AN INTEGRATED SEMICONDUCTOR MEMORY SYSTEM

H. A. Perkins and J. D. Schmidt

*Fairchild Semiconductor Research and Development Laboratory*

*Fairchild Camera and Instrument Corporation*

*Palo Alto, California*

**INTRODUCTION**

The concept of active circuit data storage (flip-flop) is as old as electronic data processing systems. The attributes of highest access speed, steady state nondestructive readout and flexibility of application have been partially offset by higher costs and higher standby power per storage bit. As a result, flip-flop storage has until recently been really only feasible for registers.

As integrated circuits have been perfected to provide a multiplicity of gates or flip-flops on each monolithic die at lower costs than traditional discrete component circuits, the always interesting possibility of an integrated all-semiconductor memory of reasonable capacity becomes exciting. Simply integrating the circuits is not enough. A maximum number of storage positions in a given device package should also require a minimum of leads. Since a memory system is the objective, the need for the usual logic level compatible interface is waived. The storage device is optimized for minimum complexity per bit of storage and greatest electrical tolerance allowance at its terminals. Interface (peripheral) circuits are required such as word bit drivers and sense amplifiers much as for magnetic core or film memory systems. However, the performance required from the peripheral circuits is much less stringent for the semiconductor memory described.

The following sections will describe the circuits, devices, packaging and system design for a random access 256-word memory of 72 bits per word.

**SYSTEM DESCRIPTION**

**System Design Goals**

The main goal to be achieved is a memory system competitive in cost and superior in performance for a certain range of applications. To achieve the goal, the storage device's complexity must not push processing technology toward low yields, must use fairly standard, easily installed packages and require a minimum number of packages for the system complement. To meet these needs, a 16-lead dual in-line package (similar to Fairchild CTμL logic family) containing 4 words of 9 bits per word (36 bits) was selected. An 8×10-inch double-sided printed circuit card readily holds 160 packages. Of this number, 128 are arranged in an array of 16×8 to provide 64 words of 72 bits each, with the balance of the packages containing word drivers.
at 2 circuits per package. Thus the 256-word memory requires 4 such cards (storage). In addition, 2 data circuit cards at 36 bits per card (bit driver and sense amplifier) and a single address register and control card are needed to complete the card complement. A power density of approximately 5 milliwatts per bit was chosen with a view to providing high speed (150-nanosecond cycle time) but moderate system power required (130 watts). The basic storage cell has wide application flexibility and the specifications which follow are one compromise to several conflicting needs. Minor modifications to the design permit optimization for higher speeds, larger capacity for one set of peripheral circuits or reduced power. The capacity of the storage device itself (36-bit package) is a function of present device and package technology and should be regarded as a first expedient.

Table 1.

<table>
<thead>
<tr>
<th>Capacity</th>
<th>256 words, 72 bits/word, 18,432 bits total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repetitive access or write cycle time</td>
<td>150 nanoseconds</td>
</tr>
<tr>
<td>Read mode</td>
<td>nondestructive</td>
</tr>
<tr>
<td>Write mode</td>
<td>jam set to one or zero (at storage device)</td>
</tr>
<tr>
<td>Read access time</td>
<td>120 nanoseconds</td>
</tr>
<tr>
<td>Data flow rate</td>
<td>480 bits/microsecond</td>
</tr>
<tr>
<td>Interface signal levels</td>
<td>+ 2 volts = one, -0.5 volt = zero</td>
</tr>
<tr>
<td></td>
<td>(compatible with Fairchild CTµL family)</td>
</tr>
</tbody>
</table>

