

NOBC

NAVAL OCEAN SYSTEMS CENTER San Diego, California 92152-5000

Technical Document 1099
June 1987

Integrated Circuit Design

V. S. Wong and J. Grinberg
Hughes Research Laboratories



Approved for public release
distribution is unlimited

The views and conclusions contained in this report are those of the authors and should not be interpreted as representing the official policies, either expressed or implied of the Naval Ocean Systems Center or the U.S. government

NAVAL OCEAN SYSTEMS CENTER
San Diego, California 92152-5000

E. G. SCHWEIZER, CAPT, USN
Commander

R. M. HILLYER
Technical Director

ADMINISTRATIVE INFORMATION

This work was performed for the Department of Defense, Ft. Meade, MD 20755. Contract N66001-84-C-0104 was carried out by Hughes Research Laboratories, 3011 Malibu Canyon Road, Malibu, CA 90265, under the direction of W.H. McKnight, Code 743, NAVOCEANSYSCEN.

Released by
J.M. Alsup, Head
Image Processing and
Display Branch

Under authority of
R.L. Petty, Head
Electromagnetic Systems
and Technology Division

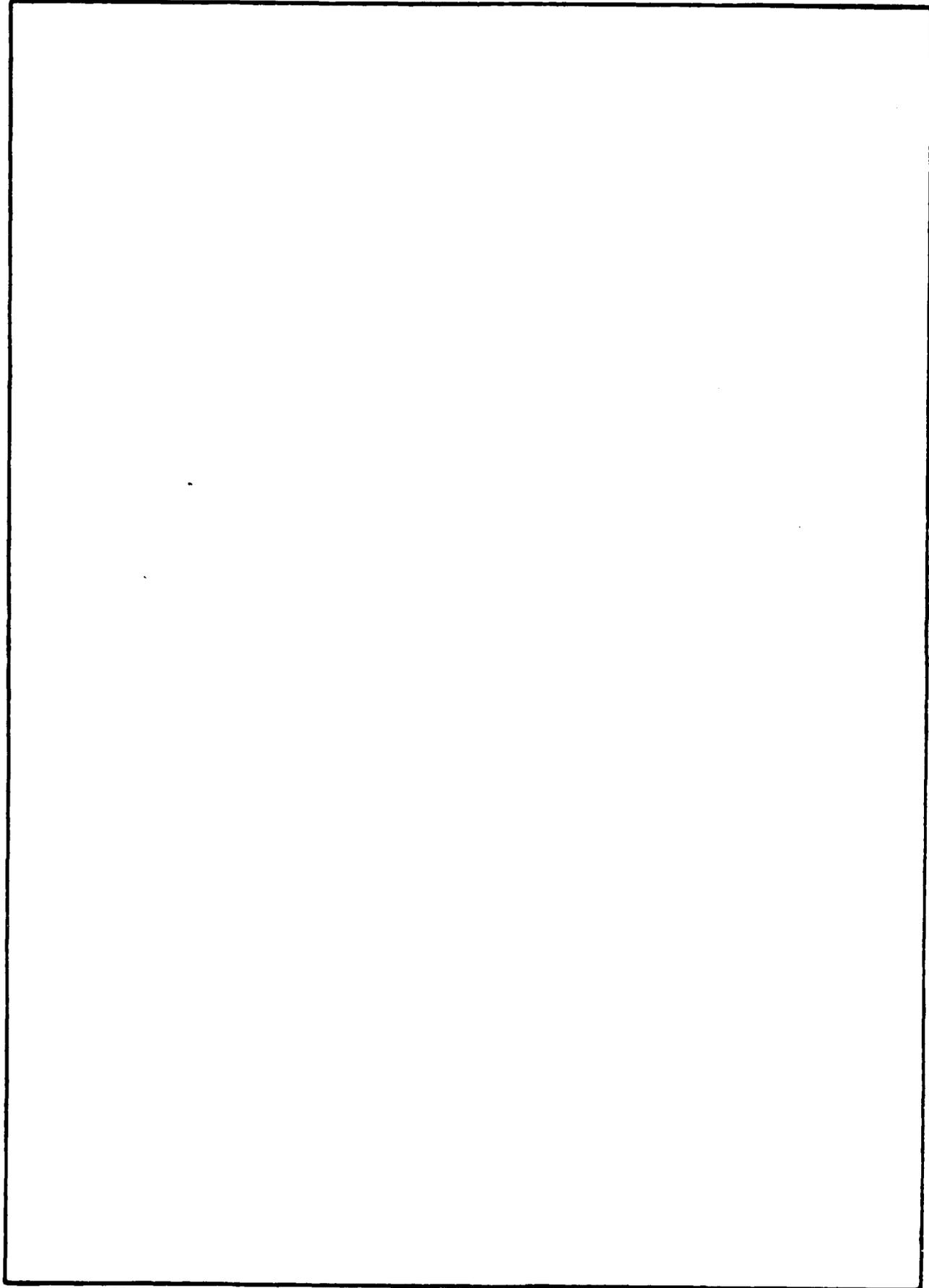
MA

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S) NOSC TD 1099	
6a. NAME OF PERFORMING ORGANIZATION Hughes Research Laboratories	6b. OFFICE SYMBOL <i>(if applicable)</i>	7a. NAME OF MONITORING ORGANIZATION Naval Ocean Systems Center	
6c. ADDRESS (City, State and ZIP Code) 3011 Malibu Canyon Road Malibu, CA 90265		7b. ADDRESS (City, State and ZIP Code) Image Processing and Display San Diego, CA 92152-6000	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Department of Defense	8b. OFFICE SYMBOL <i>(if applicable)</i> DoD-R913	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N66001-84-C-0104	
8c. ADDRESS (City, State and ZIP Code) 9800 Savage Road Ft. Meade, MD 20755		10. SOURCE OF FUNDING NUMBERS	
		PROGRAM ELEMENT NO RDDA	AGENCY ACCESSION NO DN488 839
		PROJECT NO NSA	TASK NO 740-EE93
11. TITLE (Include Security Classification) Integrated Circuit Design			
12. PERSONAL AUTHOR(S) V.S. Wong and J. Grinberg			
13a. TYPE OF REPORT Final	13b. TIME COVERED FROM Dec 83 TO Apr 84	14. DATE OF REPORT (Year, Month, Day) June 1987	15. PAGE COUNT 62
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	Field Oxide, doped N-Substrate, epitaxial silicon, sapphire substrate, silicon-on-sapphire (SOS) device, complementary metal-oxide semiconductor (CMOS), digital filter, low-power filter	
19. ABSTRACT (Continue on reverse if necessary and identify by block number)			
<p>This report describes the results of a study on the chip design of a low-power filter, using state-of-the-art CMOS technology. The filter is for speech applications and is specified to have 1024 taps with programmable weights and linear phase. The chip implementation is to have a word length of 8 to 12 bits and consume a maximum of 2.0 mA at 3.6V. Included are current capabilities of CMOS/SOS and CMOS/bulk, technologic, and in particular, the Hughes VHSIC CMOS process. The architecture of the filter is discussed and estimates are made for the power consumption, speed, device count, and projected chip size of the filter implementation. A comparison of a multiple taps and a single taps implementation of the filter is also presented in terms of power consumption and operational speed.</p>			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL W.H. McKnight		22b. TELEPHONE (Include Area Code) (619)225-7439	22c. OFFICE SYMBOL Code 743

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)



DD FORM 1473, 84 JAN

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

SECTION		PAGE
1	REPORT SYNOPSIS.....	1
	A. Introduction.....	1
	B. Summary of Results.....	2
	C. Conclusion.....	6
2	TECHNICAL ISSUES.....	9
	A. Introduction.....	9
	B. Technology Issues.....	9
	C. Architectural Issues.....	26
	D. Other Design Considerations.....	49
	E. Cost of Fabrication.....	49
	REFERENCES.....	53
APPENDICES		
	A. Toshiba 256K CMOS Static RAM.....	55
	B. Hitachi 64K CMOS Static RAM.....	57

LIST OF ILLUSTRATIONS

FIGURE		PAGE
1	Silicon-on-sapphire device structure.....	11
2	CMOS/bulk device structure.....	12
3	Total rise and fall time for interconnect vs minimum feature.....	14
4	Operating voltage range vs gate length for scaled CMOS/SOS inverter.....	16
5	Scaling relationships of various parasitic capacitance ratios to gate oxide capacitance....	19
6	Interconnect resistance scaling.....	21
7	CCD memory partitioned into four segments.....	25
8	Single-tap implementation of low power filter...	27
9	VHSIC full adder cell.....	30
10	Accumulator for adding 1024 b-bit words.....	32
11	Pipelined book multiplier.....	35
12	Reducing the power consumption of data storage by positioning into four segments.....	40
13	Output driver configuration.....	41
14	Four-tap approach for low power filter.....	47
15	Tentative schedule for processing low power filter chip.....	51

SECTION 1
REPORT SYNOPSIS

A. INTRODUCTION

We report here the results of a study on implementing a low power filter using state-of-the-art CMOS technology. The basic goal is to design a 1024 tap filter with programmable weights, that has linear phase, operates at a sample rate of 8KHz, and consumes a maximum of 2.0 mA at 3.6 V (7.2 mW power). Input data word length has been specified as between 8 and 12 bits. The work performed here is based in part on an earlier study¹ on a low power filter done at HRL for NOSC.

In Section 1.B we present a summary of the results of the study, with conclusions in Section 1.C. Section 2 contains the technical supporting details of the report. The study is divided into two main parts - Technology Issues (Section 2.B) and Architectural Issues (Section 2.C). Under Technology Issues we will review the latest developments in CMOS technology and compare the performance of CMOS/SOS and CMOS/bulk technologies. We will look at the voltage requirements of these technologies to see if 3.6 V is an acceptable power source level. In particular, the Hughes VHSIC CMOS/SOS process will be examined. In the Architectural Issues section we will analyze each component of the low power filter in terms of power consumption, speed, and gate count when implemented with the Hughes CMOS/SOS process. We will also review state-of-the-art memory chips to see if one could be suitably used for the low power filter. Lastly, in Section 2.E, we provide a preliminary estimate of the cost of fabricating the low power filter chip using Hughes CMOS/SOS technology.

B. SUMMARY OF RESULTS

1. Technology

Since an important consideration of this filter chip is extremely low power, CMOS technology is a natural candidate for implementation. CMOS circuitry dissipates mainly dynamic switching power (CV^2f power) and negligible quiescent power. Of the two CMOS technologies to choose from - CMOS/SOS and CMOS/bulk - CMOS/SOS clearly provides better speed/power performance above $1.5 \mu\text{m}$ channel length, but below $1.0 \mu\text{m}$ there is evidence that CMOS/bulk performance becomes comparable with CMOS/SOS. As channel lengths approach $1 \mu\text{m}$ and below, the lower line-to-substrate capacitance advantage held by CMOS/SOS is lost as line-to-line interconnect capacitance becomes significant. Progress is still being made to improve the speed/power performance of both SOS and bulk technologies, and it is unclear if CMOS/SOS would still be significantly better than CMOS/bulk at submicron feature sizes.

As to supply voltage requirements for state-of-the-art CMOS devices, we have reviewed the relevant literature and conclude that the current Hughes VHSIC CMOS/SOS technology will be able to operate from a 3.6 V power source.

We have also investigated using CCD technology for meeting the data storage requirements of the filter. We used a scheme that partitions the data for storage into 4 CCD registers and runs each register at 8 MHz, but only for one-fourth of the time, so that the effective operational rate is 2 MHz. When the 1500b bits (b is the word size) of data and coefficient storage is implemented using current CCD technology, we calculated that power dissipation would be $3.6b \text{ mW}$. This assumes that the lowest acceptable clocking voltage - 6V, is used. From this analysis, we see that both power consumption and required operating voltage would exceed filter specifications. Hence, we do not recommend the use of CCD technology for data storage.

