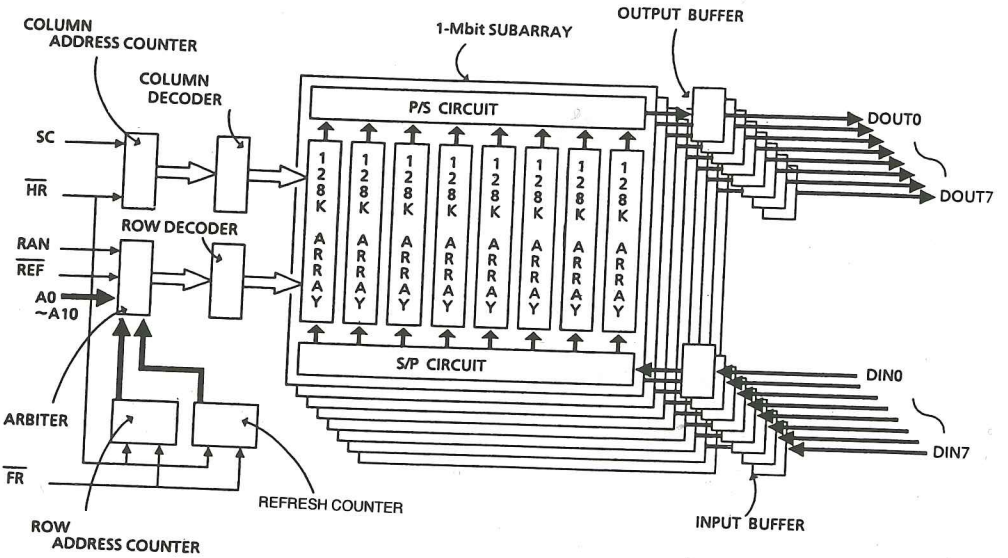


Figure 7.3 Block diagram of an 8M field memory (source: H. Kotani, 1990 [10] permission of IEEE)

of NTSC data or one second of HDTV data. It has 16 serial input ports and 16 serial output ports and runs at 100 MHz providing 1.6GB/sec of bandwidth in the system.

The internal array is similar to that shown for the 8M field memory in Figure 7.3. A timing diagram of the part is shown in Figure 7.4 [11].

The serial port clock is shown running at 100 MHz. Signals are read on the rising edge of the clock. An 80 ns RAS\ latency and 40 ns CAS\ latency for read is shown.

The merger of television and computer in the new trend to multimedia systems may mean the advent of new applications-specific television memories.

7.3 Single Port A-synchronous DRAMs for Graphics Applications

Graphics subsystems in computers have moved from being only in the high end workstations and mainframes, to being almost universally present in PCs. The requirements range from fairly simple graphics in low end PC to a much higher level of complexity in mid-range to high end PCs.

7.3.1 *Wide DRAM for Unified Memory in Low End PC*

Low end PC systems have used wide asynchronous DRAMs to obtain the bandwidth required to implement graphics features in the system. The high bandwidth of wide

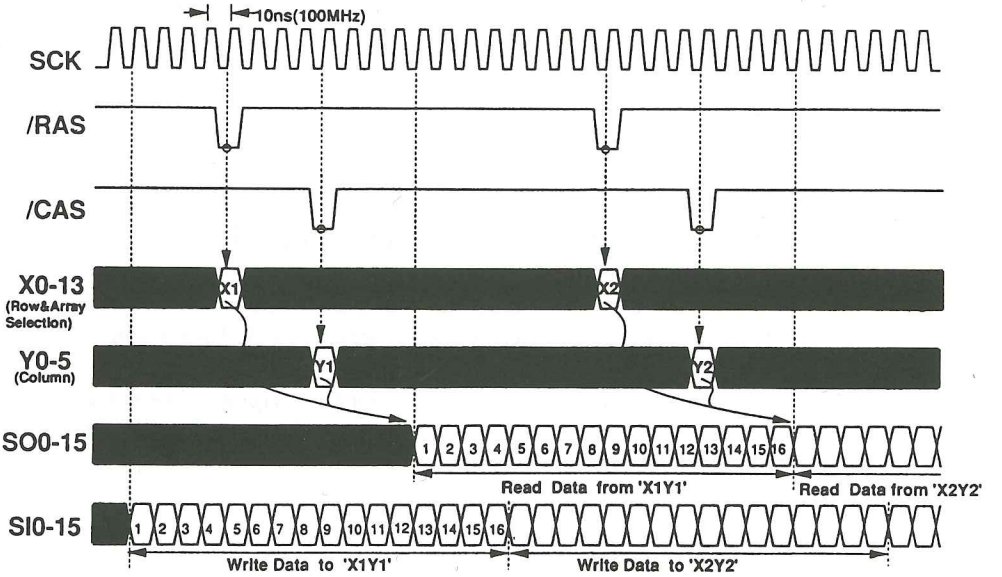


Figure 7.4 Read/write timing diagram of 256M clocked field memory (source: H. Kotani, 1994 [11] permission of IEEE)

DRAMs with fast modes such as fast page and Hyperpage (EDO) mode were described in Chapter 4. Parts such as 1M×16 DRAMs with 40 MHz Hyperpage mode cycle times are adequate for both main memory and graphics in a low end computer system.

When main memory storage and graphics memory are combined in one main memory storage location, the system is said to have a "unified memory". A unified memory using asynchronous DRAMs gives adequate bandwidth only in very low end PC systems.

7.3.2 *Wide DRAM for High Speed Printer Graphics*

Another system which needs high bandwidth memory for a graphics type of application is high speed printers. Printers have not required more than a few megabytes of memory storage along with 25–33 MHz speed and a wide interface. This means that the wide asynchronous DRAMs have also been used for this application.

For example, four 512K×8/9 fast page mode DRAMs running at 25 MHz on a 32-bit bus offers 100MB/sec datarate. This combination can be upgraded to a single 512K×32 25 MHz fast page mode DRAM, saving the cost of the four packages. This in turn can be upgraded in speed to a 512K×32 Hyperpage mode DRAM with a 33 MHz cycle time which gives 130MB/sec datarate [15]. New printers now in development may require considerably more memory which will still require the wide interface.

A pinout of a 28-pin 512K×8 DRAM TSOPII package is shown in Figure 7.5 along with that of a 70-pin 512K×32 DRAM TSOPII package. The savings in board space in going from four of the former to one of the latter are not as significant as in previous generations of upgrades, partially due to the number of power and ground pins which need to be added to keep the ratio of power and ground to DQs at 1:4 to control ground bounce.

7.4 Graphics Features on Asynchronous Single Port DRAMs

There has also been an attempt to add graphics features to the asynchronous single port DRAMs which are intended specifically for the graphics PC buffer applications. The write-per-bit function, also known as "mask write", is such a feature.

There is also a function called persistent write-per-bit, which allows a mask to persist for more than one cycle. These features are also implemented on the dual port graphics DRAMs to be described in a later section.

7.4.1 *Write-Per-Bit*

The Write-per-Bit (WPB) function [12] provides the ability to alter, or mask, some of the bits in a word while leaving other bits in the same word unaffected. If the mask is

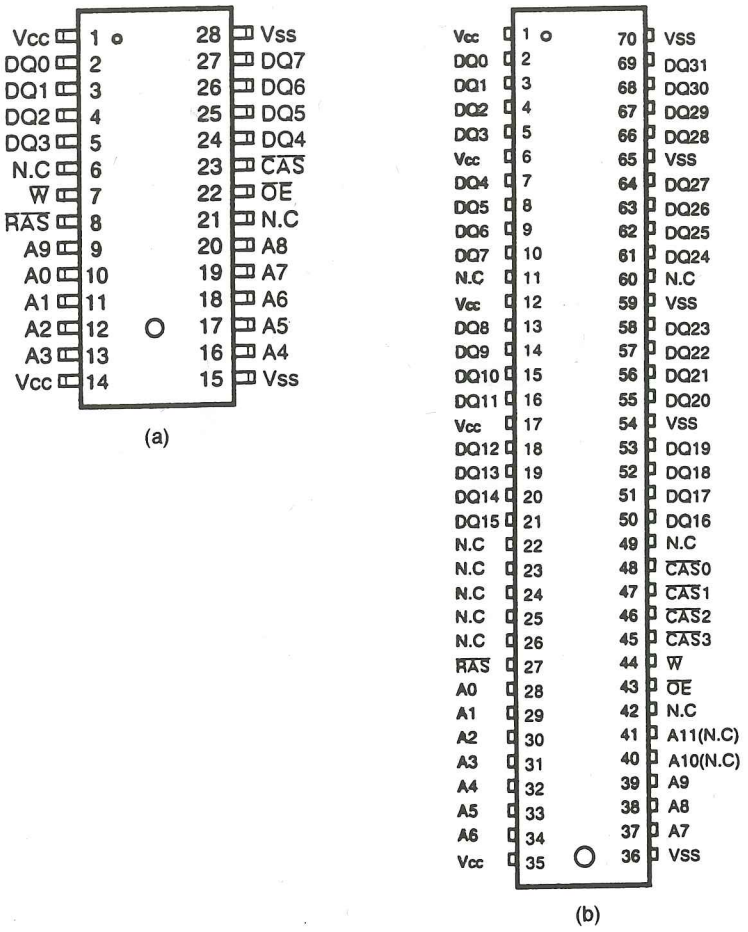


Figure 7.5 Pinout comparison of (a) 512K×8 and (b) 512K×32 DRAM packages (source: Samsung [14])

set as part of the write cycle, it can be used with no increase in cycle time over a standard read or write cycle.

Write-per-bit is implemented using a register on the data-in buffer which is latched on the falling edge of RAS\ and enabled by a low signal on the WE\ pin at a RAS high-to-low transition. A block diagram of a 256K×16 DRAM with the mask data register is shown in Figure 7.6.

This feature permits any number of bits in a word to be changed during a write cycle. In an asynchronous DRAM the mask is applied to the DQ lines and loaded into a register at the falling edge of RAS\ if the write enable (WE\) signal is low. When the DQ line is high, the corresponding bit will be written when the write cycle executes. If the DQ line is low, the bit remains unchanged.

An example of a timing diagram comparison for a simplified write-per-bit cycle and a normal write cycle is shown in Figure 7.7.

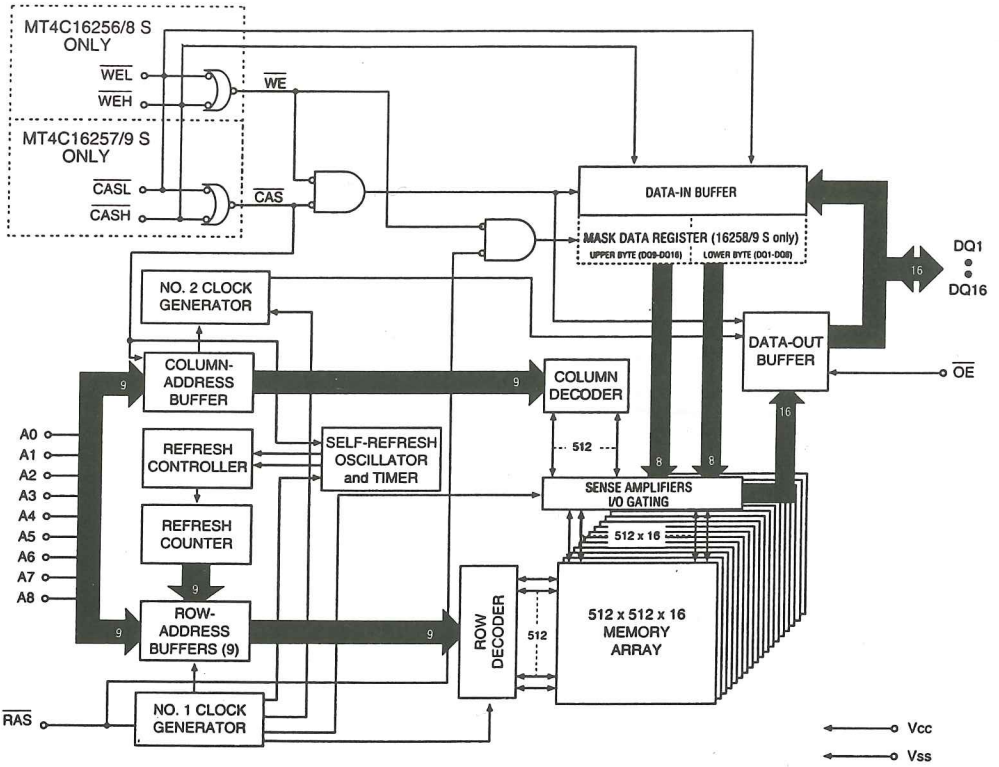


Figure 7.6 Block diagram of 256K×16 DRAM showing mask data register (source: Micron Technology [13])

Figure 7.8 illustrates the effect of the mask data on the stored data for different inputs along with the timing diagram for the masked and non-masked write for a 512K×8 DRAM [13].

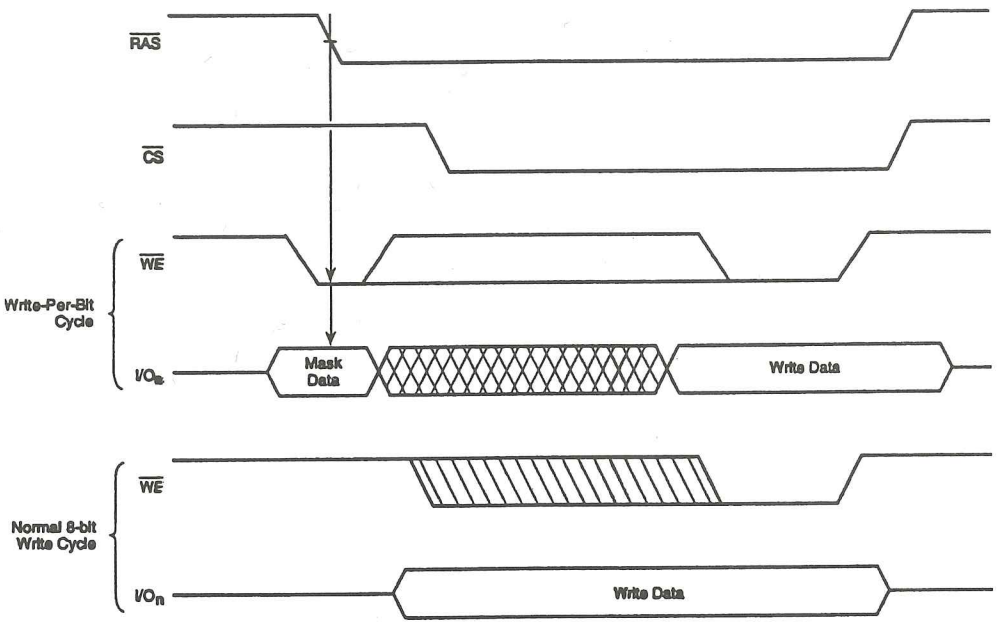
During page mode operation, a mask can be loaded at the falling edge of RAS\ and will remain set and active during a write cycle as long as RAS\ remains low. A mask is effective throughout a single page mode cycle and may not be changed during this cycle as shown in Figure 7.9 [15].

7.4.2 *Persistent Write-Per-Bit*

In systems where a single mask will be used for several cycles, some chips permit a mask to be set and persist for more than one cycle. This is referred to as "persistent write-per-bit". It has not been commonly offered on the asynchronous wide DRAMs because of the additional silicon cost involved.

7.5 Synchronous Single Port DRAMs Used in Graphics Systems

The synchronous DRAMs provide additional speed for a single port DRAM up to perhaps 200 MHz. The synchronous graphics DRAMs (SGRAMs) are a functional



I/O	Mask Data	Write Operation
$I/O_1 - I/O_8$	1 (High)	Write Bit
	0 (Low)	Disable Bit Write

Figure 7.7 Comparison of write-per-bit cycle vs. standard 8-bit write cycle (source: NEC [12])

superset of the synchronous DRAMs. They have all of the features of the SDRAMs plus some additional features useful in graphics subsystems.