System Organization

The functional components of the semi-conductor memory are similar to most random access memories as shown in the block diagram of Fig. 1. A nondestructive steady state output is obtained from the storage cell. Strictly speaking, buffer registers are unnecessary. In the memory described, registers have been included primarily as a convenience for writing in data so that the computer need not be tied up any longer than necessary. Instead of having an output level directly compatible with logic level signals, a relatively low 30-millivolt signal is sensed as a one. Although a sense amplifier is required, the advantage of relatively low-impedance sense/bit line (50 to 150 ohms) justifies its use. Such an impedance level is compatible with interconnection techniques such as printed wiring, twisted pair and coaxial cable. Yet only about 400 microamperes are required from the storage cell tending to minimize standby power needs for high speed capability.
Logic elements for the memory are standard complementary transistor micrologic units. The memory address register is formed from Dual Rank Flip-Flop circuits (Fairchild CTμL-957). The first level decoder uses dual, 4-input positive “and” gate (Fairchild CTμL-954) followed by inverters (Fairchild CTμL-952) to provide the correct signals for the second-level decoders and Word Drivers. On the first memory cycle after the memory has been idle, the address with a start command is loaded into both ranks and decoding begins immediately. After the correct word has been selected and delay through the sense amplifier completed, a gating signal transfers the output to the computer system. If an input command (write-in) is presented to the memory, data is gated into the Data Register (a Latch circuit Fairchild CTμL-968), Bit Drivers corresponding to zeros are energized and after decode delay time a Write Control signal causes the Word Driver to change the state of the storage cells to correspond with Bit Driver outputs.

For repetitive memory cycles, the earlier address is retained in the second rank of the memory address register until the first cycle is complete. The new address, if a start command is also present, is loaded in the first rank only, pending completion of the preceding cycle. Near the end of the first cycle, the new address is transferred to the second rank and, soon after the new cycle begins, the first rank is cleared. By storing the start command also in dual rank register, an asynchronous input can keep the memory cycling at maximum speed.

Special integrated circuits are the 36-Bit Storage Cell, Decoder-Word Driver, the Bit Driver, and the Sense Amplifier. (The Sense Amplifier is a Fairchild µA-710 Comparator.)

Typical voltage and current levels at the storage array interface are given in Table 2.

<table>
<thead>
<tr>
<th>Table 2.</th>
<th>( V )</th>
<th>( I_{\text{max}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word drive, read</td>
<td>1.5–2( v )</td>
<td>25 ma</td>
</tr>
<tr>
<td>Word drive, write</td>
<td>&gt; 3.0( v )</td>
<td>55 ma</td>
</tr>
<tr>
<td>Bit drive</td>
<td>-0.7( v )</td>
<td>10 ma</td>
</tr>
<tr>
<td>Sense output</td>
<td>50 mv</td>
<td>0.5 ma</td>
</tr>
</tbody>
</table>

A timing diagram is shown in Fig. 2. Times correspond to functional blocks in Fig. 1. Response of the storage bit after word drive is applied is about 40 nanoseconds. The time required to store information in a storage bit is about 50 nanoseconds. A 150-nanosecond repetitive cycle time is obtained by allowing overlap into the subsequent cycle. The sense amplifier is not limited to 256 bits on a sense line. Using a multiplexed sense amplifier (4 segments of 1024 words) a 4096 word memory of 72 bits per word should easily achieve a repetitive cycle time of 500 nanoseconds. A cycle time of 100 nanoseconds is obtained at 64 words of 72 bits. These timing estimates are projected on the basis of 5 milliwatts per bit for the particular partitioning scheme described. A higher redundancy of peripheral circuits obviously tends toward maximum speed. The other variable is the power density in the storage package. This may be modified readily by changing the applied power supply voltages or by changing the internal resistances within the device either to optimize for speed, cost or power consumption. The most significant point is that basically a single device design in the storage package will serve a large spectrum of application requirements. The goal of competitive costs is achieved mainly by producing a very large volume of essentially a single device with the attendant cost minimization realizable with microcircuitry.

**CIRCUIT DESIGN**

**Basic Memory Cell**

The particular circuit configuration that one chooses for an integrated semiconductor memory is
dependent on a number of considerations such as chip complexity and size, number of leads, power dissipation, speed, logical flexibility, fabrication capability, etc. After analyzing the complete memory system costs and considering the type of packaging to be used, we concluded that the most important criteria for selection of the circuit configuration is to minimize the number of leads required to the memory cell. This approach permits achieving the highest functional complexity possible for any given package chosen. Consistent with the above statements is the belief that the memory cell circuit should be as simple as possible even at the expense of more complex peripheral circuits if required.