2. Architectural Issues and Power Consumption

Assuming a single tap implementation for the low power filter operating at 8 MHz, the major components of the filter, along with power dissipation and device count for each component are listed in Table 1. The power calculations are based in part on a low-power Toshiba 256K CMOS static RAM chip announced at the ISSCC conference in February, 1984. The power consumption for this chip was scaled down to meet the data and coefficient storage requirements of the filter. The Hughes VHSIC CMOS/SOS process parameters were used to calculate power dissipation in the processor section. CV^2f dynamic power is assumed to be the primary source of power dissipation in this section. C is the total capacitance of each component in the processor section, V is taken as 3.6 V, and f is 8 MHz. In Table 1 the parameter b is the word size, specified as between 8 and 12 bits. The total power dissipation for the filter can be obtained by adding the power dissipated within each component, resulting in

$$P_{\text{total}} = 1.74b^2 + 246b + 18.3 \mu\text{W}.$$

Similarly, the total device count is obtained by adding the devices for each component, resulting in

$$D_{\text{total}} = 32b^2 + 9265b + 340.$$

The device count and power dissipation broken down by major components for a 10-bit filter is shown in Table 2. Data and coefficient storage requirements contribute to 89% of the total power consumption and 96% of device count. The total power consumption, device count and projected chip size for an 8, 10 and 12-bit filter are shown in Table 3. The projected chip size for the filter is obtained by estimating the area occupied by RAM and by random logic, and is given by

$$S = 51.6b^2 + 949b + 548 \text{ mil}^2.$$

Table 1. Device Count and Power for Single Tap Filter

	Device Count	Power (μ W)
Storage	9200b	241b
Adder/Acc.	42b + 340	2.26b + 18.3
Multr.	b(32b + 19)	1.74b ² + 2.16b
Output Driver	4b	0.53b
Total	32b ² + 9265b + 340	1.74b ² + 246b + 18.3

Table 2. Device Count and Power Breakdown for 10-bit Filter

	Device Count (% of total)	Power (μ W) (% of total)
Storage	92 K (96%)	2.4 (89%)
Adder/Acc.	760 (0.8%)	0.041 (1.5%)
Multr.	3.4 K (3.5%)	0.20 (7.3%)
Output Driver	40 (0.04%)	0.005 (0.19%)
Total	96.2 K (100%)	2.7 (100%)

Table 3. Total Power Consumption and Device Count for an 8, 10 and 12 bit Filter

b bits/word	Power Consumption (μ W)	Device Count	Projected Chip Size (mil ²)/(% memory)
8	2.1	76.5 K	11.4 K (59%)
10	2.7	96.2 K	15.2 K (56%)
12	3.2	116 K	19.4 K (52%)

There is a trade-off in terms of power consumption between a single tap approach and a multi-tap approach. A multi-tap approach operates at lower speed but requires more gates for implementation, whereas a single tap approach would have to operate at higher speed, but requires fewer gates. By analyzing the computational requirements of a multi-tap approach we find that the power consumption as a function of the number of taps, N , is

$$P(b,N) = 243b/N + 1.73b^2 + 2.43b + 14 \text{ uW.}$$

Here we see that only the power associated with data and coefficient storage (the first term of the equation) decreases with N . The power consumed in the processor section (the remainder of the terms) remains constant because even though the processor section is operating at lower speeds as N increases, the number of processors operating simultaneously increases with N .

The total power consumption, device count and projected chip size for a 10-bit filter is shown for 1 through 4 taps in Table 4. The device count for an N -tap filter is given by

$$D(b,N) = 9229b + 80 + N(32b^2 + 45b + 260) \text{ ,}$$

and the projected chip size is

$$S(b,N) = 876b + 129 + N(51.6b^2 + 72.6b + 419) \text{ mil}^2.$$

The first term in $D(b,N)$ and $S(b,N)$ arises primarily from the data and coefficient storage section, and the bracketed term in both equations represents contributions from the processor section.

When using a multi-tap approach, chip yield must be considered. From Table 4, one sees that going from a single to a two-tap implementation reduces the power consumption by 46%, but also increases chip size by 41%. This increase in chip size

Table 4. Power Consumption, Device Count and Chip Size for a 10-bit Filter Implemented with N Taps

Number of tap, N	Power Consumption		Device Count		Chip Size	
	(μ W)	(% dec)	#	(% inc.)	mil ²	(% inc.)
1	2.6		96.2 K		15.2 K	
2	1.4	46	100 K	4.3	21.5 K	41
3	1.0	62	104 K	8.4	27.8 K	83
4	0.82	68	108 K	12.8	34.1 K	124

may translate into a sizable reduction in fabrication yields. To improve chip yields, the filter implementation may be partitioned into two or more chips. This approach, however, would require communications between chips at processor speeds. The power consumed by chip drivers for handling inter-chip communications would in all likelihood not be compensated for by operating the filter at lower speeds.

C. CONCLUSION

The choice of a CMOS process for implementing the filter is important. The two components of the filter having special processing requirements are the RAM for data and coefficient storage and the A/D converter (if it is placed on the same chip). Choice of a technology for developing a low-power integrated RAM design is important because the RAM portion of the filter consumes most of the power on the chip. Also, the large size of the RAM makes it desirable to use a high device density technology. Our power and device density calculations for data storage were based on a 256K static RAM from Toshiba. The technology used is a two-level polysilicon, two-level metal, p-well CMOS process, with channel lengths comparable to the Hughes VHSIC SOS process. In order to meet the low power specification for the filter, a similar CMOS process may have to be used to develop a low power RAM for data storage in the filter.

In the area of A/D converter designs, significant leakage currents in SOS devices can make it difficult to achieve millivolt precision in CMOS/SOS A/D converters. However, recent work at RCA has reduced the leakage currents of SOS devices significantly. Hence it may now be possible to develop high precision A/D converters using this improved CMOS/SOS technology.

We have also examined the trade-offs between a single tap and a multi-tap approach for the filter. A single tap approach requires a higher operational speed but fewer gates to implement, whereas a multi-tap approach requires a lower operational speed but more gates to implement. A multi-tap approach would result in power savings if all the components could be integrated on one chip, but would require much higher power if two or more chips were needed for the implementation. The yield of fabricated chips becomes an important consideration when chip size is increased for multi-tap implementations. We estimate that going from a single tap to two taps for a 10-bit filter would increase chip size by 41%, but also decrease power consumption by 46%.

Given the state-of-the-art of current CMOS technology, the development of the low power filter chip must be viewed as a research effort. The low power requirement for the filter makes it desirable to implement the filter on a single chip. However, the large number of devices necessary for implementing the filter makes chip yield a primary concern. Also, to successfully meet the low power requirements of the filter, it will be necessary to integrate a high density, low power RAM technology with logic circuitry. This RAM technology, requiring advanced processing techniques, is not widely available at this time.

SECTION 2

TECHNICAL ISSUES

A. INTRODUCTION

In the preliminary report we assumed that CMOS/SOS would be the technology for implementing the low power filter. In this report we will consider the trade-offs between CMOS/SOS and CMOS/bulk more critically. We will also examine the operating voltage requirements of CMOS technologies to see if 3.6 V is an acceptable operating voltage. The Hughes VHSIC CMOS/SOS process is a good example of what is possible in state-of-the-art CMOS technologies, so we will present the performance parameters of this process as being representative of what can be achieved in CMOS/SOS technologies today. We will also analyze power consumption requirements if CCD technology is used for data storage on the filter chip.

Using the performance parameters for the Hughes VHSIC CMOS/SOS process, we will obtain speed, power consumption and device count estimates for the adder/accumulator, multiplier and output drivers on the filter chip. Hughes does not have a current effort to develop a 1.25 μm static RAM chip, so to obtain power and device count estimates for data and coefficient storage on the filter chip, we will examine CMOS static RAM chips currently available in the market. We will also look at the trade-offs between using a single tap and a multi-tap implementation for the filter chip.

B. TECHNOLOGY ISSUES

In the following sections we will review state-of-the-art CMOS technologies and compare CMOS/SOS and CMOS/bulk in terms of power consumption, speed and device technology. We will also examine operating voltage requirements for these technologies,

and interconnect parasitic capacitances as feature sizes shrink to submicron dimensions to see how significant they are in comparison to gate capacitances. Specifically, we will present the performance characteristics for the Hughes VHSIC CMOS/SOS process. Finally, we will investigate the use of CCD technology for data storage.

1. CMOS Technology

In the preliminary study it was assumed that CMOS/SOS would be used for fabricating the low-power filter. It is generally acknowledged that CMOS/SOS is lower power and faster than CMOS/bulk technology at channel lengths greater than about 1.0 μm . Below 1.0 μm , however, there is evidence that CMOS/bulk becomes competitive with CMOS/SOS. Traditionally CMOS/SOS is used for radiation-hard applications, whereas CMOS/bulk is usually used where chip yield and cost effectiveness is a factor. In choosing a technology for implementing the low power filter, some consideration should be given to developing a low-power, high-density staticRAM and an A/D converter (if it is to be on the same chip) using that technology. These are two crucial components of the filter, and the successful implementation of these components could depend on the choice of a technology.

As is well known, CMOS/SOS technology derives its superiority from the fact that there is virtually no capacitance from interconnects to the non-conducting sapphire substrate (Figure 1). This low capacitance translates into higher switching speeds and lower dynamic power dissipation - two important considerations in choosing a technology. CMOS/bulk, on the other hand, does exhibit considerable interconnect capacitance to the substrate (Figure 2). First, line-to-substrate capacitance arises from metal and poly-silicon wiring over field oxide. This capacitance is usually about an order of magnitude less than gate oxide capacitance. Second, there is diffused

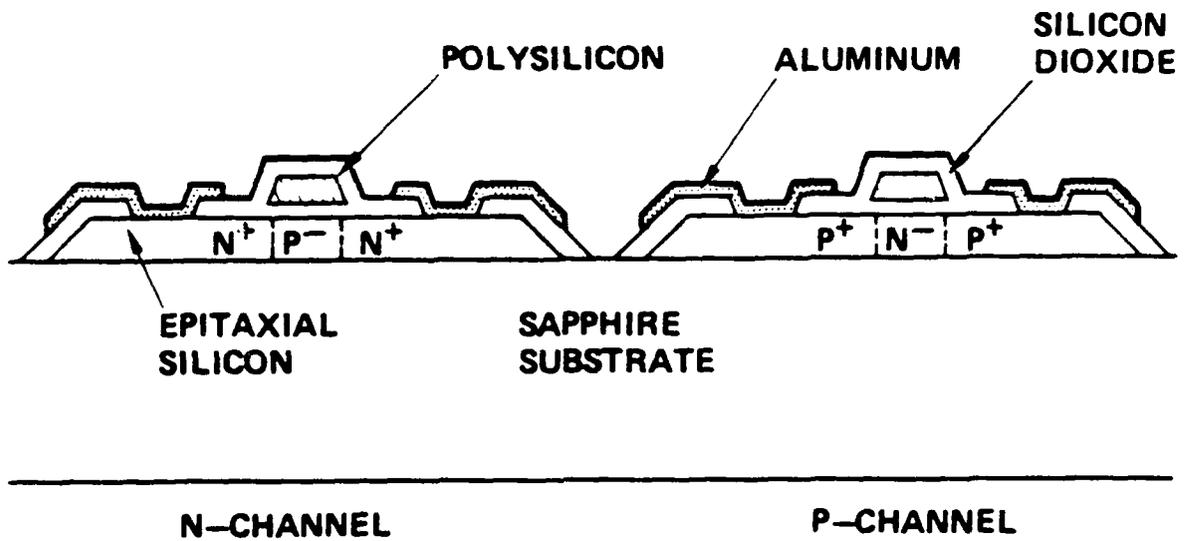


Figure 1. Silicon-on-sapphire device structure.