The earliest synchronous graphics DRAM was a 4Mbit part developed by a few vendors such as United Memories, Fujitsu, and Hitachi. An 8M SGRAM, which is functionally identical to the 4M, is being offered by many vendors and is expected to become the first major SGRAM part.

7.5.1 4M Synchronous Graphics DRAMs

The 4Mbit standard $\times 16$ SDRAM is 3.3 V with an LVTTL interface and/or 5 V with a TTL interface. It is single ported and runs up to 66 MHz providing 132MB/sec of bandwidth. Graphics features supported included an eight-bit Block Write and a special function pin (DSF pin) to select between a standard SDRAM function and the SGRAM functions [17].

The block diagram of a 4M synchronous graphics RAM is shown in Figure 7.10(a) and a pinout is shown in Figure 7.10(b). The SGRAM uses two banks internally which can provide a high speed sustained burst.

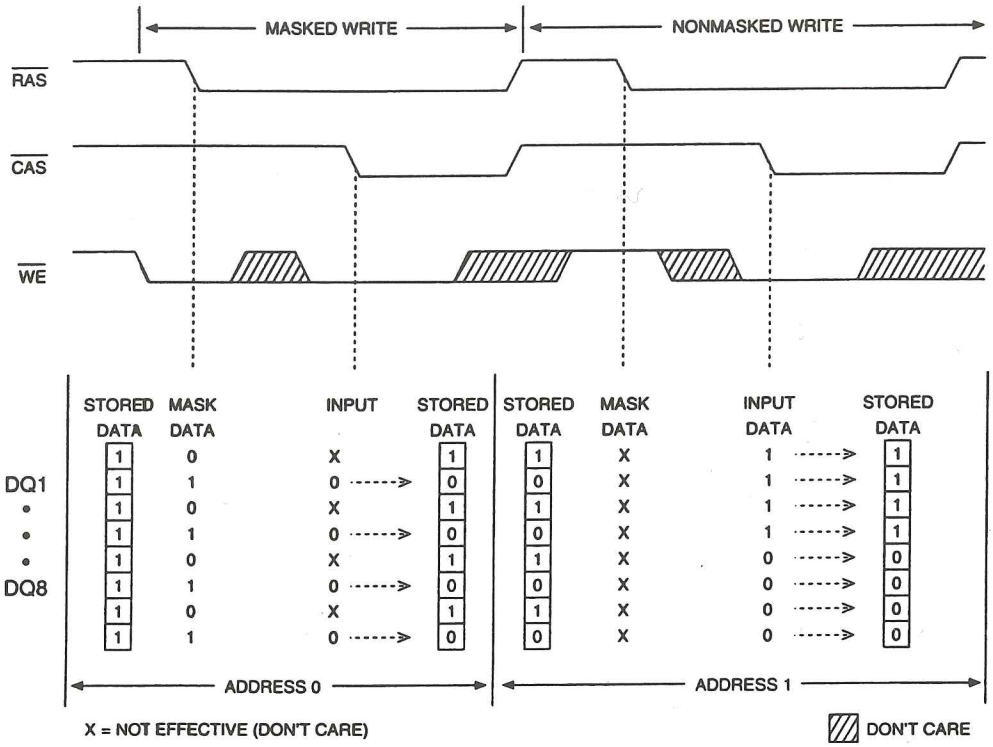


Figure 7.8 Illustration of effect of WPB mask on stored and input data (source: Micron Technology [13])

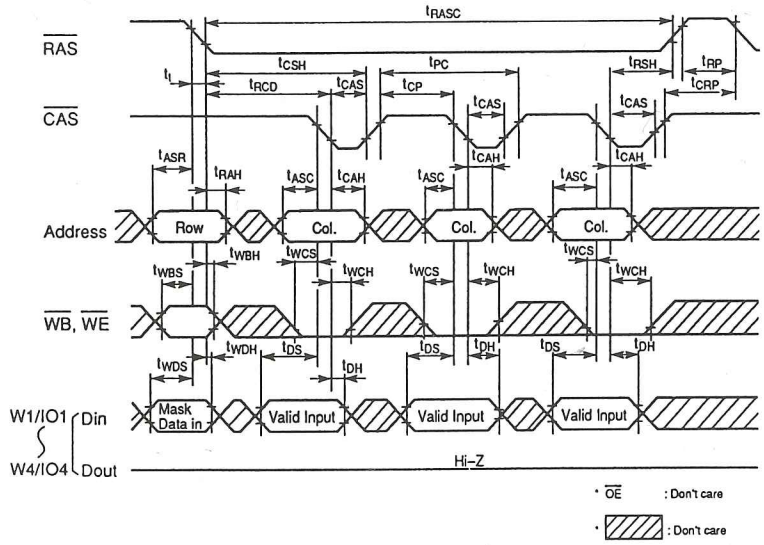


Figure 7.9 Write-per-bit during a fast page mode early write cycle (source: Hitachi [15])

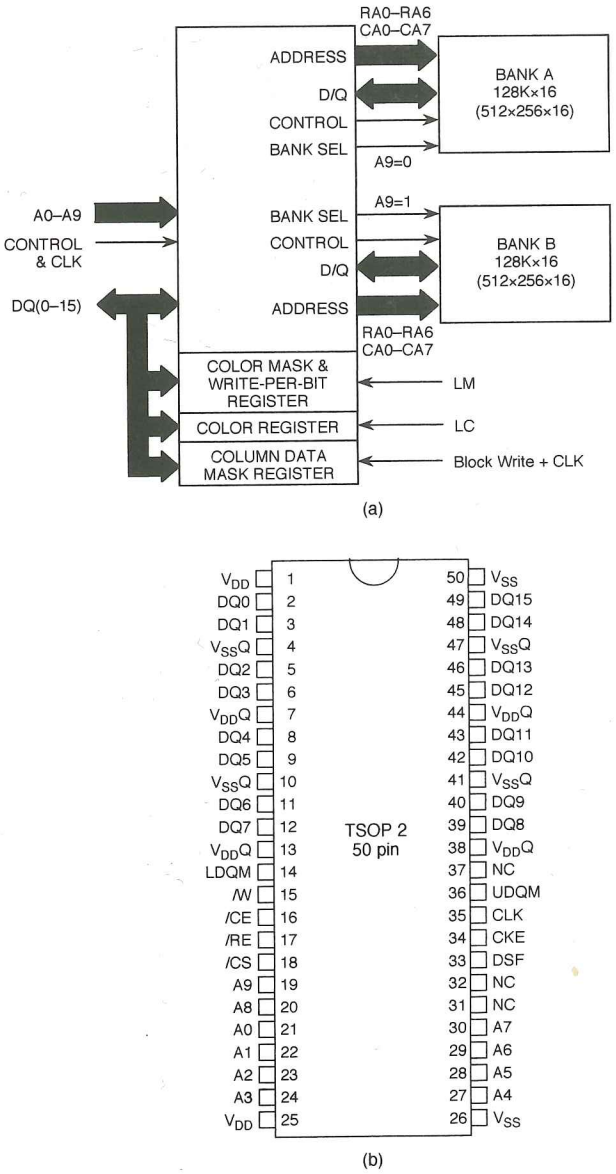


Figure 7.10 4M synchronous graphics DRAM: (a) functional block diagram; (b) pinout (source: United Memories)

The DSF pin is intended to implement the graphics features on this part. If the DSF pin is not connected, the part is intended to be a 4M SDRAM version of the standard 16M SDRAM. Graphics features include block write, and block write with auto precharge which will be described further in the next section.

7.5.2 8M SGRAMs

The 8M synchronous graphics DRAM is a standard part which is being sourced by many suppliers. It is a single port SDRAM with 3.3 V LVTTL interface. Various of the SGRAMs are also expected to offer other high speed interfaces such as HSTL or T-LVTTL. The organization is 128K (131 072) × 2 banks × 32 bits. That is, it has two banks each of which are organized 128K×32. The speed has been specified up to 100 MHz providing 400MB/sec datarate.

The standard package is the 100-pin plastic QFP with 14 × 20 mm^2 package dimensions and 0.65 mm pin pitch. The 8M SGRAM has a pin rotation which is similar to that of the 4M SGRAM. It will also be pin compatible with the 16M SGRAM which is organized 512K×32 and is in development at several suppliers.

The pinout and pin definitions are shown in Figure 7.11 [9].

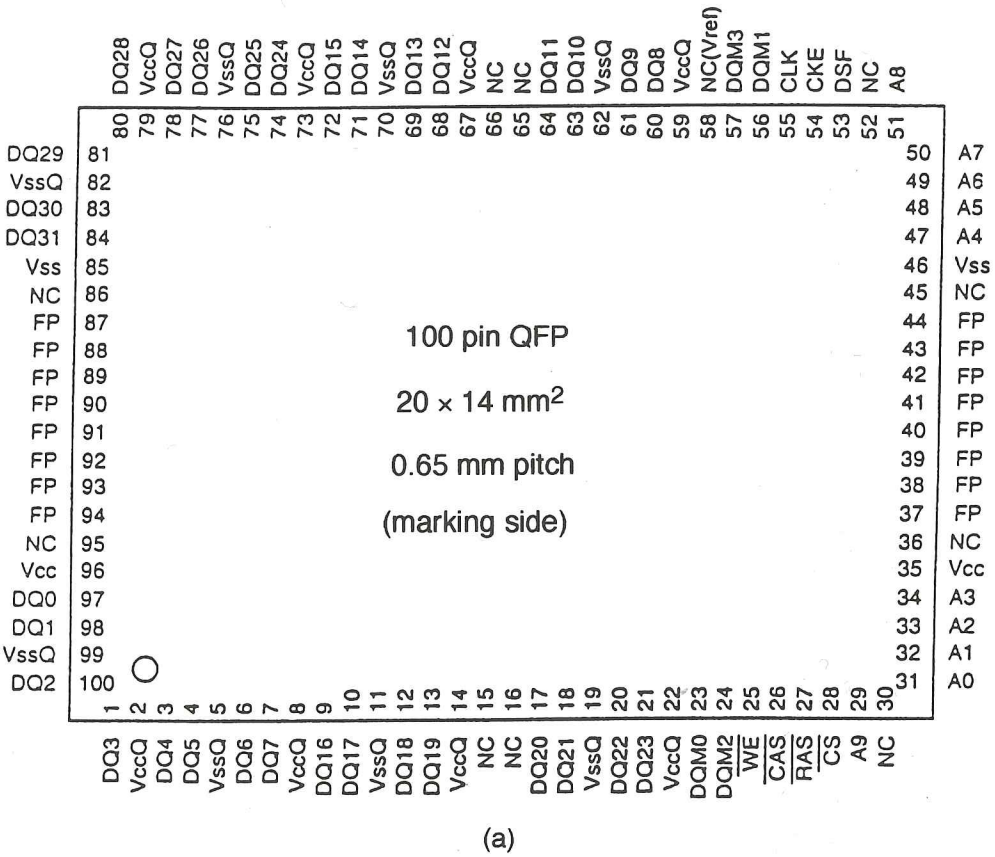


Figure 7.11 8M synchronous graphics RAM: (a) pinout; (b) pin definitions (source: NEC [9])

SYMBOL	FUNCTION
A0 – A9	ADDRESS INPUTS
A0 – A8	ROW ADDRESS INPUTS
A0 – A7	COLUMN ADDRESS INPUTS
A9	BANK SELECT
DQ0 – DQ31	DATA INPUTS / OUTPUTS
CS\	CHIP SELECT
RAS\	ROW ADDRESS STROBE
CAS\	COLUMN ADDRESS STROBE
WE\	WRITE ENABLE
DQM0 – DQM3	DQ MASK ENABLE
DSF	SPECIAL FUNCTION ENABLE
CKE	CLOCK ENABLE
CLK	SYSTEM CLOCK INPUT
VCC	SUPPLY VOLTAGE
VSS	GROUND
VCCQ	SUPPLY VOLTAGE FOR DQ
VSSQ	GROUND FOR DQ
FP	FLOATING PIN (WITH INTERNAL CONNECT TO VBB)

(b)

SOURCE: NEC [9]

Figure 7.11 *(continued)*

7.6 Special Graphics Features on SGRAMs

Special graphics features included on the 8M SGRAMs [5, 29] beyond the basic SDRAM features include masked block write, and mask write which includes the write-per-bit function. These features are standardized.

To the command functions present in the normal SDRAM mode register, the SGRAM has added Special Mode commands which control Color and Mask Registers which have also been added.

The Color Register is used in block writes and the Mask Register is used in mask writes (write-per-bit). The Mode Register with a Special Mode Register section blocked out is shown in the block diagram of an 8M SGRAM in Figure 7.12. Also shown are the Color Register and Mask Register.

The Command Truth Table of the SGRAM contains the standard command functions of the SDRAM and the Special Mode Register command functions for color and mask operation of the SGRAM as shown in Figure 7.13.

The Special Mode Register is controlled by the DSF (designated special function) control pin.

If the DSF pin is low (inactive), the 8M SGRAM operates similarly to a JEDEC Standard SDRAM. For standard SDRAM operation, the addresses A0–A8 are row addresses when the active command is given.

When CAS\ is active, addresses A0–A7 are column addresses and address A8 enables and disables the auto-precharge function. Address A9 is the bank select, BA. For BA low, Bank 0 is selected, and for BA high, Bank 1 is selected.

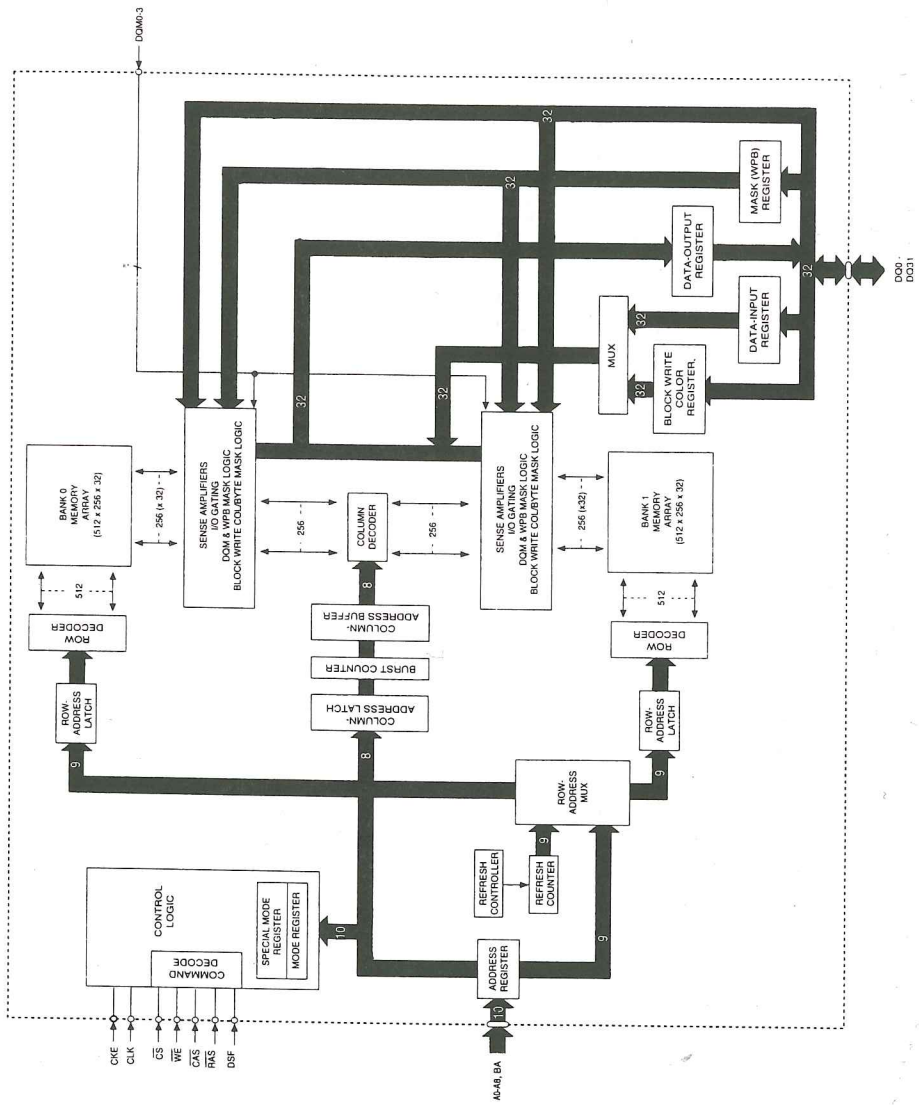


Figure 7.12 Block diagram of an 8M SGRAM (source: Micron Technology [5])

