1998

Current mode sense amplifiers applied to dual port register files by Larry R. Fenstermaker.

Larry R. Fenstermaker
Lehigh University

Fenstermaker,
Larry R.

Current Mode Sense Amplifiers Applied to Dual Port Register Files

January 10, 1999

CURRENT MODE SENSE AMPLIFIERS APPLIED TO DUAL PORT REGISTER FILES

by

Larry R. Fenstermaker

A Thesis

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

The Department of Electrical Engineering and

Computer Science

Lehigh University

November 24, 1998

This thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science.

December 2, 1998

Date

_____

Thesis Advisor

_____

Chairperson of Department

ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to Professor Michael Schulte for his assistance in the preparation of this manuscript. In addition, the author wishes to acknowledge Kevin O'Connor, for his initial work and his helpful suggestions.

TABLE OF CONTENTS

LIST OF FIGURES

Abstract

CURRENT MODE SENSE AMPLIFIERS APPLIED

TO DUAL PORT REGISTER FILES

by Larry R. Fenstermaker

Two types of sense amplifiers, quasi-current mode and full-current mode, are designed and applied to a standard CMOS register file using the 0.6 micron Lucent Technologies process. Compared to a typical register file that uses a current mirror type of sensing, the read access time is reduced by a factor of three and the power dissipation is reduced by a factor of four when the size of the register file is 128 words. For larger register files, more significant reductions in access time and power dissipation are achieved. A full description of both current mode sensing techniques is given, along with a description of the architecture and circuit modifications to the register file.

Chapter 1

INTRODUCTION

Register files are small, fast memories, usually with multi-port access [1], that are becoming increasingly popular as embedded structures in very large scale integrated circuit (VLSI) designs. Although register files can be as simple as a bank of flip-flop registers, their high usage on chips shows a need for efficient structures that are more complex [2]. This added complexity is necessary for area reduction and performance improvements.

The read operation of a dual port register file, the operation most likely to limit overall performance, is the focus of this work. In particular, the sensing operation of the read cycle is investigated, with two non-conventional techniques being compared to conventional memory read operations for speed and power. These techniques are applied to the register file compiler in the Lucent Technologies 0.6μm CMOS standard cell library. In order to better understand the sense amplifier application, the architecture and operation of the Lucent register file is described in chapter 3. While area reduction is not the main purpose for this work, one of the goals is to maintain the current register file size.

Both of the sense amplifiers that are introduced here utilize current mode sensing, although one of the methods, the cross-coupled enabled sense amp, operates in a

quasi-current mode. A major portion of this thesis is devoted to giving a complete description and analysis of both types of sensing techniques. A theoretical description along with circuit design, analysis and simulation are included for each type of sensing. This is followed by a description of the practical application of the two sensing methods to a dual port register file.

The final portion of this paper is dedicated to a comparison between the different sensing techniques, with comments about the advantages and disadvantages of each. An extension of these techniques to future technologies concludes the paper.

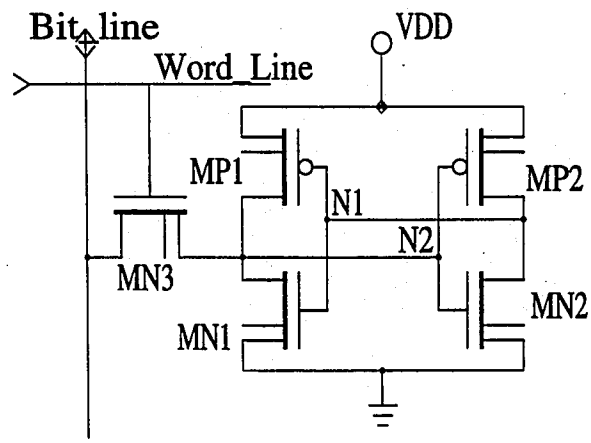

Figure 1. Single Access Memory Latch.

Register files consist of banks of latches, or memory cells, for data storage and some control circuitry to access the contents of the latches. They are accessed directly as a first level of storage required by certain hardware. The latch is usually a pair of cross coupled inverters, as shown in Figure 1. This type of latch is static, rather than

dynamic, because it has load elements, PMOS transistors MP1 and MP2, that offset charge leakage at the drains of the storage transistors.

Because they are compact and area efficient, register files are used to save area and power over groups of single standard cell flip-flop latches where large memories are still impractical either because of area constraints or required access speeds. Small register files, usually less than 24 words, are not very practical since the overhead associated with the control circuitry negates any area savings gained by the small size of the memory cell. In that case, separate standard cell flip-flops are used instead.

Register files are used extensively in communications circuits and even more extensively in computer circuits. Typical CPU's, for example, have many architectural features that require register files. In addition, simultaneous reading from and writing to the register files saves quite a bit of time [3] making dual, or multi-port, register files that allow simultaneous read and write access very useful devices. Register files are also used in interface circuits, such as FIFOs, where multiple ports are also favored. With the increasing importance of these types of circuits, the dual port register file is becoming a very popular circuit.

The use of register files in Lucent Technologies VLSI application specific integrated circuits (ASICs) has increased dramatically over the past few years mostly due to the increased development of ASICs for computer circuits. The amount of ASIC chip area devoted to register files has increased more than five times in the past five years.

Register files are now used on 65% of all new ASIC designs and an average of 30% of the chip area is taken up by register files on those chips [2].

Improved communications and computer architectures have also resulted in the need for larger register files. Data transfers are being made in blocks, whose sizes are increasing, mostly because improved circuits speeds allow processors to handle more data at a time. The Lucent Technologies 0.6 μm CMOS standard cell register file has a maximum of 128 words. New circuit techniques are necessary to expand this.

Chip speeds are also increasing dramatically. The typical ASIC operating frequency has increased from less than 60MHz to over 300MHz in the past five years. It is imperative that library components, especially highly used circuits, such as register files, keep pace.

In addition to size and speed, power consumption has become a critical factor in ASIC chip design. Because of shrinking design rules, more transistors and functions, are being packed into chips. Improved packaging techniques are also allowing for larger overall chip die area. Keeping power to a minimum is essential to keep temperatures within operating range.

Additionally, because of the proliferation of portable consumer products, low power consumption is important to maintain battery life and meet the lower voltages required by battery-operated circuits. While lower voltages help to reduce power somewhat, the increased number of transistors per chip combined with higher frequencies provide great incentive to reduce power consumption wherever possible.

The critical timing operation in the Lucent Technologies standard cell register file is any type of read from memory. For example, in the 0.6 μm CMOS technology, the 128 word deep by 32 bit wide register file can be written in 9.1 ns. at 5.0 volts and in 12.3 ns. at 3.3 volts, but the access time is 11.9 ns. at 5.0 volts and 17.3 ns. at 3.3 volts [4].

Clearly, in order to reduce power consumption through voltage reduction, the read path of the register file must be improved. The largest delay in the read path is from address decode to data out. Therefore, that is the focus of this work.

Chapter 2

REGISTER FILE SENSING TECHNIQUES

Inverter Driver

One of the simplest methods to get data out of a memory is to directly drive the output through a selector, which is an integral part of each memory cell. Figure 1 shows a single transistor, MN3, as a selector, referred to as an *access device* or access transistor. Typically a complete row of memory cells is selected by the output of an address decoder. This signal is referred to as a *word line.*

The word line in Figure 1 is used to access the memory through MN3 during both a read and a write. This single transistor could be duplicated, providing separate read and write access to the memory cell, as shown in Figure 2, where MN3 provides dedicated read access and MN4 provides dedicated write access to the memory cell. Word lines that are dedicated to selection for reading only, as opposed to writing and reading, are referred to as *read word lines.* Similarly, word lines that are dedicated to write access are referred to as *write word lines.*

A row of memory cells usually represents a full word of memory. A row of memory can represent any sized word, and is not limited to a 16-bit or a 32-bit word. For

example, the Lucent standard cell register file word size ranges from a minimum of one bit per word to a maximum of 36 bits per word [4].

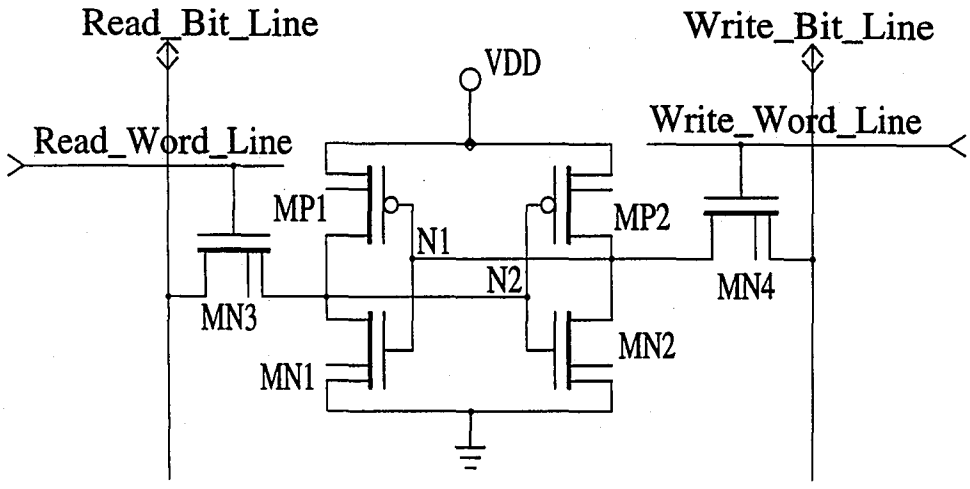

Figure 2. Dual Access Memory Latch.

Figure 3 shows a typical array layout for a 4-bit wide memory that is four words deep, where each block represents an individual memory cell. Each row corresponds to a different word and each column corresponds to a bit position in each word.

A single word line, or a pair of read and write word lines, is connected to all cells in the row. When a single word is selected, the whole word that is output consists of four bits, one from each column. Rows can be added to the bottom of the array to expand the number of words. Columns can be added to the left of the array to expand the number of bits per word.

Address decoders are conveniently located to the left or right in order to drive word lines that cross each bit in a word. Data, in and out, is most easily accessed from the top or bottom by the bit lines shown in Figure 3.

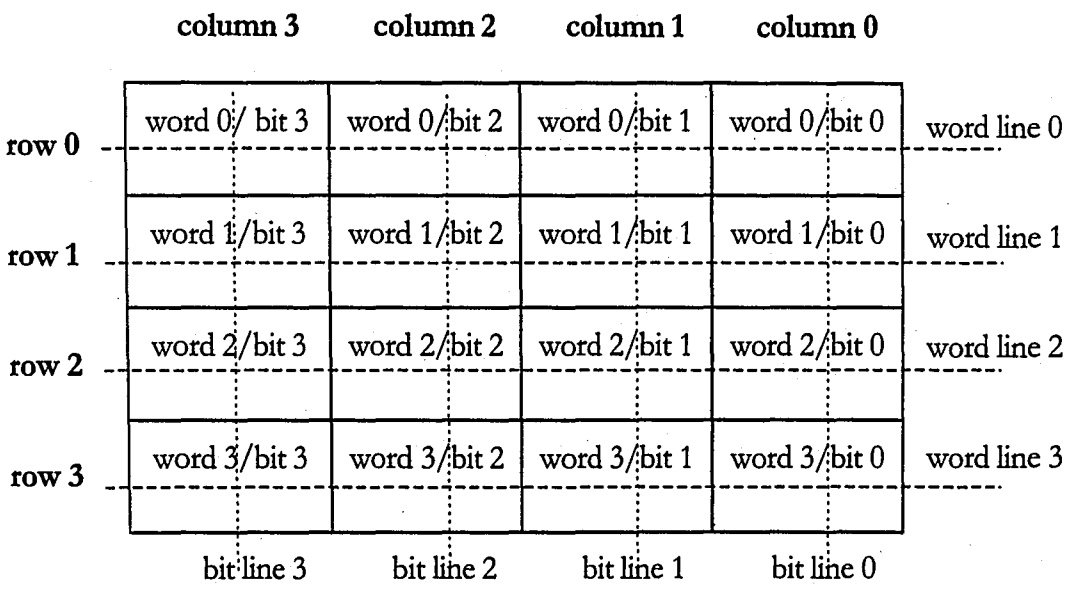

Figure 3. Memory Core Layout.

In order to prevent routing from dominating the width of the memory core, the outputs from all access devices in a column are tied to a single bus called a *bit line*, as shown in Figure 3. This prevents the need for separate vertical routes from each memory cell that would saturate the horizontal space with an increase in the number of words.

The size of the memory cell determines the overall size of the entire register file and is usually the largest percentage of the total area. The memory core takes up an area equal to the number of memory cells required. Without special column muxing, as discussed later in this paper, each row of memory is made up of one cell per bit width of the

register file and there are as many rows as there are words in the register file. The total area required for the memory core of the register file is:

$$A = W * N * C \quad (1)$$

where A is the total memory core area, W is the number of words, N is the number of bits per word and C is the area of a single memory cell. The size of the memory cell, therefore, influences both the width and the height of the memory core. As an integral part of the memory cell, the selector, access device in most cases, must be as small as possible.

Without additional circuitry, the output of the selector, the access transistor in the memory cell, is the data out from the memory cell, the value of the bit line. While this is simple, it is not very fast. In order to minimize area and to allow the memory cell to be easily written, transistor sizes in the memory core are minimum size with very little drive strength.

The simplest and most area efficient selector is a minimum size single pass transistor that is enabled by a word line, as shown in Figure 1. The drive strength of a pass transistor is proportional to the amount of current that can be passed through it; a function of the transistor width and type. NMOS devices have greater drive strength than PMOS devices of the same size [5], and are usually used for the single transistor selector. Signals driven directly by the memory cell through a pass transistor are very slow.

The small drive strength deficiency of the memory cell can be overcome by buffering the signal either before or after the access devices through an inverter driver. A buffer placed before the access devices would be a part of the memory cell and would increase the memory cell size, influencing both the height and width of the register file. Additionally, the conductivity of the access device would limit signal development through the access device.

A buffer placed after the access device, if not attached to individual memory cells, must be shared by all the memory cells in a column, i.e. the same bit of each word. The output of the access device from the selected memory cell is loaded down by the capacitance of the drains of all the access devices in that column, a number of transistors equal to the number of words in the register file. As the number of words in the register file gets large, the capacitive loading on the bit lines increases until it is difficult to drive by an individual memory cell through an access device.

Increasing the width of the access devices to enhance the drive will increase the capacitive loading on the bit line by an amount linearly proportional to the number of words times the increase in width. This method of using an inverter as a buffer to drive the memory output directly will work for small word sizes, where the limit depends on the technology, but a point of diminishing returns is reached when increasing the size of the access devices to drive the outputs of larger register files.

For larger register files, the minimum rise/fall time of the bit lines is limited by the access devices to the memory cell. As indicated by the simulation results of the 0.6μm CMOS register file in Figure A-3, these times can be quite large.

Despite these long transition delays, there are several methods to reduce the actual time to output data. Because NMOS transistors are inherently faster than PMOS transistors [5], bit lines can be pulled high with large dedicated PMOS transistors, a process called precharging, before each read operation. The bit lines are then disconnected from the power supply by turning off the PMOS devices when the memory cell is accessed.

Since the bit lines always start at a high before a read, the switch point of the inverter buffer can be adjusted to a higher voltage level. The time to switch the buffer is reduced because the buffer does not need to wait until the bit line drops to the voltage of the normally lower switch point. This method can, however, introduce some risk. If the switch point is too high, simple noise on the bit lines could be seen as a logic change and false outputs could occur.

Differential Sensing

Special analog circuits, called sense amplifiers are used to allow the output to safely and consistently switch with smaller than normal bit line swings. These circuits sense small voltage differentials between bit lines connected to both sides of the memory cell and therefore require dual access devices in the memory cell, as shown in Figure 4, where a single word line controls the access to both access devices. Similar to the single transistor access cell in Figure 1, this cell architecture requires the read and write operation to share the access devices.

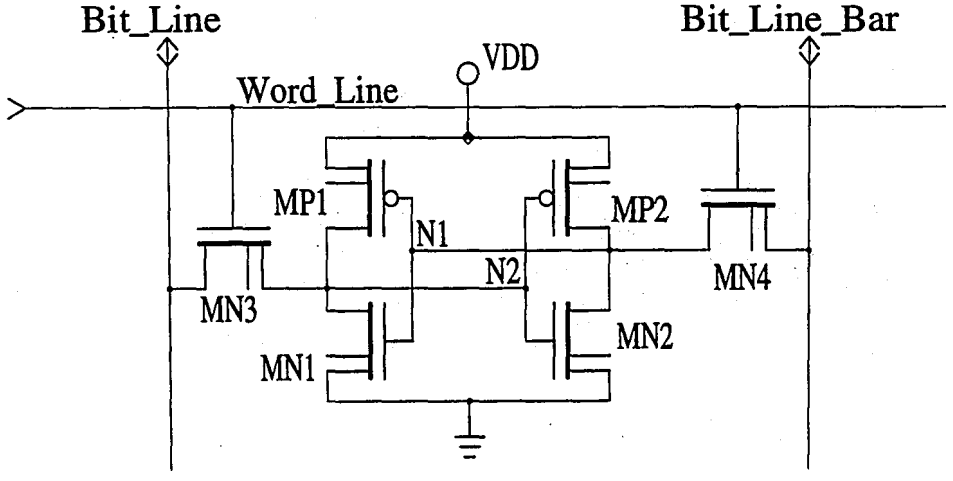

Figure 4. Dual Side Access Memory Latch.

Dual access devices couple opposite sides of the memory cell to separate bit lines that can be compared to determine the sense of the memory cell data. Pairs of bit lines are used to indicate both the non-inverted(bit) and the inverted data(bit bar) memory content. The inversion is referenced relative to the input data of the register file.