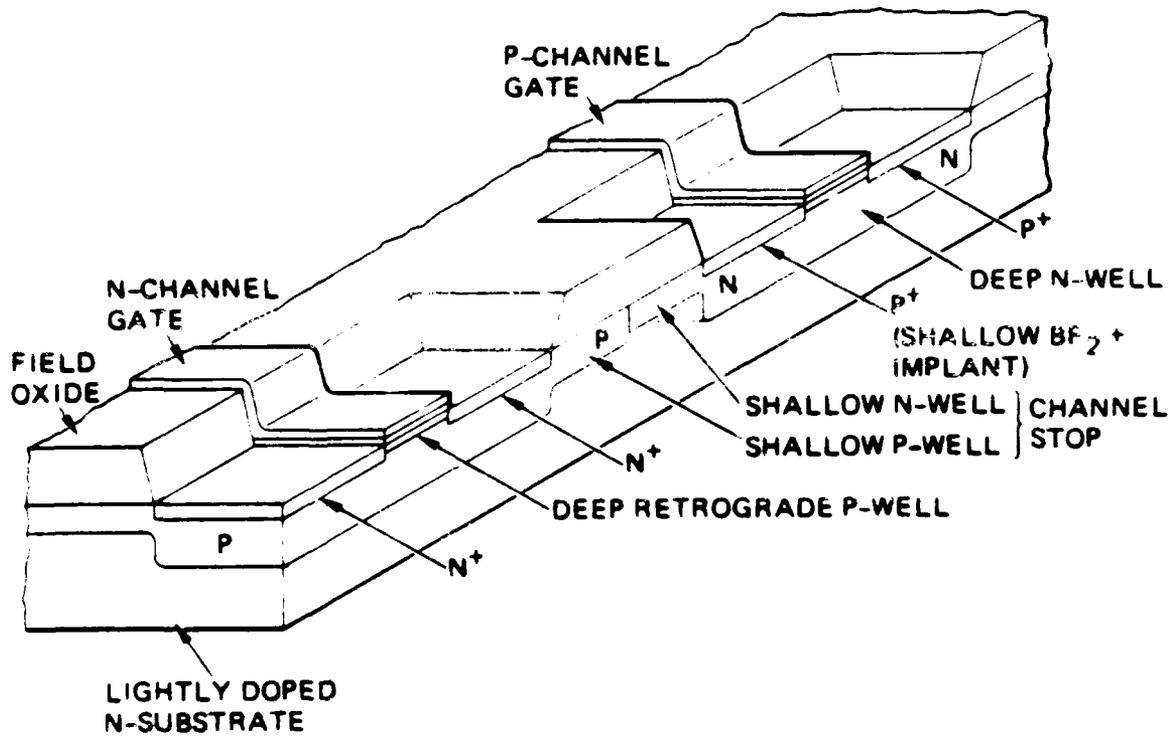
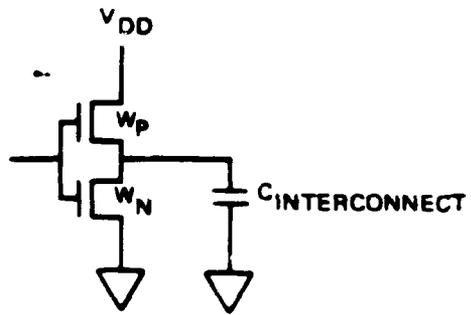
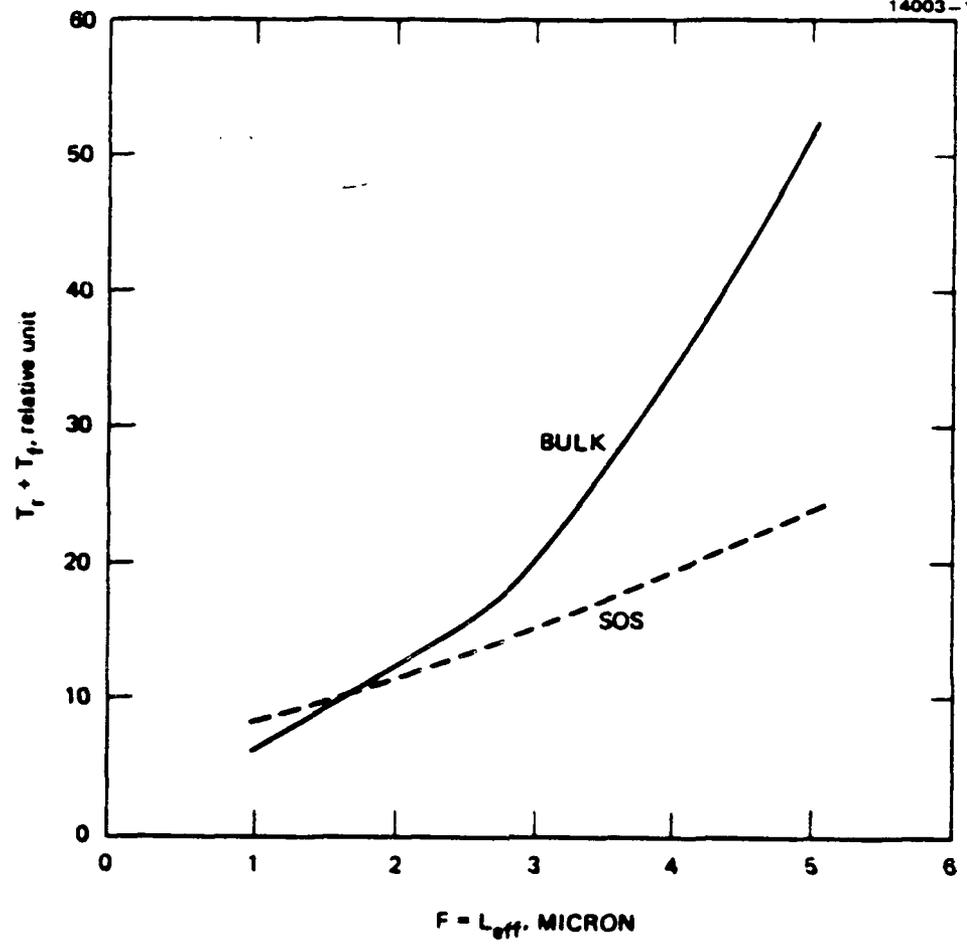


Figure 2. CMOS/bulk device structure.

line capacitance, primarily at the sidewall adjoining the field oxide, with a typical value of about 4×10^{-4} pF/ μm . The presence of greater capacitance in CMOS/bulk circuits results in lower performance of bulk circuits at device dimensions greater than $1.5 \mu\text{m}$. This is evident in Figure 3 which compares the switching speed of a CMOS/SOS and a CMOS/bulk inverter as a function of minimum feature size.²

As feature sizes are decreased to less than $1.5 \mu\text{m}$, the performance of CMOS/bulk circuits become comparable to that of CMOS/SOS circuits. First, the mobility of carriers in bulk devices are greater than those of SOS devices. Because silicon grown on sapphire contains many more defects than bulk silicon, the scattering of carriers reduces their mobilities in SOS devices to less than that of bulk devices. The effect of this is a higher drive current for bulk devices than SOS devices at equal channel lengths. Second, interconnect line-to-line capacitance becomes significant as feature sizes decrease below 1μ (see Section 2.B.). Since line-to-line capacitance is present in both SOS and bulk technologies, the advantage of lower line-to-substrate capacitance enjoyed by SOS circuits is lost at submicron dimensions. As shown in Figure 3, the switching speed of a CMOS/bulk inverter is comparable to that of an equivalent SOS inverter at less than $1.5 \mu\text{m}$. Moreover, the study at Hewlett-Packard² comparing the performance of ring oscillators on SOS and bulk technologies at a channel length of $1.3 \mu\text{m}$ concluded that both speed and power dissipation were about the same.

Improvements are still being made in both CMOS/SOS and CMOS/bulk technologies. New crystal growth techniques have reduced the defects in silicon grown on sapphire, thereby increasing the mobilities of SOS devices.⁷ The latchup problem, particularly severe in CMOS/bulk circuits as feature sizes approach $1 \mu\text{m}$, is also being solved. Recent studies at Toshiba indicate that CMOS/SOS circuits may still enjoy a speed/power advantage over CMOS/bulk circuits at submicron dimensions.⁸ How much of a performance advantage CMOS/SOS



$$\begin{aligned}
 W_p &= W_n \\
 L_p &= L_n \\
 \beta_R &= 1
 \end{aligned}$$

$$T_r + T_f \propto \frac{C_T}{I(\text{PMOS})} + \frac{C_T}{I(\text{NMOS})}$$

E.S., B.A., C.C., G.H.
 CICO, 5/80
 C.S. W., K.R., J.M., B.B.
 HPL, 5/80

Figure 3. Total rise and fall time for interconnect vs minimum feature.

circuits will still have at submicron dimensions is unclear at this time. The Hughes VHSIC program has interest in both CMOS/SOS and CMOS/bulk technologies, although current CMOS/SOS development at 1.2 μm channel lengths is at a more advanced stage.

2. Power Supply Voltage for CMOS Filter Chip

One of the specifications for the low-power filter was that it operate from a 3.6V power supply. Current VHSIC CMOS technology is targeted at 5.0V operation. This may be partly because of a desire to maintain voltage compatibility with other digital logic families (particularly TTL), most of which operate at 5.0V. In this section we will consider whether it is possible to operate current and future VHSIC CMOS technologies at 3.6V.

Some work has been done within the Hughes VHSIC program in studying the operating voltage range of CMOS circuits as device dimensions are scaled down. The result of this work is shown in Figure 4. As device dimensions are scaled down, there are several phenomena that limit the operating voltage range of CMOS circuits. Among these are device punchthrough, oxide breakdown, junction breakdown, device turn-on threshold, and excess thermal generation. From the VHSIC study, the two factors that limited MOSFET operation when device dimensions were optimally scaled were identified.

At the high end, junction breakdown occurred when drain-to-substrate potentials exceeded the breakdown voltage. This phenomenon is well understood as avalanche breakdown in reverse biased pn junctions. When electric fields within the junction reach a critical value (around 3×10^5 V/cm), carrier impact ionization will cause a rapid increase in current flow through the junction. Because substrate doping tends to increase as device dimensions are scaled down (for threshold compensation and to decrease depletion widths), critical breakdown fields are

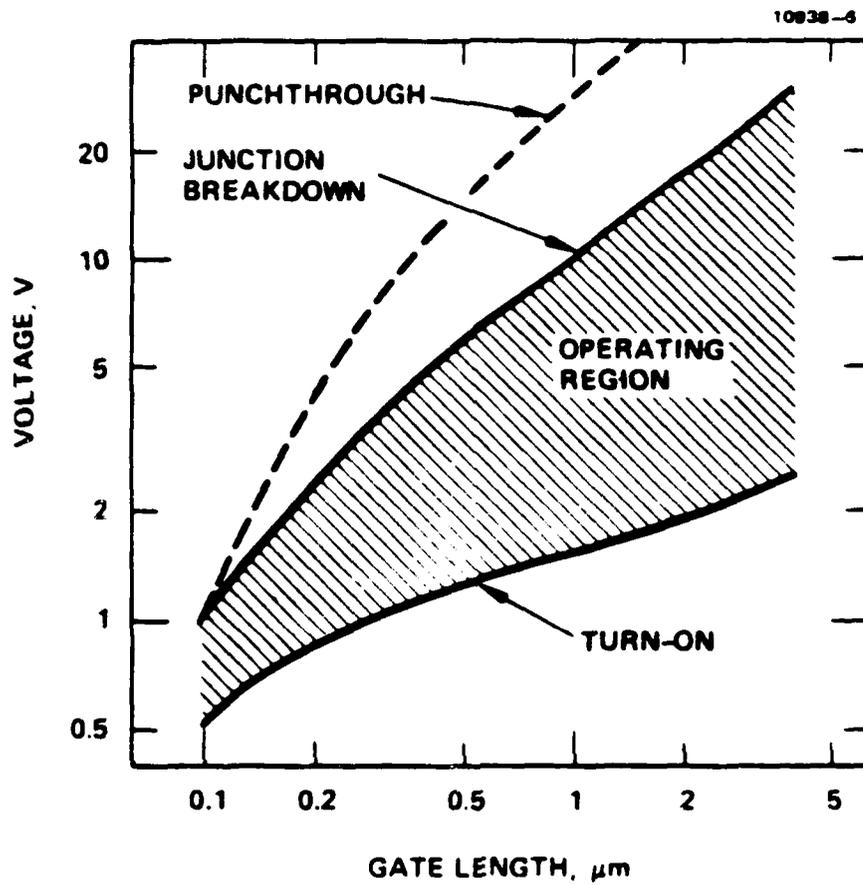


Figure 4. Operating voltage range vs gate length for scaled CMOS/SOS inverter.

reached at lower drain potentials. Hence, operating voltages will have to be decreased as device dimensions are scaled down to prevent junction breakdown.

At the low end, device operation is limited by the turn-on voltage of devices. This parameter can be controlled to some extent by changing the substrate doping by ion implantation and gate oxide thickness. To maintain a satisfactory noise margin, however, it is desirable to use an operating voltage above the turn-on voltage by several times the thermal voltage, kT/q .

In Figure 4 the limits of operating voltage for CMOS/SOS technology is plotted against the gate length of devices. The shaded area represents the acceptable operating voltage range for circuits. As can be seen from the figure, current VHSIC technology at $1.25 \mu\text{m}$ can be operated safely at 3.6V. Moreover, we believe that future VHSIC submicron technologies down to $0.5 \mu\text{m}$ feature sizes can be operated at 3.6V without any problem.