NAME (FUNCTION)	$\overline{CS}$	$\overline{RAS}$	$\overline{CAS}$	$\overline{WE}$	DSF	DQM	ADDR	DQs
COMMAND INHIBIT (NOP)	H	X	X	X	X	X	X	X
NO OPERATION (NOP)	L	H	H	H	L	X	X	X
ACTIVE (Select bank and activate row)	L	L	H	H	L	X	bank/row	X
ACTIVE with WPB (Select bank, activate row and WPB)	L	L	H	H	H	X	bank/row	X
READ (Select bank & column and start READ burst)	L	H	L	H	L	X	bank/col	X
WRITE (Select bank & column and start WRITE burst)	L	H	L	L	L	X	bank/col	VALID
BLOCK WRITE (Select bank & column and start BLOCK WRITE access)	L	H	L	L	H	X	bank/col	MASK
PRECHARGE (Deactivate row in bank or banks)	L	L	H	L	L	X	Code	X
BURST TERMINATE	L	H	H	L	L	X	X	Active
AUTO REFRESH or SELF REFRESH (enter SELF REFRESH mode)	L	L	L	H	L	X	X	X
LOAD MODE REGISTER	L	L	L	L	L	X	OpCode	X
LOAD SPECIAL MODE REGISTER	L	L	L	L	H	X	OpCode	VALID
Write enable/output enable	-	-	-	-	-	L	-	Active
Write inhibit/output High-Z	-	-	-	-	-	H	-	High-Z

Figure 7.13 Command truth table for 8M SGRAM (source: Micron Technology)

If the DSF pin is high (active), the 8M SGRAM graphics operations are active. These include the "Special Mode Register Load" cycle, and the various masked write and block write functions.

7.6.1 *Load Special Mode Register Cycle*

When all control pins are low and the DSF pin is high, the Special Mode Register is loaded using inputs A0–A8 and the bank select BA. The "Load Mode Register" command can be issued when both banks of the SGRAM are idle.

A block diagram of the Special Mode Register definition is shown in Figure 7.14.

7.6.2 *Load Mask Register*

During a "Load Special Mode Register" cycle, A5 controls the Mask Register. If A5 is "0", the Mask Register is unchanged. If A5 is "1", the Mask Register is loaded with the new data applied to the DQs. The Mask Register then acts like a per DQ bit mask during Masked Write and Masked Block Write Cycles. The mask register will retain this data until it is loaded again or until the power is turned off.

7.6.3 *Load Color Register*

Similarly during a "Special Mode Register Load" cycle, A6 controls the 32 bit color register. If A6 is "0" the Color Register is unchanged and if A6 is "1" and the special

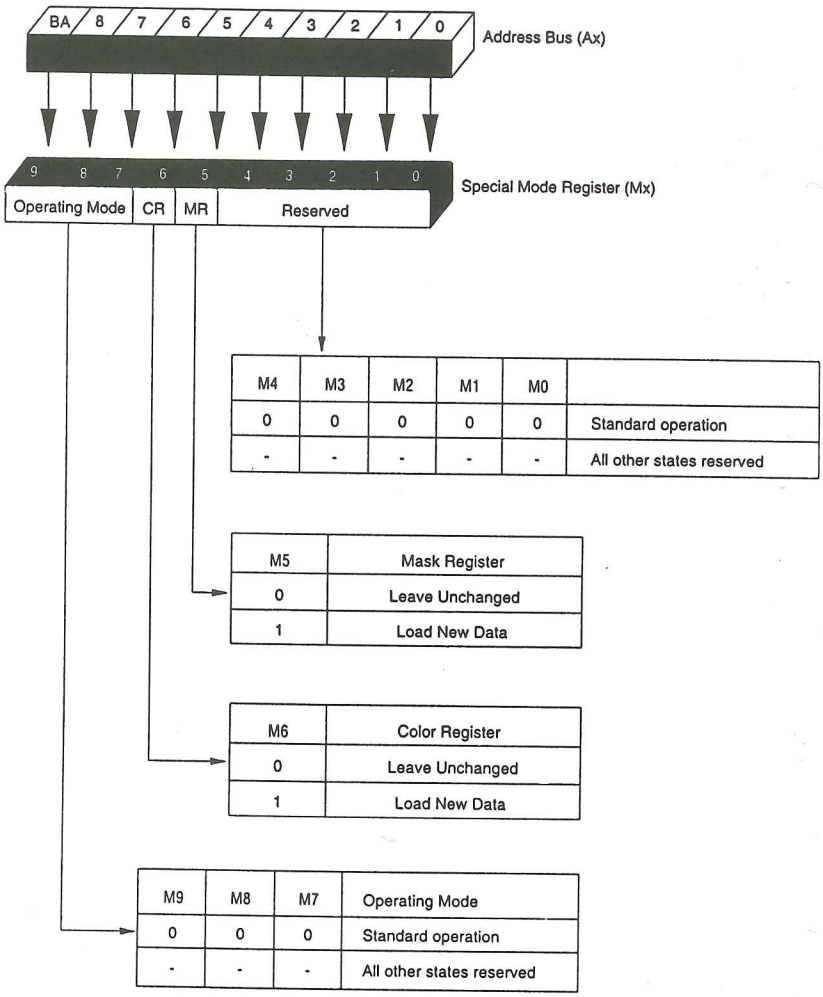


Figure 7.14 Special mode register definition (source: Micron Technology [5])

mode register load conditions are in effect, the Color Register is loaded with the data applied to the DQs. The Color Register then supplies the data during Block Write cycles. It will retain this data until reloaded or power is turned off.

The Load Special Mode Register command can be given when both banks are idle. It can also be given if either or both banks are active but no Read, Write, or Block Write access is in progress.

Other Special Mode Register inputs include A0–A4 which are all set at zero for standard operation. Other configurations of A0–A4 are reserved for future definition. A7, A8, and BA indicate the operating mode.

A truth table and timing diagram for a "Special Mode Register Load" cycle is shown in Figure 7.15.

On the rising edge of the clock, CS\, RAS\, CAS\ and WE\ are low, and DSF is high. The register data on A0–A8 and the BS is sampled and the Mask or Color

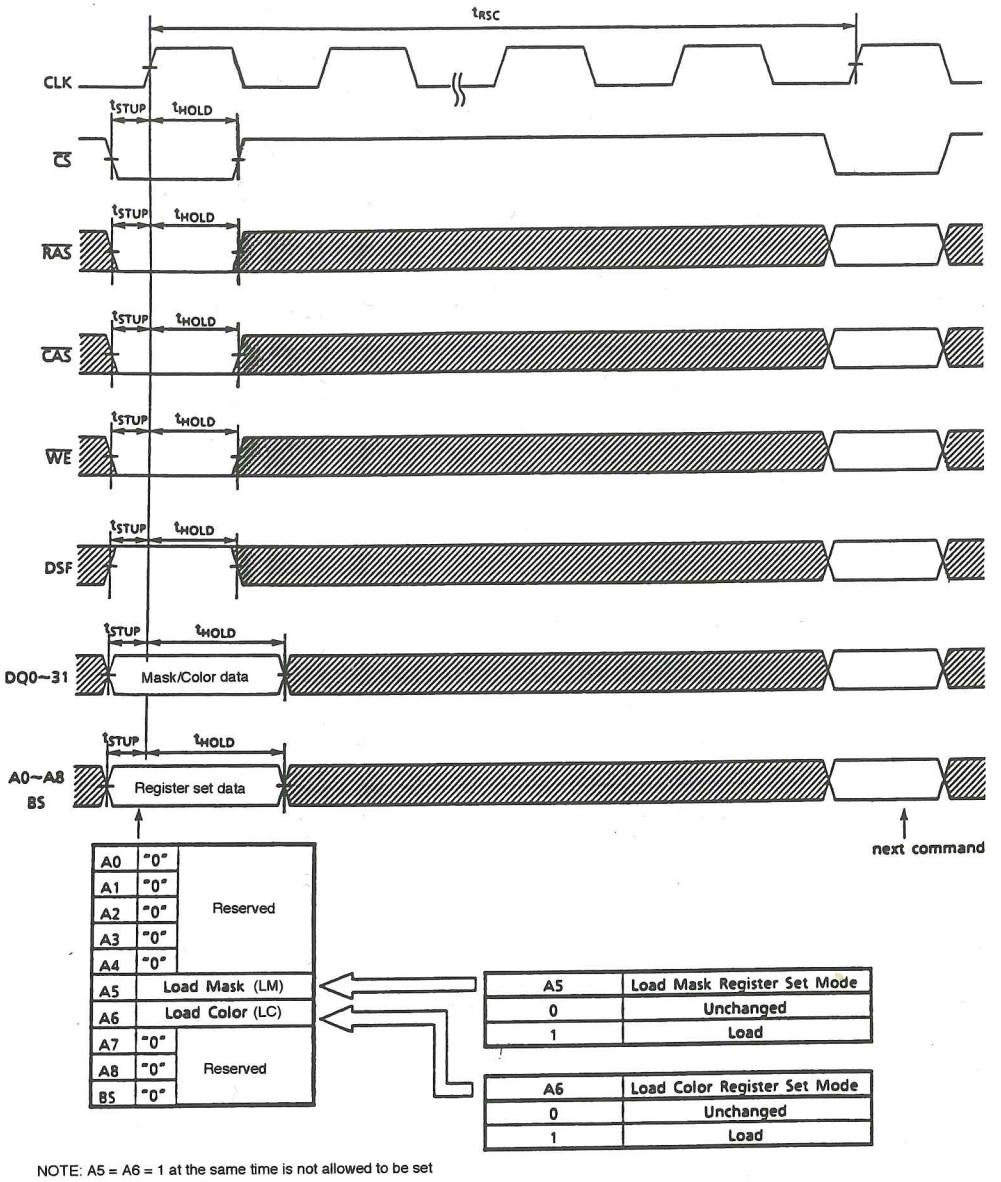


Figure 7.15 Special mode register load timing diagram (source: Toshiba)

Register data on DQ0–31 is sampled. A5 = A6 = 1 is not allowed, that is, the Color and Mask Registers can not be set on the same cycle.

7.6.4 *Active Graphics Commands*

Commands with DSF high are also shown in the Command Truth Table for activate Masked Write (Write-per-Bit) and activate Block Write. Otherwise the DSF pin is a

"don't care", although some manufacturers recommend that it be held low for compatibility with future, as yet undefined, special modes.

7.6.5 *Masked Write-Per-Bit*

The "Active with Mask Write (Write-per-Bit)" command is similar to an active command during which the write-per-bit is activated. Any subsequent write or block write cycles to the selected bank or row will be masked in accordance with the contents of the 32-bit Mask Register, the DQM signals, and, in the case of a Block Write command, the column/byte mask information from the color register [5].

The write-per-bit data acts as an I/O (DQ) mask. It uses the bits in the Mask Register to mask various data inputs applied to the DQ pins during the write cycle as illustrated in Figure 7.16. This figure uses for illustration only the first I/O (DQ) byte (first eight bits of the 32-bit I/O).

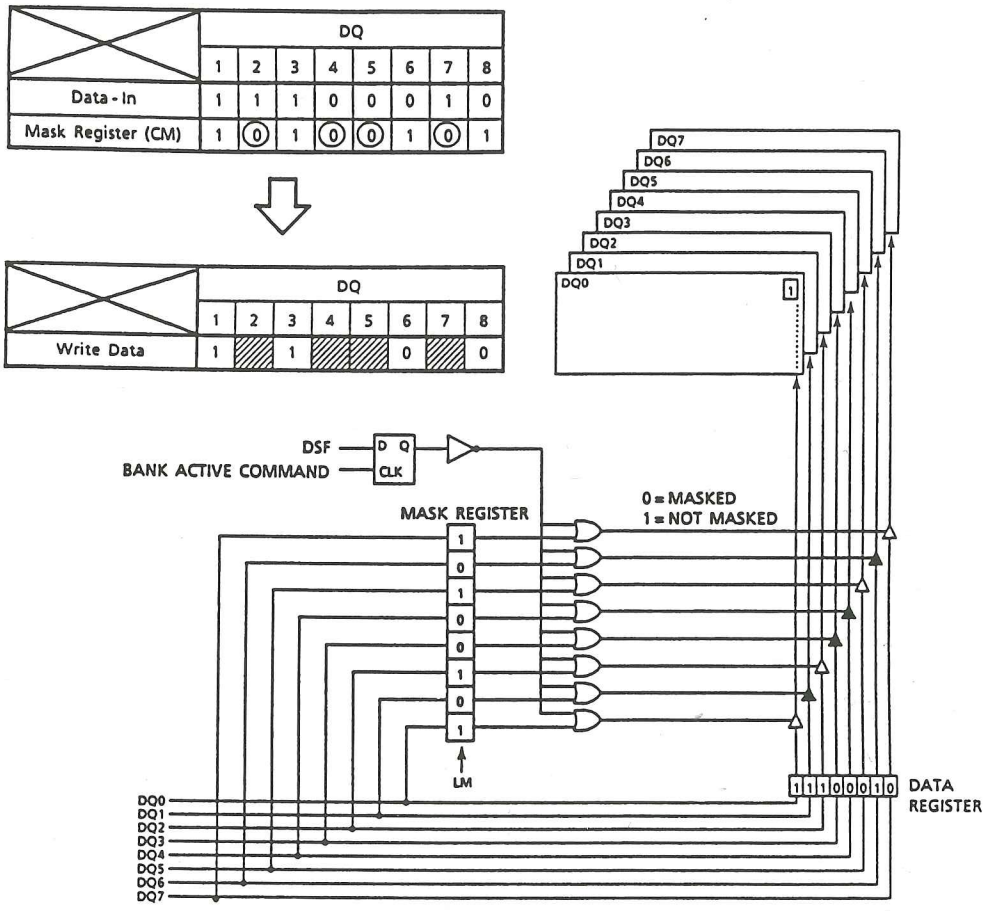


Figure 7.16 Masked write (write-per-bit) (source: Toshiba)

Write-per-bit can be burst. If DSF is low when the Bank Active command is executed, then the write-per-bit mask will be disabled for that cycle. If DSF is high at the Bank Active command the mask will be enabled and the DQ inputs will be masked.

7.6.6 Block Write

The "Block write" command is used to write a single 32-bit word into a block of eight consecutive column locations in one row. These column locations are designated by a starting column addresses (A3C–A7C) and the bank select. The single 32-bit data value comes from the color register which must have been previously loaded. This is illustrated in Figure 7.17.

In this figure we see that a "1" from the least significant bit of the color register (LC) is stored in each of the eight column locations in the DQ0 array. Similarly a "0" from the next bit of the color register is stored in the DQ1 array, etc.

The data from the color register is masked by the data from the mask register where "1" in the mask register enables the data from the color register and "0" in the mask register disables it. For example the "0" in the LM+2 location of the mask register disables the "1" applied from the LC+2 location of the color register so that the column-byte in DQ2 has a "0" written as shown in Figure 7.16.