Differential sensing techniques employ *sense amplifiers* to sense small voltage differences between bit (non-inverted) and bit bar (inverted) lines to determine if the data in the memory, as reflected on the bit lines, is a logic 1 or a logic 0. A single-ended differential sense amplifier, like the current mirror circuit shown in Figure 5, is most commonly used in static memories.

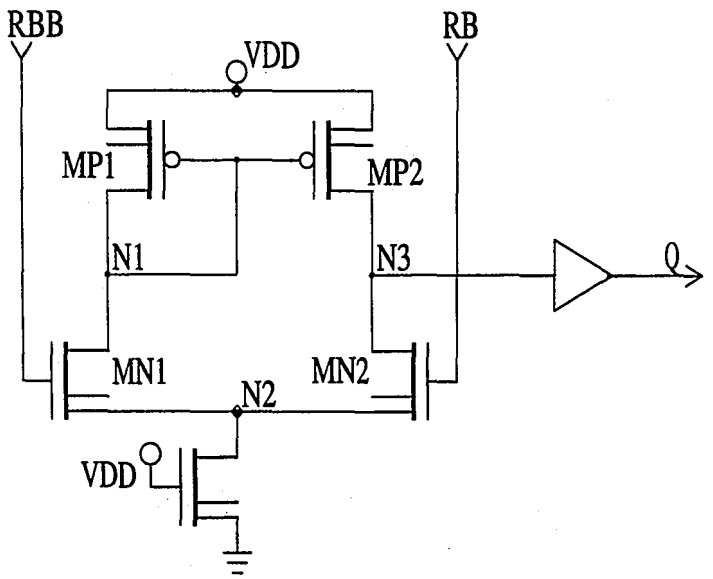

Figure 5. Current Mirror Voltage Differential Sense Amp.

In order for this type of sense amplifier to work, both bit lines must be precharged to a high voltage level, at or near input supply voltage level, Vdd. Referring to Figure 5, with both bit lines(RB and RBB) high, transistors MN1 and MN2 are on and nodes N1 and N3 are balanced at some intermediate voltage. One word line then goes high to select a row of memory cells, turning on the access devices in memory cells of that row. The memory cell begins to pull-down the bit(RB) or bit bar(RBB) line, depending on the contents of the memory cell. If RB is pulled low, MN2 is turned off and N3 is

charged to the full high, logic "1" state through MP2. If RBB is pulled low, MN1 is turned off, N1 is kept high, and N3 is pulled low through MN2 indicating a logic "0".

This type of sensing is very fast and is a great improvement over the single inverter buffer [6], but can be a source of DC current leakage [7].

Start-up Locked Sense Amp

As a step toward meeting the need for low power circuits, particularly those that cannot tolerate DC leakage, such as battery-operated circuits, the differential sense amp in the Lucent standard cell register file was replaced by the start-up locked sense amplifier. This sense amplifier, shown in Figure 6, is based on a cross coupled NOR gate structure [7]. This sense amp has the advantage over the current mirror of reaching full VDD output levels no matter which logic level is sensed.

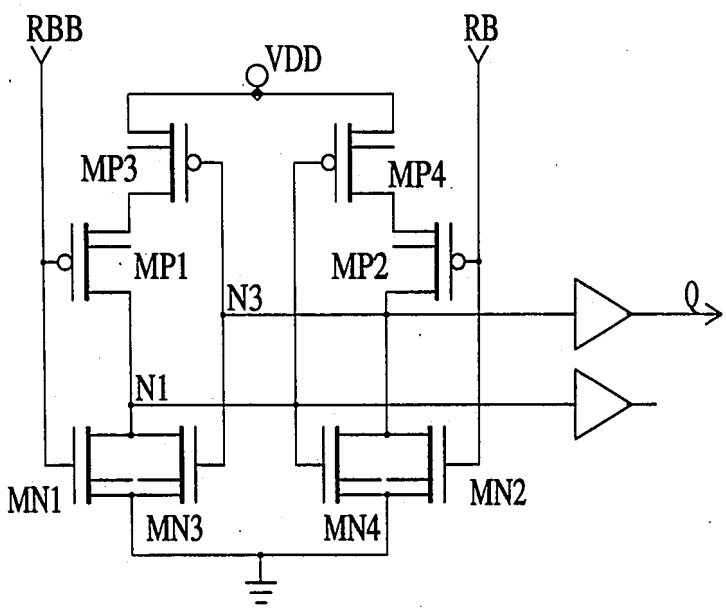

Figure 6. Start-up Locked Sense Amp.

Similar to the current mirror sense amp, this sense amp requires precharged bit lines before the sensing operation can begin. As the voltage on one of the bit line pairs is pulled low by the memory cell, through the access devices, the PMOS/NMOS pair

controlled by that bit line(for example MP1 and MN1 controlled by RBB in Figure 6) will cause the output of that transistor pair to switch from a low to a high.

Of course the low going bit line has to drop below the switch point of the transistor pair before the output will switch, but the series PMOS devices(MP3 and MP4 in Figure 6) vary the drain voltage to the pull-up side of the inverter pair, thereby enhancing the regenerative nature of the latch by decoupling the large capacitance of the bit lines from the signal development. The parallel NMOS devices(MN3 and MN4) insure an initial locked startup state to prevent DC leakage current before a first read takes place.

The outputs are buffered in order to give a high strength output signal without loading the developing nodes of the sense amp. A similar first stage buffer is applied to both sides of the sense amp, even though the output is taken from only one side. The buffer stage on the opposite side is a dummy that balances the load on the developing nodes to prevent favoritism in signal development; a logic " 1" should develop as fast as a logic "0".

Other than the elimination of DC power, this sense amp was designed to perform with characteristics similar to the current mirror differential amplifier [7], this being slightly faster and with lower operating power consumption. In order to minimize the impact to the layout generator, the area of this sense amp was also kept similar to the current mirror amplifier. As such, this is an improvement over the current mirror type of sense

amp and is used in the comparison of the new sense amp circuits that are the subject of this paper.

Chapter 3

THE DUAL PORT REGISTER FILE

In order to fully understand the application of the circuits presented in this paper, a brief description of the operation of the Lucent standard cell register file is presented here. The block diagram of the register file shown in Figure A-1 is typical of register files throughout the industry.

The circuit consists of an array of true dual port, eight transistor, memory cells, as shown in Figure 7, selected by separate address decoders for the read and for the write operation. This memory cell has dedicated read and write access devices that access both sides of the memory.

Input data and write addresses are latched by the clock edge while data is transferred to the memory core. These latches make the write operation synchronous with respect to the clock. This clock, however, can run at any frequency, up to the defined maximum. This signal actually need not be a clock of regular frequency, so long as the minimum pulse width conditions are met according to the data sheets given in the standard cell catalog [4].

The read operation is completely isolated from the write operation, except that if the contents of the memory cell are changed by a write, the outputs reflect the new data if

the read address is pointing to the memory location that was written. Dedicated pairs of bit lines for reading (RB and RBB) and for writing (WB and WBB) insure the isolation between the read and the write operation. These can be seen in the memory cell schematic in Figure 7.

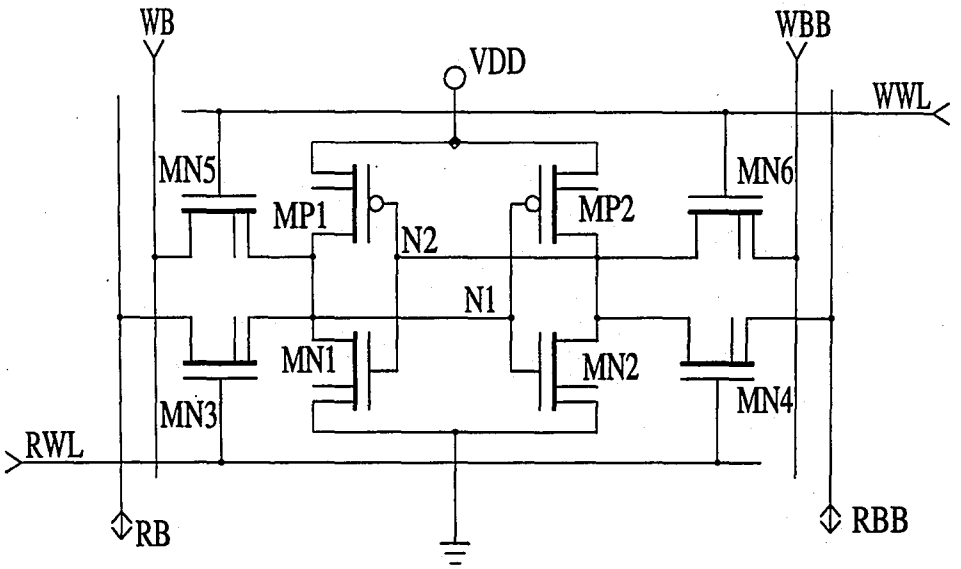

Figure 7. True Dual Port Memory Latch.

Unlike the write operation, the read is completely asynchronous in that it takes place whenever a change occurs in the read address, or, as mentioned previously, if the data in the cell that the read address is pointing to changes because of a write. A read from a change of read address begins when an address transition detector (ATD) on any of the read addresses detects an address change and generates a pulse that causes the read bit lines to be precharged. To speed up the read operation, the read bit lines are not precharged to full VDD.

While the bit lines are precharged, all read word lines are pulled low to protect the contents of the memory while the read address is decoded. The trailing edge of the precharge pulse brings up the word line that is indicated by the output of the decoded address. This allows the memory cell to pull one of the read bit lines low, reflecting the state of the memory. The sense amplifiers, one per bit or column, project this data to the outputs through the output buffers.

Since the read word lines remain high at all times except during precharge, any change to the memory due to a write will cause the read bit lines to change state. Eventually the change of state of the read bit lines will cause the state of the sense amp to flip, and the new data will be reflected on the outputs. This operation is not preceded by a precharge of the bit lines and, as a result, some of the bit lines must be pulled high as well as low by the memory cell through the read access devices in the memory.

Since PMOS transistors are weaker than NMOS transistors, and since PMOS transistors are used as pull-ups in the memory cell, the low to high read bit line transition is particularly slow. Fortunately, however, the read address is already decoded and enough time is saved by not waiting to decode the address so that the time required to read new data after a write is comparable to the time required to read data from memory after a change of read address.

Chapter 4

DIFFERENTIAL CURRENT-MODE SENSE AMPLIFIERS

Overview

One of the major disadvantages of the differential sense amps discussed so far is that one of the bit lines is always allowed to be pulled completely to ground, and then charged to near VDD in order to operate. Although larger transistors can be used to speed up the precharge operation, this tends to consume considerable power if the capacity of the bit lines is large, such as for register files with a large number of words. There is also some physical limit to the size of the precharge transistors used for precharge.

In order to eliminate the need for large voltage swings in the high capacitance bit lines to perform a read, designers of large memories have developed differential current sensing techniques. Unlike the sense amplifiers described previously, current-mode sense amplifiers, such as the one described by Blalock and Jaeger [8], sense current differences between currents coming from the memory cell and a reference source rather than sensing voltage differences.

In addition to eliminating the need for large current swings on the bit lines, these techniques minimize the affects of the capacitance associated with bit lines. Current

sensing methods have been thoroughly investigated for noise immunity and performance by Blalock and Jaeger [8] and by Seevinck [9], among others. These circuits were developed for use in large memories as pre-amplifiers, the outputs of which were further amplified through standard voltage amplifiers.

Originally developed for large DRAM technology, Seevinck has further demonstrated their usefulness in large SRAM's as well. All of these applications have been used as the first stage amplifiers for large synchronous memories with single access cells like the ones in Figures 1 and 2. None of the applications have been used in an asynchronous memory as a single amplifier, nor with a dual port memory cell, such as the ones in the Lucent standard cell register file.

Cross-coupled Enabled Sense Amp Circuit Description

The cross-coupled enabled sense amp shown in Figure 8 eliminates the precharge portion of the read cycle by using clamping devices, MP6 and MP7, to keep the bit lines in a precharged state. The cross-coupled latch is made up of the pull-up devices MP1 and MP2 and the pull-down devices MN1 and MN2.

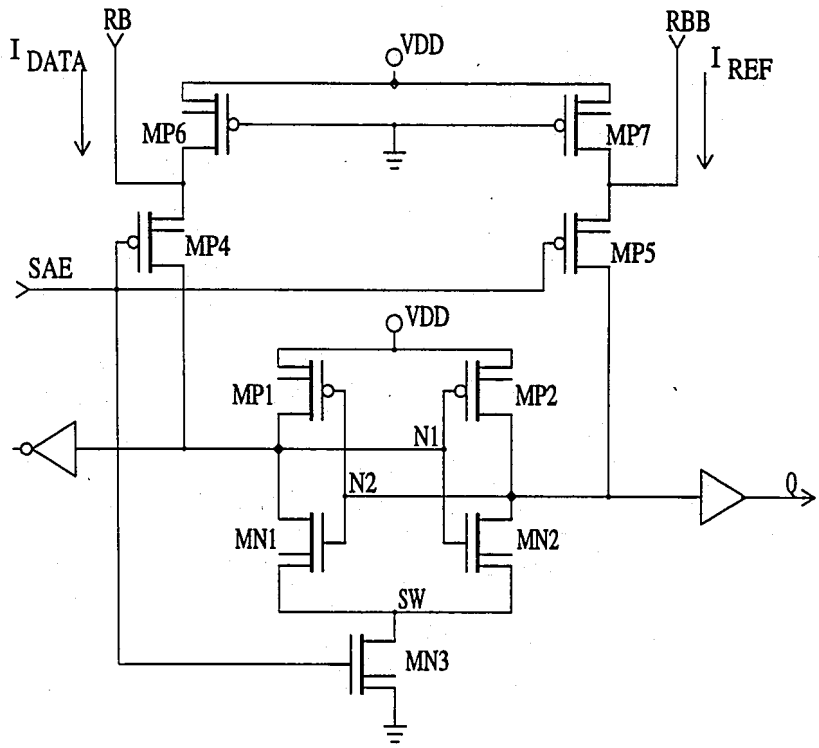

Figure 8. Cross-coupled Enabled Sense Amp.

A pair of PMOS devices, MP4 and MP5 in Figure 8, allow the bit lines to be disconnected from the nodes of the sense amp. Further, MN3, an NMOS device, can be used to disconnect the sense amp cross-coupled inverters from ground. Since this device is controlled by the same signal, SAE, as the bit line connection PMOS devices,

the bit lines are disconnected from the sense amp when the sense amp is connected to ground and are connected to the sense amp when the sense amp is disconnected from ground. Because of this disconnection of bit lines from the sense amp during output signal development, this circuit is not fully current-mode as described in previous works.

A series of inverters are used as buffer stages to drive heavy output loads. The first inverter in the string is small with an identical inverter added to the opposite node of the sense amp for balance.

This sense amplifier circuit uses nine transistors, besides the transistors in the buffers, one more than the 8 transistors in the start-up locked sense amp and four more than the five transistors in the current mirror sense amp. Two of the transistors, however, are used to clamp the bit lines high and replace the large precharge transistors used with the other types of sense amp circuits. Therefore, the total number of transistors for this circuit is actually less than the number required in the start-up locked circuit and only two more than the current mode type of sense amp, a number easily compensated for by the smaller size of the clamping transistors as opposed to the precharge transistors.

Cross-coupled Enabled Sense Amp Operation

The cross-coupled enabled sense amp is shown attached to a dual port memory cell in Figure 9. The memory cell is repeated one time for each word of the register file. Refer to the read data path in Figure 9 during the following description of operation for the sense amp.

Until a read occurs, the sense amp enable signal, SAE, is kept high. During this time, the sense amp is enabled, connected to ground, and remains in a latched state with the cross-coupled inverters settled to opposite states, depending on the result of the previous read. As previously mentioned, the bit lines are disconnected from the sense amp by the sense amp access devices, MP4 and MP5, at this time and are kept high by the clamping devices, MP6 and MP7.

When a read occurs, the memory read access devices, MN4 and MN5 are turned on by a selected read word line and the sense amp enable signal, SAE, is pulled low. There is no critical timing between the word line going high and SAE going low. Turning on the access devices to the memory causes one of the pair of read bit lines, RB or RBB, to be pulled to a lower voltage level by the pull-down device, MN8 or MN9, of the side of the memory cell that is low.

This low side of the read data path has a DC current component, I_{DATA}, through one of the clamping devices, the bit line, a memory access device, and the pull-down device of the inverter in the memory cell which holds that memory node low. The high side has essentially no current flow since the drain to source voltages is zero across all

devices in the equivalent high side path. This can be a reference current, I_{REF}, to the sense amp circuit, as described by Blalock and Jaeger [7].