3. Interconnect Parasitic Capacitances as a Function of Feature Size

In the preliminary study on the low power filter, it was assumed that the parasitic capacitances arising from interconnects were negligible compared to gate capacitances. We will examine this issue in greater detail in this section.

In a study performed for NOSC in 1980 entitled "Develop Submicron Devices,"³ the parasitic capacitance arising from wiring interconnects as device feature sizes were scaled down was considered. In this study it was assumed that all dimensions scale linearly with x , the field oxide thickness was 12.5 times the gate oxide thickness (t_g), and the width and spacing of wiring was 1.5 times the gate length, L_{ch} . Under these conditions the gate capacitance, given by,

$$C_g = \epsilon_{ox} * L_{ch} * W_{ch} / t_g$$

decreases directly with L_{ch} .

Figure 5 shows the interconnect parasitic capacitances relative to gate capacitances as device dimensions are scaled down. There are two primary sources of wiring capacitances - line-to-substrate (C_s) and line-to-line capacitances (C_m). Let us first consider the case of short wiring interconnects within cells, assuming they have an average length of $8L_{ch}$. As seen from the figure, the ratio of $C_s(8L_{ch})/C_g$ would be constant, since the lengths of these wires would scale directly with gate lengths. This ratio is found to be almost unity. Therefore, line-to-substrate capacitance is significant in CMOS/bulk technology. This capacitance, however, is negligible in CMOS/SOS technology because there is no conducting substrate. From the figure, the ratio $C_m(8L_{ch})/C_g$ for short wires is seen to be an order of magnitude less than 1 for gate lengths above $1 \mu m$, and increase to exceed 1 below $1 \mu m$. From this we can conclude that for CMOS/SOS technology gate capacitance would be dominant at a VHSIC gate length of $1.2 \mu m$, but below $1 \mu m$ line-to-line capacitance becomes significant.

Now let us consider the case of wiring interconnects with dimensions comparable to chip size, L_c . These long interconnects are evident in regular chip designs such as memory chips and programmable logic arrays. As can be seen in Figure 5, long wiring capacitances tend to dominate over other capacitances. In the line-to-substrate case, the ratio $C_s(L_c)/C_g$ is about 100 at $5 \mu m$ gate lengths and increases rapidly to exceed 10^4 at submicron dimensions. In the line-to-line case, $C_m(L_c)/C_g$ is about 1 at $5 \mu m$ and increases to be comparable to $C_s(L_c)/C_g$ at submicron dimensions. At $1.2 \mu m$ $C_s(L_c)$ is about 10^3 greater than C_g and $C_m(L_c)$ is about 10^2 greater than C_g . The contribution to total chip capacitance from long wires will be significant in regularized structures where there are many long wires spaced closely together, such as in memory chips. In random logic chips, however, the collective contribution from gate capacitances usually dominates total chip

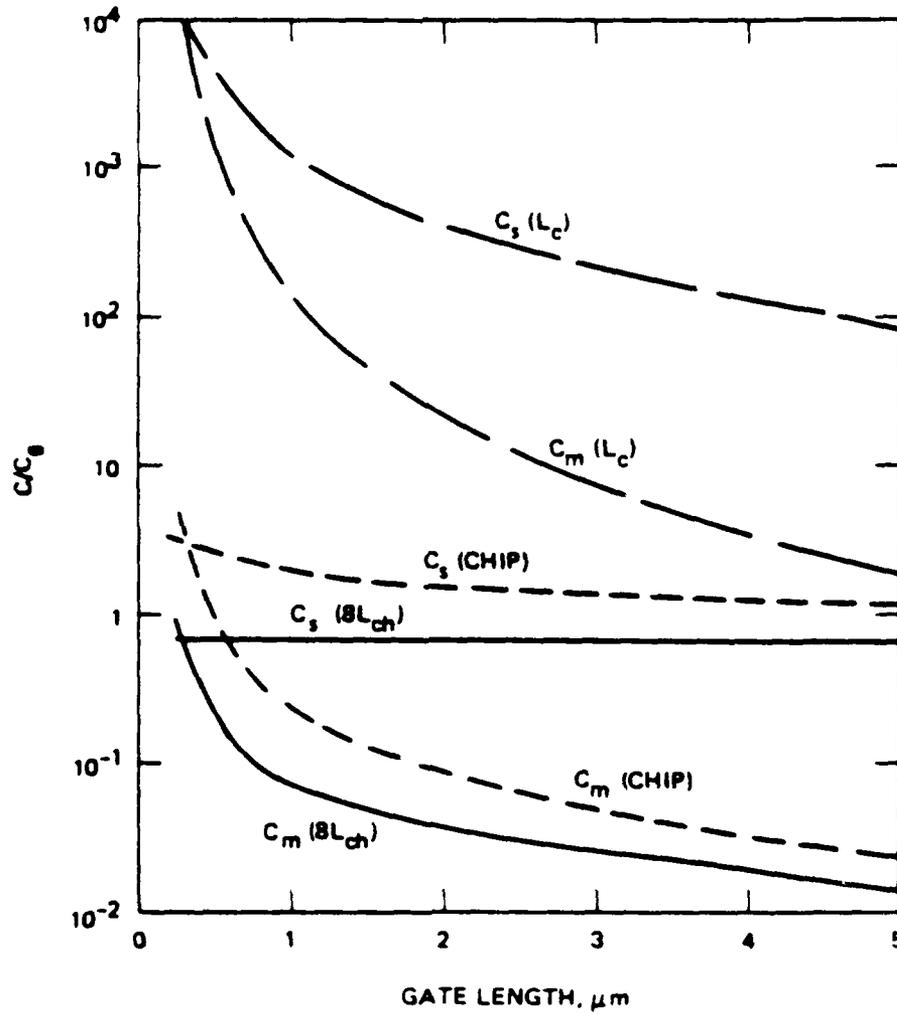
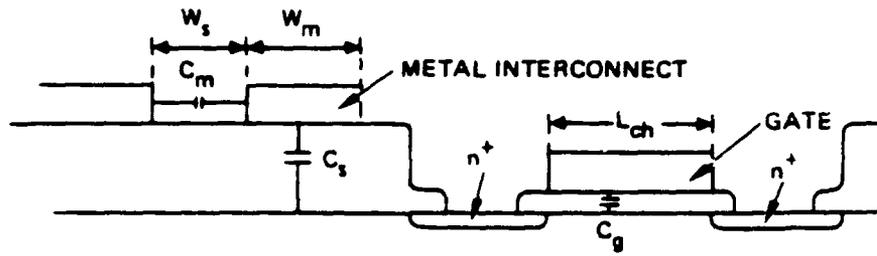


Figure 5. Scaling relationships of various parasitic capacitance ratios to gate oxide capacitance.

capacitance. Figure 5 shows the total line-to-substrate ($C_s(\text{chip})$) and line-to-line ($C_m(\text{chip})$) wiring capacitance for a chip dominated by short wiring interconnects.

In the above analysis it was assumed that all feature sizes scaled linearly with x . In practice, however, this would not be desirable; technology limitations would prevent the scaling down of certain feature sizes before others. For example, as the gate oxide is scaled below 150 Å, breakdown mechanisms begin to occur, causing leakage currents across the oxide, and reducing the reliability of devices. Below 50 Å, direct quantum mechanical tunneling of electrons across the oxide occurs. Hence, for MOS devices to be useful at submicron dimensions, the gate and field oxide thickness would have to be scaled down less than linearly. Effectively, gate and wiring capacitances would decrease more than linearly with x . As wire widths are scaled down, however, the resistance of these wires increases (Figure 6). At submicron dimensions, the resistance of these wires would be great enough to introduce considerable RC time delay in the propagation of signals. Also, as wire widths are scaled down, electromigration failure becomes more prominent. These two phenomena dictate that wire widths (and spacing) would have to be scaled down less than linearly. This would result in greater line-to-substrate capacitance (in CMOS/bulk circuits) and less line-to-line capacitance than if scaling were done linearly.

Another source of interconnect capacitance in CMOS/bulk is diffusion line capacitance. The contribution from diffusion line capacitance tends to increase as feature sizes decrease. This is because as feature sizes decrease, substrate doping is increased for threshold compensation, resulting in smaller depletion regions at diffusion junctions, and hence, greater capacitance. The exact relationship between diffusion capacitance and the scaling down of feature sizes is left for further study.

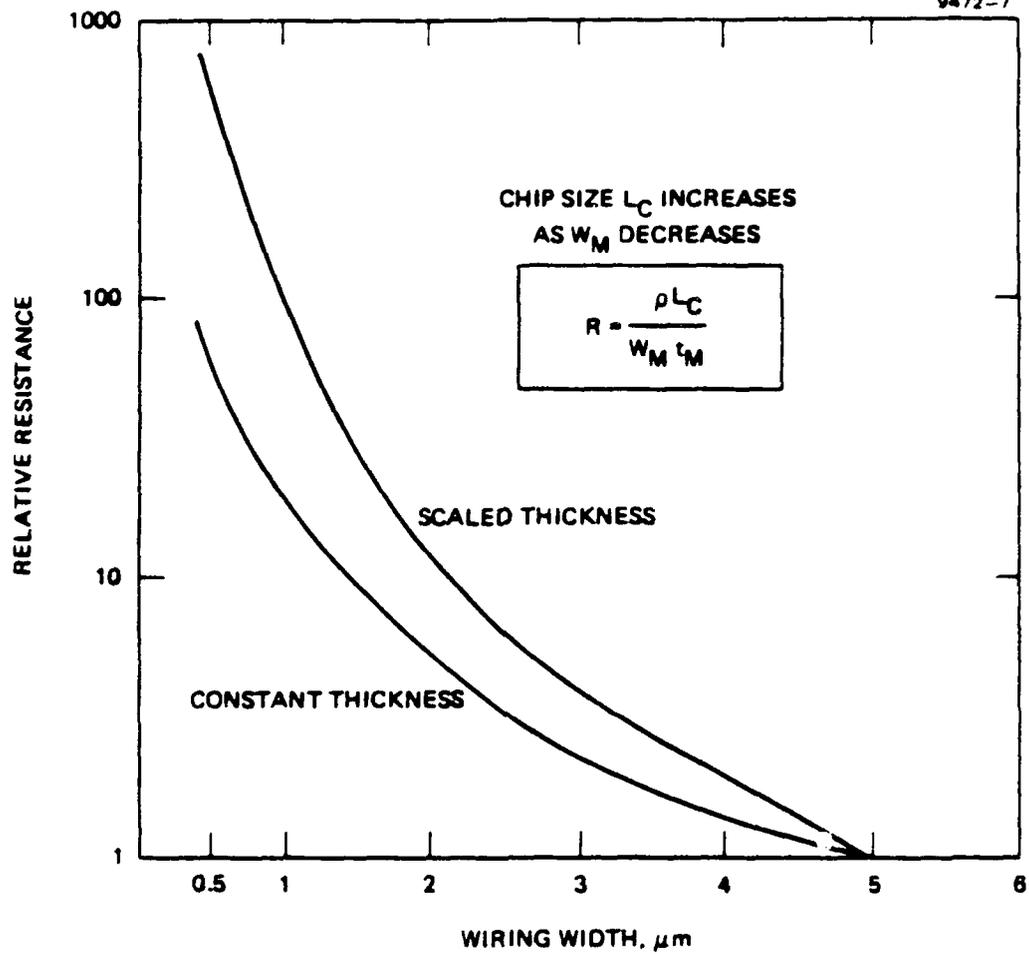


Figure 6. Interconnect resistance scaling.

This section has summarized the results of a study on wiring capacitances as feature sizes are scaled down linearly. Although feature sizes do not scale linearly in practice, we will use the results of this study in our report and assume that gate capacitance is dominant in CMOS/SOS circuits at the VHSIC gate length of 1.2 μm .

4. VHSIC Technology

The current emphasis of the VHSIC program at Hughes is on 1.2 μm CMOS/SOS technology. This technology has been demonstrated with fabrication of a 72,000 device correlator chip for the VHSIC program. Research is also underway to develop a sub-micron SOS process at either 0.75 μm or 0.5 μm feature size. The projections are that a submicron technology will be available in 1987. There is also interest at Hughes Newport Beach in CMOS/bulk technology where a 3.0 μm process is available at this time. Also, a 1.2 μm CMOS/bulk process is currently being developed. In this section we will summarize the performance characteristics of current VHSIC 1.2 μm technology as applied to the low-power filter design.