The data on the DQs when the Block Write command occurs can be used to mask specific column-byte combinations within the block. DQM signals are applied to all eight columns for each byte DQM (0–3). This is sometimes referred to as a "column data mask". For example, in Figure 7.16, the CM+1 bit of the column of data stored in DQ0 is masked by the column data mask so the "1" written here from the color register does not appear.

Figure 7.18 illustrates the effect of the Mask Register, Color Register and Column Data Mask for columns 0–7, the lower byte.

A timing diagram for a page mode block write for CAS Latency 3 is shown in Figure 7.19. In cycle 1, Bank 1 in activated. In cycle 3, a Block Write is activated for Bank 1. The Bank Select pin (BS) is high indicating bank selected and A8 is low indicating Bank 1. DSF is high activating the color and mask registers for the I/O (DQ). Column addresses A3–A7 select the first block of eight columns to be written out of 32 blocks in the row. The DQM provide the mask for the column byte input on the DQ.

7.7 Clock Enable on SGRAM

The Clock Enable (CKE) truth table is unchanged from the SDRAM. It indicates the logic state of CKE at clock edge n and the state at the immediately prior state $n-1$. This truth table was shown previously in Figure 6.24.

7.8 Current State Truth Table on SGRAM

While the Current State Truth Table was included in the Command Truth Table, it is broken out in Figure 7.20 for closer examination. The current state is the state of the SGRAM immediately prior to clock edge n.

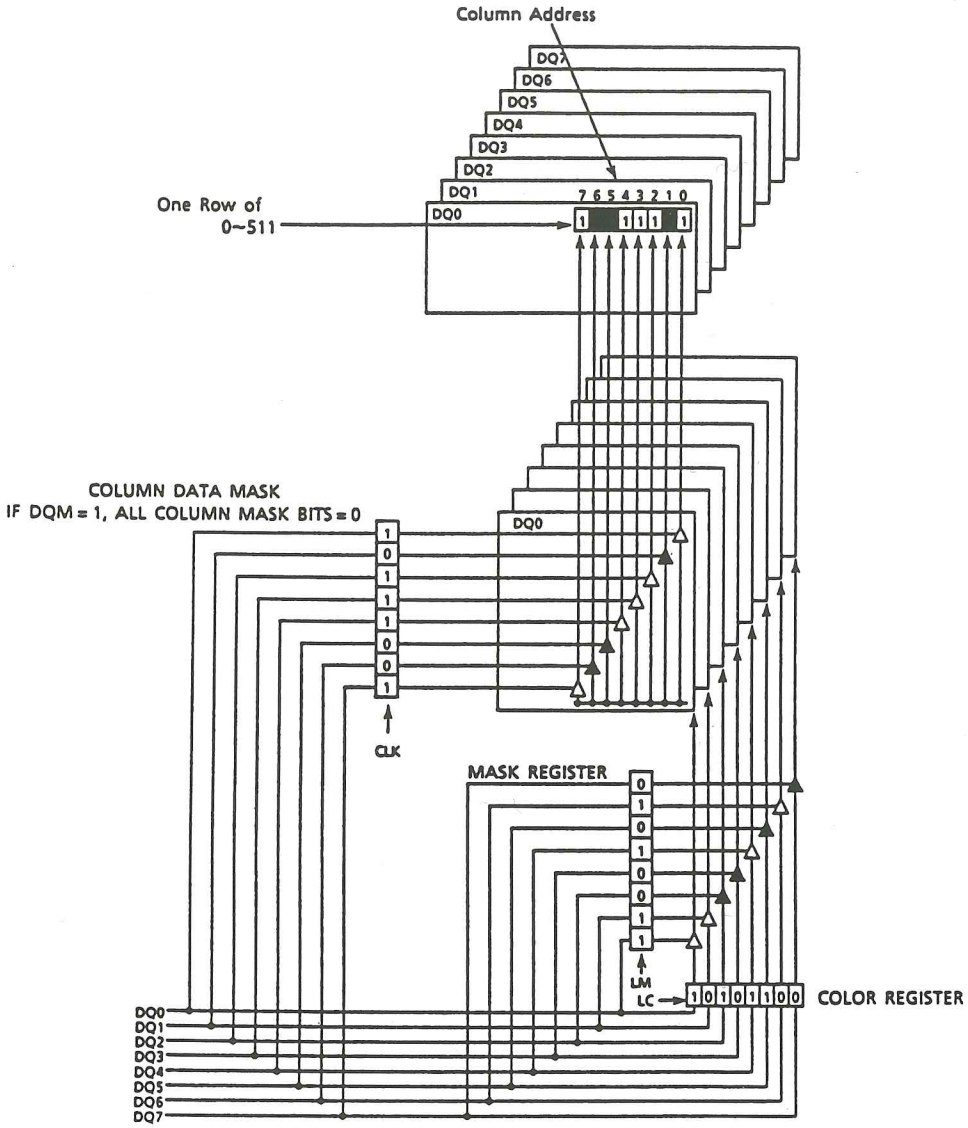


Figure 7.17 Illustration of the effect of a block write command (source: Toshiba)

The current state describes the commands issued to a specific bank. An idle bank is one that has been precharged with t_{RP} met. "Row Active" means the bank has been activated and t_{RCD} met, but no accesses are in progress. "Read" or "Write" means a read or write burst has been initiated but has not yet terminated.

The time for the underlying DRAM to perform the action initiated must be met in all cases. For example, "Precharge" starts with the clocking in of the precharge command and ends after time t_{RP}. So when activating a row, t_{RCD} must be met,

	DQ							
	1	2	3	4	5	6	7	8
Mask Register (LM)	1	1	⓪	⓪	1	⓪	1	⓪
Color Register (LC)	1	0	1	0	1	1	0	0
Column Data Mask (CM)	1	⓪	1	1	1	⓪	⓪	1

	DQ							
	1	2	3	4	5	6	7	8
Column 0	1	0			1		0	
Column 1								
Column 2	1	0			1		0	
Column 3	1	0			1		0	
Column 4	1	0			1		0	
Column 5								
Column 6								
Column 7	1	0			1		0	

Figure 7.18 Lower byte mask/color registers and column mask (source: Toshiba)

and when refreshing t_{RP} must be met. The "Burst Terminate" command is not bank specific but affects any Read or Write burst in progress in either bank.

7.9 Other Single Port Graphics DRAMs

Various other single port DRAMs which are primarily targeted at graphics subsystems have been announced. They include parts from Mosys, Rambus, Hitachi, Mitsubishi, and Neomagic among others. Brief descriptions of these alternative single port DRAMs follow.

7.9.1 Syncronous protocol DRAM from Rambus, Inc.

A synchronous DRAM type developed by Rambus, Inc. has been made in various densities from 4Mb to 18Mb.

This DRAM has a proprietary interface which permits data transfer on a ×8 or ×9 wire bus at a peak transfer rate of 500MB/sec per DRAM [7]. Multiple DRAMs can be on a single bus [19]. Each acts as a slave in responding to bus transactions initiated by a proprietary master device which is a specialized ASIC. The parts are in vertical

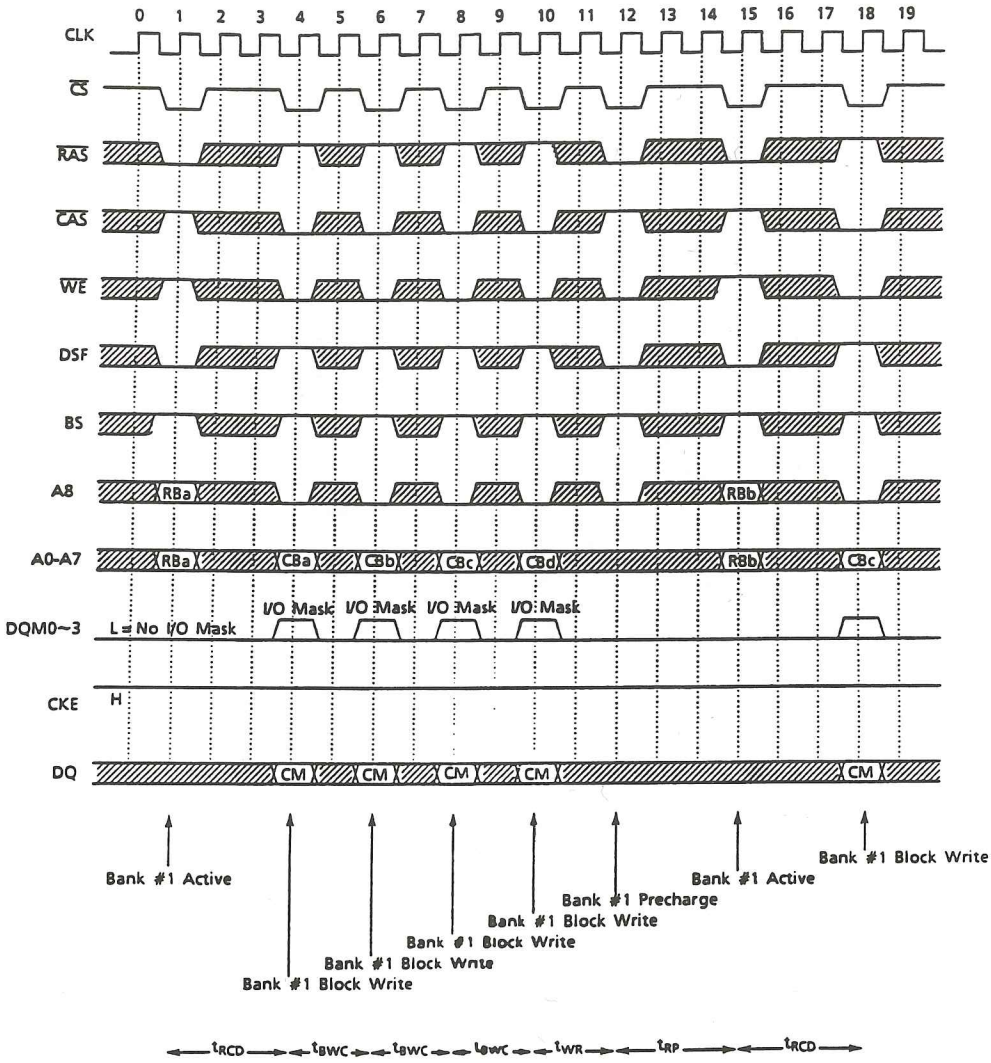


Figure 7.19 Timing diagram for a page mode block write (source: Toshiba)

single-in-line or in surface mount packages which can be lined up on a bus, or "channel". The channel consists of 29 wires, as shown in Figure 7.21 including 8 or 9 data, 8 ground, 5 power, 2 clock, a VREF and various control pins [28].

The "master" controller chip sends out request packets to all the "slave" DRAMs on a single bus channel. All the DRAMs on a channel receive the request packet which contains the address and control information specifying the requested operation. Each DRAM examines the packet. The one which is addressed sends back either an "acknowledge" or a "busy" signal, called a NACK, to the controller.

If the DRAM which is addressed is busy, then it sends back a "NACK" signal to the controller and the controller tries the access again later. If the addressed DRAM is

CURRENT STATE	$\overline{CS}$	$\overline{RAS}$	$\overline{CAS}$	$\overline{WE}$	DSF	COMMAND/ACTION
Any	H	X	X	X	X	COMMAND INHIBIT (NOP/ continue previous operation)
	L	H	H	H	L	NO OPERATION (NOP/ continue previous operation)
Idle	L	L	H	H	L	ACTIVE (Select bank and activate row)
	L	L	H	H	H	ACTIVE w/WPB (Select bank, activate row and WPB)
	L	L	L	H	L	AUTO REFRESH
	L	L	L	L	L	LOAD MODE REGISTER
	L	L	L	L	H	LOAD SPECIAL MODE REGISTER
Row Active	L	H	L	H	L	READ (Select bank and column and start READ burst)
	L	H	L	L	L	WRITE (Select bank and column and start WRITE burst)
	L	H	L	L	H	BLOCK WRITE (Select bank & column and start BLOCK WRITE access)
	L	L	H	L	L	PRECHARGE (Deactivate row in bank or banks)
	L	L	L	L	H	LOAD SPECIAL MODE REGISTER
READ (AUTO-PRECHARGE DISABLED)	L	H	L	H	L	READ (Select bank and column and start new READ burst)
	L	H	L	L	L	WRITE (Select bank and column and start WRITE burst)
	L	H	L	L	H	BLOCK WRITE (Select bank & column and start BLOCK WRITE access)
	L	L	H	L	L	PRECHARGE (Truncate READ burst, start precharge)
	L	H	H	L	L	BURST TERMINATE
WRITE (AUTO-PRECHARGE DISABLED)	L	H	L	H	L	READ (Select bank and column and start READ burst)
	L	H	L	L	L	WRITE (Select bank and column and start new WRITE burst)
	L	H	L	L	H	BLOCK WRITE (Select bank & column and start BLOCK WRITE access)
	L	L	H	L	L	PRECHARGE (Truncate WRITE burst, start precharge)
	L	H	H	L	L	BURST TERMINATE

Figure 7.20 Current state truth table (source: Micron Technology)

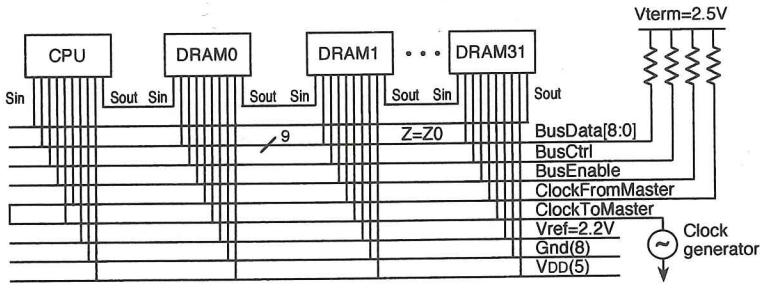


Figure 7.21 Illustration of the configuration of a 4.5Mb 500Mb/sec protocol DRAM (source: N. Kushlyama (1993) [28] with permission of IEEE)

not busy, it returns an acknowledge signal then, the read or write transaction that was requested proceeds.

If the DRAM is available and a write was requested, the master follows the request packet with a data packet which is written into the sense amplifiers if the row corresponding to that data is currently on the sense amplifiers. This is called a write hit and is shown in the timing diagram in Figure 7.22(c) [28].

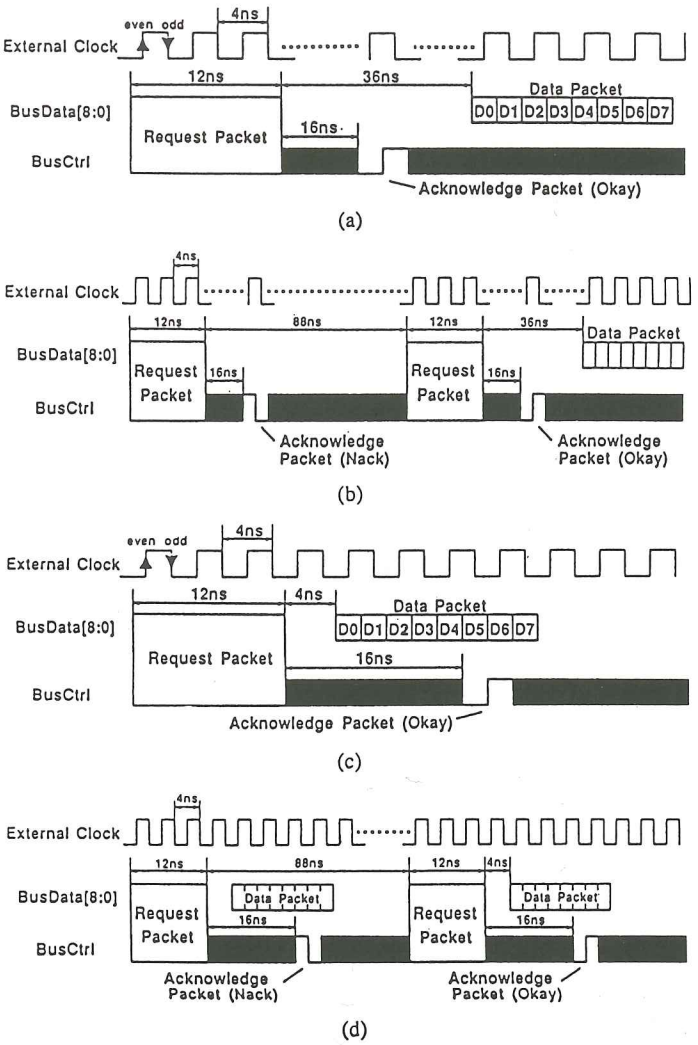


Figure 7.22 Timing diagrams for the 4.5Mb synchronous DRAM from Rambus, Inc.: (a) read hit; (b) read miss; (c) write hit; (d) write miss (source: N. Kushlyama (1993) [28] with permission of IEEE)

If the DRAM is available, a read is requested, and the data stored at the address requested is present on the sense amplifiers; this is called a "read hit" and is shown in the timing diagram in Figure 7.22(a). The DRAM sends the requested data on to the master over the special channel. The master then converts the high speed data on the channel back to the standard interface used in the system. In cases where the DRAM is available and there is a "hit" the datarate can be as high as 500MB/sec for the burst of data accessed. This figure is for the burst and does not allow for bus protocol overhead and initial latency.