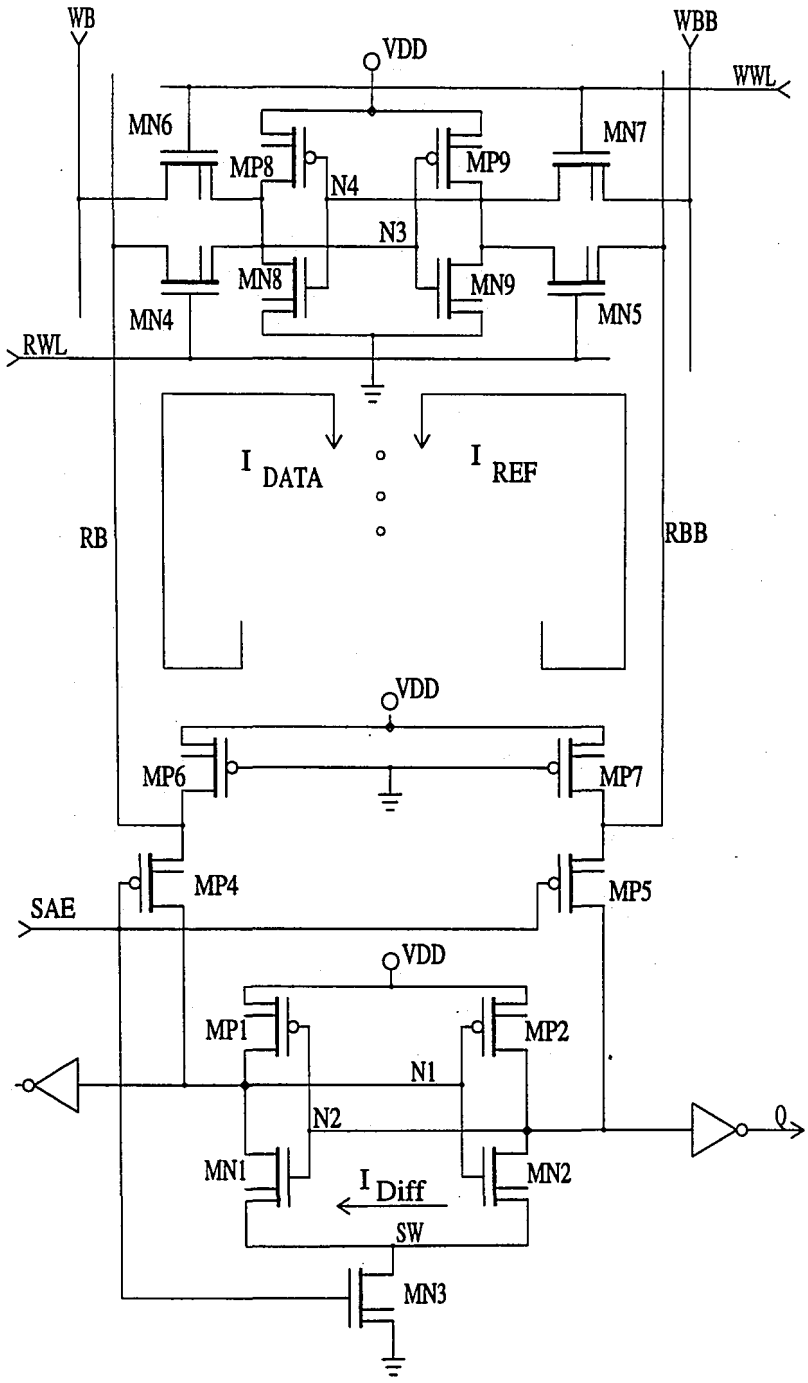

Figure 9. Data Path With Cross-coupled Enabled Sense Amp and Single Cell

Although Figure 9 indicates that I_{DATA} goes through MP6, RB, MN4 and MN8, which indicates that N3 is low and that a "0" is stored in the cell, with WB the same sense as the input data, and I_{REF} goes through MP7, RBB, MN5 and MN9. The references change if a "1" is stored in the memory cell. In that case, the current through RB is zero and used as the reference current, and RBB conducts the data current.

With SAE low, the sense amp is disconnected from ground and, since the bit lines are high, the output nodes of the sense amp, N1 and N2, are pulled high through the sense amp access devices, MP4 and MP5. At this time the sense amp is in a preset state with both outputs held high, or near high, depending on the voltage level of the low bit line.

The current differential between I_{DATA} and I_{REF} is projected onto the output nodes of the sense amp through the sense amp access devices. A current similar to I_{DATA} is developed in the adjacent sense amp output node through the corresponding sense amp access device and pull-down transistor in the latch.

With the sense amp enable device, MN3, turned off this differential current is developed through node SW and will cause an equal but opposite current in the opposing pull-down devices in the sense amp. Referring to Figure 9 for example, if RB is conducting current, I_{DATA}, a source to drain current I_{DIFF} will flow in MN2 and an equal drain to source current will flow in MN1. This negative current through MN1 will cause the impedance looking into the drain of MN1 to be a negative resistance.

A short time later, when the output nodes of the sense amp reach their maximum potential, the sense amp is enabled by SAE going high. With SAE high, the sense amp output nodes are disconnected from the bit lines by MP4 and MP5 turning off. MN3 is turned on simultaneously, connecting the drains of the NMOS devices of the cross-coupled inverters to ground. This cross-coupled pair of inverters are in a metastable state and act as a high-gain positive feedback amplifier [7]. The sense amp will quickly settle to a stable state with the side that sees the negative resistance on the pull-down transistor, N1 in the case of the currents indicated in Figure 9, going low.

With the sense amp disconnected from the bit lines, the read word lines can now be returned low, allowing the bit lines to be pulled to full VDD by the clamping devices. This completes the read cycle.

This operation can be seen in the simulation waveforms in Figure A-4, where the read word lines are VRWL0 and VRWL1, and the bit lines are VRB0, VRBB0 and VRB1 and VRBB1. The output Q1 correspond to bit column 1 with read bit lines VRB1 and VRBB1 and the output Q0 is bit 0 associated with read bit lines VRB0 and VRBB0. In this simulation, the sense amp enable, VSAE, goes low before the word line goes high.

Transistor Sizing for the Cross-coupled Enabled Sense Amp

The conductance of the MOS transistor is proportional to the width to length ratio of the gate. The drain to source current is a function of this ratio:

$$I_{ds} = kW/2L(V_{gs} - V_t)^2 (1 + \lambda V_{ds}) \qquad (2)$$

where I_{ds} is the drain to source current, W is the gate width, L is the gate length (also referred to as effective channel length), k is a process constant, λ is a physical constant for silicon, V_{gs} is the gate to source voltage, V_t is the threshold voltage, and V_{ds} is the drain to source voltage of the MOS transistor. With the gate length at the minimum for the process technology, the only way for designers to increase the conductance of the MOS device is to increase the width [10]. Therefore, the transistor sizing talked about in this section will inherently be referring to gate width and assume that gate length is minimum, unless specifically stated otherwise.

Since the outputs of the sense amp are buffered, the drive strength of the cross-coupled inverters in the sense amp latch need not be very large. The speed at which regeneration occurs will be determined by the size of these transistors. Normally they are kept at a modest two times the size of minimum transistors for a technology. Since they are only required once per bit column, their size will have a one dimensional affect on the size of the register file and these devices can be made larger to gain signal development speed. Making them too large slows down the sense amp preset operation.

The sense amp enable and access devices are driven by the same signal, SAE in Figure 9. In fact, this same signal drives all of the devices in every sense amp in the register file. For a register file that is 32 bits wide, this signal would have to drive the gates of 64 of these devices, two per sense amp. Keeping the access and enable devices small allows the sense amp enable signal to develop faster in large bit size register files.

If these devices are too small however, the operation of the sense amp is hindered. The preset operation, for example, depends on the charge sharing of the bit lines and the nodes of the sense amp through the sense amp access devices. These devices are usually minimum size to reduce the loading on the sense amp enable signal, but increased to speed up the preset operation. Note that, since NMOS transistors are generally twice as conductive as PMOS transistors, the size of the sense amp enable device, MN3 in Figure 9, only needs to be half the size of the pull-up devices in the sense amp latch, MP1 and MP2 in Figure 9. This allows the sense amp enable transistor to be minimum, if the sizes of the transistors in the latch are twice minimum.

Finally, the size of the bit line clamping devices determines the bit line voltage differential, a parameter that is important to noise immunity of the bit lines; not noise immunity of the sense amp. Some of the factors that may affect the bit line are inter-nodal coupling from some other signal, such as the write bit lines, or power or ground noise.

The voltage offset of the bit line at the sense amp access device can be calculated by looking at the current loop circuit of the bit lines as a voltage divider, where the

voltage of the bit line is proportional to VDD times the ratio of the sum of the impedance of the memory access device plus the impedance of the memory pull-down device plus the resistance of the read bit line divided by the sum of the impedance of the clamping device, the impedance of the memory access device, the memory pull-down device and the bit line resistance:

$$V_{RB} = [(Z_{MP4} + Z_{MN8} + R_{RB}) / (Z_{MP6} + Z_{MN4} + Z_{MN8} + R_{RB})] * V_{DD} \qquad (3)$$

where V_{RB} is the voltage of RB, Z_{MP6} is the impedance of MP6, Z_{MN4} is the impedance of MN4, Z_{MN8} is the impedance of MN8 and R_{RB} is the bit line resistance of RB in Figure 9. This is the actual voltage of the data read bit line and must be compared to the voltage of the reference bit line, which is pulled to VDD.

The read access devices in the memory cell need to be as small as possible for several reasons. Similar to the sense amp enable signal, the word line drives the gates of the memory access devices for every bit in a row, or word. In order to make the word access fast, the loading on this line should be kept minimum.

As mentioned earlier, the capacitance on the drain of the access device is a major portion of total capacitance of the bit line. Although the current-mode sense amp reduces the affects of bit line capacitance, they are not completely eliminated. Bit line voltage differential for bit line noise immunity is still affected by the bit line capacitance.

Another reason for small memory access devices is to protect the memory cell from destructive reads by charge sharing with the bit lines. Since the bit lines are clamped high, there is a large amount of charge on the bit lines that can be shared with the memory when the access devices are turned on. This charge sharing is the principle by which DRAM memory is written. The ability of the memory to resist a change of state when the access devices are turned on is called the memory margin and is determined, according to Prince [1], by the ratio of the access device to the pull-down device size. In order to maintain a constant margin, the pull-down devices must be increased proportional to the access devices. Increasing the pull-down devices makes it harder to write, increasing the write cycle time.

Full Current-mode Sense Amp Circuit Description

The cross-coupled enabled sense amp described in the previous sections was quasi current-mode. Although the sensing operation detected current differentials, the circuit developed the output signals with the bit lines disconnected from the sense amp. The circuit described in this section is full current-mode and develops the output signal with the bit lines connected.

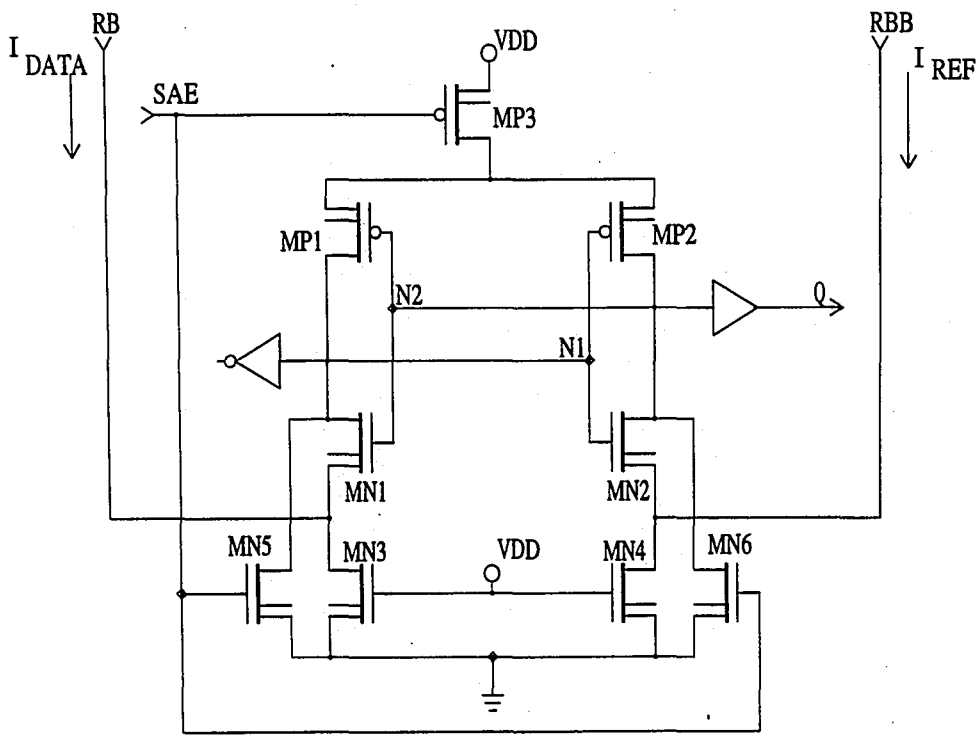

Figure 10. Full Current-mode Sense Amp.

Similar to the cross-coupled enabled sense amp, the full current-mode sense amp circuit uses clamping devices to prevent large voltage swings on the bit lines. The clamping devices in the full current-mode sense amp, however are NMOS devices,

MN3 and MN4 in Figure 10, that clamp the bit lines to ground. Clamping the bit lines to ground gives the bit lines better noise margin over clamping to VDD because the operation is now guaranteed to work in common mode [1].

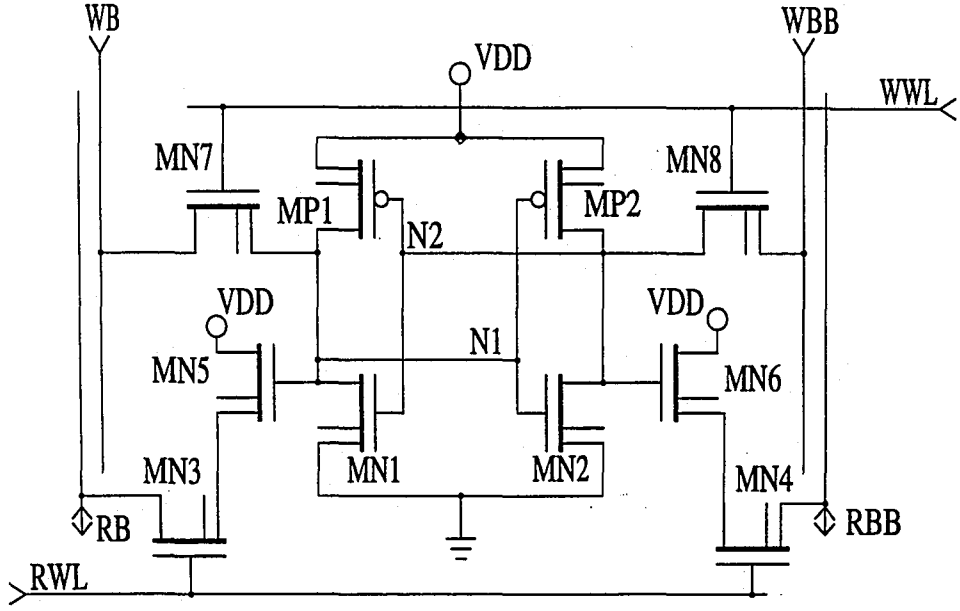

Figure 11. Dual Port Memory Cell with Isolated Pull-up Read.

The pull-down nature of this clamp requires a modified memory cell with a pull-up read path to VDD. This modified memory cell is shown in Figure 11. In this cell, the read access devices, MN3 and MN4, are connected to sources of a pair of NMOS devices, MN5 and MN6, whose gates are controlled by the internal memory nodes, and whose drains are connected to VDD. These two transistors are additional to the traditional eight transistor dual port memory cell shown in Figure 7.

The latch circuit in this sense amp is made up of the pair of cross-coupled inverters, MN1, MN2, MP1 and MP2 in Figure 10. There are also three devices that enable the

sense amp circuit. The PMOS transistor, MP3 in Figure 10, connects the cross-coupled latch to VDD when the sense amp enable signal, SAE, goes low. Also controlled by SAE, two NMOS devices, MN5 and MN6 in Figure 10, connect the output nodes of the sense amp to ground.

This sense amp also has a chain of inverters to buffer the output, with the first stage mirrored on both sides of the sense amp for balance.

Full Current-mode Sense Amp Operation

The data path for the full current mode read operation is shown in Figure 12 with the modified memory cell and the full current-mode sense amp connected through the read bit lines, RB and RBB. Similar to the data path of the cross-coupled latched sense amp read data path, the modified memory cell is repeated once for each word in this data path.

The operation of the full current-mode sense amp is similar in timing but opposite in sense to the cross-coupled enabled sense amp. For the full current-mode sense amp, the sense amp enable signal, SAE in Figure 12, is low when the sense amp is latched. When this signal goes high, the sense amp is reset, as opposed to the preset condition of the cross-coupled enabled sense amp.

SAE going high begins the read cycle of this circuit. Referring to Figure 12, when SAE goes high, the sense amp latch is disconnected from VDD by MP3. At the same time, MN5 and MP6 turn on to pull the latch nodes, N1 and N2, low. The sense amp is then in the reset condition.

Simultaneous to SAE going high, the read word line, RWL, for the selected row of memory goes high turning on the memory read access devices, MN9 and MN10. One of the pairs of memory read pull-up devices, MN11 or MN12, will already be turned on by the high side of the memory, N1 or N2.