Table 5 summarizes the feature sizes and electrical parameters of the Hughes 1.2 μm CMOS/SOS process.⁴ The minimum drawn gate length is 1.4 μm , resulting in a channel length of 1.2 μm after lateral diffusion at source and drain are taken into account. The threshold voltages of the p and n-channel devices are nearly identical at 1.2V. With gate oxide thickness at 400 \AA , the gate oxide capacitance is 8.6×10^{-4} pF/ μ^2 . From Section 2.B.3, we can assume that gate capacitance will be the primary source of power dissipation in CMOS/SOS circuits. For a minimum geometry device of 2 μm x 1.2 μm , the capacitance per gate is

$$C_g = 8.6 \times 10^{-4} \text{ (pF}/\mu^2) \cdot 2 \times 1.2 \text{ (}\mu^2) = 2.1 \times 10^{-3} \text{ pF/gate.}$$

Table 5. Hughes VHSIC CMOS/SOS Process Parameters

<u>Minimum Dimensions</u> (μ)			
Transistor Length	1.4	(Leff = 1.2)	
Transistor Width	2.0		
Metal Width	2.4		
Metal Spacing	2.6		
Polysilicon Width	1.4		
Polysilicon Spacing	2.2		
<u>Nominal Thickness</u> (\AA)			
Silicon	5000		
Poly Silicide	5500		
Field Oxide	5000		
Metal	7500		
Gate Oxide	400		
<u>Electrical Parameters</u>			
<u>Contact Resistances</u> (ohms) for $2\mu \times 2\mu$ contact area			
	<u>Max</u>	<u>Typical</u>	
Metal/N ⁺	100	50	
Metal/P ⁺	30	10	
Metal/Poly	5	2	
N ⁺ /Metal/P ⁺	250	100	($2\mu \times 4\mu$ contact area)
<u>Sheet Resistances</u>			
N ⁺ Si	100	40-60	
P ⁺ Si	200	100-115	
Poly Silicide	5	2.5-4.5	
Metal	0.05	0.04	
<u>Electrical Parameters</u>			
V _{TN}	1.2 V.	V _{TP}	-1.2 V.
C _{OX}	8.6×10^{-4} pF/ μ^2		
μ_N	300 cm ² /V-sec	μ_P	170 cm ² /V-sec
g _N	62 μ hos	g _P	35 μ hos
R _{ch,n}	16 Kohms/sq.	R _{ch,p}	28.6 Kohms/sq.

The corresponding dynamic power dissipation for a minimum geometry device operating with a supply voltage of 3.6V at 8MHz is

$$\begin{aligned} P_g &= 2.1 \times 10^{-3} \text{ (pF/gate)} (3.6\text{V})^2 8 \times 10^6 \text{ (MHz)} 50\% / 2 \\ &= 0.054 \text{ uW/gate.} \end{aligned}$$

The 50% factor arises from assuming that the device changes state every other cycle. We will use these values for C_g and P_g in Section 2.C for calculating the power consumption of components for the low power filter.

From Table 5, the channel resistance in the linear region is 16 K Ω /sq for an n-channel device and 28.6 K Ω /sq for a p-channel device. These values are much greater than the interconnect and contact resistances shown in the same table. Therefore, in estimating the speed of circuit components for the filter in Section 2.C, we will assume that channel resistance and gate capacitance contribute the most to propagation delay in the circuits.

5. CCDs for Data Storage

In the preliminary study of the low-power filter chip, it was discovered that the memory for storage of data and filter coefficients (about 1.5b Kbits total) consumed a significant amount of power. In this section we will consider the use of CCD shift registers to see if power for data storage can be minimized.

Since most of the power dissipated in a CCD shift register is CV^2f dynamic power, we will try to minimize this power by assuming a design based on four 256-word shift registers connected as shown in Figure 7. Every 1/(8KHz) second the B switches connect the four shift registers into one long 1024-word shift register, and a new datum is inserted at INPUT into the shift register. The B switches are then flipped the other way so that data within each of the four shift registers can

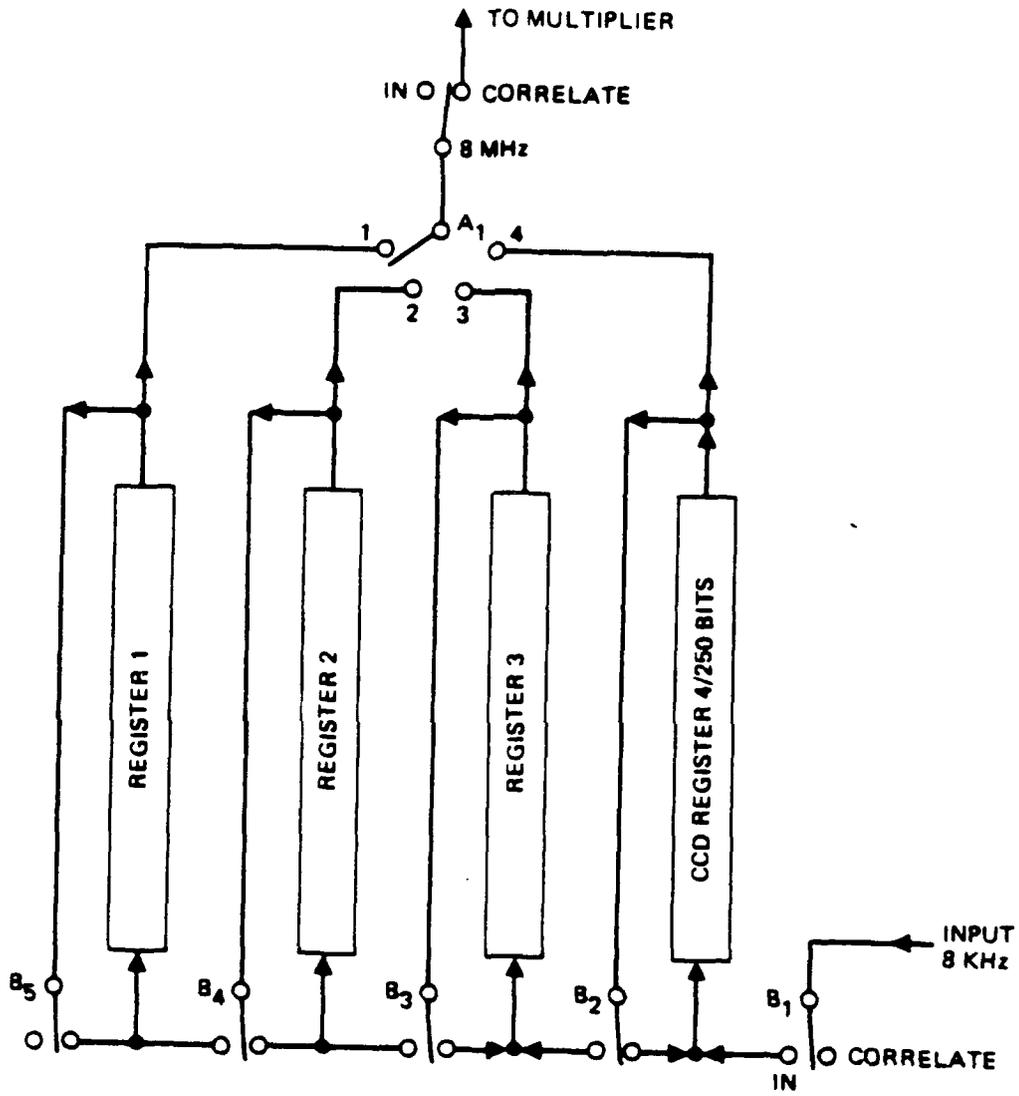


Figure 7. CCD memory partitioned into four segments.

circulate within themselves. However, only one of the 256-word shift register is circulating at one time, and as it does so, data is fed through switch A at 8MHz to the multiplier for correlation. The shift rate for each shift register is 8MHz, although each shift register stops for

$$\frac{256}{8 \text{ MHz}} \times 3 \sim 96 \mu\text{s}$$

while data is being fed from the other three shift registers. Effectively, each shift register is working at 2 MHz. The shift register for coefficient storage would be treated in a similar way.

Table 6 lists the parameters for CCD technology developed at Hughes. Using a minimum line width of 2.5 μm and a cell size of 10 μm x 10 μm results in an array area of $1.5 \times 10^5 \text{b } \mu\text{m}^2$ for 1500b bits of data storage. This translates to 50b pF of electrode capacitance. The minimum operating voltage of these CCD circuits is 6V. Hence, the power required to operate this memory array is

$$P = 50\text{b pF} \cdot (6.0\text{V})^2 \cdot 2\text{MHz} = 3.6\text{b mW}.$$

For a minimum word size of 8 bits, the power consumed would be approximately 29 mW, exceeding the low power filter requirement. Also, the need to run the CCD array at a minimum of 6V tends to rule out the use of CCD technology for the low power filter.

C. ARCHITECTURAL ISSUES

In the preliminary report we advocated a single tap approach for the low power filter (Figure 8). In the following sections we will estimate the operational speed, power dissipation, device count, and chip size of a single tap filter

Table 6. Hughes CCD Technology

Minimum Line Width	2.5 μ
Cell Size	10 μ x 10 μ
Oxide Thickness	1000 \AA
Total Array Area (1500b bits)	1.5 x 10 ⁵ b μ^2
Total Capacitance	50 b pF
Power Consumption (@ 6V, 2 MHz)	3.6 b mW

14003-8

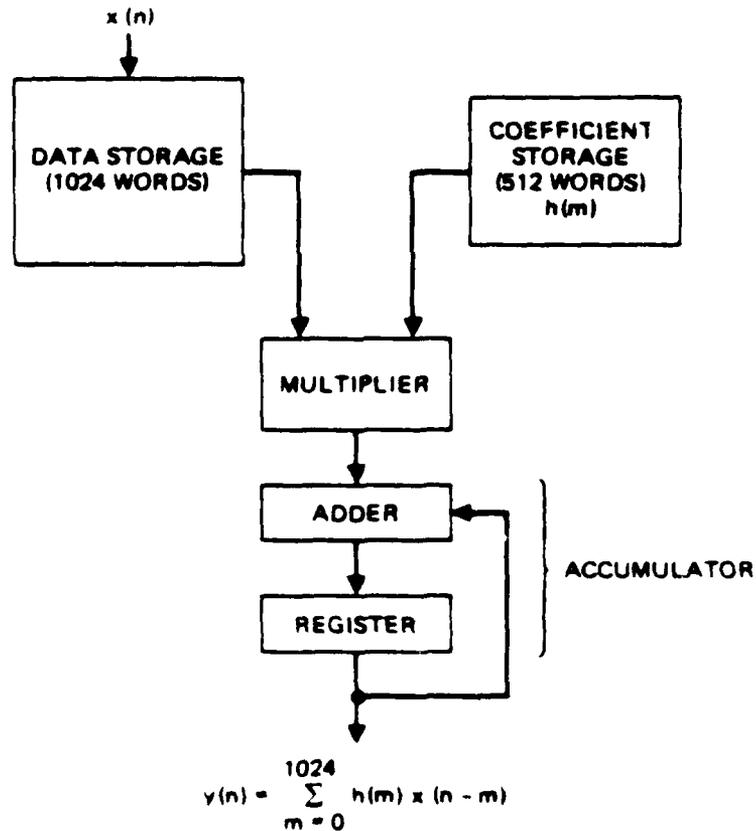


Figure 8. Single-tap implementation of low power filter.

when implemented using current Hughes VHSIC CMOS/SOS technology. We will use the technology parameters presented in Section 2.B.4 of this report. In Section 2.C.6 we will examine the trade-offs of a multi-tap implementation for the filter.

In estimating the power dissipation of the filter components, we will make two simplifying assumptions. First, we will assume that parasitic capacitances for CMOS/SOS technology, including line-to-line interconnect capacitances and interconnect crossover capacitance, is small compared with gate capacitances. From the data in Figure 5, this is a reasonable assumption. In this figure we see that line-to-line interconnect capacitance at $1.2 \mu\text{m}$ is an order of magnitude less than gate capacitance. Second, we will assume that in the case of RAMs, power consumption is proportional to the size (number of bits and chip area) of the RAM. This is not strictly true, since in CMOS RAMs a considerable amount of power is dissipated in driving the capacitances of long data and address lines. These capacitances do not scale linearly with RAM size. However, we will scale down power consumption for state-of-the-art CMOS static RAM chips to obtain first order estimates of power dissipation for data and coefficient storage in the filter. Making these two assumptions for power dissipation will enable us to estimate total power consumption for the filter without laying out the components first.