If the DRAM is available when a read or write is requested but the correct row address is not currently on the sense amplifiers then there is a "miss". The DRAM proceeds to access the addressed row and when the data is on the sense amplifier either sends it to the master (in the case of a read), or writes the data packet to the sense amplifier (in the case of a write). Figure 7.22(b) shows the case of a DRAM available with a read miss and access, and Figure 7.22(d) shows the case of a DRAM available with a write miss and access. The penalty for a miss is the need to incur the RAS\ access time of the DRAM to access the requested row. For misses, initial access was 152 ns for a read or 120 ns for a write for the early 18Mb density part [19].

In the case where the DRAM is not available when first requested and then a miss occurs when it is available, the resulting access time can be significantly delayed.

Since standard DRAM technology underlies the part, there is an initial latency before the peak transfer rate is achieved during which the various DRAMs are accessed and the sense amplifiers filled.

The 500MB/sec peak transfer rate during a burst is equivalent to that which could be achieved with four 60 MHz 1M×16 SDRAMs on a 64-bit bus, or two 120 MHz 1M×16 SDRAMs on a 3-bit bus. For the same datarate, smaller granularity is therefore possible with this DRAM.

This DRAM has a proprietary synchronous protocol for fast block-oriented transfers and a proprietary low signal swing interface to the proprietary bus. There are on-chip registers which permit flexible addressing control. There are two cache lines per DRAM with each cache line being 1KB each. The 18M DRAM has on-chip RAS\, CAS\ and refresh logic. It is packaged in a proprietary vertically oriented 32-pin single-in-line surface mount plastic package [7].

New pins on the chip include the bus data pins for request, write and read protocols. The data lines, which are active low, carry the request packet with the address, operation codes and the count of the bytes to be transferred. The receive clock (RxClk) is aligned with incoming request and write data packets. The transmit clock (TxClk) is aligned with the data being sent out on reads and in acknowledge packets. The reference voltage (V_{Ref}) is the logic threshold voltage for the low swing signals. The bus control is a control signal to frame packets, transmit opcodes and to acknowledge requests. The bus enable is a control signal which is pulsed to power-up the bus. The daisy chain input (S_{In}) and output (S_{Out}) pins are used to reset the daisy chain input and output. They are active high [7].

A graphics controller for this DRAM has been designed by Cirrus Logic and may result in this part being used in some 1MB frame buffer graphics applications in PCs [20].

7.9.2 *Multiple Bank DRAM from Mosys*

On all of the DRAMs considered to this point, latency for the first access has been a problem. Column access time for data already on the sense amplifiers of the single port DRAMs has been getting faster through the use of wider, faster interfaces and faster access modes.

The SDRAM and the Rambus, Inc. DRAM introduced the concept of reducing the latency for the first access by interleaving two banks. If both banks already have a

row of data active, and the burst that the processor wants next is on one of these rows, then there is no delay. If the requested row is not active, then the latency for the first access may be only partially hidden.

The addition of graphics features integrated into the DRAM has provided some help with the graphics operations requiring random bit accessability.

A DRAM from Mosys, Inc. has attempted to solve these problems by integrating many small DRAM banks onto a single chip all connected to a fast common bus internal to the chip and controlled on chip for clock skew.

The DRAM RAS\ latency problem is reduced by the small size of the DRAM banks and the column latency is reduced by the very short CAS lines to the bus. A schematic block diagram is shown in Figure 7.23.

The interface is synchronous. All signals enter and exit the chip at the control end.

Speed is gained because a small DRAM is inherently faster than a larger DRAM in the same technology due to reduced wordline capacitance and shorter internal wiring. For example, the organization of the banks of the ×32 Mosys DRAM is 256 × 32 × 32 compared with the organization of a 16M ×16 DRAM which is 1024 × 1024 × 16. The 256 × 32 banks used in the Mosys DRAM in a 16M technology are individually faster than 2048 x 1024 bank of the 16M DRAM.

The RAS access time of this multi-bank DRAM in a 0.5 micron technology is 36 ns, CAS access time is 12 ns and there is a 6 ns burst cycle time (166 MHz). The external bus is 16 bits wide. Both clock edges are used for synchronous transfer of 32 bits per clock cycle, so peak bandwidth for burst accesses is 668MB/sec.

The speed is maintained by using a fast metal bus across the chip. Clock skew is avoided on the chip by having the internal bus terminated on the I/O side of the chip. Byte level write mask capability is provided.

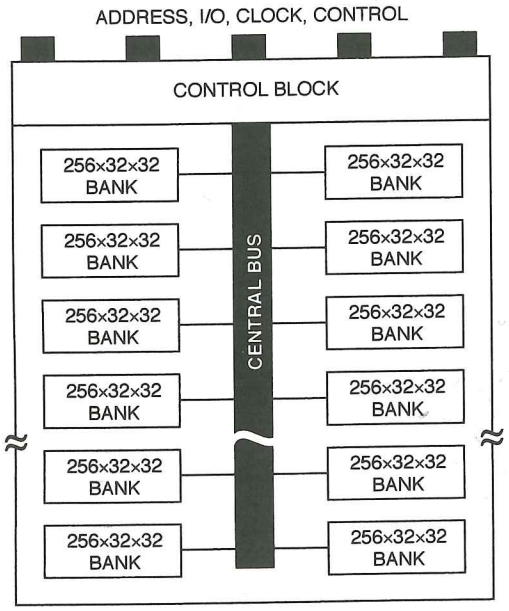


Figure 7.23 Block diagram of Mosys DRAM chip concept (source: Mosys)

Not only are the individual banks faster, but the existence of more than the two banks present in the 16M SGRAM and the Rambus, Inc. DRAM increases the probability of a hit on an accessed row, thereby reducing the frequency of occurrence of time-consuming row accesses. This increases the average bandwidth of the chip.

Improvements in average bandwidth of the multibank DRAM over the two-bank SDRAM and the one-bank EDO DRAM are shown in Figure 7.24.

The Mosys DRAM up to 576K×32 is offered in a 128-pin plastic QFP or a 50-pin PLCC.

Mosys attempted to minimize the problem that all applications-specific DRAMs have – increased silicon area and hence cost.

First, they noted that the various frame buffer applications always waste memory since frame resolutions and DRAM densities never quite match. For example a 1024×768×16 screen resolution requires 12.39Mbits of memory storage. If a 16M DRAM is used then 3.71Mbits of memory are wasted, that is, the 16M is 25 percent too large for this application.

If a Mosys DRAM with 50 256K banks is used, the amount of data storage is 12.8M. This is enough to service the basic application and still permit the chip to remain within the chip size of the 16M DRAM. One drawback is that graphics subsystems designers frequently find a use for the additional memory.

Secondly, Mosys noted that if a few extra banks are included, then yield can be increased by substituting good redundant banks for defective banks, thereby increasing the overall chip yield. A drawback is the chip size must be larger by the size of the redundant banks. This can be a compelling argument in the early years of a high density DRAM when the production yields are low. It may be less so when the technology is mature and the basic yields are high.

The main benefits of this part are expected to be low pin count, minimum memory configuration, and the low latency.

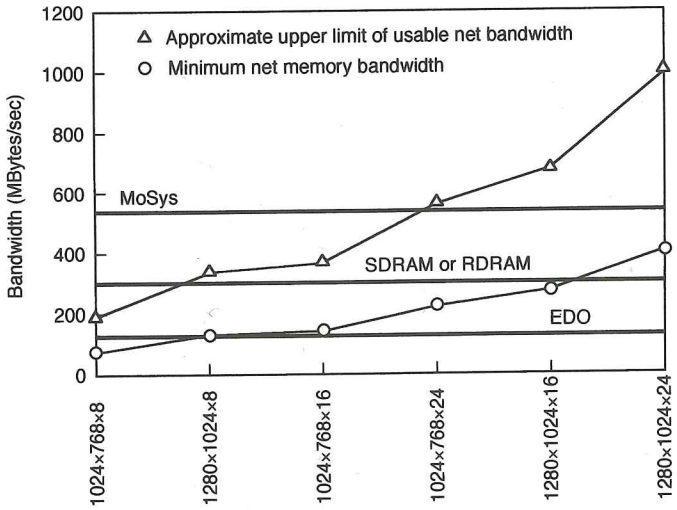


Figure 7.24 System bandwidth requirement and minimum net bandwidth by DRAM type (source: Mosys)

Notice in Figure 7.24 that even the bandwidth of the Mosys DRAM is not claimed to service frame buffers with resolutions of 1280 × 1024 × 16 and higher. These higher end systems fall into the workstation category and have historically been serviced by the multiport video DRAM (VRAM) or by large banks of interleaved DRAMs.

The requirement for graphics buffers with bandwidth beyond that possible with single port DRAMs still exists. It appears to be moving in two directions, those systems continuing to use multiport DRAMs, and those considering applications-specific memories.

In the next section multiport DRAMs are discussed, and in the following section some of the new applications-specific DRAMs now appearing are discussed.

7.10 Overview of Multi-Port Graphics DRAMs (VRAMS)

VRAMs attempt to compensate for the slow speed of the DRAM in video display applications by having both a random port and a serial port. The random port operates like a conventional DRAM. The serial port provides continuous refresh to a raster display screen. High bandwidth transport of information occurs between the two memories.

There are three basic VRAM operations. These are asynchronous random read and write access of the parallel RAM port, high speed clocked access of the serial SAM port, and transfer of data between any row in the RAM and the SAM (Serial Access Memory).

The RAM and SAM ports can be independently accessed at any time except during an internal transfer between the two memories. Some VRAMs provide bidirectional transfer of data between the RAM and the SAM, while others provide only for transfer from the RAM to the SAM. All VRAMs have high speed serial read capability and some have serial write capability.

Since the high end graphics market is small, the number of DRAM suppliers who make VRAMs is small. At the 4M density there are at least three suppliers of VRAMs including TI, IBM, and Toshiba.

7.11 An Introduction to VRAMs, the 4M VRAMs

A block diagram of a typical 4M VDRAM [21] which is organized 256K × 16 is shown in Figure 7.25.

Shown are the 256K × 16 RAM and upper and lower transfer gates of the split 256 × 16 Serial Data Register (SAM). The RAM array is organized 512 × 512 × 16. Extended page mode, which is similar to Hyperpage (EDO) mode, is used for RAM accesses. The SAM has serial read and write capability and data can be transferred between any row in the RAM and the SAM. Each half of the split SAM register is 128 × 16 bits. Mask operations are supported by a write-per-bit function, a color register, block write control logic, and flash write control logic.

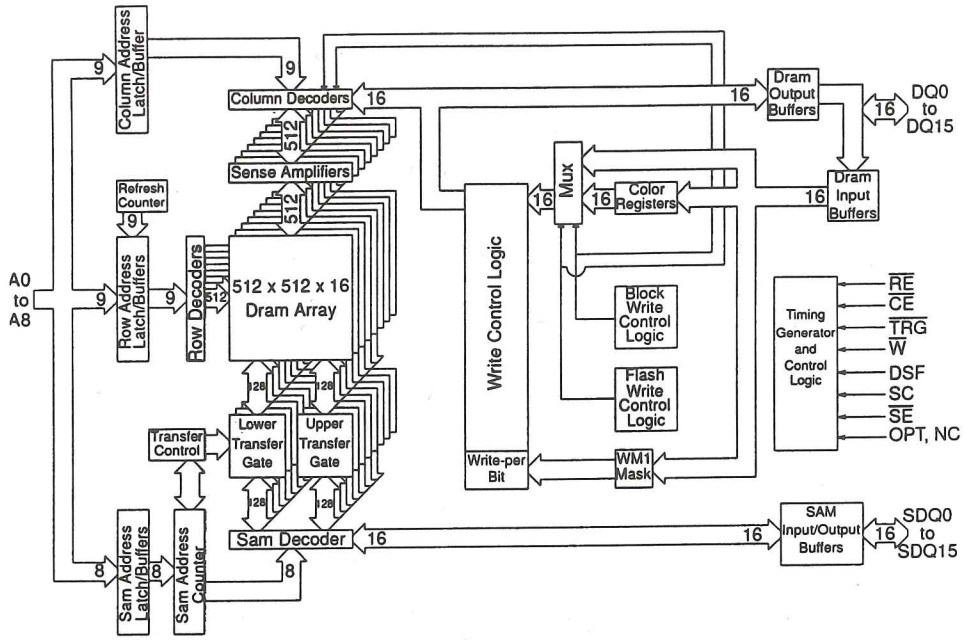


Figure 7.25 Block diagram of a typical 4M VDRAM (source: IBM [21])

A pinout of a 4M VDRAM in a TSOPII and SOG (SSOP) package is shown in Figure 7.26.

The DQ pins are the RAM I/O's, the SQ pins are the SAM outputs, the SC pin is the serial clock and the SE\ is the serial enable.

A TRG\ pin selects either the data transfer operation or the DRAM operation. During DRAM operation, the TRG\ is held high as RAS\ falls and acts as an output enable for the RAM DQ pins. For transfer the TRG\ is held low.

The WE\ pin enables the "write-per-bit" or mask function when it is held low as RAS\ falls, and acts as the Write Enable if held high as RAS\ falls.

The DSF pin is the special function control for the CBR refresh, the split register transfer, and the various mask functions. Mask functions include block write, non-persistent write-per-bit, persistent write-per-bit, and load color register.

The QSF pin is a special flag output pin that indicates which half of the SAM is being accessed. When QSF is low the serial address pointer is accessing the least significant 128 bits of the SAM and when QSF is high it is accessing the most significant 128 bits.

7.12 RAM Operations

Asynchronous RAM port operations include normal random access read, early and late write, extended page mode read and write, read–modify–write, load mask register, and load color register and flash write.

Figure 7.27 shows a functional table for random access operation.