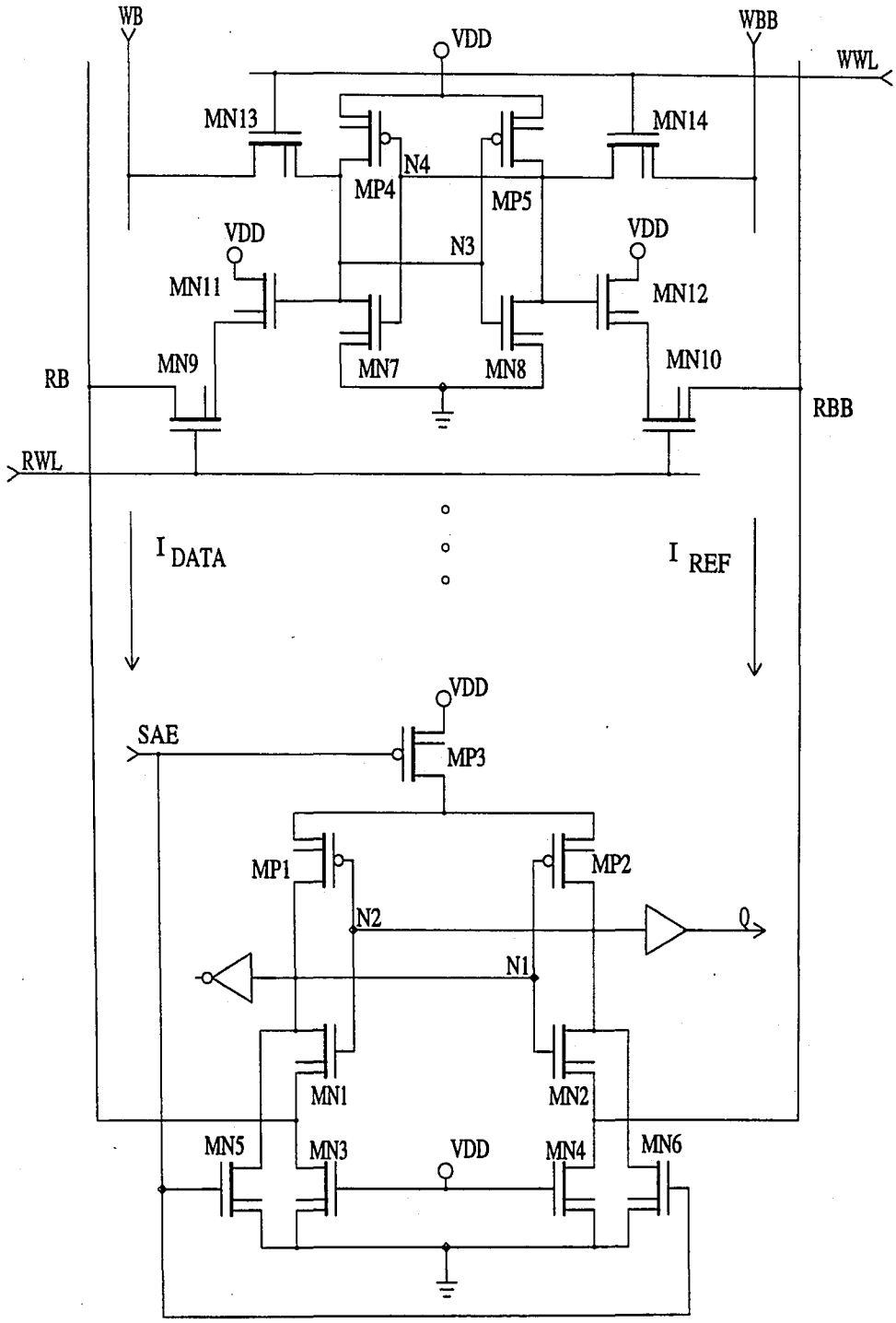

Figure 12. Data Path With Full Current-mode Sense Amp and Single Cell

This memory access will cause a current, I_{DATA}, in one of the read bit lines, RB or RBB, depending on the sense of the memory cell. The opposite bit line will have zero current because the corresponding read pull-up device will be turned off by the low side of the memory. Once again, the zero current can be used as a reference current, I_{REF}, for the sense amp.

Figure 12 is labeled with memory node N3 as a logic 1 and I_{DATA} in read bit, RB. In this case the current flow will be from VDD through MN11, MN9, RB, and MN3 to ground. If a logic 0 were stored in the memory, the data current would occur in the path with RBB.

Following the labeling in Figure 12, a current differential is developed between the current through the sense amp clamping devices, MN3 and MN4. When SAE goes low, the cross-coupled latch composed of MN1, MN2, MP1 and MP2 operates as a positive feedback amplifier. With the output nodes reset to an equal potential, the latch is in a metastable state. The difference current causes the positive feedback of the latch to amplify this difference and drive the outputs to a latched state.

When the outputs of the sense amp have fully developed, the read word line can once again be brought low, turning off the DC current path through the bit line. This completes the read cycle of the full current-mode sense amp.

This operation can be seen in the simulation waveforms in Figure A-5, where the read word lines are labeled VRWL0 and VRWL1 and the read bit lines are signals VRB0,

VRBB0, VRB1 and VRBB1. The output, Q0, corresponds to column 0 and the output, Q1, refers to column 1. The sense amp enable signal is VSAE.

Transistor Sizing for the Full Current-mode Sense Amp

As per the discussion on transistor sizing for the cross-coupled enabled sense amp, this discussion will focus on scaling the transistor widths, with lengths fixed at minimum unless specified otherwise. The speed at which regeneration occurs within the latch depends upon the load on the output nodes and the drive strength of the inverters that make up the latch. Since the output of the latch is buffered before the final output, the loading on the internal nodes can be kept very small. Similar to the cross-coupled sense amp, the size of the latch devices, MN1, MN2, MP1 and MP2 in Figure 10, can be kept at twice the minimum transistor width for a technology. Any increase in size for these transistors would increase the load on the outputs.

The size of the sense amp reset transistors, MN5 and MN6 in Figure 10, determines the speed at which the sense amp is reset between read operations. These devices can be slightly smaller than the latch pull-down devices because they need to transfer the same amount of charge as the pull-down devices, but they do not have to remove additional charge from the pull-up devices, since the sense amp enable device, MP3 in Figure 10 is turned off at the same time that these devices are turned on. Additionally, increasing the size of the reset devices to speed up the reset will add to the load of the sense amp enable signal, SAE in Figure 10, which, again is connected to every sense amp, one per bit.

The sense amp enable transistor, MN3 in Figure 10, must be large enough to conduct the same amount of current as the pull-up devices in the latch and should therefore be twice the minimum width or larger, depending on the size of the pull-ups in the latch.

The clamping devices in this sense amp, along with the bit line resistance, and the read access and the read pull-up devices in the memory cell determine the bit line voltage differential, which is important for bit line noise immunity. Again , this voltage offset can be calculated by analyzing the voltage divider circuit consisting of the impedance of the clamping device, the read bit line resistance, the memory read access device and the memory read pull-up device:

$$V_{RB} = [(Z_{MN3} / (Z_{MN3} + Z_{MN9} + Z_{MN11} + R_{RB})] * V_{DD} \quad (4)$$

where V_{RB} is the voltage of the bit line, RB, Z_{MN3} is the impedance of the clamping device, MN3, Z_{MN9} is the impedance of the read access device, MN9, Z_{MN11} is the impedance of the read pull-up device, MN11, and R_{RB} is the bit line resistance of RB in Figure 12. This voltage is then compared to the voltage of the reference bit line, which is at ground.

The read access and read pull-up devices can be sized independent of the size of the memory latch, since they do not couple charge from the bit lines to the memory nodes. This is a major advantage of this memory cell over the typical 8 transistor cell. The read pull-up devices add some capacitive loading to the memory nodes, equal to the gate capacitance of the read pull-up devices, and need to be kept small so as not to increase the size of the memory core. However, the pull-ups and the access devices

can be large enough to provide sufficient current to the sense amp, provided the read access devices are not so large that the gate capacitance overloads the read bit line drivers.

Chapter 5

CIRCUIT SIMULATIONS

Overview

Analog simulations were performed on the start-up locked sense amp, the cross-coupled enabled sense amp and the full current-mode sense amp using the Lucent Technologies internal mixed mode simulator ADVICE and the process models from the Lucent Technologies 0.6 μm CMOS technology. ADVICE is very similar to SPICE, the most commonly used mixed mode simulator. All simulations were run under nominal conditions, 25°C and nominal processing at both 3V and 5V.

The circuits that were simulated were similar to the one shown in Figure 9, with additional memory cells added by scaling a single memory cell to represent the loading affects of multiple words of memory. Simulations were run on data paths whose columns were 128, 256, 512 and 1024 words high, i.e. the memory cell was scaled 128, 256, 512 and 1024 times. Only two memory cells in each column were actually accessed by word lines, the remaining memory cells were represented by a single scaled memory cell. The memory cells that were accessed each contained opposite logic values, i.e. a logic "1" and a logic "0".

Scaling was accomplished by representing the memory cell as a subcircuit and using the automatic scaling command in ADVICE. This command takes a parameter that represents the number of times that the subcircuit is to be repeated, and analyzes the subcircuit elements as if the subcircuit were repeated as many times as the parameter. Capacitors and diodes, for example, are treated as parallel elements, and resistors are treated as series elements.

The memory cell and the sense amplifier circuits were laid out using real geometries and design rules from the Lucent 0.6 μm CMOS technology. The Lucent graphics editor, GRED, was used to accomplish the layout manually. ADVICE descriptions of the circuits, including transistors and routing parasitics, were extracted from the layout artwork using the Lucent automatic extraction tool, GOALIE. A high level description of the circuit connectivity, including the scale factors to represent multiple words, was put together manually.

The input stimuli for the simulations were piece-wise linear voltage sources that went full rail, from 0V to VDD, with a rise or fall time of 1 ns. The 1 ns. rise and fall times were selected because it was felt that these were achievable slopes with this technology. The input stimuli was applied directly to the circuit; buffers were not used. The timing and sense of the input stimuli were customized to match the requirements of the particular sense amp circuit.

Power analysis was performed using a special circuit consisting of a current controlled current source driving an RC network as shown in Figure 13. The current source was controlled by the supply current, I_{VDD} which represents the total current for the circuit.

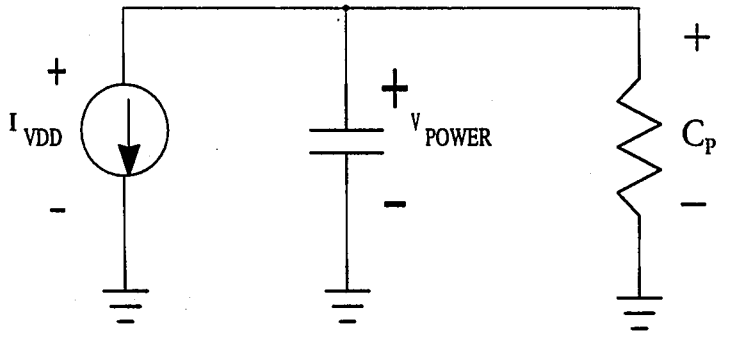

Figure 13. ADVICE Power Meter Circuit.

The current from the controlled source is used to charge a capacitor, C_P in Figure 13. The voltage across the capacitor increases linearly with the amount of charge dumped onto it by the current source. By adjusting the value of the capacitor, resistor and the current source, the voltage can be made to represent the integrated current, which mathematically is the same as the power consumed by the circuit over the time period of the simulation. The resulting power, as represented by VPWR, is shown in Figures A-4 and A-5.

Start-up Locked Sense Amp Simulations

The simulations of the start-up locked sense amp were run for comparison only, and reflect a portion of the read characteristics of the REGFILE as reported in the standard cell catalog [4]. The resulting simulation waveforms are shown in Figure A-6 and Figure A-7.

Referring to Figure A-6, the input stimuli for the circuit are the precharge control voltage, VPC, and the read word lines, VRWL0 and VRWL1. Each of the word lines accesses one of the two memory cells containing actual data. Only one read word line at a time is high, representing a read of only one cell at a time.

For the simulations of this circuit, the bitlines were precharged, VPC high, until a word line was brought high to access the memory. The first read was repeated so as to be sure that none of the internal nodes were at some random initial start up value.

The circuit outputs, VQ0 and VQ1, are the buffered outputs shown as Q in the sense amp circuit in Figure 6. The internal sense amp output nodes, VQ0_N1, VQ0_N3, VQ1_N1 and VQ1_N3, are also shown in the simulation results.

The simulations results shown in Figure A-6 are for a 5.0 volt supply voltage, V_{DD}. Figure A-7 shows the same simulation with V_{DD} at 3.0 volts. As stated earlier, the bit lines are not charged to full V_{DD}, but are precharged through NMOS devices and are an N-channel threshold voltage less than V_{DD}.

The power simulation results in Figure A-4 for the 5V operation show how power increases whenever an operation is performed. The incremental power increases from these waveforms match the times of the operations in the waveforms in Figure A-6. The operation waveforms were not included along with the power plots because they would have distorted the power results, being a different scale. The power simulation waveform is presented only to show how the power meter works and to show the variation of power across the size of the column.

Cross-coupled Enabled Sense Amp Simulations

The simulation results for the cross-coupled enabled sense amp are shown in Figures A-8 and A-9, where once again the read word lines are VRWL0 and VRWL1. For these simulations, however, instead of a precharge input signal, this sense amp circuit requires a sense amp enable input, shown as VSAE. The simulation results shown in Figure A-8 are for V_{DD} at 5.0 volts; the simulation results shown in Figure A-9 are for V_{DD} at 3.0 volts

The timing of the input stimuli is also different from that of the start-up locked sense amp circuit. The sense amp enable signal is low until some time after one of the read word lines goes high. The timing of the sense amp enable signal in these simulations was kept the same for all word sizes to compare output development time. In reality, the sense amp circuits with less words per column could have been enabled much sooner than those with fewer words per column.

The power simulation results for the 3V simulation are shown in Figure A-5. Notice that, since this sense amp circuit is faster than the start-up locked circuit, the time scale is reduced in the cross-coupled enabled simulation, for the same three read access cycle. The resulting power seems to be greater for this simulation, but since the power for CMOS circuits is dependent on the frequency [10]:

$$P = C * V^2 * F \qquad (5)$$

the normalized power, or power per MHz, for this circuit turns out to be less than that of the start-up locked circuit. One additional note to the power is that there is very little variation in power from the 128 word to the 1024 word column.

Full Current-mode Sense Amp Simulations

The input stimuli for the full current-mode sense amp simulations are very similar to those of the cross-coupled enabled sense amp, except that the sense amp enable for the full current-mode sense amp is the inverse of the enable for the cross-coupled enabled. The frequency of operation and the relative timing of the input signals is the same.

The simulation results are shown in Figure A-10 and Figure A-11 for V_{DD} at 5V and 3V, respectively. Notice that the bit lines in these simulations remain very close to ground, unlike the bit line voltages of the cross-coupled enabled sense amp, where the bit line voltages remained near V_{DD}.

Chapter 6

APPLICATION TO THE LUCENT REGISTER FILE

Circuit Modifications for the Cross-coupled Enabled Sense Amp

As stated previously, current-mode sense amps have already been used in memories. The application of a current-mode sense amp to a memory, therefore is not unique. However, what distinguishes this application from other memory applications is that:

1) the register file memory is an 8 transistor true dual port cell, and

2) the register file is asynchronous.

The first point is both an asset as well as a liability. The dual port nature of the memory cell provides the cell with separate read and write access. As mentioned in chapter 4, the memory margin depends on the ratio of the size of the access device to the size of the pull-down device. Since the dual port cell has both read and write access devices, if both are active simultaneously, calculations of the memory margin must include both access devices. The memory margin is, therefore, the ratio of the sum of the widths of the read and write access devices to the width of the pull-down device.

Of course, the write access device is usually turned on with the intent of overwriting the cell, and a lower margin actually helps to speed up the write. However, there are

occasions when the write access device is turned on without the intent to overwrite the cell.

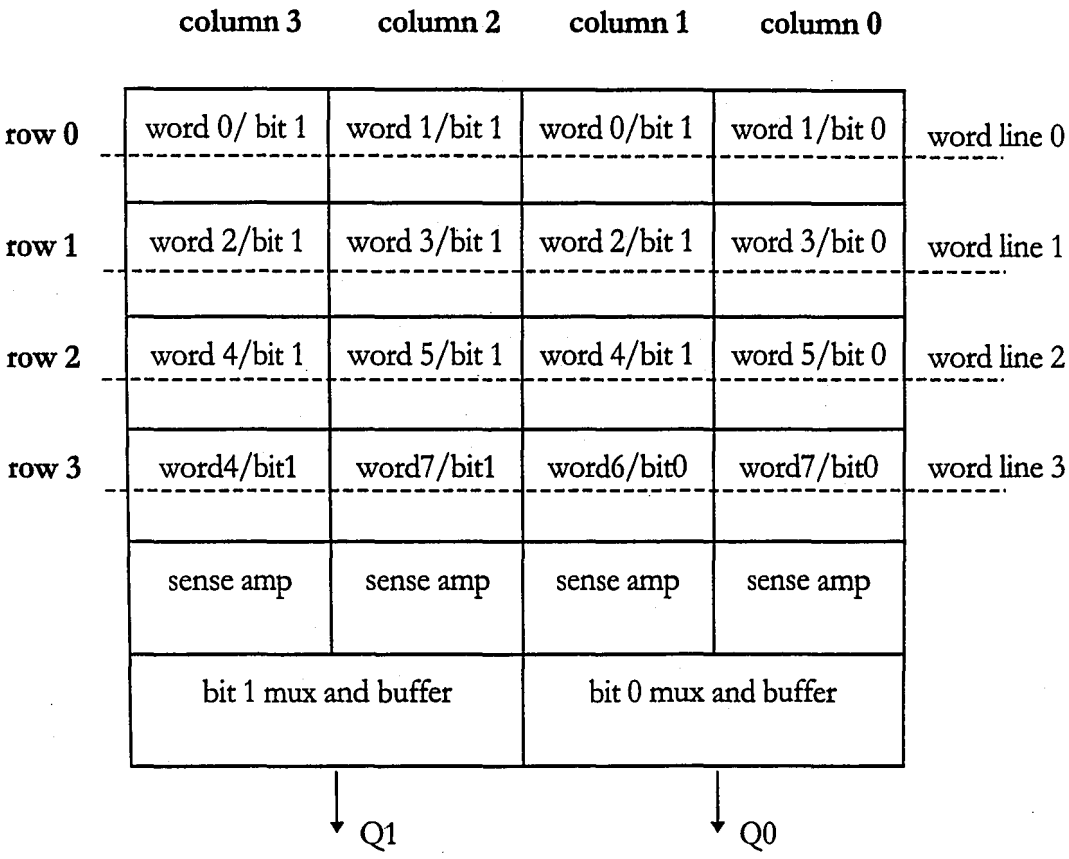

Figure 14. Memory Bit Map for a 2 Column Multiplexed Core

For example, even though the write cycle is fast, the write word line is turned on with the clock, and could remain on for a full half clock cycle. Once the memory has been written, the cell must be stable and the cell margin is important. Since the read operation is asynchronous, a read could occur immediately after the write has already been accomplished and before the write access device is turned off. In that case, the

cell margin depends on the size of both access devices. For that reason, it is important to size the memory devices according to both access devices being turned on simultaneously.