In obtaining power consumption estimates for the filter, we will assume that cell designs utilize minimum geometry devices. Cells in the Hughes VHSIC library are usually designed to drive large capacitive loads, and hence dissipate more power and operate at higher speeds. In estimating the speed and power for a filter design using minimum geometry devices, the VHSIC values for power dissipation and operational speeds will be scaled down accordingly. Also, in calculating the power dissipation of cells in the filter, we will assume that only half the electrical modes in the cell change state during any clock cycle. This provides a conservative estimate for power dissipation, since probably fewer than 50% of the modes change state

every cycle. For the purpose of estimating power dissipation for the filter, however, we will use the conservative value calculated using $1/2 CV^2f$ (50%).

1. Adder/Accumulator Configuration

In this section we will present the basic adder and accumulator configurations to be used in the low-power filter. The adder is a crucial element used repeatedly in the accumulator and multiplier sections of the filter. We will obtain estimates for the power, speed, and silicon area of the adder and accumulator in terms of current VHSIC 1.2 μm CMOS/SOS technology.

The full adder cell being used in the Hughes VHSIC program is shown in Figure 9.⁵ There are a total of 26 devices in the cell, with channel widths ranging from 5 μm to 17 μm . For the low power filter, however, we will assume a minimum geometry device design with channel width of 2.0 μm . Using the parameters in Table 5 for the VHSIC 1.2 μm CMOS/SOS process, the total gate capacitance was calculated to be

$$\begin{aligned} C &= 8.6 \times 10^{-4} \text{ (pF}/\mu^2) \times 2\mu \times 1.2\mu \times 26 \\ &= 0.054\text{pF.} \end{aligned}$$

Total power dissipation for the cell is given by

$$\begin{aligned} \text{Power} &= \frac{1}{2} CV^2f. \text{ (50\% duty cycle), } \quad C=0.054\text{pF, } V=3.6\text{V, } f=8\text{MHz} \\ &= 1.4\mu\text{W.} \end{aligned}$$

The 50% duty cycle arises from assuming that only half the nodes in the cell change state every cycle.

The performance of this cell has been simulated as part of the VHSIC effort. From the simulations it was found that the propagation delay from input to the SUM output was approximately 8.5ns. Propagation delay from C_{in} to C_{out} for a cell was 3.6ns. These propagation delay times would be longer if minimum geometry devices were used in the cell.

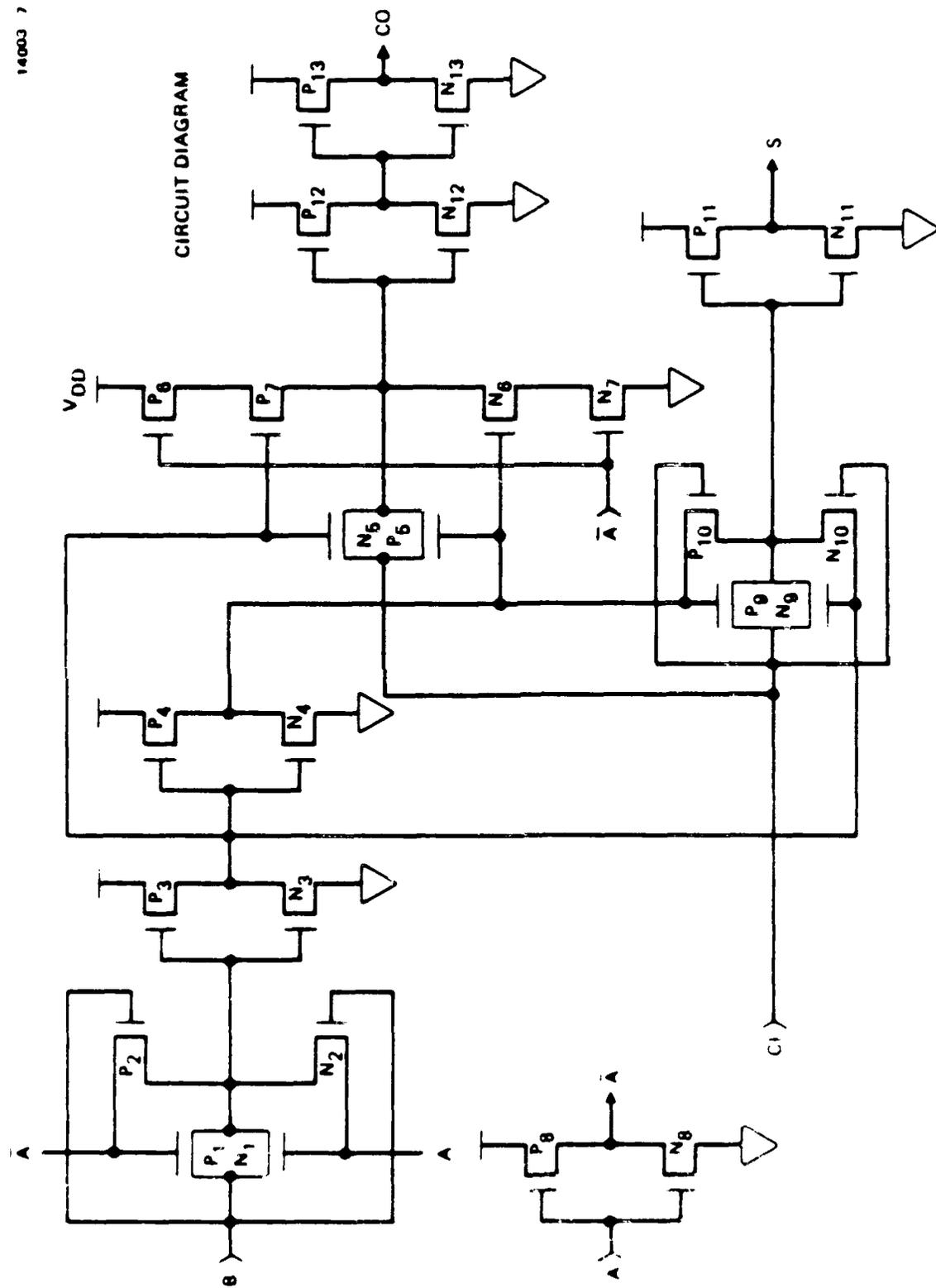


Figure 9. VHSIC full adder cell.

Accumulator

The output of the multiplier has to be summed for 1024 cycles to obtain one convolution point. To perform this summation, an accumulator with $b+10$ adders is provided, as shown in Figure 10. The outputs of these adders are captured in shift registers and fed back to the adders every clock cycle to be added to the next output from the multiplier. The maximum carry propagation through the accumulator using VHSIC device geometries for $b=12$ bits is

$$T_c = 22 \times 3.6\text{ns} = 79.2\text{ns},$$

much less than the cycle time for the filter. This propagation delay would increase if minimum geometry devices are used to implement the adders. If this delay becomes longer than the filter cycle time, carry-lookahead techniques may be used to decrease the delay. Table 7 lists the number of devices in the accumulator and the power dissipation when minimum geometry devices are used. The total power dissipated by the adder/accumulator section is

$$P_{\text{acc}} = 2.26b + 18.3 \text{ uW},$$

and the total device count is

$$D_{\text{acc}} = 42b + 340 .$$

2. Multiplier Configuration

In this section we will consider the multiplier for the low-power filter in terms of the Hughes VHSIC 1.2 μm technology. There are several possible multiplier configurations (tree, array, ROM-based), but the configuration based on a modified Booth's algorithm seems to be best from the standpoint of power

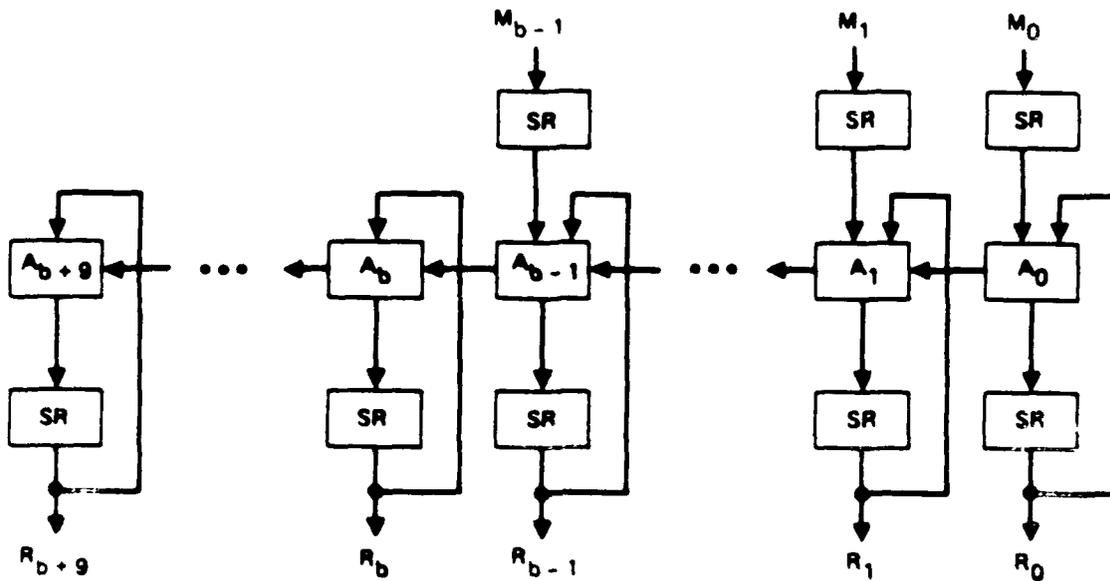


Figure 10. Accumulator for adding 1024 b-bit words.

Table 7. Device Count and Power Dissipation for Accumulator

Cell Type	† Cells	Devices per cell	Total † Devices	Power (μ W)
Adder	$b + 10$	26	26 ($b + 10$)	1.4 ($b + 10$)
SR	2 ($b + 5$)	8	16 ($b + 5$)	0.86 ($b + 5$)

dissipation, device count, speed, and layout geometry. This algorithm uses a radix 4 method to examine the multiplier word 3 overlapping bits at a time. Partial products are accumulated in half the number of steps as necessary in other schemes. Moreover, negative numbers are handled as well, in 2's complement form. Table 8 shows the method for accumulating partial products based on examining 3 bits of the multiplier.

The Booth multiplier, implemented in pipeline fashion, is shown in Figure 11. The incoming multiplier words are stored in shift registers at the right, and decoded 3 overlapping bits at a time by the Booth Decoders (BD's). Depending on the value of the 3 bits, the control lines to the Select circuits are activated to add either 0, X, 2X, -X or -2X to the partial product. The actual addition is performed in ripple carry form using the full adder described in the previous section. Note that only $(b/2-1)$ rows of adders are needed to accumulate the partial products.

The worst case propagation delay through one stage of the pipeline is the sum of set-up time for the Booth decoders, select circuits, and carry propagation through the adders:

$$T_{\text{stage}} = T_{\text{bd}} + T_{\text{sel}} + 12 * T_{\text{carry}}$$

This propagation delay through one stage would be 63ns if VHSIC geometry devices were used in the design, but would be longer if minimum geometry devices were used. We do not anticipate that one pipeline stage delay will exceed the filter cycle time even if minimum geometry devices are used in the design.