After investigating both linear selection and coincident selection circuits as well as separate data in and data out buses, we decided on a linear selection cell with a common data in-data out bus as shown in Fig. 3. The circuit is a conventional emitter-coupled binary except that the right-hand load resistor has been omitted and this node tied to ground. This forces the left-hand collector output node (point A) to vary from −0.7 volt (defined as a zero) to +0.7 volt (one state) under quiescent conditions. These voltages are determined only by the transistor parameters and are as follows:

\[
V_{one} = +V_{be \, sat} - V_{ce \, sat} \quad \text{and} \quad V_{zero} = -V_{be \, sat} + V_{ce \, sat}
\]

Transistor \(T_3\) acts as an input/output switch and is driven by a bi-amplitude positive pulse as shown on the word select line. Consider first the read operation. A lower amplitude positive pulse forward biases the \(T_3\) base to emitter turning the transistor on. If a one is stored (+0.7 at A), \(T_3\) acts as a normal transistor and a positive output appears across \(R_L\). Its amplitude is determined by the ratio of \(R_L\) to \(R_1 + R_E\) and the magnitude of \(+ V_{cc} V_{ce \, sat}\). The purpose of \(R_E\) is to prevent point A from going closed to ground that +200 mv even with \(D_0\) shorted to ground and minimum \(V_{ce \, sat}\) on \(T_3\).

If a zero is stored, \(T_3\) acts as an inverse transistor and a slight current (dependent on inverse beta) flows into node A, tending to produce a negative output signal. Thus, a zero or a negative level at \(D_0\) corresponds to a zero and a positive level corresponds to a one. When reading a zero, the base collector junction of \(T_3\) is forward biased forcing base current into node A. In order not to write a one into the cell, this current must be limited and this in turn is the reason for the bi-amplitude word select pulse. To write a one, transistor \(T_4\) is left off and the higher amplitude word select pulse is applied. This forces sufficient current into node A to cause \(T_1\) to turn off and \(T_2\) to turn on, thus writing in a one. To write a zero, \(T_4\) is saturated, pulling the data in-data out line negative and thus pulling node A negative when \(T_3\) is saturated.

In designing the circuit, three primary constraints must be met to satisfy the conditions that the cell performs correctly at d-c or low frequency. These design constraints are:

1. Minimum write current is sufficient to write a one.
2. Maximum read current does not write in a one if the F/F is in the zero state.
3. Minimum read current is sufficient to saturate the gate transistor (\(T_3\)).

In considering the design problem, it is soon realized that there are many more degrees of freedom in the design than there are constraints on design conditions which must be met. Thus one must use many secondary design considerations and some judgment in order to arrive at a unique solution. Our approach was to consider first the word driver circuit and determine roughly what read and write voltage levels and spreads could be realized. From this preliminary work, all terminal voltages were specified for the cell (i.e., word line and data line voltages for both reading and writing). The next step is to define the power supply voltages and the approximate power level of the cell. From the standpoint
of wide tolerances in the cell, it is desirable to have $+V_{ee}$ and $-V_{e}$ both large in magnitude compared to $V_{be}$, and $R_1$ and $R_2$ large so that the node A and the emitter node are fed from a constant current source. However, this increases both the power dissipation and the size of the circuit in addition to yielding a somewhat slower circuit due to the RC time constant at node A. Consideration of these tradeoffs lead to fixing $+V_{ee}, -V_{e}$ and $R_2$.

The three worst case equations for the three primary constraints are:

1. $V_A = \left[ \frac{R_3}{R_1 + R_3} \right] \left[ \frac{V_w - V_{be}}{R_3} + \frac{V_{cc}}{R_1} \right] - \left( \frac{V_e - V_{be \text{ sat}1}}{R_2} \right)$

(Use high temperature values for transistor parameters)

2. $V_A = \left[ \frac{R_3}{R_1 + R_3} \right] \left[ \frac{V_w - V_{be}}{R_3} + \frac{V_{cc}}{R_1} \right] - \left( \frac{V_e - V_{be \text{ sat}1}}{R_2} \right)$

(Use low temperature values for transistor parameters)

3. $V_A = (R_E + R_L) \left[ \frac{V_R - V_{be \text{ sat}1}}{R_3} + \frac{V_{cc}}{R_1} \right] + V_{ce \text{ sat}3}$

where the bar above a parameter indicates the maximum value and a bar below indicates the minimum value. $V_w$ is the "read" voltage on the word select line and $V_w$ is the "write" voltage at the same point. All other notations are defined in Fig. 3.