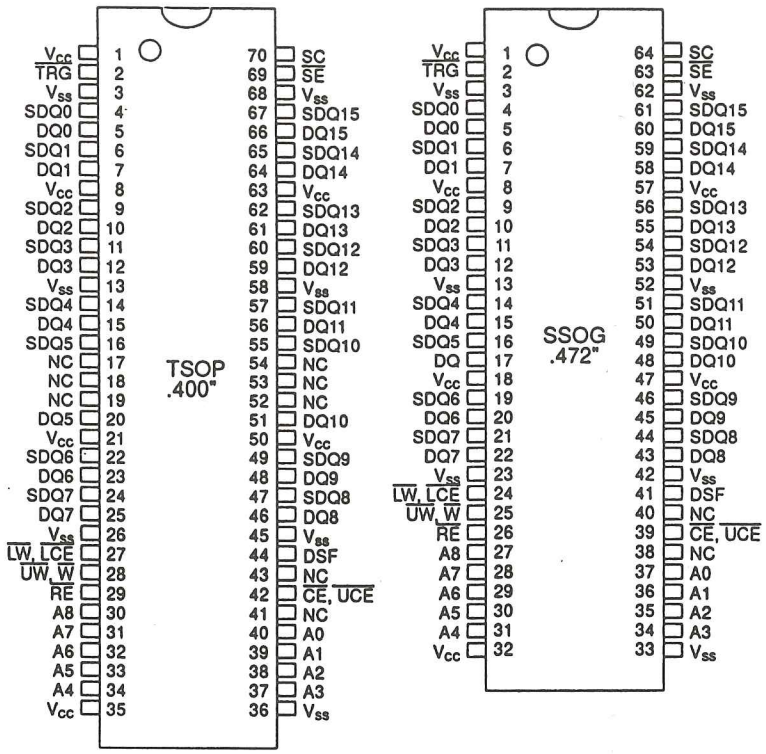


Figure 7.26 Pinout of a typical 4M VDRAM (source: IBM [21])

7.12.1 *Extended Read and Write Mode*

The extended read and write mode functions are similar to the Hyperpage mode on the single port RAM in that the rising edge of CAS\ no longer controls the outputs during a burst. This is illustrated in the timing diagram for fast page mode read cycle shown in Figure 7.28.

4M VRAMs have extended read and write mode functions on the RAM port that run as fast as 25 ns cycle time [21].

7.12.2 *Random Port Mask Functions*

The mask functions are similar to those discussed for the wide asynchronous DRAMs with graphics features. The difference between non-persistent write-per-bit and persistent write-per-bit follows.

Non-persistent write-per-bit (also called "new mask mode") permits a mask to be loaded via the 16 DQs on the falling edge of RAS\ on any write cycle. After the mask is latched, valid input on the DQ pins will be masked for that cycle only.

Persistent write-per-bit (also called "old mask mode") requires a "load mask register" cycle to load the mask into the internal mask register. This valid input data is

Menu Code	$\overline{RE}$				$\overline{CE}$	Address		DQ_i			FUNCTION
	$\overline{CE}$	$\overline{TRG}$	$\overline{W}$	DSF	DSF	$\overline{RE}$	$\overline{CE}$	$\overline{RE}$	$\overline{CE}$	$\overline{CE}$, $\overline{W}$	
CBR	0(5)	X	1(4)	0	-	X	-	X	-	-	$\overline{CE}$ before $\overline{RE}$ Refresh
CBRS	0(5)	X	0(3)	1	-	Stop	-	X	-	-	$\overline{CE}$ Before $\overline{RE}$ Refresh and mode set (2)
CBRN	0(5)	X	1(4)	1	-	X	-	X	-	-	$\overline{CE}$ Before $\overline{RE}$ Refresh without mode reset
ROR	1	1	1	X	-	Row(1)	-	X	-	-	$\overline{RE}$ Only Refresh
LCR	1	1	1(4)	1	1	Row(1)	X	X	X	Color	Load Color Register
LMR	1	1	1(4)	1	0	Row(1)	X	X	X	Mask	Load Mask Register
RT	1	0	1(4)	0	X	Row	TAP	X	X	X	Read Transfer
MWT	1	0	0(3)	0	X	Row	TAP	WPBM	X	0	Write Transfer (Masked)
SRT	1	0	1(4)	1	X	Row	TAP	X	X	X	Split Read Transfer
MSWT	1	0	0(3)	1	X	Row	TAP	WPBM	X	X	Split Write Transfer (Masked)
RW	1	1	1(4)	0	0	Row	COL	X	X	Data Input	Read Write Cycle (No Mask)
RWM	1	1	0(3)	0	0	Row	COL	WPBM	X	Data Input	Read Write Cycle (Masked)
BW	1	1	1(4)	0	1	Row	COL A3-A8	X	-	ADDR mask	Block Write Cycle (No Mask)
BWM	1	1	0(3)	0	1	Row	COL A3-A8	WPBM	X	ADDR mask	Block Write Cycle (Masked)
FWM	1	1	0(3)	1	X	Row	X	WPBM	X	X	Flash Write Cycle (Masked)

Notes:

1. Row address needed only for refresh operation to the selected row. Otherwise this is a don't care.
2. This cycle is used to put the chip into special modes. The A_i at $\overline{RE}$ fall select the desired mode.
3. Either $\overline{W}$ is 0.
4. Both $\overline{W}$ are 1.
5. Either $\overline{CE}$ is 0 on Dual CE parts.

Figure 7.27 4M VRAM function table for RAM operation (source: IBM [21])

then masked from the internal register and any inputs on the DQ pins on the falling edge of RAS\ are ignored. A CBR\ refresh cycle is required to reset this cycle.

The block write is also similar to that discussed for the single port DRAMs. Data from the color register is used to write up to 64 bits of data simultaneously to one block of the memory array. The block is implemented as 4 columns $\times$ 4 DQs for some 4M VRAMs and as 8 columns $\times$ 8 DQs or 8 $\times$ 16 for others.

Either persistent or non-persistent write-per-bit can be applied to the block write cycle.

Figure 7.29 shows a timing diagram for a load-color-register cycle, followed by a block-write cycle with no write mask, and a block-write cycle where the mask is loaded and used [22].

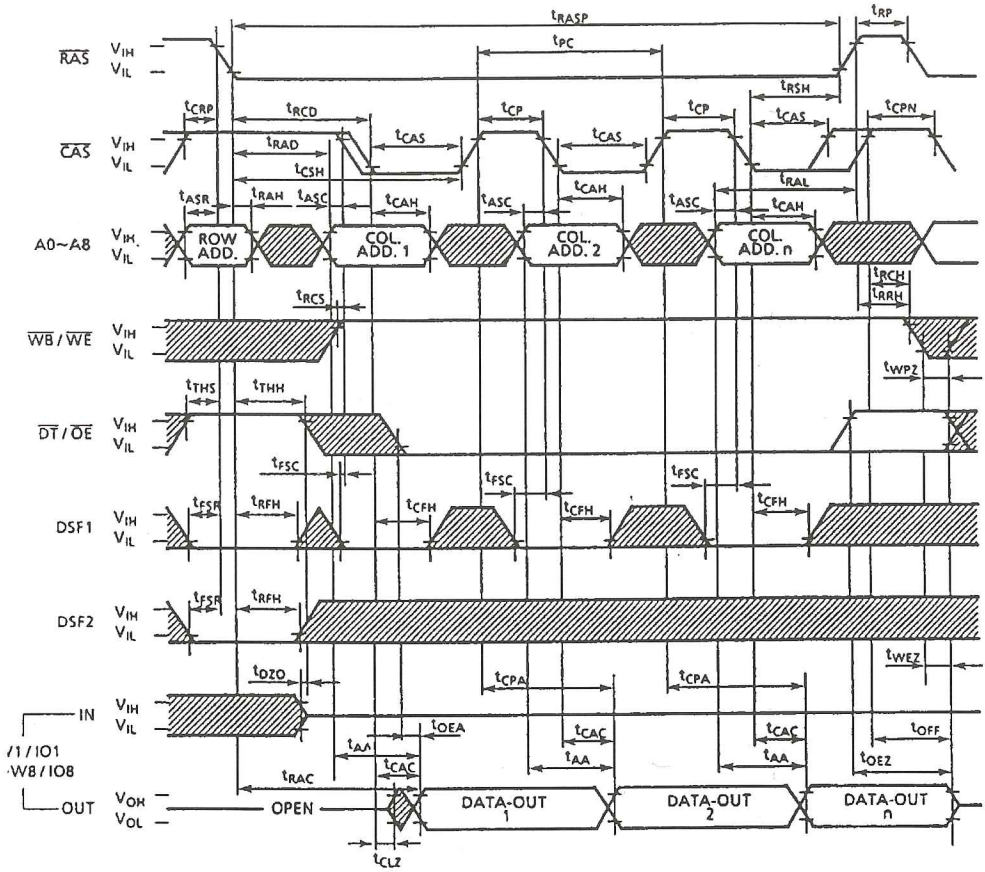
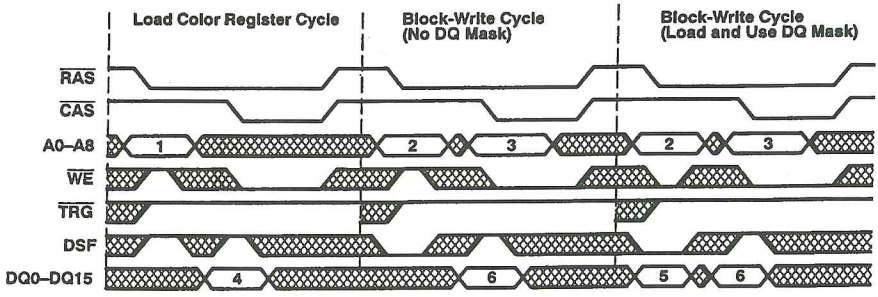


Figure 7.28 4M VRAM extended fast page mode read cycle (source: Toshiba [2])



Legend:
1. Refresh Address
2. Row Address
3. Block Address (A2–A8)
4. Color Register Data
5. DQ Mask Data: DQ0–DQ15 are latched on $\overline{RAS}$ falling edge.
6. Column Mask Data: DQ_i–DQ_{i+3} (i=0,4,8,12) are latched on either the first $\overline{WE}$ falling edge or the falling edge of $\overline{CAS}$, whichever occurs later.

= Don't Care

Figure 7.29 4M VRAM timing diagram showing a masked block-write cycle (source: Texas Instruments [22])

7.12.3 Flash Write

Flash Write is available on some of the 4M VRAMs. It is a special RAM cycle which lets the data in the color register be written into all the memory locations of a selected row. Each bit of the color register corresponds to one of the RAM I/O blocks, so a Flash Write cycle writes to a plane of the RAM array corresponding all the I/O locations in a single row. It can be masked by a write-per-bit function on an I/O basis. Figure 7.30(a) illustrates schematically an array plane written by a single masked Flash Write cycle (lower byte only) and Figure 7.30(b) shows a simplified Flash Write timing diagram.

Flash Write is used for fast plane clear operations. If a different row address is specified for each Flash Write cycle, the entire array can be cleared in 512 Flash Write cycles. Assuming a cycle time of 130 ns, a plane clear operation can be completed in less than 66.6 microseconds.

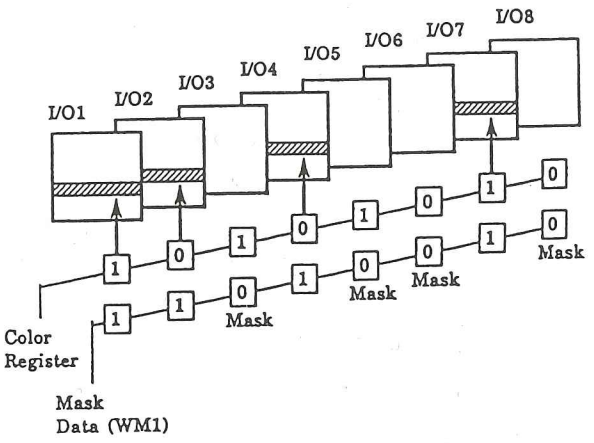


Figure 7.30 Illustration of a single Flash Write operation: (a) array plane for single masked Flash Write cycle; (b) simplified Flash Write timing diagram (source: Toshiba [2])

7.13 Transfer Operations between the RAM and SAM

For transfer operations between the RAM and the SAM, there are two cases, one for the 256 × 16 SAM, and one for the 512 × 16 SAM.

7.13.1 256 × 16 SAM

The SAM illustrated in the block diagram in Figure 7.25 is 256 × 16 bits and one row of the RAM array is 512 × 16 bits. During a transfer operation from the RAM to the SAM, the 256 × 16 bits from half of a 512 × 16 bit row in the RAM are transferred to the SAM.

The transfer can be done either as a full register operation or as a split register operation.

For a full register transfer read operation, nine row addresses (A0–A8) are latched at the falling edge of RAS to select one of the 512 rows for transfer. Address A8 selects which half of the row is to be transferred. A0–A7 select one of the SAM's 256 available tap points from which the serial data is read out. A full register transfer read is illustrated in Figure 7.31(a).

In a split register transfer read operation, the serial data register is split in half as illustrated in Figure 7.32(b). The lower half contains bits 0 to 127 and the upper half contains bits 128 to 255. The advantage of the split register operation is that while one half of the SAM data is being read out of the SAM port, the other half can be loaded from the memory array. Split register transfer provides the means for continuous loading of the SAM without interrupting the continuous flow of data from the SAM port by alternating the halves of the SAM that are being loaded and read.

7.13.2 512 × 16 SAM

The 4M VRAMs having a 512 × 16 SAM load an entire 512 × 16 row of the RAM into the SAM during a transfer read operation as illustrated in Figure 7.32(c). A split register transfer read in this case is from one half of the RAM row to one half of the SAM register as ilustrated in Figure 7.32(d).

Figure 7.32 shows the timing diagram for split register transfer read timing for a 256 × 16-bit SAM with Tap Point "M" indicated by addresses A0–A6, A7 is "don't care", and A8 identifies the specified half of the row.

Pointer control permits definition of the starting locations of the serial port in a multiport DRAM and simplifies the control logic required for scrolling and hardware windows. The tap pointer is a counter that defines the starting point in the serial data register into which data is entered. The data is entered serially and wraps around when the end of the register is reached.

7.14 SAM Operation

Data is shifted out of the SAM registers at the rising edge of the serial clock (SC) with the Serial Enable (SE\) held low.

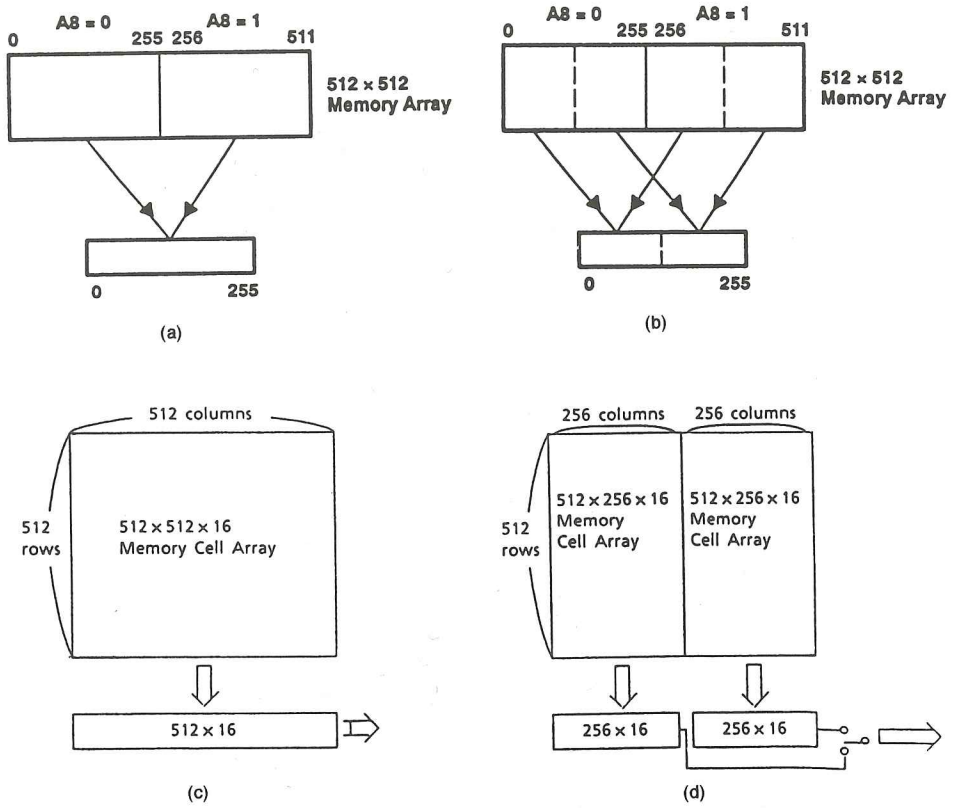


Figure 7.31 Register transfer read (RAM to SAM): (a) 256-bit SAM – full register transfer (b) 256-bit SAM – split register transfer (source: TI [18]; (c) 512-bit SAM – full register transfer; (d) 512-bit SAM – split register transfer (source: Toshiba [2])

The timing diagram for a simple serial read cycle is shown in Figure 7.33.