Table 9 lists the number of cells and devices used in the pipelined Booth multiplier. The power dissipation shown for each cell type is based on $CV^2f(50\%)/2$ dynamic power, where C is the total gate capacitance in the cell, V is 3.6V, and f is 8MHz. The 50% factor is included assuming only half the nodes in the cell change state each cycle. The total power consumed by the multiplier is

Table 8. Modified Booth's Algorithm for Accumulating Partial Products

$Y_{i+1} Y_i Y_{i-1}$	Add to Partial Product
0 0 0	0
0 0 1	X
0 1 0	X
0 1 1	2X
1 0 0	-2X
1 0 1	-X
1 1 0	-X
1 1 1	0

X - Multiplicand

Table 9. Device Count and Power Consumption for Components of Pipelined Booth Multiplier

	Devices Per Cell	# Cells	Total # of Devices	Power (μ W)
Shift Register	8	$\frac{b}{2}(\frac{5}{2} b-1)$	$4b(\frac{5}{2} b-1)$	$0.22b(\frac{5}{2} b-1)$
Select	18	$b^2/2$	$9b^2$	$0.48b^2$
Booth's	98	$b/2$	$49b$	$3.8b$
Full Adder	26	$b(\frac{b}{2} - 1)$	$26b(\frac{b}{2} - 1)$	$1.4b(\frac{b}{2} - 1)$

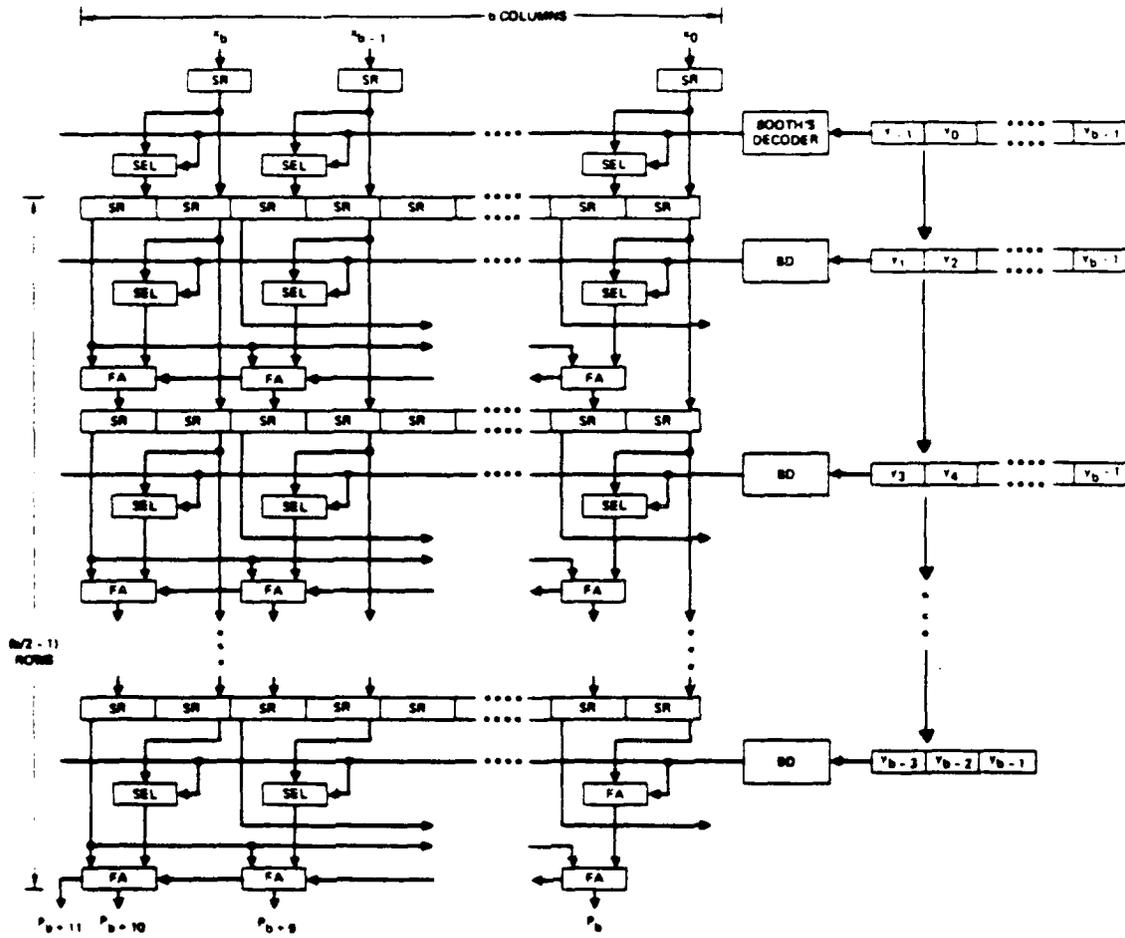


Figure 11. Pipelined booth multiplier.

$$P_{\text{mult}} = b(1.74b + 2.16) \text{ uW} ,$$

and the total device count for the multiplier is

$$D_{\text{mult}} = b(32b + 19).$$

3. Data and Coefficient Storage

In the preliminary study of the low-power filter it was estimated that the data storage section would consume the most power on chip. This estimate was based on a Hughes 16K static RAM fabricated using 2.5 μm CMOS/SOS technology about 3 years ago. For a more accurate estimate of speed and power dissipation of current CMOS memory chips, we surveyed the papers presented at the most recent ISSCC conference (held February 1984). These papers are representative of what can be achieved in memory design today.

Table 10 shows in summarized form the characteristics of the low power RAMs presented at the ISSCC conference.⁶ Most of these RAMs use CMOS technology, and the effective gate lengths are comparable to current Hughes VHSIC gate lengths (1.2 μm). The access times of the static RAMs are in general better than those for dynamic RAMs because charge sensing and refreshing of dynamic memory cells require a longer cycle time. The access times of the static RAMs would meet the requirements of the low power filter easily, but the longer cycle times of the dynamic RAMs would be a problem.

The lowest power RAM in Table 10 is a 256K CMOS static RAM developed by Toshiba (Appendix A). The gate lengths are 1.2 μm for n-channel devices and 1.5 μm for p-channel devices (current Hughes VHSIC technology uses 1.2 μm for both n and p-channel devices). Access time for this chip is 46ns, much less than the 125ns required in the low power filter. Active power dissipation measured at 1MHz is 10mW, and standby power is 0.03mW. If we scale the active power according to the requirements for the

Table 10. Performance Characteristics of Current RAM's

Company	Size	Tech	L_{eff} (n-ch/p-ch) (μ)	Cell Size (μ)	Access Time (ns)	Cycle Time (ns)	Active Power (mW)	Standby Power (mW)	Power/Kb (mW)
<u>Static RAMs</u>									
Toshiba	256Kb	CMOS	1.2/1.5	11 x 13.5	46		10 @ 1 MHz	0.03	0.04 @ 1 MHz
Inmos	64Kb	CMOS	1.5/1.7	12.2 x 23.9	30		250	60	3.9
NEC	64Kb	CMOS	1.5/1.8	14.9 x 19.3	25		350	15	5.5
Hitachi	64Kb	CMOS	1.3/1.3	8.0 x 16.0	22		70 @ 1 MHz	<1	1.1 @ 1 MHz
Toshiba	64Kb	CMOS	1.2/1.5	18.0 x 20.0	28		225	110	3.5
<u>Dynamic RAMs</u>									
NTT Atsugi	1Mb	CMOS	0.5/0.9	3.7 x 5.4	140	350	250	5	0.24
NEC	1Mb	NMOS	1.0	5.5 x 8.0	120	300	290 @ 3 MHz	15	0.28
Hitachi	288Kb	CMOS	2.0	6.8 x 13.6	100	100	400 @ 10 MHz	10	1.4 @ 10 MHz
IBM	256Kb	NMOS	1.8	5.7 x 18	80	180	300 @ 3 MHz	25	1.2 @ 3 MHz
Hitachi	1Mb	NMOS	1.6	3.0 x 7.0	90	260	300 @ 3.8 MHz	10	0.29 @ 3.8 MHz

low power filter (data storage = 1.5b Kbits, voltage supply = 3.6V, speed = 8Mhz), we obtain

$$P = \frac{1.5b}{256} \cdot \left(\frac{3.6}{5.0}\right)^2 \cdot \frac{8}{1} \cdot 10 \text{ mW} = 0.243b \text{ mW.}$$

From this calculation, we can see that a static RAM with 1500b bits of storage could be designed using state-of-the-art CMOS technology that would meet the specifications of the low power filter. Such a RAM would require approximately 9000b devices to implement, assuming a design utilizing 6 devices/cell.

We should provide a note of caution here. The technology used for fabricating the low power Toshiba RAM chip is a two-level polysilicon, two-level metal, p-well CMOS process. Each memory cell utilizes 4 transistors and 2 polysilicon resistor loads. These polysilicon resistors require tight processing tolerances so that the resistances would be relatively temperature invariant, and uniform resistances are maintained across the wafer. This requires fairly advanced processing techniques that may not be widely available yet. Without a CMOS process similar to this, it may be difficult to develop a RAM that would meet the low power specifications of the filter.

In comparison, the next lowest power consuming RAM presented at the ISSCC conference was a 64K CMOS static RAM from Hitachi (Appendix B). The gate length for both n- and p-channel devices in this chip was 1.3 μm . The power consumption measured for this chip at 8MHz was about 150mW (Figure 5, Appendix B). If we scale this down for the low power filter's requirements, we obtain

$$P = \frac{1.5b}{64} \cdot \left(\frac{3.6}{5.0}\right)^2 \cdot 150 \text{ mW}$$
$$= 1.8b \text{ mW.}$$

Obviously, this would exceed the filter specification for power consumption. We stress here again that developing a RAM that would meet the low power specification for the filter is critically dependent on the availability of a suitable CMOS technology.

Another scheme for lowering the power consumption of the data storage section further is to partition the memory into segments and activate one segment at a time, when data is needed from it. This scheme is illustrated in Figure 12. Since dynamic power is proportional to capacitance, C , and assuming that capacitance is proportional to area, A , then decreasing the area of the memory activated at one time by 4 would result in lowering the power consumption by a factor of 4 as well. Each segment of the partitioned memory still operates at 8MHz when that segment is activated, but as far as power consumption is concerned, the effective operational rate is 2MHz.

4. Output Drivers

Output drivers can consume a significant amount of dynamic power because of the relatively large off-chip capacitance they have to drive. In this section we will consider the power dissipated by output drivers in terms of current Hughes VHSIC technology. This will determine whether it is possible to divide the implementation of the filter into two or more chips and yet maintain low power consumption.

Figure 13 shows a typical CMOS output driver handling an off-chip capacitance of 20pF. The drive ratio of the output devices have to be large enough to ensure that rise and fall times of the off-chip signal is satisfactory. Here, they are shown with 100:1 ratios, adequate to drive 20pF loads in less than 30ns. Another beefed-up inverter is used to drive the considerable capacitance at gates P1 and N1 (calculated to be about 0.25pF). The drive ratios of P2 and N2 are chosen to be 10:1, presenting a gate capacitance of about 0.025pF to the previous stage.

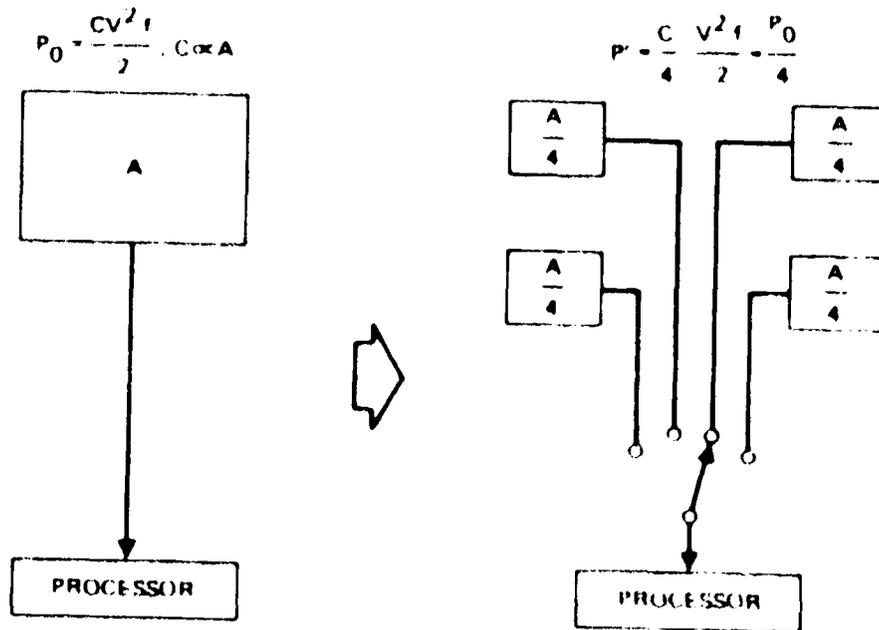


Figure 12. Reducing the power consumption of data storage by partitioning into four segments.