Additionally, when the number of words of memory is very large with respect to the number of bits per word, the aspect ratio of the register file becomes very tall and narrow. This irregular aspect ratio can cause problems when the register file is placed in a VLSI chip.

It is sometimes desirable to divide a bit, a single column, of memory into two or more columns, with the input and output of each column multiplexed to a single bit. In order to accomplish this, the word line for a particular row accesses only one column at a time. For example, if the memory is multiplexed into two columns, the leftmost of two adjacent columns would be accessed for even words and the rightmost column of the pair would be accessed for odd words. Going across a row, every two memory cells would belong to a bit in a word. A two column multiplexed memory is shown in Figure 14.

The most common method to prevent overwriting the memory in a column that is not selected, is to hold the write bit lines high, in that column, while the write word line turns on the access devices to the entire row, including the column that is not to be overwritten. Since writes are accomplished by pulling the memory cell low [1], the memory cell in the column with the bit lines held high would not be affected. If a read were to occur simultaneously at this word, the charge coupling would be the equivalent

of the charge of the read bit lines and write bit lines coupled to the memory and this would be the worst case for memory margin.

While decreased memory margin is a problem with the dual port cell, the dedicated read bit lines is an asset to the application of the clamped bit line current-sense amp. Single access memories share a single set of bit lines for both reading and writing by time multiplexing the operations. Write operations require one of the bit lines to be fully grounded. For a read to occur after a write, both bit lines must be pulled high, the process previously referred to as precharging.. This negates the advantage of the clamping devices in the sense amp. With separate read bit lines, it is not necessary to precharge the bit lines before a read, since they are never pulled low for a write operation.

The asynchronous nature of the register file must also be dealt with. So long as reads and writes occur at different words, there is no problem. The read bit lines and word lines are completely separate from the write bit lines and word lines allowing the two operations to continue without interference from each other. Except that, when a read and write take place at the same word, the value of the output depends on whether or not the new data was in the cell when the read occurred. Even worse, if the cell is in the process of being written, the data in the memory could be indeterminate when the read occurs. The register file circuit must be modified to determine what data to output in that case.

The dual port register file described in chapter 3 was modified for this project according to the block diagram shown in Figure A-2. Several additions to the original register file, shown in Figure A-1, were made, but the modifications were minor and the original size of the register file block was not changed.

The precharge pulse generated by the address transition detector in the original register file was used advantageously as a sense amp enable signal for the cross-coupled enabled sense amp circuit. Recall from the operation of the cross-coupled enabled sense amp that the sense amp enable had to pulse high only long enough to preset the sense amp. Since this operation is faster than precharging the bit lines, the pulsewidth of this signal can be reduced.

In addition to the sense amp preset action, the read word line must be turned on prior to enabling the sense amp. The address transition pulse was also used to enable the read word line, shown logically by the OR gate in Figure A-2.. With these modifications in place, the modified address transition pulse simultaneously turned on the read word line and preset the sense amp. The trailing edge of the pulse enabled the sense amp, and the output data was generated some time after the trailing edge of the pulse.

In order to solve the problem of simultaneous reading and writing to the same word, a comparator was added to the circuit, the ADM block in Figure A-2, to compare the value of the read and write address. This circuit is composed of XNOR logic gates to

compare equivalent bits of each address, an AND gate to combine the results, and a pulse generator to generate a pulse if a match occurs.

The pulse generated by this address match detector is used to perform a read in the same way as the address transition detect pulse from the read addresses. Further, the match pulse is combined with the address transiton pulse so that, if the read address has just changed due to the start of a read cycle, the sense amp preset condition will be extended until the new data has been written to the memory and is projected onto the read bit lines.

In this manner, whenever a write occurs at the same word that the read address is pointing to, the new data is read. This operation matches the original functionality of the register file.

These minor changes were made to the Lucent 0.6 μm. standard cell register file. Various sizes of the modified register file were manufactured on a single test chip to verify the design. When putting together the final circuits for the test chip, two additional modifications were realized.

Because of the dependence on very minute current differences which determine the sense of the data being read, the sense amp must be very balanced. In addition to the elements of the sense amp being balanced, the bit lines must also be balanced. Minor differences in the routing capacitance or the access device diodes connected to the bit lines can cause a major imbalance when multiplied by the number of words per column.

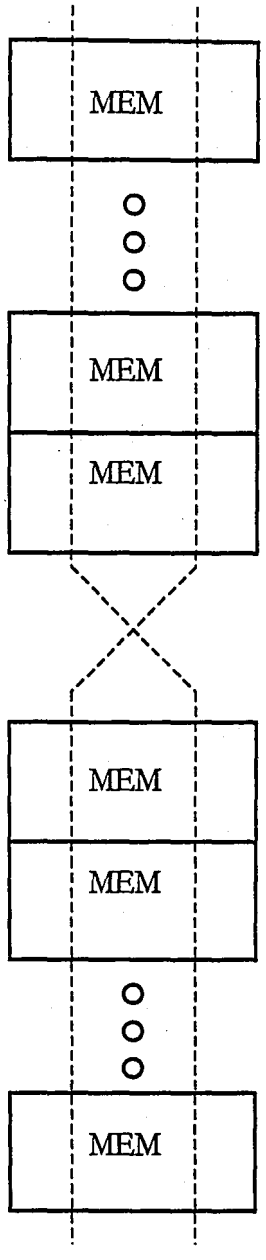

Figure 15. Memory Column with Twisted Bit Line Pairs.

In order to minimize imbalances in the bit lines along a column, the bit lines were connected to opposite sides of the memory for half of the column. The bit lines were physically crossed midway through the column, separating the top and bottom of the column as shown in Figure 15. In order to maintain the correct sense of the data, the write bit lines were also crossed at the center of the column.

Circuit Modifications for the Full Current-mode Sense Amp

In addition to the circuit modifications that were made for the cross-coupled enabled sense amp, the full current-mode sense amp circuit requires modifications to account for the inverted sense of the sense amp, including a modified memory cell with an isolated read as shown in Figure 11. One minor modification that is needed is the inversion of the enable signal to the sense amp.

The memory cell modification, however, is not minor. Two additional devices are required in a cell, which is already as compact as possible. This addition is possible, however, with only a slight increase in the size of the memory cell. The eight transistor dual port memory cell used previously has an odd number of pairs of like transistors. There are two matching PMOS devices that can be placed next to each other, but then there are three pairs of matching NMOS devices that cannot be easily arranged to minimize the area of the cell.

The arrangement of transistors of the original register file cell is shown in Figure A-12. There is room to fit the additional two transistors next to the pair of isolated NMOS devices. Some additional area is required for routing. The layout of the transistors in the modified cell is shown in Figure A-13.

Despite the additional area, this cell has several advantages. Since the read bit lines are not connected to the internal nodes of the memory through an access device, no charge sharing takes place during a read. The read is therefore completely isolated

from the write. Because of this, the devices in the cross coupled latch can be minimized, which will speed up the memory write operation.

Another advantage of this cell is that the devices in the read access path can be adjusted independently, except that increasing the size of the memory pull-up devices that are controlled by the internal nodes of the memory adds loading onto the memory nodes. Having the ability to independently adjust the size of the read devices simplifies the adjustment for sufficient bit line separation, which guarantees the noise margin of the read bit lines.

Unfortunately, these modifications to the memory cell, which affect the overall size of the cell, require adjustments to the peripheral cells of the register file in order to match the memory cell. Therefore, this sense amp was not included on the test chip with the cross-coupled enabled sense amp.

Chapter 7

SUMMARY AND CONCLUSIONS

The current mode sense amps showed a marked improvement in increased speed and reduced power over the voltage differential start-up locked sense amp. A comparison of the speeds is shown in the graph of Figure 16.

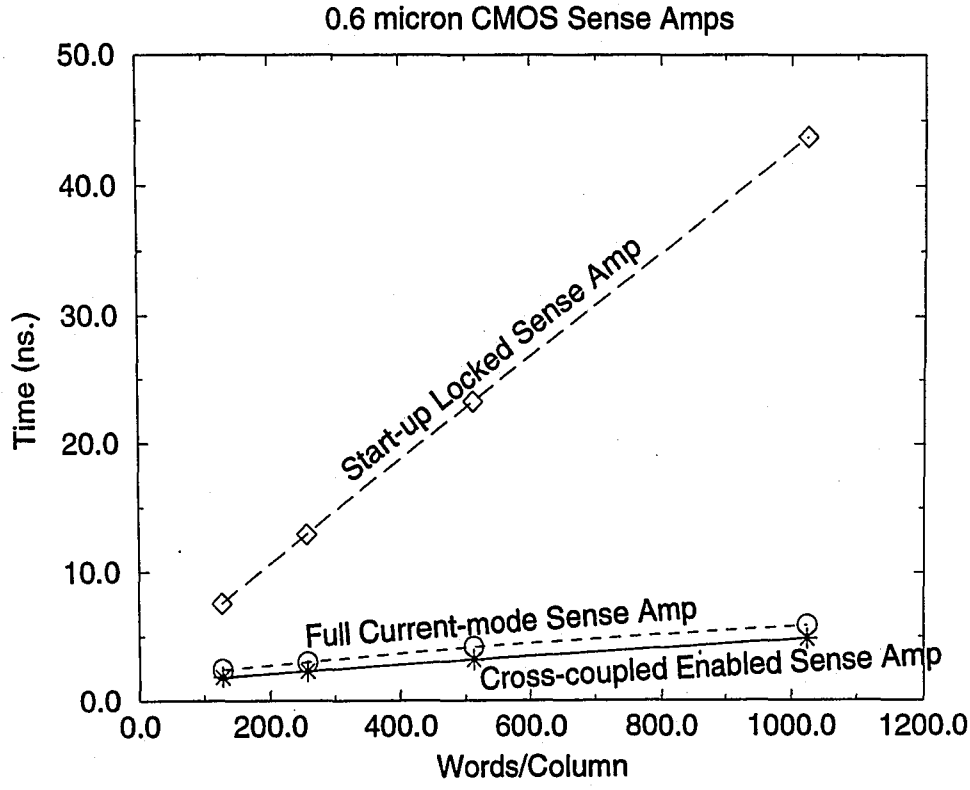

Figure 16. Read Access Time Comparison.

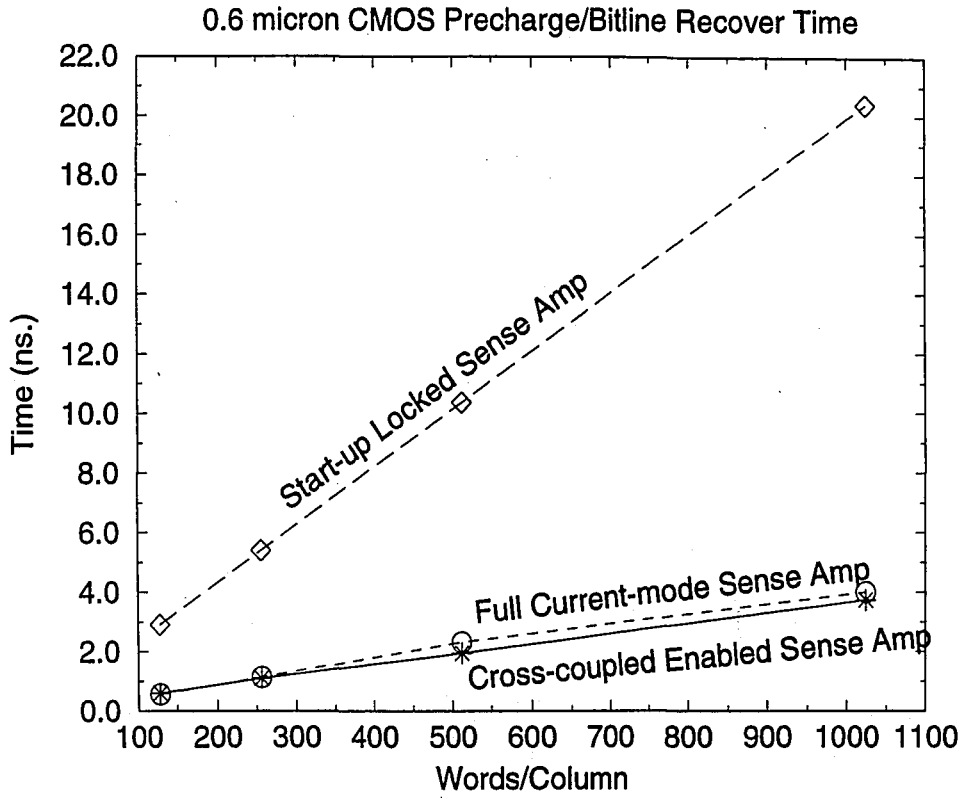

Figure 17. Precharge Time Comparison.

Figure 17 shows a comparison of the precharge portion of the read cycle for each of the sense amp circuits. For the current-mode sense amp circuits, this is the time for the read bit lines to recover to full rail, V_{DD} or ground. The evaluation time of the sense amp circuits is compared in Figure 18.

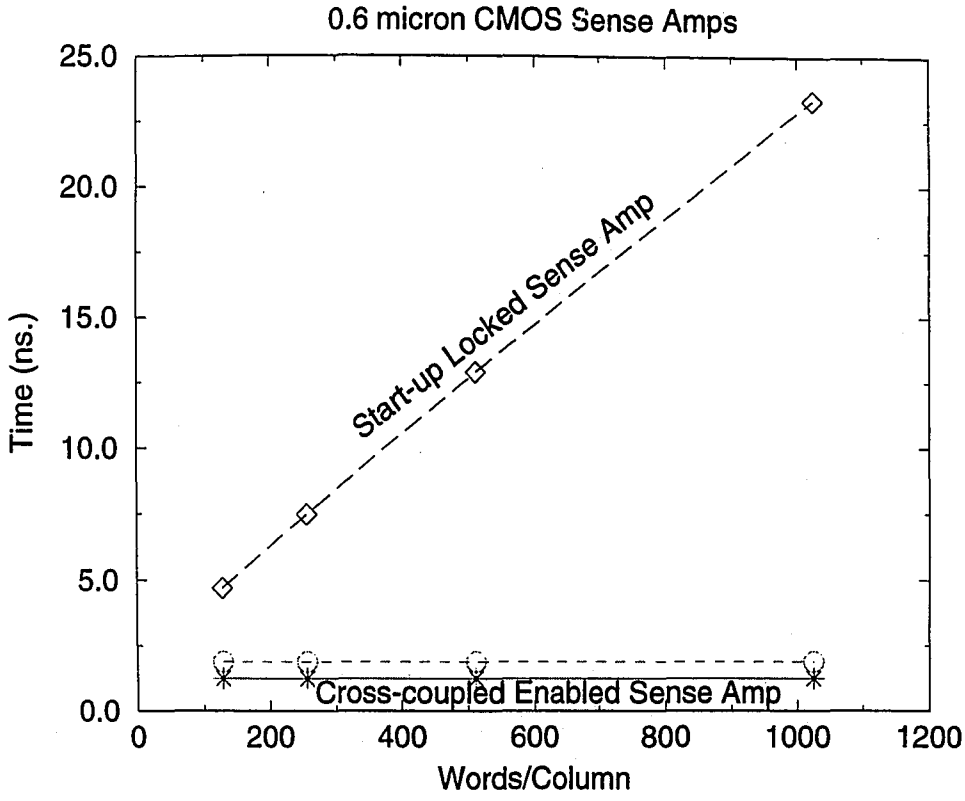

Figure 18. Sense Amp Evaluation Time Comparison.

While the speed was increased, the power per MHz was reduced, as shown in Figure 19. These two factors resulted in a remarkably better power-speed product for both current-mode sense amplifier circuits, as shown in Figure 20.

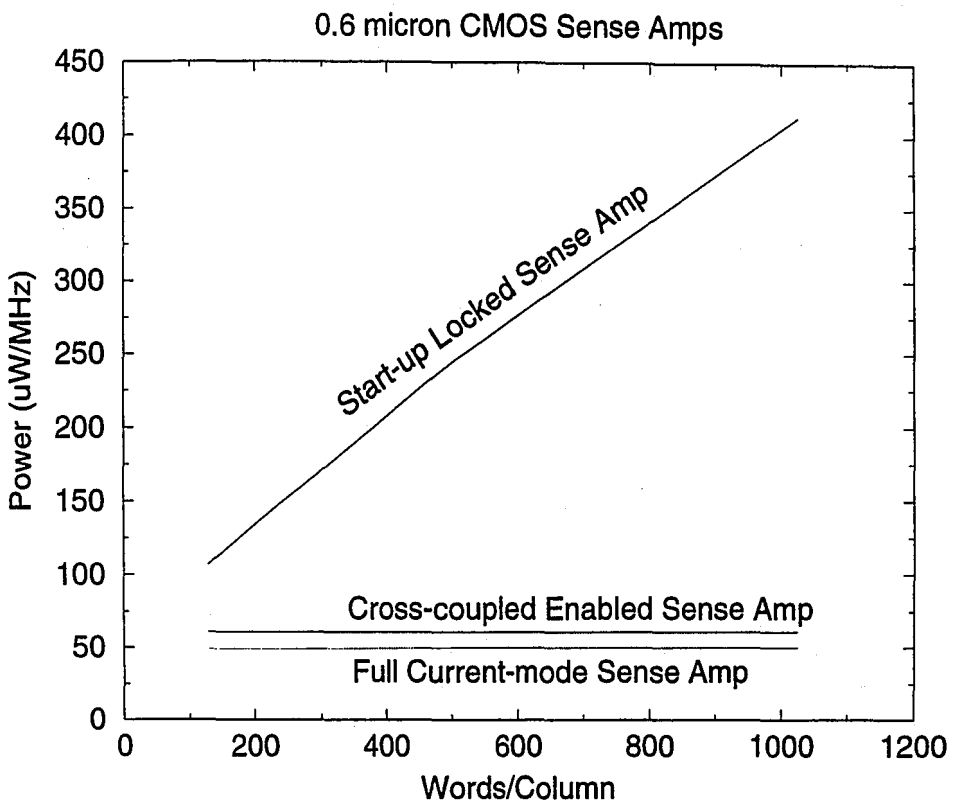

Figure 19. Power Comparison.