The TRG\ is "don't care" except when RAS\ falls when it must be high to avoid initiation of a register-data transfer operation.

The SAM of a 4M VRAM can run as fast as 66 MHz [21].

7.15 Video DRAM Standards and Market

The JEDEC JC42.3 Standards Committee has specified mandatory and optional features for the 2M and 4M VRAM.

Features which are required on all 2M and 4M VDRAMs which claim to be JEDEC standard include read transfer cycle, split read transfer cycle, read write cycle (non-mask), read write cycle with new mask data, and CAS\ before RAS\ refresh.

Mandatory optional features of which, if any of the features are used, all must be used, include CBR refresh/stop, mask write transfer, block write with no mask, block write cycle, flash write with mask, and load color register.

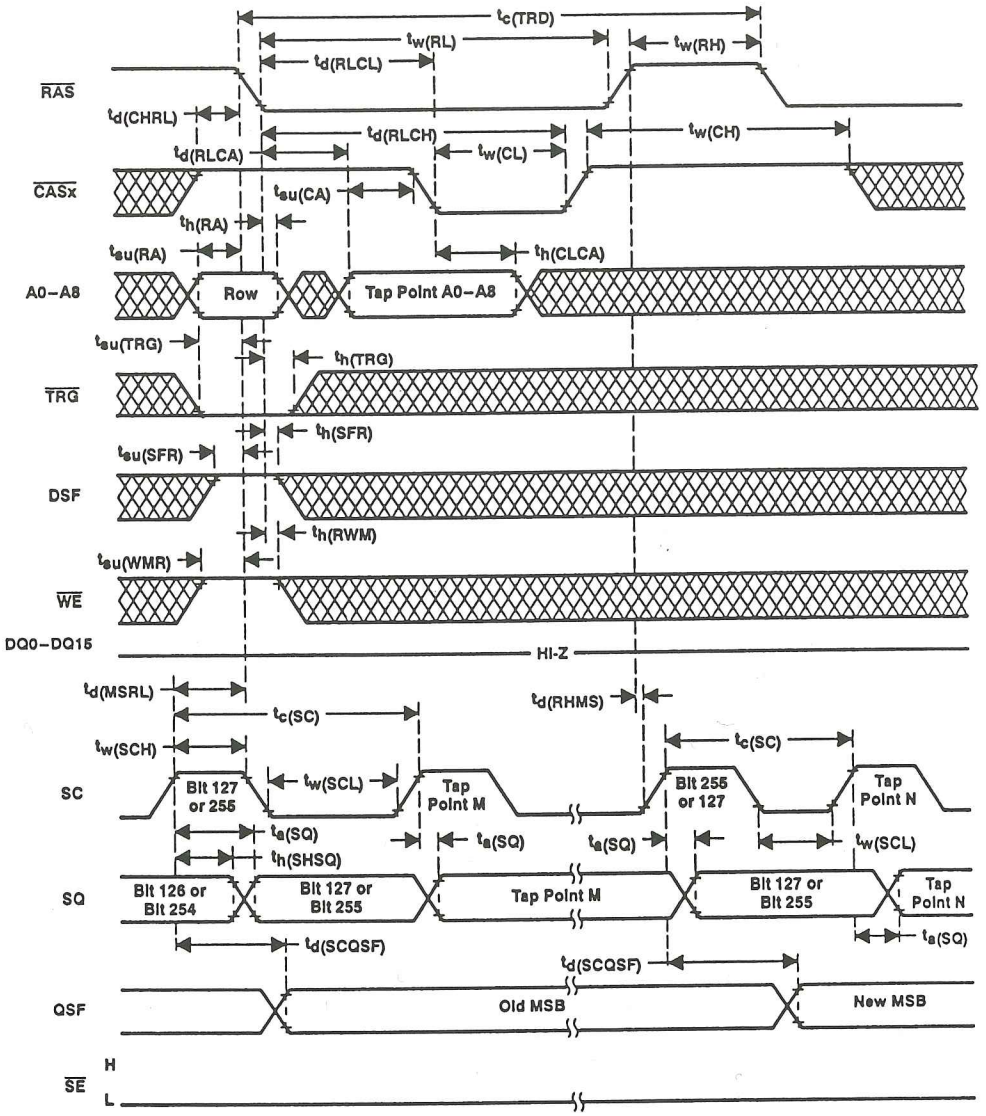


Figure 7.32 Split register transfer read timing (source: Texas Instruments [18])

Beyond the mandatory optional features, there are optional features which are at the discretion of the manufacturer. These include split write transfer, load mask register, pipelined fast page mode and extended out fast page mode, and a logic function set cycle.

The major drawback of the multiport RAMs has been the additional silicon, and hence cost, they incur over the silicon of the standard DRAM. The VRAM size has ranged up to 40 percent over that of a comparable DRAM. The additional features were a major part of this size increase.

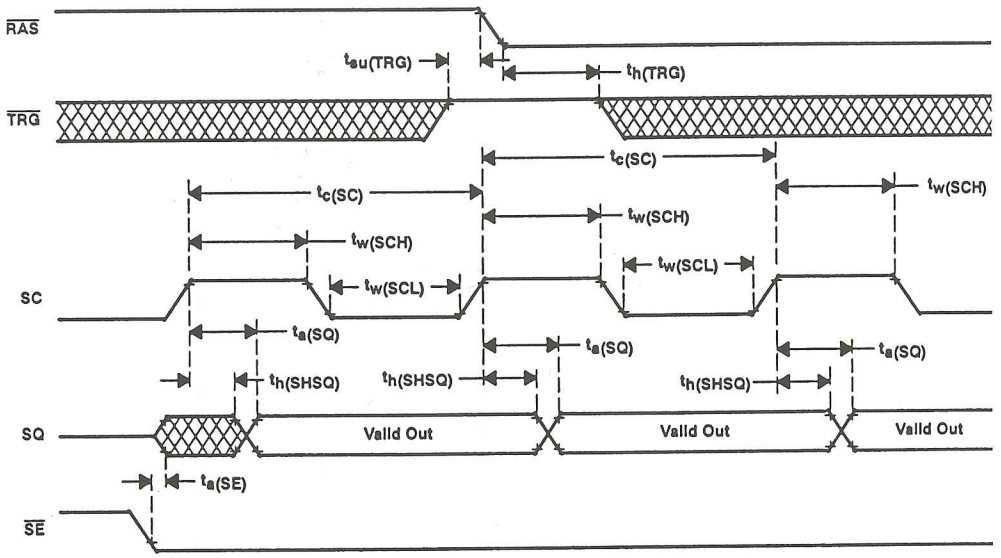


Figure 7.33 4M VRAM serial read operation (source: Texas Instruments [18])

In response to the competition from high bandwidth single port DRAMs, VRAMs have been simplified so that the silicon adder is now normally between 7 and 15 percent. An example of such simplification is going from a 512-bit to a 256-bit SAM. At a lower cost the VRAMs still provide higher bandwidth in a high end frame buffer environment than a single port DRAM.

7.16 8M Video DRAM

7.16.1 Samsung "Window RAM"

An example of a simplifed VRAM is the 8M VRAM from Samsung [23] which they call the "Window RAM". This has been the only 8M density VRAM to appear to date. A block diagram is shown in Figure 7.34.

The 256K×32 RAM consists of 32 I/O blocks of 512 × 512 arrays. The split SAM is two small 64 × 16 SAMs in parallel with 2:1 multiplexed 16-bit serial output for speed.

A minimum number of features are included. Mask function is handled with an eight column block write which allows a 32-byte block of data to be transferred at one time. There is a two-bit color register and bit and byte mask capability. Scroll and "aligned block move" operations, also known as bit block transfer (BitBLT), are done with four 256-bit latches which are used as temporary storage for internal DRAM read cycles. Serial read with split register read transfer is provided. A drawback of this part is that the display refresh overhead is higher than other VRAMs due to the narrower 256-bit load path to SAM.

The part is packaged in a 120-pin PQFP. The pinout is shown in Figure 7.35.

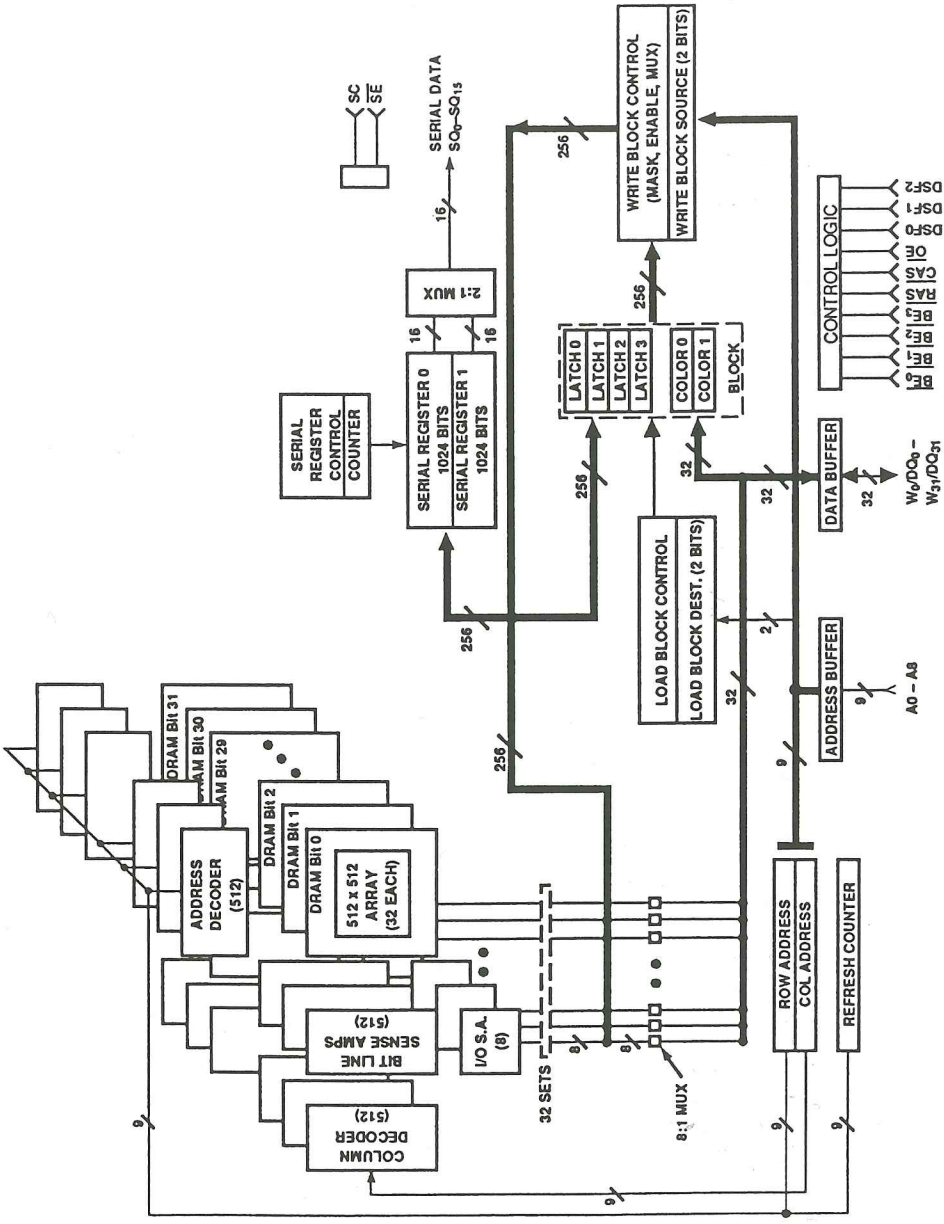


Figure 7.34 8M VRAM ("Window RAM") block diagram (source: Samsung [23])

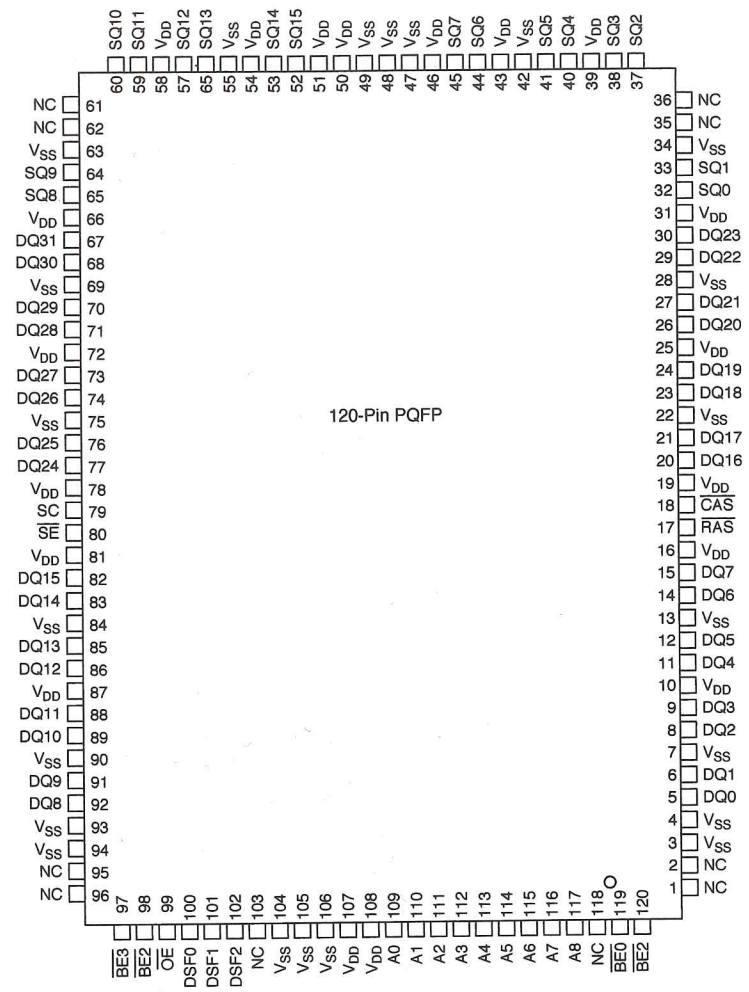


Figure 7.35 Pinout for 8M VRAM (Window RAM) (source: Samsung [23])

The four Byte Enable (BE\) pins select I/O bytes with BE0 corresponding to DQ0–DQ7, BE1 to DQ8–DQ15, etc. BE\ high disables the I/O byte, and BE\ low disables it. Color register selection is made by D(i) with D(i) = 1 selecting color register 0 and D(i) = 2 selecting color register 2. Simplified RAS\ and CAS\ control truth tables are shown in Figure 7.36.

An output enable (OE\) control permits control of the data out during extended data out read and write cycles since the rising edge of CAS\ no longer controls the I/O. During a new row access the state of OE\ when RAS\ falls controls mask register update: if OE\ = 1 the mask register content is updated, if OE\=0 the cycle uses previously loaded mask data. A read–write cycle timing diagram is shown in Figure 7.37.

The color and mask register timing signals are also indicated for masking during the read–write cycle. For further details of timing the reader should refer to the vendor specification for this product.