The following values were used in the design equations:

Resistor tolerance . . . . . . . $\pm$ 30%
Power supply tolerance . . . . . . . $\pm$ 5%
"Read" word select voltage tolerance . . . . 0.25 volt
Minimum difference between "read" and "write" word select voltage . . . 1 volt
$V_A$ (write 1) . . . . . . . . . . . . . = 0 volt
$V_A$ (read 0) . . . . . . . . . . . . . = -200 mv
$V_A$ (read 1) . . . . . . . . . . . . . = +200 mv

The voltage levels at node A were chosen so that there would be no change in the state of the cell during reading. By using the above values in Eqs. (1) and (2), $R_1$ and $R_3$ can be uniquely determined in terms of $R_2$ which was previously chosen. The values of $R_1$ and $R_3$ along with $R_L$ were then used in Eq. (3) and $R_E$ was determined. The analytical results were verified by breadboarding all worst configurations using kit-integrated parts and determining the points at which failures occurred.

The decision was made to package the memory circuits in Fairchild's new dual in-line configuration having 14 or 16 leads available. Since the memory cell requires . . . power supply contacts (+ $V_{ee}, -V_e$, and ground), 13 signal leads are available with the 16 pin package. These 13 pins can be used in a near optimum fashion by organizing the memory chip as a 4-word array of 9 bits per word. This gives 36 bits of storage with 4 word select lines and 9 data in-data out lines required. The chip size is 60 by 80 mils and utilizes two layers of metal to solve the crossover problem. Fig. 4 is a photomicrograph of the 36 bit memory array.

Peripheral Circuits

Final Decoder and Word Driver. In order to minimize the number of gates required for word decoding, the decoding has been broken down into two levels called a first decoder level and final decoder. The final decoder is a two-input "and" gate whose output ties to the word driver circuit. Thus for a memory of $2^n$ words, there are $2^n$ of these circuits required. The first level decoding is performed by $2 \times 2^n/2$ "and" gates with $n/2$ inputs each. For $n=8$, there are 256 final decoder and driver circuits and 32-4 input "and" gates for the first level decoder. The fan out required of the first level decoder is simply $2^{n/2}$ with this decoding scheme.

The final decoder and word driver circuit is shown in Fig. 5. It consists of a current mode gate providing a "negative and" function with the non-inverting output directly coupled to a common emitter inverter stage which in turn drives an emitter follower output stage. To make the input levels compatible with the first decoder output, the reference transistor in the C.M. stage ($T_3$) is tied one diode drop above the group. This also gives some temperature compensation to the circuit. The input resistors are used to suppress any tendency of the circuit to oscillate. The bi-amplitude output is achieved by controlling the voltage on the write control bus (WCB). If this line is tied to ground, a voltage divider is created and the lower amplitude output (read) results. If the bus is opened, the high-
Figure 4. 36 Bit memory circuit (60 x 80 mils).

Figure 5. Final decoder and word driver.

er amplitude output (write) results. Again the purpose of diode D2 is to provide temperature compensation. The output stage is designed to drive a 72 bit or less length word which requires about 50 ma current for writing.

Again the circuit was worst-case designed using the same tolerances as previously listed for the basic memory cell. In addition, the output transistor was designed to withstand momentary shorting to the minus supply voltage. The circuit was designed to operate within the correct levels at a junction temperature of 110°C and was actually tested at this ambient temperature with no special heat sinks on the transistors.