While both current-mode sense amps showed a marked improvement in the power-speed product, this is not without some disadvantages. The layout of the circuit must be done very carefully, with particular attention to balance. Additionally, since the voltage differential is so small, the bit lines are susceptible to noise. Very small amounts of coupling from external sources, such as nearby wire routes, could cause the sense of the bit lines to be disturbed. For that reason, the read bit lines should be separated from other logic routes, such as the write bit lines. They should also be shielded wherever possible.

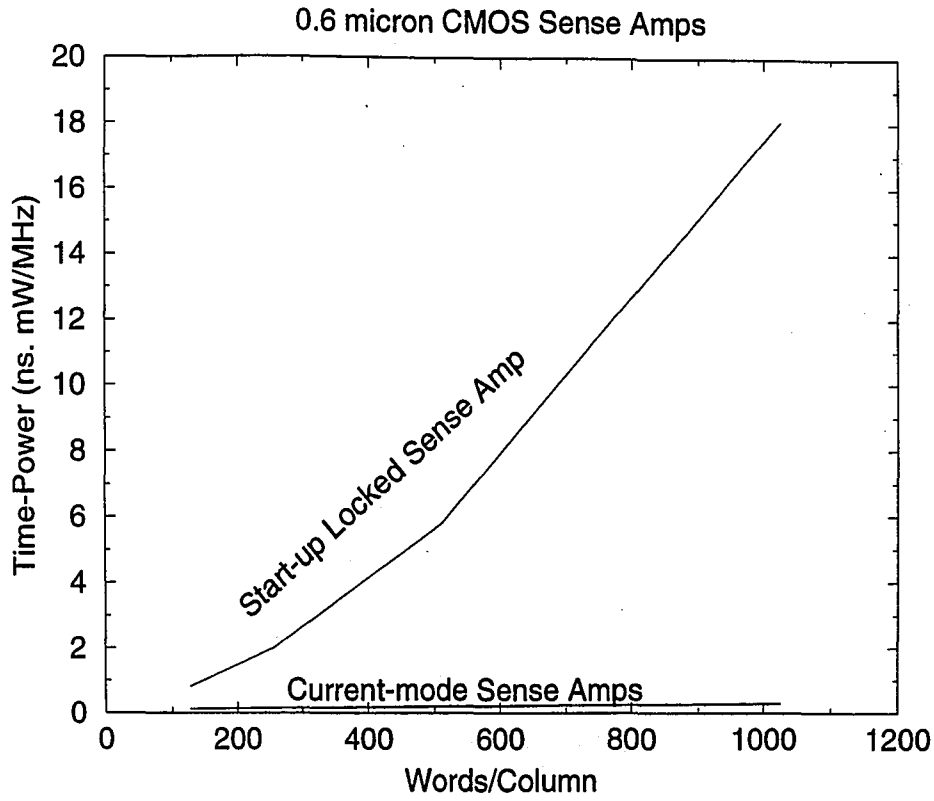

Figure 20. Speed-power Product Comparison.

These circuits are ideally suited for smaller technologies. As technologies shrink, parasitics tend to dominate performance. The low sensitivity to parasitics demonstrated by this work, clearly shows that current-mode sense amps are necessary in future technologies. Also as power is becoming an important element in future technologies, these current-mode sense amps seem to be an ideal solution.

REFERENCES

1. B. Prince. *Semiconductor Memories*, 2d ed. John Wiley & Sons, 1991.
2. L. Fenstermaker. "Macrocell Usage Statistics". *Lucent Technologies Internal Memorandum*, March 16, 1998
3. D. Patterson & J. Hennessy. *Computer Architecture A Quantitative Approach,* 2d ed. Morgan Kaufmann Publishers, Inc, *1995.*
4. *AT&T Microelectronics HS600C &LP600C CMOS Standard-Cell Libraries Data Book*, April 1994
5. S. Sze. *Semiconductor Devices Physics and Technology*, John Wiley & Sons, 1985.
6. M. Annaratone. *Digital CMOS Circuit Design*, Kluwer Academic Publishers, 1986.
7. K. J. O'Connor, "Modification to 0.9µm REGFILE Macrocell to Obtain Zero DC Power Dissipation," *AT&T internal memorandum*, May17, 1990.
8. T. N. Blalock and R. C. Jaeger, "A High-Speed Clamped Bit Line Current-Mode Sense Amplifier," *IEEE Journal of Solid-State Circuits*, Vol. 26, No. 4, April 1991.
9. E. Seevnick, et. al, "Current-Mode Techniques for High-Speed VLSI Circuits with Application to Current Sense Amplifier for CMOS SRAM's," *IEEE Journal of Solid-State Circuits,* Vol.26, No. 4, April 1991.
10. N. H. E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design*, Addison-Wesley, 1994.