$\overline{CAS}$ ↓ *2	$\overline{RAS}$ ↓ *1							Mnemonic Code	Function
$\overline{BE}_{3-0}$	$\overline{CAS}$	$\overline{OE}$	DSF2	DSF1	DSF0	RA_{8-0}	W_{31-0}		
X	0	X	X	X	0	X	X	RST	Reset Cycle
X	0	X	X	X	1	X	X	CBR	CBR Refresh
⇑ (Note 2)	1	0/1	X	X	0	Row	WPB Mask	RW	New Row Initiation for any RW Cycle
⇑ (Note 2)	1	0/1	X	X	1	Row	WPB Mask	TEST	Test Cycle

(a)

				$\overline{CAS}$ ↓ * 3										Mnemonic	
$\overline{BE}_{3-0}$ (Note 2)	DSF2	DSF1	DSF0	CA 8	7	6	5	4	3	2	1	0	W_{31}/DQ_{31-} W_0/DQ_0	Code	Function
⇑	1	0	0	X	X	X	X	X	X	X	0/1	0	Pixel Color Data	RCR	Read Color Reg 0 or 1
⇑	1	0	1	X	X	X	X	X	X	X	0/1	0	Pixel Color Data	LCR	Load Color Reg 0 or 1
⇑	1	0	0	X	X	X	X	X	X	X	X	1	Mask Data	RMR	Read Mask Reg
⇑	1	0	1	X	X	X	X	X	X	X	X	1	Mask Data (WPB)	LMR	Load Mask Reg
X	0	0	0	|←Column Address→|					X	0	0		X X X X	UFBR	DRAM to $Latch_0$
				|←Column Address→|					X	0	1		X X X X		DRAM to $Latch_1$
				|←Column Address→|					X	1	0		X X X X		DRAM to $Latch_2$
				|←Column Address→|					X	1	1		X X X X		DRAM to $Latch_3$
⇑	0	1	1	|←Column Address→|					X	0	0			UFBWL	$Latch_0$ to DRAM
				|←Column Address→|					X	0	1				$Latch_1$ to DRAM
				|←Column Address→|					X	1	0		← Byte Mask →		$Latch_2$ to DRAM
				|←Column Address→|					X	1	1				$Latch_3$ to DRAM
⇑	0	0	1	|←Column Address→|					X	0	0		← Byte Mask →	UFBW8	From Color Reg 0 to DRAM
				|←Column Address→|					X	0	1		← Byte Mask →		From Color Reg 1 to DRAM
				|←Column Address→|					X	1	0		←Col Reg Select →		C0 = Fgrd, C1 = Bkgrd to DRAM
				|←Column Address→|					X	1	1		←Col Reg Select →		C0 = Bkgrd, C1 = Fgrd to DRAM
X	0	1	0	|←Column Address→|					X	X	0		X X X X	SRT	Split Read Transfer
X	0	1	0	|←Column Address→|					X	X	1		X X X X	SRTR	Split Read Transfer with SAM Pointer Reset
⇑	1	1	0	← Column Address →									$D_{OUT}(Q_{31-0})$	UFR	Ultra Fast Page Read Cycle
⇑	1	1	1	← Column Address →									$D_{IN}(D_{31-0})$	UFW	Ultra Fast Page Write Cycle

⇑ = Byte Control

(b)

Figure 7.36 Simplified functional truth tables for 8M VRAM: (a) RAS\ control; (b) CAS\ control (source: adapted from Samsung [23])

This 8M VRAM can do graphics at the rate of a 1.6GB/sec fill and 0.64GB/sec aligned Bit BLT. The timing for a 10-pixel vector draw is 4.1M vectors/second and for 7 × 8 character draw the rate is 4.5M characters/second.

The read/write bandwidth of the ×32 RAM with a 20 ns page cycle time is 200MB/sec. The "ultra fast" page mode on this part is similar to an output enable controlled Hyperpage (EDO) mode.

7.17 8M and 16M Synchronous VRAMs

An 8M synchronous VRAM has been standardized by JEDEC, but has not to this point been developed. At the present time any attempt at a multiport synchronous

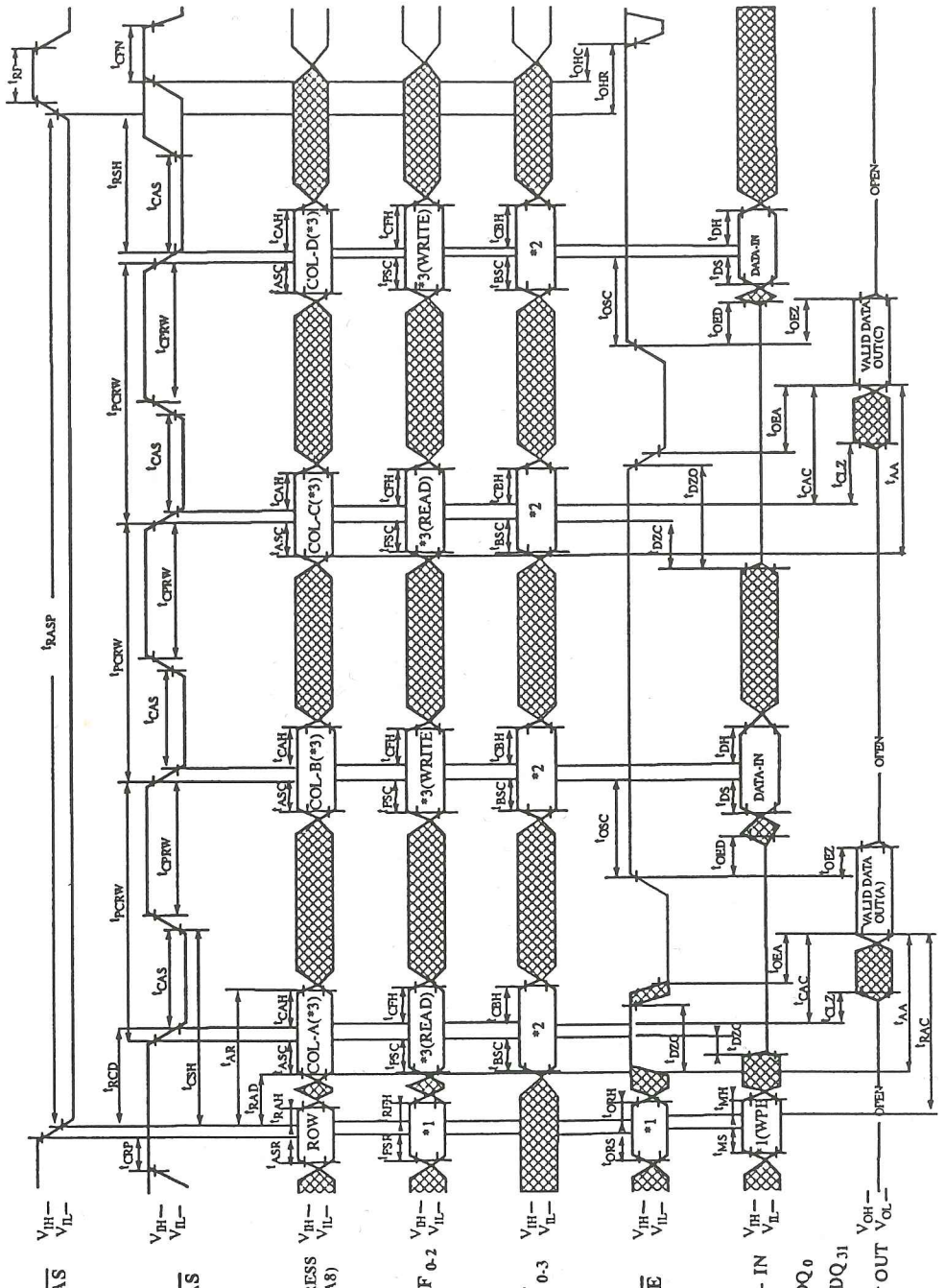


Figure 7.37 Output enable controlled (EDO) read–write cycle (source: Samsung [23])

DRAM is more likely at the 16M density level. Such a part would be expected to continue the trend to faster RAM access and stripped-down features to reduce cost adders over the fast single port DRAMs.

7.18 Triple Port VRAM

A triple port RAM has been offered by Micron [13]. While the part is no longer actively being marketed, it is still in use with a small user base and will be mentioned in passing. This 1M part has a 256K×4 DRAM with 20 ns fast page mode access, and two 512 × 4 bidirectional SAMs. Data can be accessed at the RAM port, at either of the SAM ports, and transferred bidirectionally between the DRAM and either SAM. All three ports can be operated independently except during a transfer of data between ports.

The bidirectional aspect of the RAM and SAM have made this part of interest in data communications and networking applications as well as for high end graphics processing.

7.19 VRAMs with Z-buffers

A few graphics RAMs have been developed for three-dimensional graphics applications. 3D computer graphics has been used primarily in engineering workstations, but is now becoming more popular in PCs for games and virtual reality applications. Two DRAM chips that are targeted at this applications area are described in this section.

7.19.1 3D-RAM

A 3D frame buffer DRAM with very fast rendering to 400 million pixels per second has been developed by Mitsubishi [25] primarily for the engineering workstation environment.

Standard DRAMs or VRAMs have been less than adequate in this application due to the additional bandwidth needed to process the large amount of data required for color and depth. Assuming a graphics workstation has 1280 × 1024 pixel screen resolution plus 24 bits per pixel for color (plus eight for color blend) and 32 bits for depth per pixel, the total bit requirement for a single frame is 10.2MB of memory.

The normal memory operation in 3D pixel rendering is the read–modify–write. This is because the old Z value must first be read. The old and new Z value are compared and the result written back to the memory. Comparison, for example, could entail deciding which data, old or new, is in front in the Z-direction since this is the information which must remain visible.

The 3D RAM combines on one chip a VRAM subsystem consisting of a four-bank DRAM and two SAM registers, a triple ported SRAM cache subsystem, a blend logic

unit, and a compare logic unit. The SAM outputs refresh the screen. The cache subsystem consists of a tag cache, a data cache, and a compare unit.

New data moves from the processor into the compare and blend units. The new data is blended with the data in the cache, the blend is compared to the old data, and the result is transferred to the SRAM cache and thence to the DRAM banks. From the DRAM banks, the new frame is transferred to the SAM register and out to the video display. A 4 × 8 × 8 block transfer is available on the 3D RAM, but it is rare to change all 256 bits in the block in one block write.

A schematic block diagram showing the the 3D RAM internal data path in a "mostly write" situation is shown in Figure 7.38.

The 3D RAM handles data at a rate of 400M-pixels/sec.

While 3D RAM configurations for PC applications have been discussed [26], to date none has been indicated to be in production.

7.20 Integrated Frame Buffers

Another direction that the DRAM frame buffer can go is to fuller integration of the DRAM with the graphics controller and the RAMDAC. One such circuit, from Neomagic, Inc. has been developed [27].

This chip combines a super VGA (SVGA) graphics accelerator, RAMDAC, frequency synthesizer and local bus interface (PCI, VL bus) with an 8Mbit DRAM. The chip takes advantage of the wide internal bus of the DRAM by having a 128-bit datapath between the DRAM and the controller.

Even at a comparatively slow speed, the bandwidth of such a chip can be significant. For example, running at 30 MHz on a 128-bit bus, such a chip would have a datarate of 500MB/sec. This is double the rate of a 66 MHz ×32 SGRAM running at 66 MHz (264MB/sec). The slower speed can also reduce power consumption, making this type of chip satisfactory for portable PC applications.

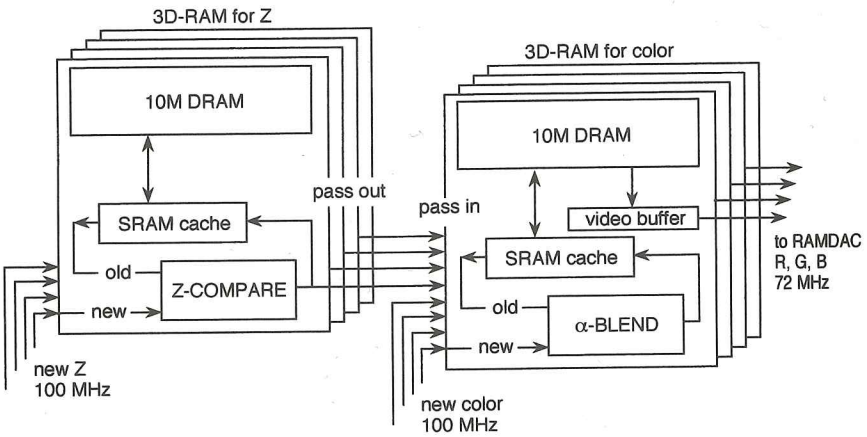


Figure 7.38 Schematic block diagram of 3-D RAM data path (source: Mitsubishi [26])

Bibliography

1. JEDEC Standard 21C.
2. Toshiba, *Application Specific DRAM Data Book*, 1994.
3. Frame buffer architecture, *NEC Memory Products Databook, Volume 1, DRAMs, DRAM Modules, Video RAMs*, 1993.
4. 256K×4 Field Memory, TMS4C1060-30, *Texas Instruments MOS Memory Databook*, 1993.
5. Micron Technology, *8M SGRAM preliminary datasheet.*
6. Toshiba, *8M SGRAM preliminary datasheet.*
7. Toshiba, *Specialty DRAM Databook.*
8. United Memories, *4M SGRAM Preliminary Data Sheet.*
9. NEC (uPD481850), *8M SGRAM Preliminary Datasheet.*
10. H. Kotani, *et al.*, A 50 MHz 8-Mbit video RAM with a column direction drive sense amplifier, *Journal of Solid State Circuits*, Vol. 25, No. 1, February 1990, 30–35.
11. H. Kotani, *et al.*, A 256-Mb DRAM with 100 MHz serial I/O ports for storage of moving pictures, *Journal of Solid State Circuits*, Vol. 29, No. 11, November 1994, 1310.
12. *NEC Memory Products Data Book, Volume 1 2, DRAMS, DRAM Modules, Video RAMs*, 1993.
13. Micron Technology, *Specialty DRAM Graphics/Communications Data Book*, 1993.
14. Samsung, *DRAM Databook*, 1995.
15. Hitachi, *DRAM Databook*, 1993.
16. Prince, B., *Semiconductor Memories*, 2nd Edition, John Wiley & Sons, 1992.
17. United Memories, *4M Synchronous Graphics DRAM Preliminary Datasheet*, 1994.
18. Texas Instruments, *MOS Memory Databook*, 1995.
19. R. Myravaagnes, Rambus memories shipping from Toshiba, *Electronics Products*, November, 1993
20. R. Wilson, Cirrus unwraps Rambus graphics line, *Electronic Engineering Times*, 26 June, 1995, 16.
21. IBM Microelectronics, *4Mb VRAM Data Sheet and Applications Notes*, 1994
22. Texas Instruments, *4-Megabit Video RAM, Application Guide*, 1993.
23. Samsung Semiconductor, *256K × 32 CMOS Window RAM Preliminary Datasheet*, September 1993.
24. Micron Technology, *1995 DRAM Data Book.*
25. K. Inoue, *et al.*, A 10Mb 3D frame buffer memory with Z-compare and alpha-blend units, *IEEE International Solid-State Circuits Conference*, 1995, 302.
26. T. Watanabe, *et al.*, 3-D CG media chip: an experimental single-chip architecture for three-dimensional computer graphics, *IEICE Trans. Electron.*, VOL. E&&C, No. 12, December 1994, 1881.
27. Calle R., Graphics IC for portable PCs brings frame buffer on chip, *Electronics Products*, May 1995, 103.
28. K. Kushiyama, *et al.*, A 500 Megabyte/s Data-Rate 4.5M DRAM, *IEEE Journal of Solid State Circuits*, Vol. 28, No.4, April 1993, 409.
29. Samsung, 128 × 32bit × 2 Bank Synchronous Graphics RAM preliminary data sheet.