In order to keep the integrated circuit relatively simple, only two such circuits were put on one chip and single layer metal was used. The chip is 45 mils on a side, has a nominal power dissipation of 250 mw and is shown in Fig. 6. The chip is packaged in the 14 lead dual in-line header and two such packages are required to service the four words contained in the memory package.

Data Circuits. The sensing problem with this memory cell is strictly a compromise solution based on system size, speed and complexity. Basically, the output of the memory cell during reading is a given current (about 0.5 ma) and this current may be driven into any impedance from zero on up as long as it is referenced to ground. Thus an ideal method of sending would be to use a common base stage which would terminate the line in a low impedance and provide voltage gain with good bandwidth. Unfortunately, the polarity is such that a PNP transistor is required and good PNP common base stages are not easily integrated. For this reason, and also to limit the number of special integrated circuits that have to be designed, we decided to build the prototype system using a conventional integrated amplifier, the μA710. All 256 words in the memory system are tied to a common data line. This can be done since the leakage current of the gating transistor in the memory cell and its emitter capacitance are both very low. The data lines are then termi-
nated in a 150Ω resistor and the 0.5 ma sensing current produces about 75 mv output signal nominally. This signal is amplified to the logic level by the μA710 and a CTμL-956 buffer is used as an output gate and line driver as shown in Fig. 7.

![Figure 7. Data circuits.](image-url)

A CTμL dual latch element is used as a data input register. This circuit is simply a flip-flop with single rail input and output and a gate on both the input and output. Its output cannot pull negative enough on the data line to write in a zero and thus a single common emitter transistor tied to –0.8 volt is required on the output of the latch. Saturating this transistor when a “write” pulse is present on the word select line will write in a zero and, if the transistor is turned off, the data line will be referenced to ground and a one will be written in.

In the preliminary design, data line recovery time was recognized as an important system parameter. The capacity associated with the line was estimated as 0.2 pf per emitter times 256 plus about 20 pf wiring for a total of 70 pf. Thus the RC time constant would be about 10 nanoseconds with a 150 ohm termination. This appears to be a reasonable number.

With regard to packaging, all circuits are housed into dual in-line packages with two latches, two data drivers, two μA710’s and two buffers per package. In addition, a special interface circuit will be provided so that the memory will be compatible with any logic levels commonly used. Thus it require five of the dual in-line packages to service two bits of data.
PACKAGE DESIGN

Printed Circuit Boards

All the circuits are packaged in the dual in-line 14 or 16 lead configuration which, in turn, are inserted and flow-soldered in a two-sided printed circuit board. Fig. 8 is a photograph of the storage board which is 8" x 10". The boards are standard 1/16" G-10 material with 1 oz. copper. The plated through holes are 30 mils in diameter with a 40 mil land. The copper runs are 20 mils wide with a minimum clearance of 20 mils. The power supply voltages and ground are interconnected on the two sides of the board to form an interleaved grid pattern. This helps to reduce the self inductive and resistive drops in the lines.

Interconnection Techniques

The 7 pc cards are interconnected by means of a two sided mother board which also functions as the base of the specially developed connector. Fig. 9 shows a photograph of the top side of the mother board with some of the contacts installed. The signal interconnection pattern between boards is contained on this side of the board. This is also the side into which the contacts and pc boards are inserted. The other side of the mother board contains power supply busing and solder contacts.

Fig. 10 shows a photograph of the connector assembly. It consists principally of three parts: the mother board, the contacts, and a set of cams. When the cams are in the relaxed position, the pc cards may be freely inserted or withdrawn since they do not touch the contact fingers on the board. After a board is inserted the cans are rotated, causing the contacts to exert a wiping action on the fingers and thus make contact. The contact pressure is designed to wipe through any oxide films that may be built up on the pc card.

The same mother board and connector is designed to be used as the equipment interface connector in order to minimize hand wiring required in the memory system. Fig. 11 shows a photographs of the interface connector card with sub-miniature coaxial cable connected to it. Two such cards are required to complete the approximately 160 connections to the system. The cables are secured with a hood and clamp piece that feeds them out the back end of the card cage in a bundle.
Card Cage

Figure 12 shows a photograph of the partially assembled card cage. This mechanical assembly provides two primary functions. They are mechanical support and guiding of the individual boards into the mother board and connector assembly, and providing proper cooling for the assembly via a small fan and a controlled air flow path. In addition, the card cage will house the right angle drives that allow rotating the cams from the front of the unit. Thus it will be possible to completely service the unit from the front, i.e., withdraw and insert cards or insert an extender board and contact assembly.