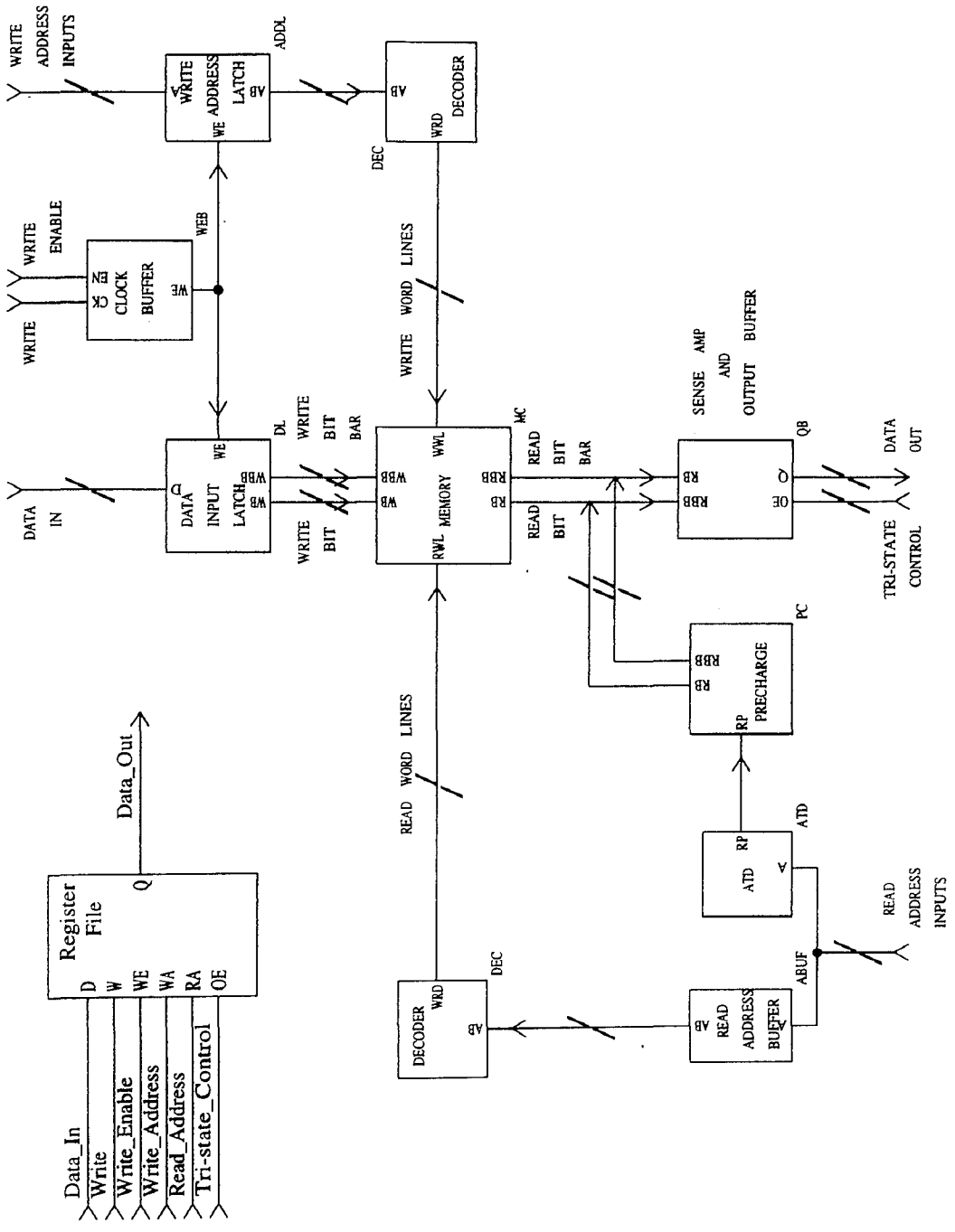

Figure A-1. Register File Block Diagram

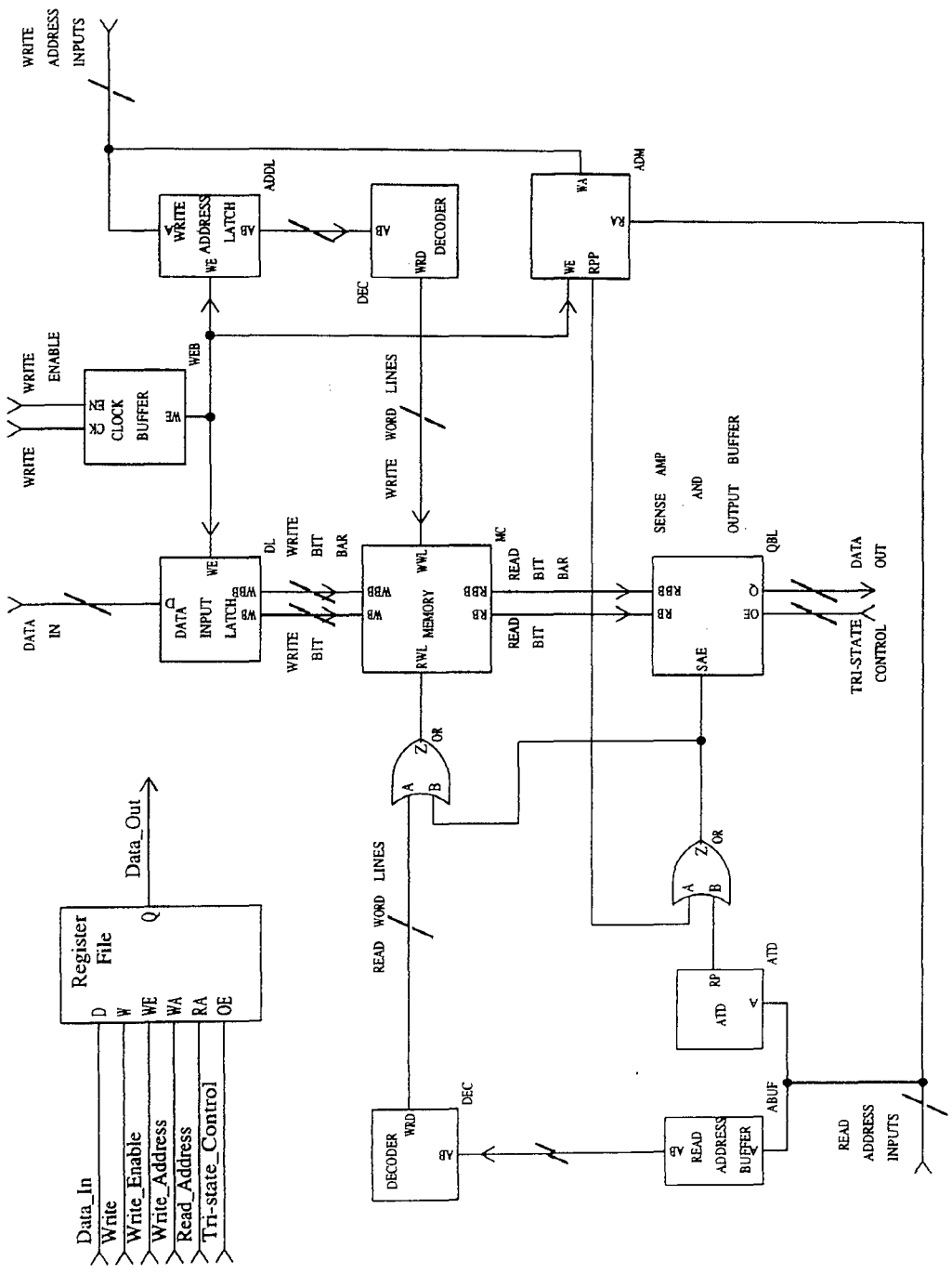

Figure A-2. Register File with Current-mode Sense Amp

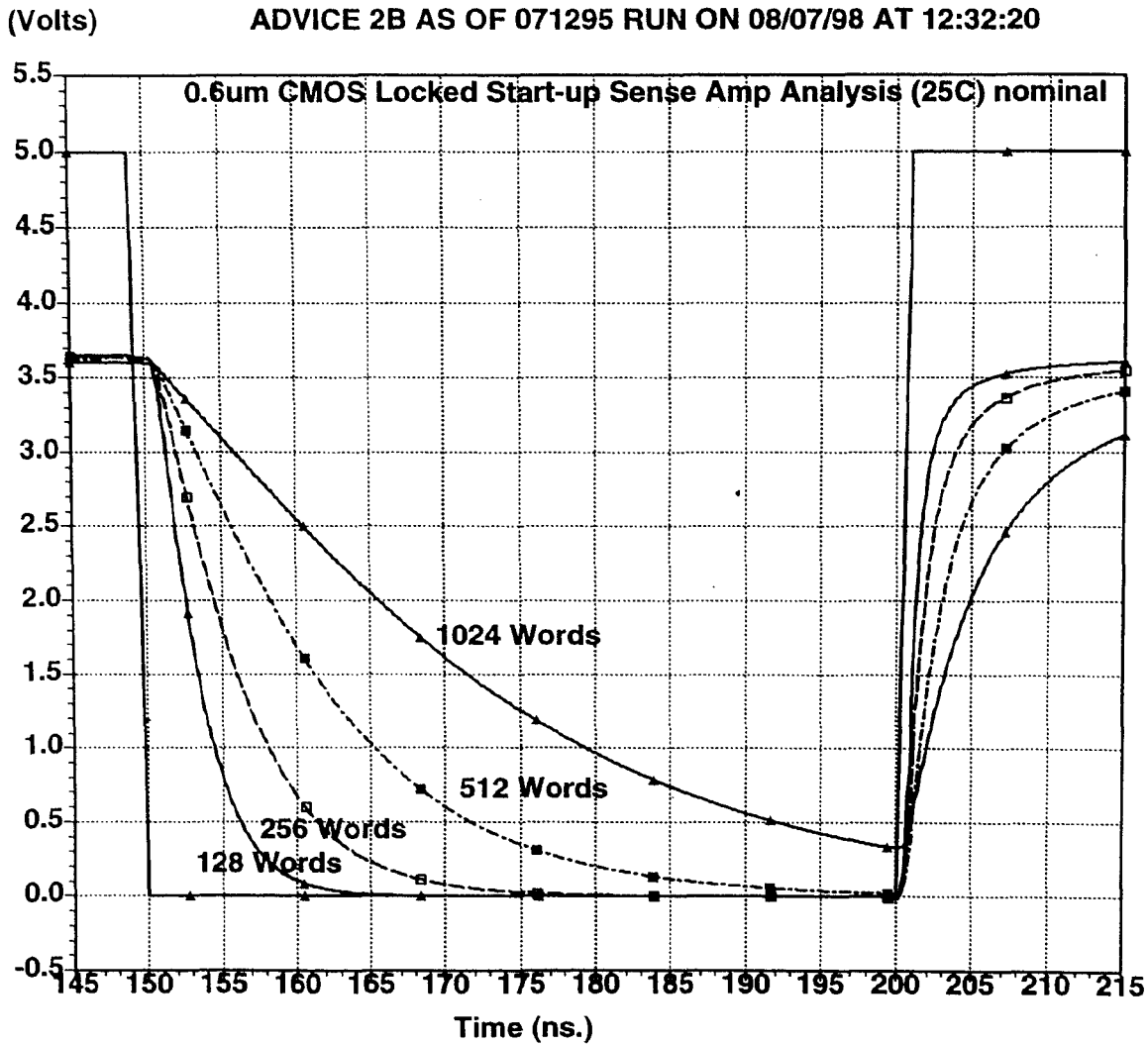

Figure A-3. Bit Line Development for Various Number of Words in a Column

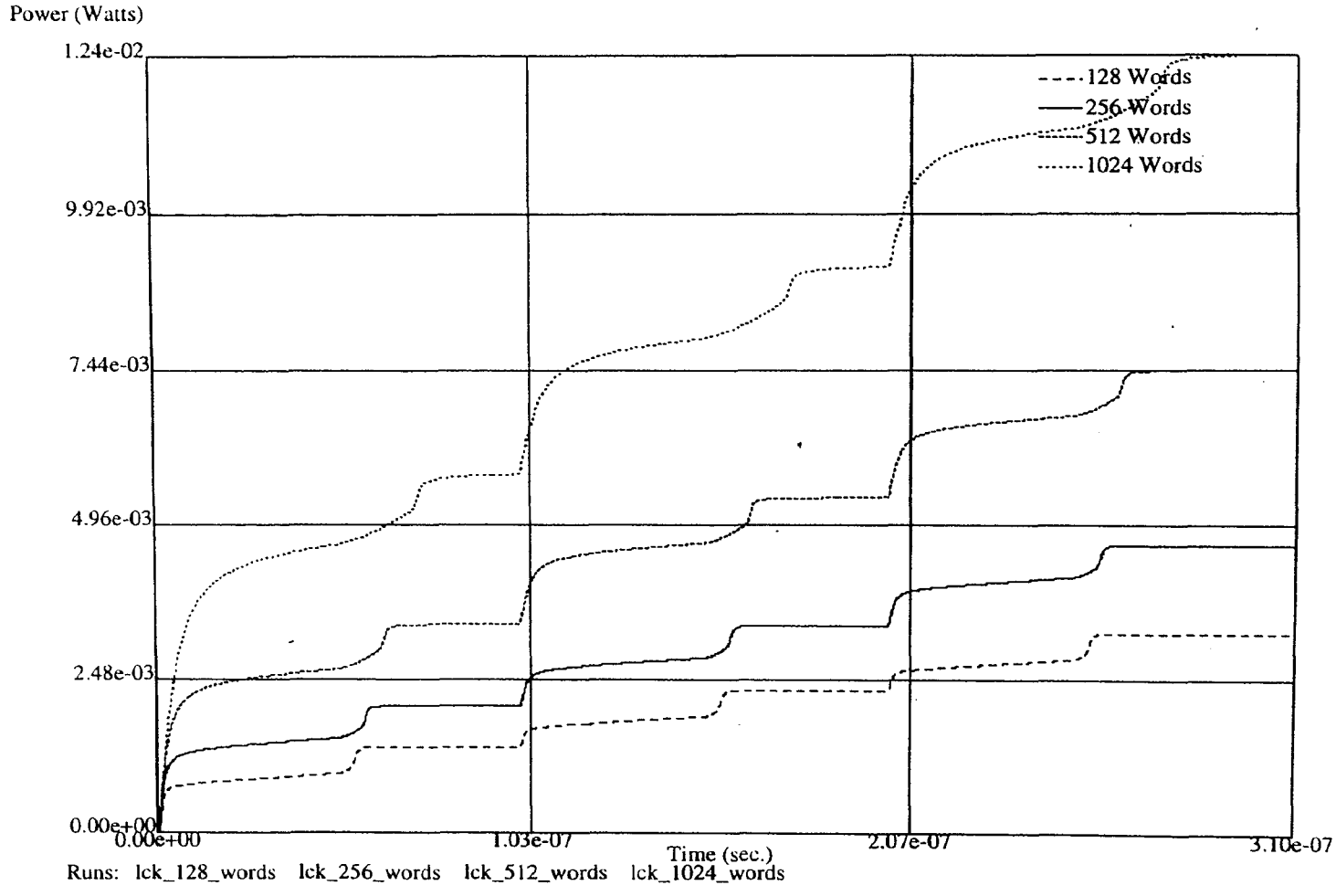

Figure A-4. Power Results for the Locked Start-up Sense Amp at 5V

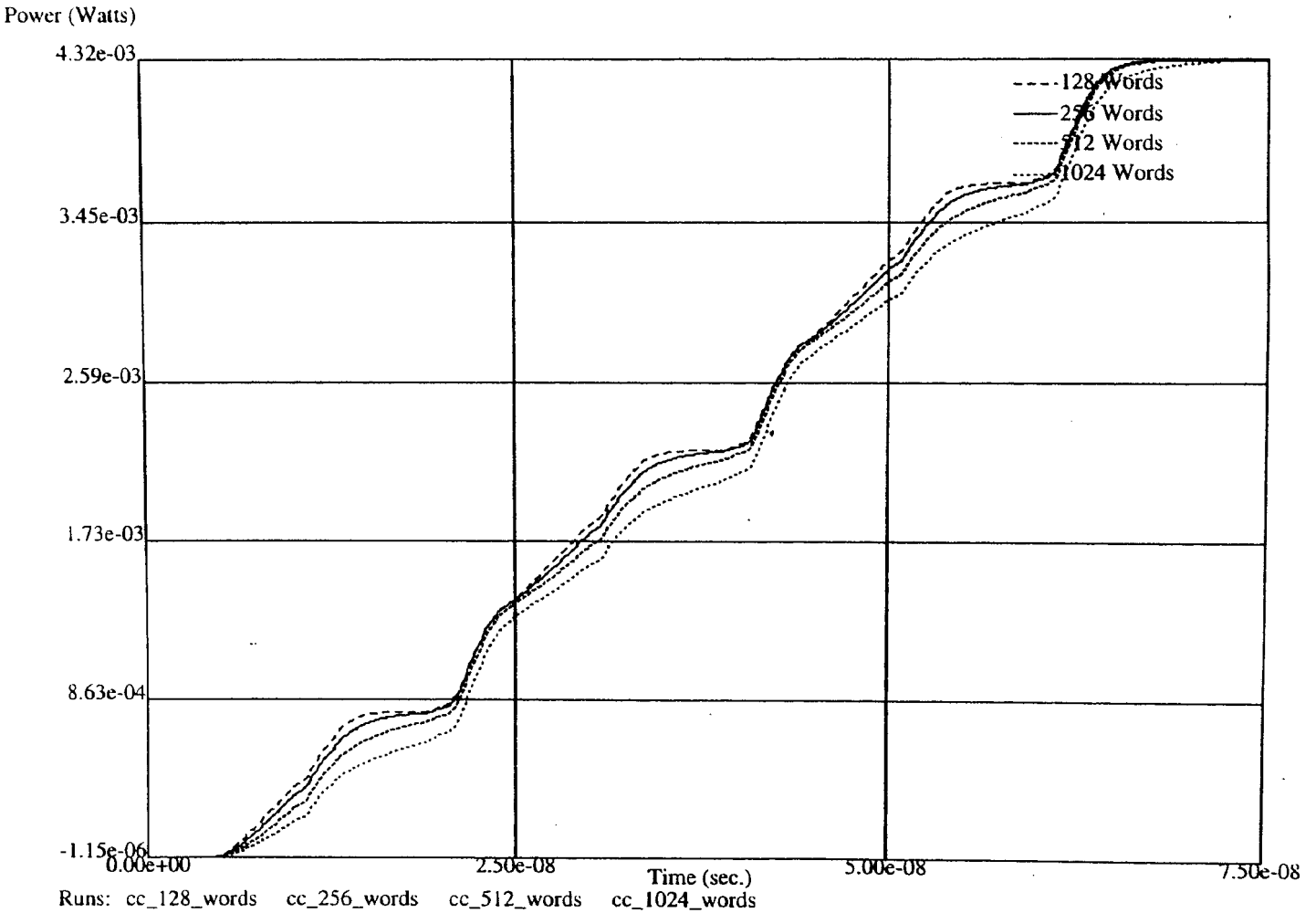

Figure A-5. Power Results for the Cross-coupled Enabled Sense Amp at 3V

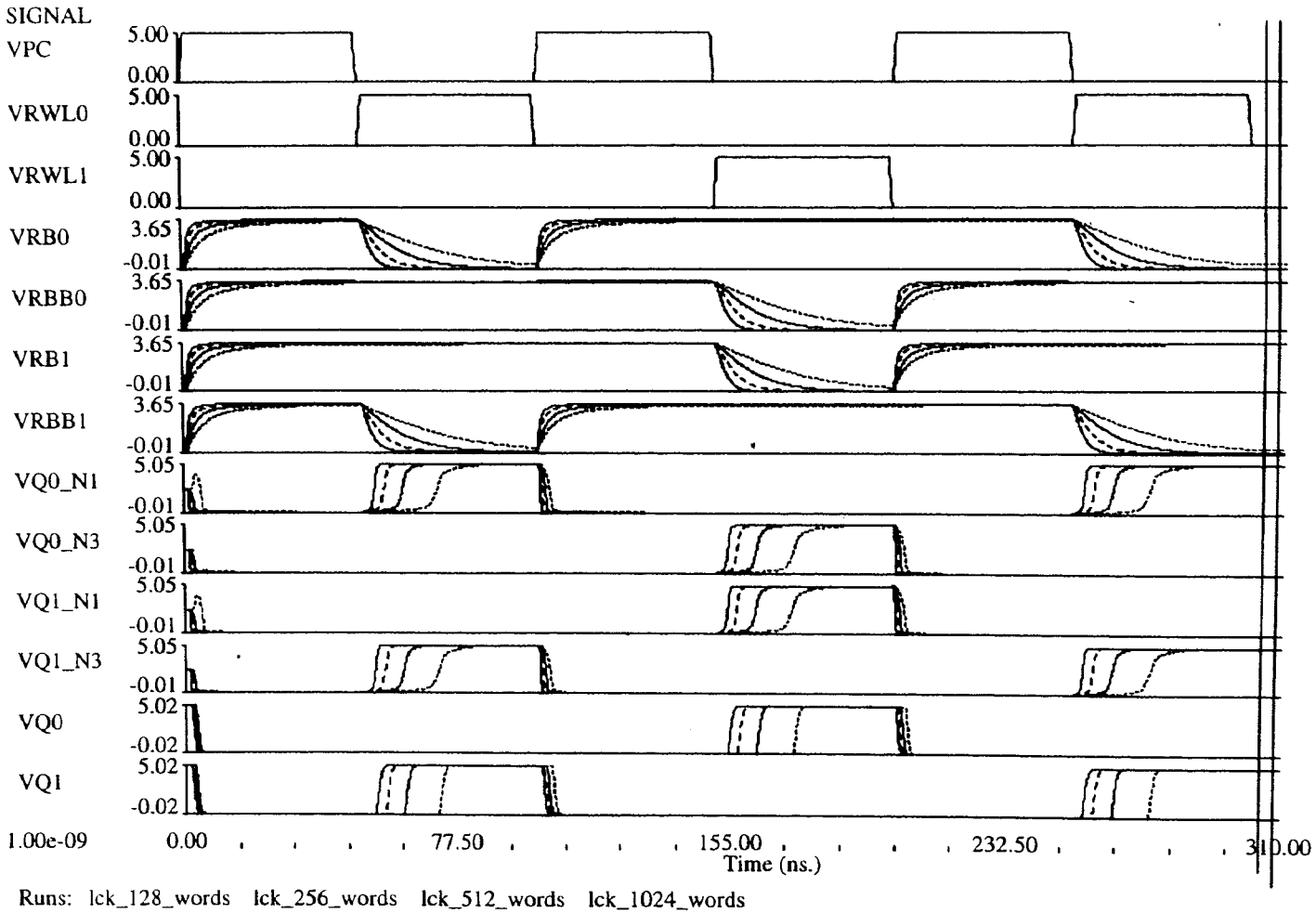

Figure A-6. Voltage Results for Locked Start-up Sense Amp at 5V

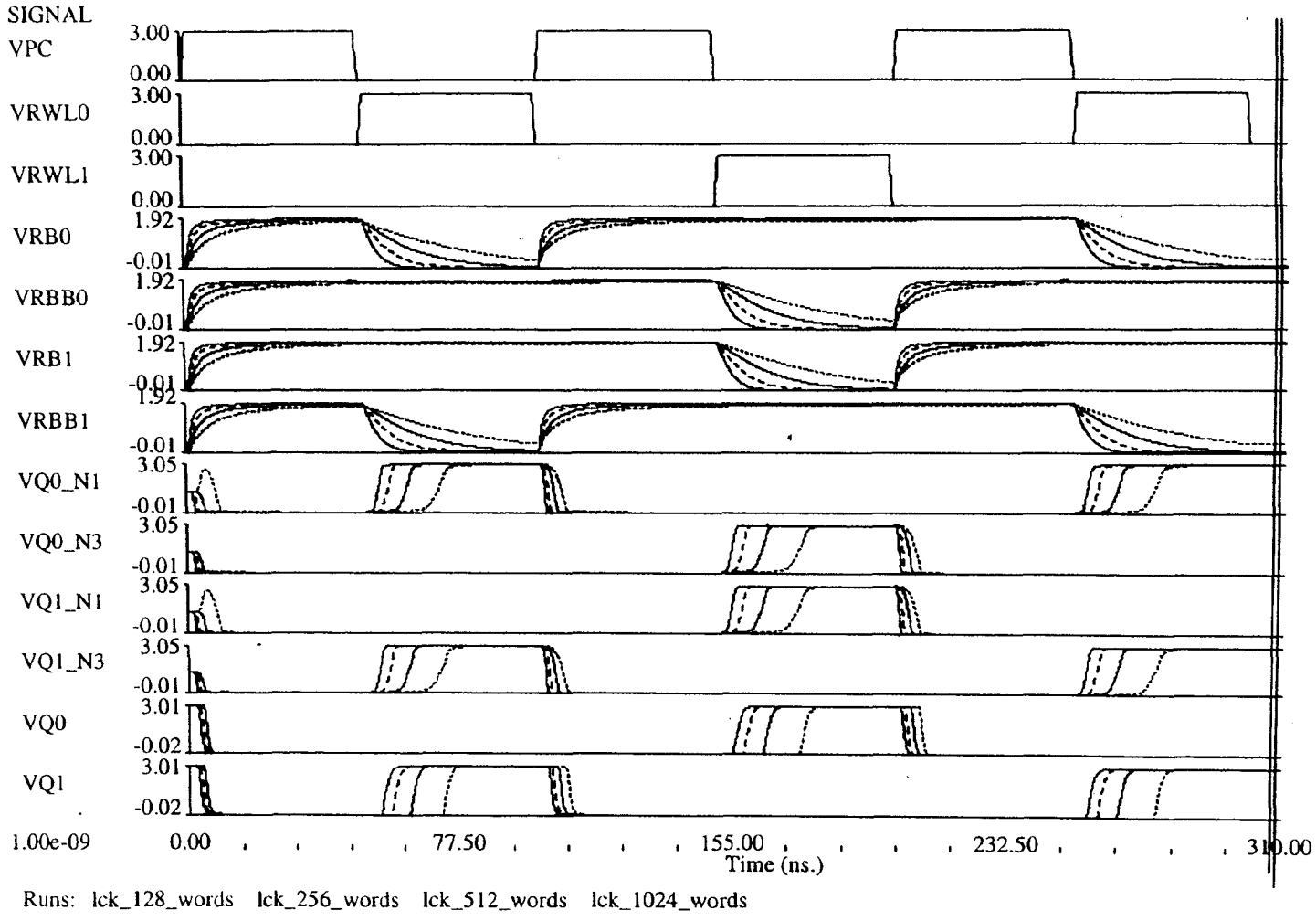

Figure A-7. Voltage Results for the Locked Start-up Sense Amp at 3V

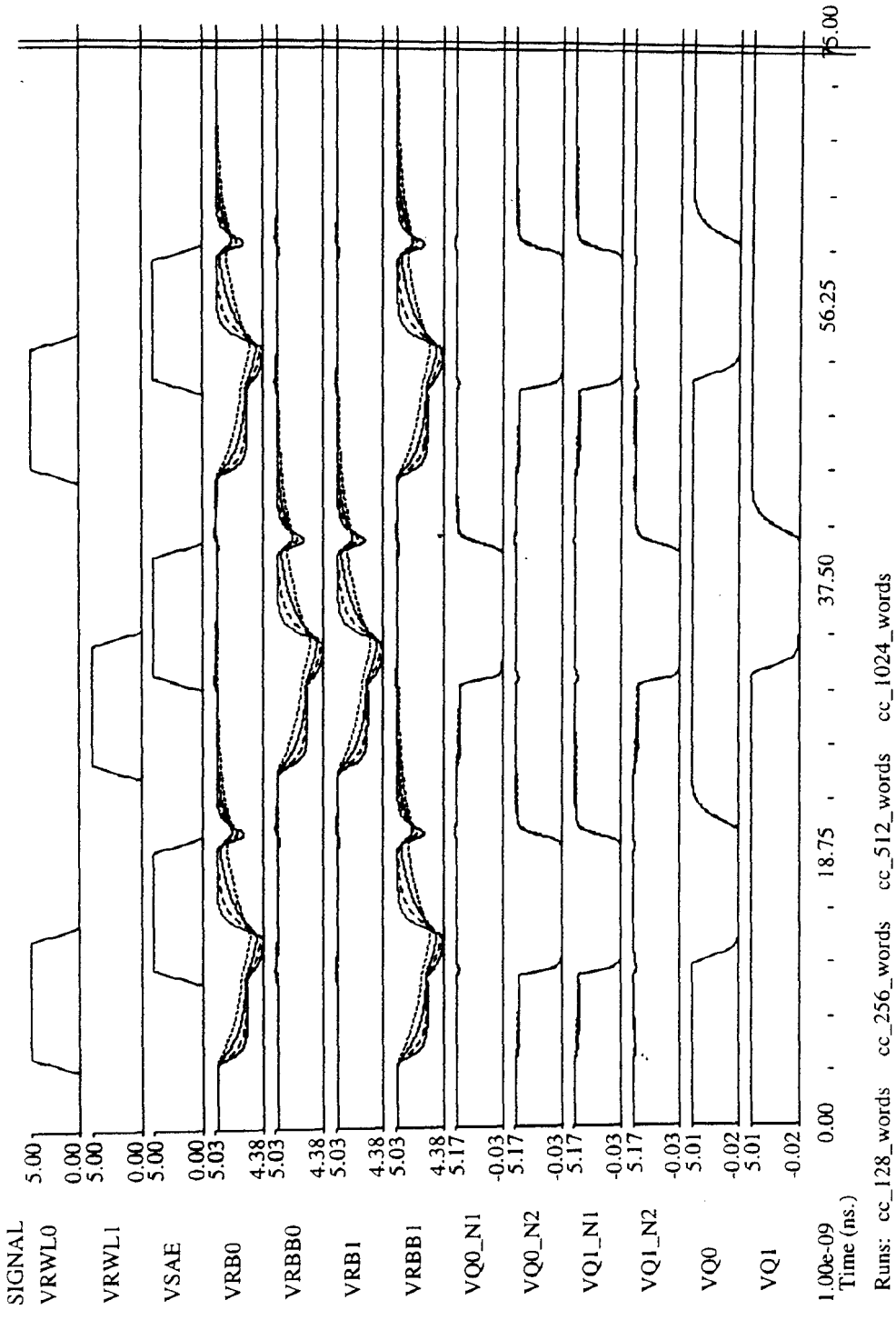

Figure A-8. Voltage Results for Cross-coupled Enabled Sense Amp at 5V

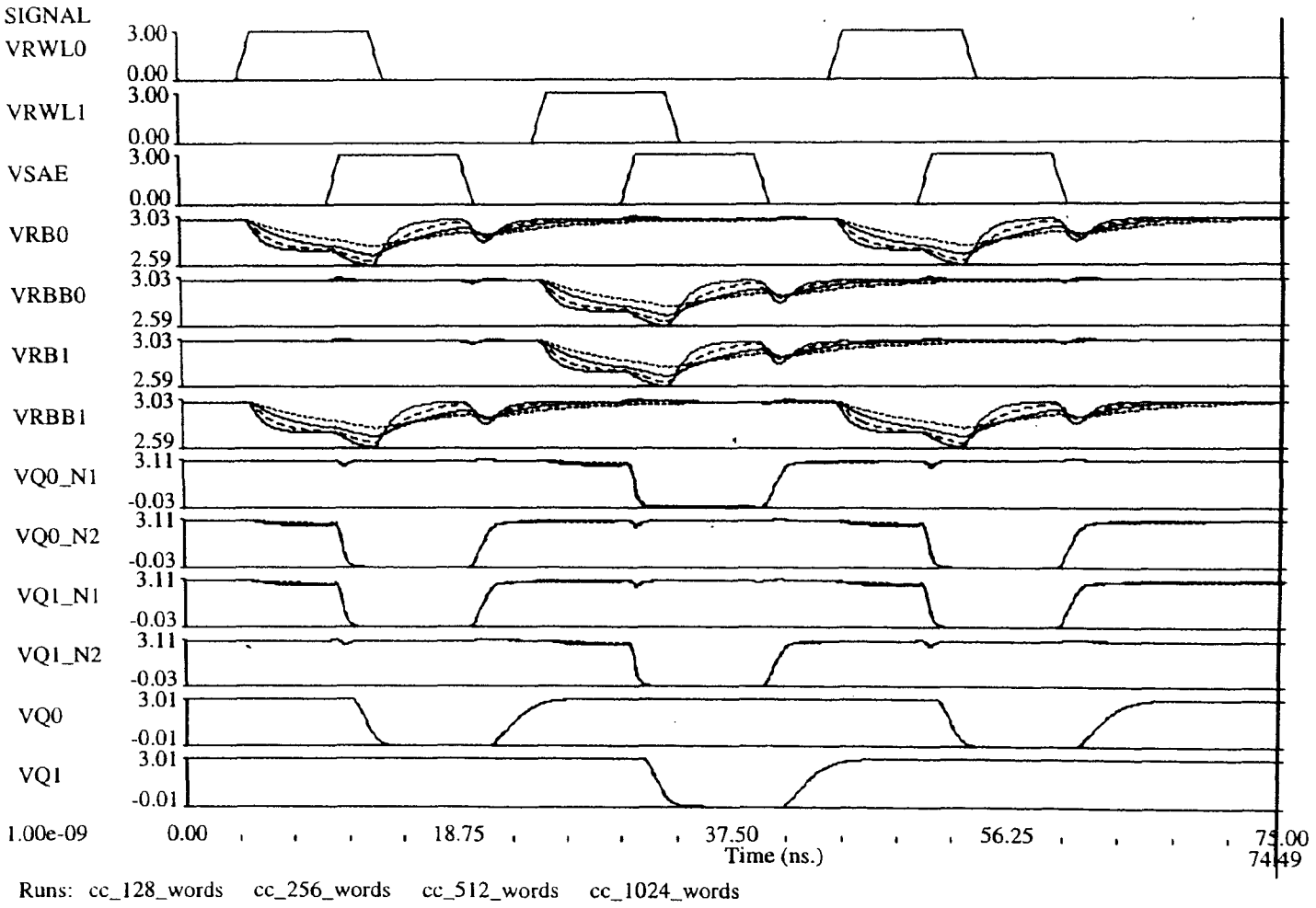

Figure A-9. Voltage Results for Cross-coupled Enabled Sense Amp at 3V

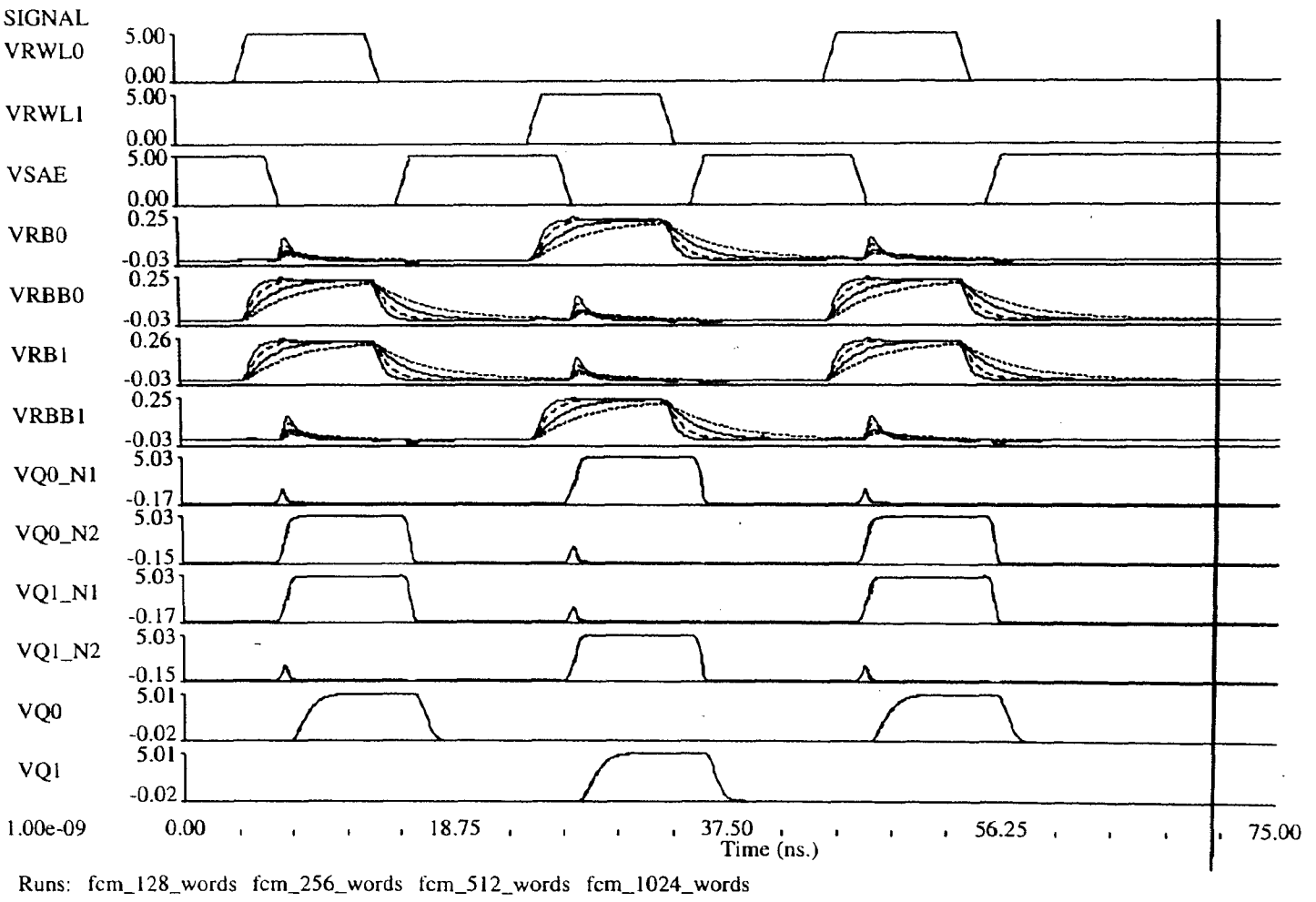

Figure A-10. Voltage Results for Full Current-mode Sense Amp at 5V

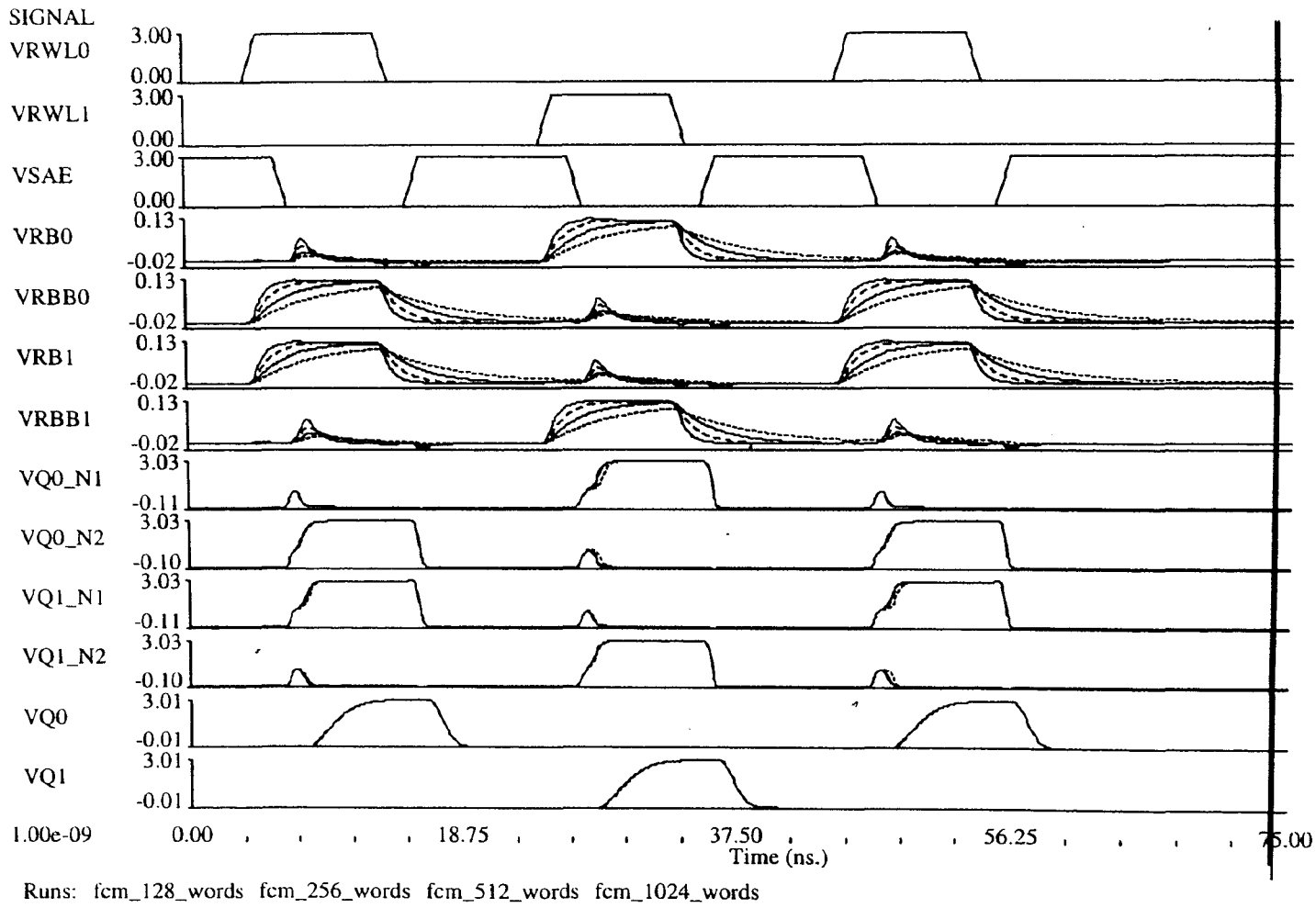

Figure A-11. Voltage Results for Full Current-mode Sense Amp at 3V

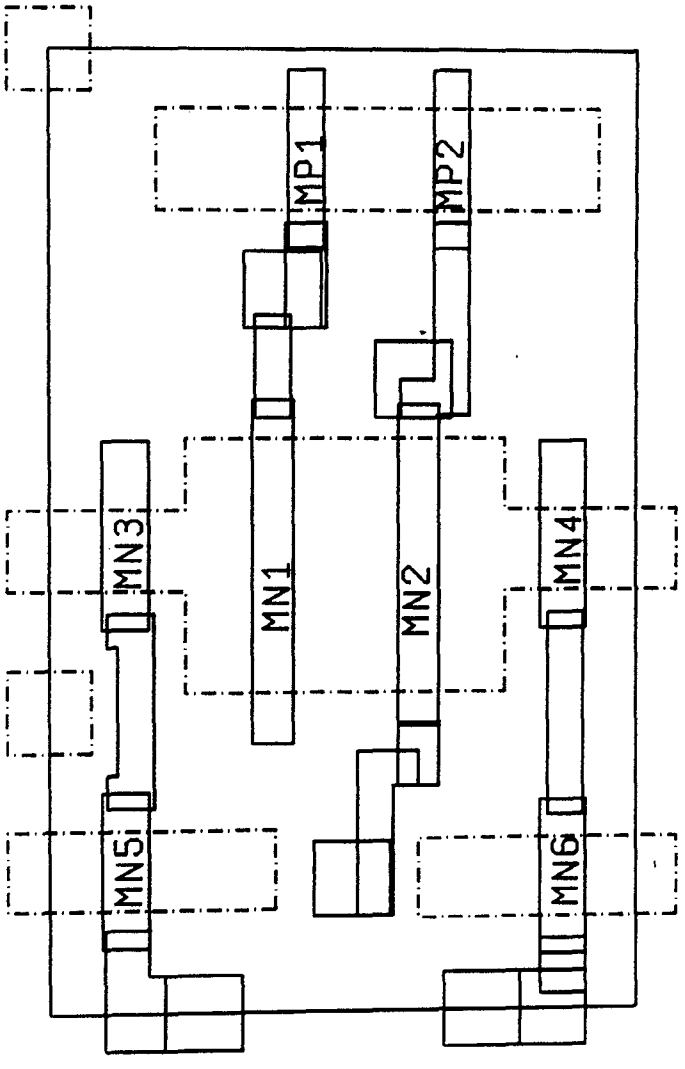

Figure A-12. Transistor Layout of 8 Transistor Dual Port Memory Cell

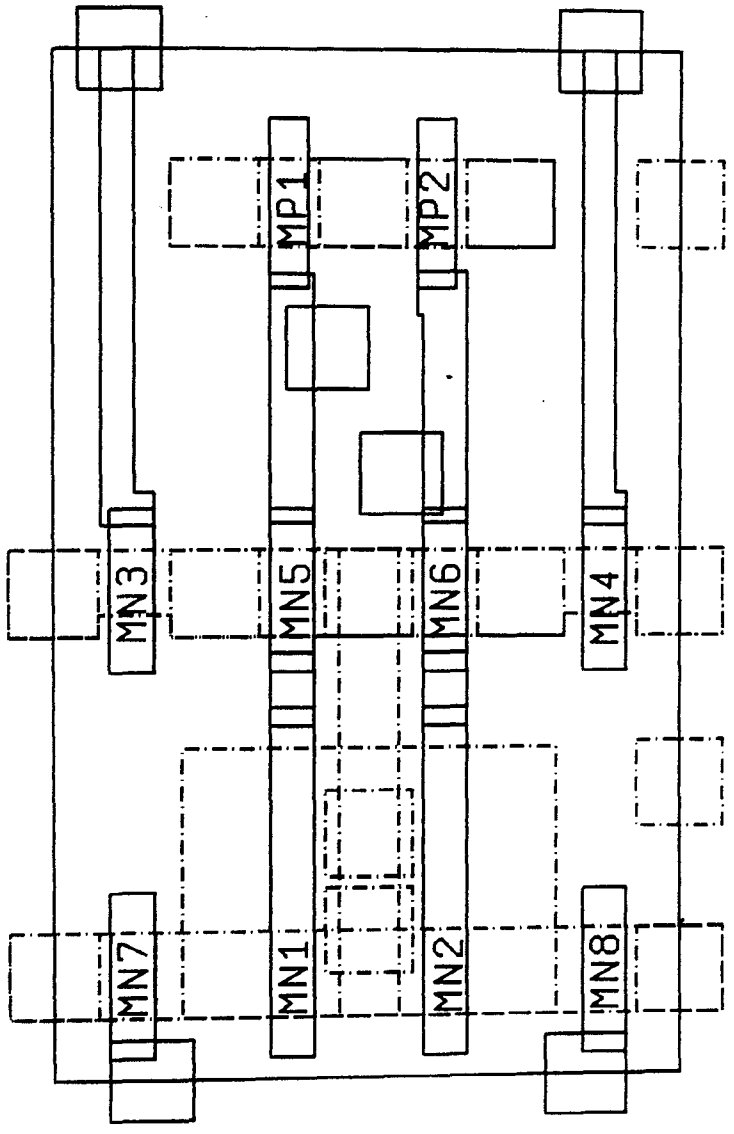

Figure A-13. Transistor Layout of 10 Transistor Dual Port Memory Cell

BIOGRAPHY

Larry Fenstermaker is a married grandfather and father of two. Born in East Allen Township, Northampton County, Pennsylvania on December 28, 1951, he is the son of Margaret James. After graduating from Northampton Area Joint Senior High School in June of 1969, he attended Lehigh University. He received an Associate in Science in Engineering from Northampton County Community College in December, 1979 and a Bachelor of Science in Electrical Engineering from Lafayette College in June, 1986.

As a member of the Technical Staff at Lucent Technologies since 1984, Larry is currently a circuit designer in the ASIC Memory and Compilers Department of the Analog and Memory Core Laboratory in the Microelectronics Division. He has presented a paper entitled "A Low-Power Generator-Based FIFO using Ring Pointers and Current-Mode Sensing" at the 1993 International Solid State Circuits Conference and has been awarded a patent for a "FIFO With Word Line Match Circuits for Flag Generation".

END OF TITLE