Overall Unit and Power Supply

The card cage is bolted to a standard 7" rack panel. The panel has a cutout to permit servicing the unit. The depth of the assembly will be about 11", allowing the power supply to be placed directly behind the memory unit in the same rack space. The power supply requirements are approximately as follows:

<table>
<thead>
<tr>
<th>Voltage</th>
<th>Current</th>
<th>Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ 12 V</td>
<td>0.7 amps</td>
<td>± 10%</td>
</tr>
<tr>
<td>+ 4.5 V</td>
<td>7 amps</td>
<td>± 5%</td>
</tr>
<tr>
<td>+ 2 V</td>
<td>5 amps</td>
<td>± 5%</td>
</tr>
<tr>
<td>- 0.8 V</td>
<td>1 amp</td>
<td>± 10%</td>
</tr>
<tr>
<td>- 2 V</td>
<td>3 amps</td>
<td>± 10%</td>
</tr>
<tr>
<td>- 5 V</td>
<td>15 amps</td>
<td>± 5%</td>
</tr>
</tbody>
</table>

OPERATING CHARACTERISTICS

A pulse program consisting of write one, read one, write one, read zero, and then repeat has been used for the testing. The oscillographs of Fig. 13 show some of the waveforms for this system. A is the negative output pulse from the first decoder and inverter as it appears on one of the decode input lines on the storage card. These are all essentially "open" lines so the reflections present are reasonable.

B shows the output of the final decoder and word driver as it appears on the word select line. The first pulse is the read zero, followed by write one, etc. Mid-amplitude delay through the circuit is observed to be about 20 nanoseconds. The long tail off (about 50 nanoseconds) on the trailing edge of the pulse is due to stored charge in the base of the gating transistors. The cycle time shown here is 150 nanoseconds.

C shows the output of the sense amplifier. As ex-
pected, a signal is produced by the write "1" opera-
tion as well as the read "1" operation. Mid-amplitude
delay from word line to data out is about 30 nano-
seconds.

Figure 10. Connector assembly.

Figure 11. System interface connector assembly.
To demonstrate possible crosstalk that exists on the bit line, the repetition rate was slowed down and a zero and one superimposed. Figure 14(a) shows the word line and 14(b) the bit line waveforms. For this case, the two adjacent bit lines of both sides of the central bit had complement data written in and read out at the same time as the central bit line was being operated on. The photo is of the central line and the small amount of crosstalk at the leading and trailing edges is apparent.

CONCLUSIONS

The results obtained have shown the feasibility of producing integrated circuit memory systems which offer advanced performance and design simplicity. Such memories may be expected to have an important place in applications where high speeds at low costs for moderate storage capacity are needed. Our experience has shown that the engineering of an integrated semiconductor memory system is much easier than that required for a thin magnetic film memory of comparable performance.

As semiconductor fabrication techniques continue to advance, further improvement in semiconductor memory systems is certain.

ACKNOWLEDGMENTS

Many persons have contributed in one way or another to this project and all cannot possibly be named. However, special thanks are due to the Digital Integrated Circuit Section (R. Seeds) for fabricating the storage array and the Digital Systems Research Department (R. Rice) for many of the packaging concepts and, in particular, for the design of the mother board connector assembly. We are also indebted to J. Friedrich for the design and fabrication of the word driver circuit, and H. Zinschlag for the design of the data circuits.
Figure 13. Memory system waveforms (50 nanoseconds/square, 2 volts/square — except as noted).
   a. Decode input pulse.
   b. Word line.
   c. Sense amplifier output.

Figure 14. Superimposed one and zero outputs (40 nanoseconds/square).
   a. Word line (2 volts/square).
   b. Bit line (100 mv/square).