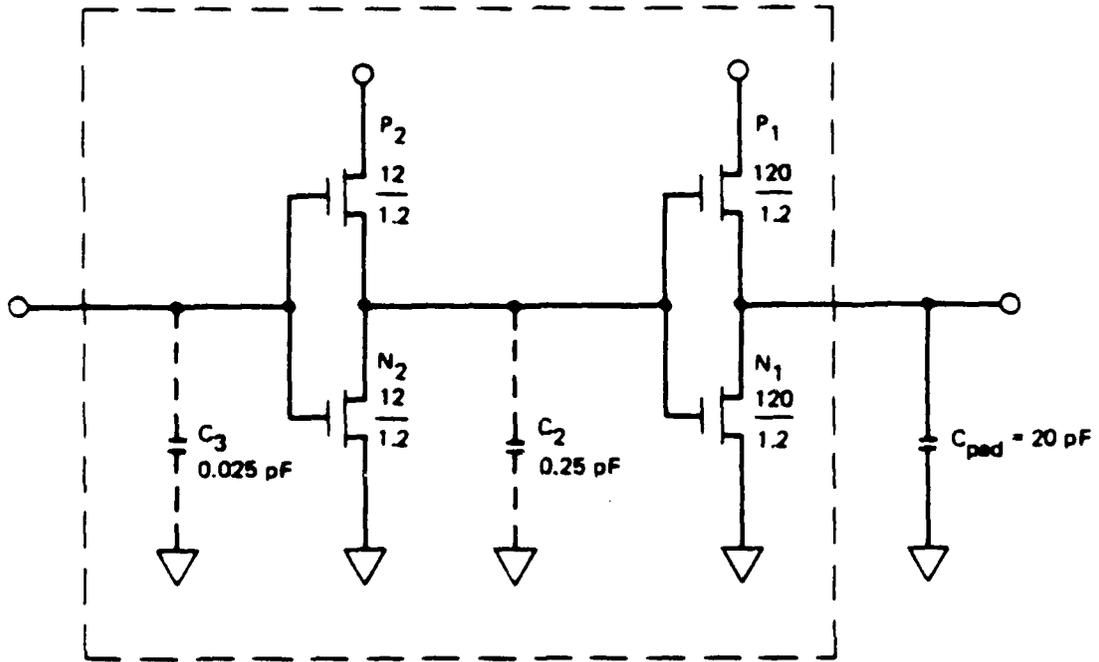


Figure 13. Output driver configuration.

Dynamic power dissipation is calculated using $CV^2f/2$. For the case where the filter is partitioned into two or more chips (for example, if the RAM were on a chip separate from the multiplier (accumulator), the output drivers handling signals between these chips would have to operate at processor speed, (i.e., 8MHz for a single tap implementation). The power consumed by a single such driver would then be

$$P = 20.275pF \cdot (3.6V)^2 \cdot 8MHz \cdot (50\% \text{ duty cycle}) / 2 \\ = 0.53mW$$

Twelve such drivers operating in parallel (actually more than 12 would be needed to handle communications between data storage and processor for a 12-bit implementation) would consume more than 6mW. Comparatively, if the filter were totally integrated on one chip, the output drivers would operate at 8KHz, and the power consumed by each would be 0.53uW. Also, considerably fewer output drivers would be needed in an integrated chip approach, since no inter-chip communication would be necessary. From this analysis, we strongly favor the single chip approach.

5. Total Power Requirements, Device Count, and Chip Size

In this section we will summarize the results of preceding sections and obtain estimates for total power consumption, device count and chip size for the low power filter. The cell types used in the filter are listed in Table 11, together with the power dissipation and device count for each cell. Table 12 shows the breakdown for the major components of the filter in terms of cell types. The power consumption and device count of each major component is summarized in Table 13. Total power consumption for the filter is obtained by adding up the power consumed by each component, resulting in

$$P_{\text{total}} = 1.74b^2 + 246b + 18.3 \mu W$$

Table 11. Power Dissipation and Device Count for
Cells Used in Low Power Filter

Cells	Power @ 8 MHz (μ W)	Device Count
Shift Register	0.43	8
Full Adder	1.4	26
Select Cell	0.96	18
Booth Decoder	7.5	98
RAM	0.16	6

Table 12. Power Dissipation and Device Count for Functional Blocks of Filter,
 SR-Shift Register, FA-Full Adder, SEL-Select, BD-Booth Decoder,
 RAM-Random Access Memory, OD-Output Driver

Functional Block	# of Cells						Device Count	Power (μ W)
	SR	FA	SEL	BC	RAM	OD		
Data Storage					1024b		6100b	160b
Coefficient Storage					512b		3100b	81b
Adder/Accumulator	$2(b + 5)$						$16(b + 5)$	$0.86(b + 5)$
		$b + 10$					$26(b + 10)$	$1.4(b + 10)$
Multiplier	$\frac{b}{2}(\frac{5}{2}b - 1)$						$4b(\frac{5}{2}b - 1)$	$0.22b(\frac{5}{2}b - 1)$
			$b(\frac{b}{2} - 1)$				$26b(\frac{b}{2} - 1)$	$1.4b(\frac{b}{2} - 1)$
Output Drivers (8 kHz)			$b^2/2$				$9b^2$	$0.48b^2$
				$b/2$		b	49b	3.8b
							4b	0.53b

Similarly, the total device count is obtained by adding up the total number of devices in each component:

$$D_{\text{total}} = 32b^2 + 9265b + 340 \text{ devices.}$$

This device count does not include circuitry for timing, control, and memory address generation. To allow for these extra components, an extra 5 to 10% must be added to the total device count calculated above. The estimate for total power consumption must be similarly increased to account for the extra circuitry.

The power consumption and device count for a 10-bit filter broken down by major components is shown in Table 2. One should note that data and coefficient storage make up about 90% of the device count and power consumed. Table 3 lists the total power consumption and device count for an 8, 10 and 12 bit filter using the equations for P_{total} and D_{total} just derived.

To estimate the size of the filter chip we have to estimate the area taken up by RAM and the area taken up by random logic. To estimate the device density for random logic, we note that the Hughes VHSIC correlator chip contains 72,000 devices and measures $315 \times 368 \text{ mil}^2$. The resulting device density for this chip is $0.62 \text{ device/mil}^2$. We will assume that this is a representative density for random logic fabricated using $1.2 \mu\text{m}$ technology. To estimate the device density for memory arrays, we note that the Toshiba 256K RAM mentioned in Section 2.3.4 contains approximately

$$256K \text{ cells} \cdot 4 \text{ devices/cell} = 10^6 \text{ devices.}$$

on a chip measuring $6.68 \times 8.86 \text{ mm}^2$ ($263 \times 349 \text{ mil}^2$). The device density for this chip is $10.9 \text{ devices/mil}^2$. The chip size for the filter can be estimated by adding the area occupied by RAM and the area occupied by the processor:

$$\begin{aligned}
 S &= 9200b/10.9 + (32b^2 + 65b + 340)/0.62 \text{ mil}^2 \\
 &= 51.6b^2 + 949b + 548 \text{ mil}^2.
 \end{aligned}$$

The projected chip size for an 8, 10 and 12-bit filter is listed in Table 3. Approximately half the chip area is taken up by data and coefficient storage.

6. Single Tap Versus Multiple Taps

In our preliminary study we chose a single tap implementation (Figure 8) for the low power filter. A single tap implementation involves buffering the incoming data samples in memory and then multiplexing the data and filter coefficients at high speed into the multiplier/accumulator portion of the filter. Whereas the data samples are acquired at 8KHz, this scheme requires the processor section to be run at 8MHz. A multi-tap approach could be run at lower speeds, but would require more gates to implement. Since dynamic power dissipation in CMOS technology is directly proportional to total capacitance and speed of operation, there is a trade-off between using a multi-tap and a single tap approach.

We will now try to estimate the trade-off between using a single tap and a multi-tap approach in terms of power consumption. Figure 14 shows a four-tap implementation for the low power filter. From this figure we see that even though data and coefficient storage become partitioned as the number of taps, N, increases, total data and coefficient storage remains the same. Power consumed by the RAM can be estimated in the same way as in Section 2.C.3:

$$P = \frac{1.5b}{256} \cdot \left(\frac{3.6}{5.0} \right)^2 \cdot \frac{f}{N} \cdot 10 \text{ mW} = 0.243b/N \text{ mW}.$$

Here f is the operational speed for the single tap approach, 8MHz, and f/N is the operational speed for an N-tap filter. The number of multipliers and adders, would increase with the number of taps, so the total capacitance for these components would be

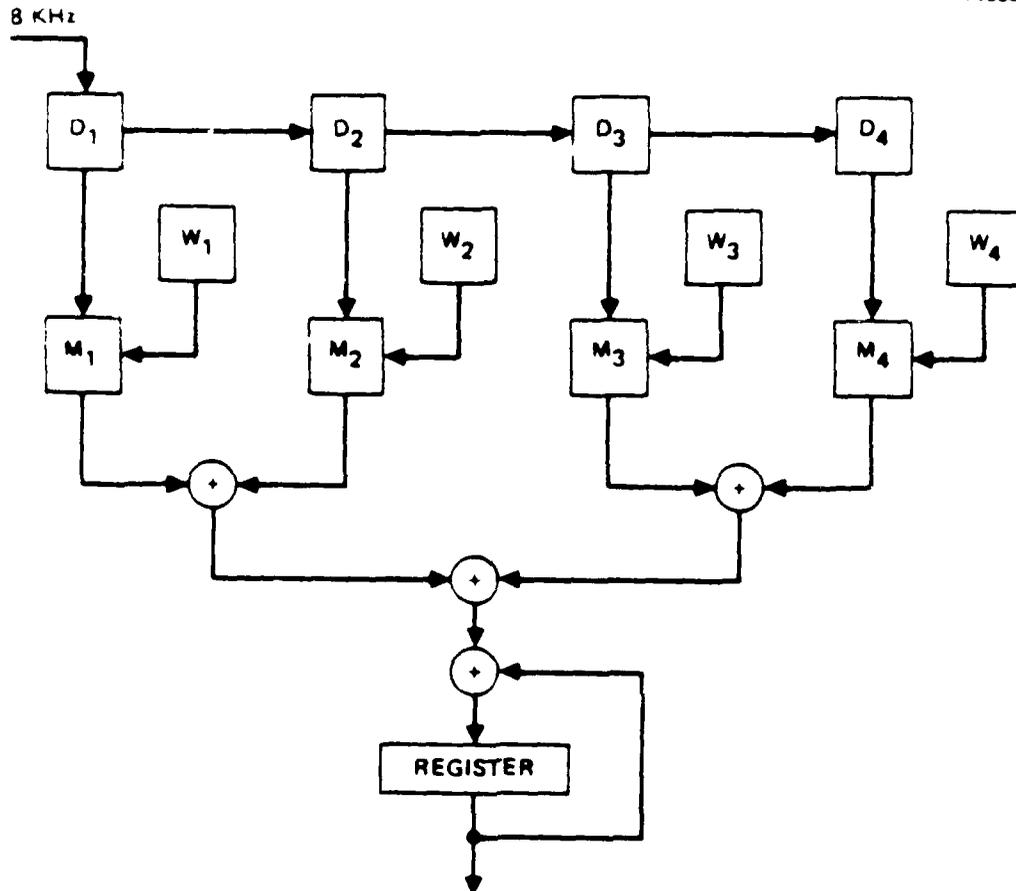


Figure 14. Four-tap approach for low power filter.

$N(C_m + C_a)$, where C_m and C_a are the total capacitances associated with a multiplier and an adder, respectively. The power dissipated by these components would be

$$P = N(C_m + C_a)V^2 (f/N)(50\%)/2 = 25.6 \times 10^6 (C_m + C_a).$$

Combining this with power dissipation for the RAM and results from sections 2.C.1 and 2.C.2, total power for an N-tap filter is given by

$$P(b,N) = 243b/N + 1.73b^2 + 2.43b + 14 \text{ uw.}$$

Hence, total power for an N-tap filter shows a decrease only with the data and coefficient storage component. This component, however, consumes 90% of filter power, and any power savings here would be significant.

Table 4 lists the power consumption, device count and chip size for a 10-bit filter implemented with 1 through 4 taps. Device count is obtained with the help of Table 12,

$$D(b,N) = 9220b + 80 + N(32b^2 + 45b + 260) ,$$

and chip size obtained in a similar manner to section 2.C.5,

$$\begin{aligned} S(b,N) &= 9200b/10.9 + [20b + 80 + N(32b^2 + 45b + 260)]/0.62 \\ &= 876b + 129 + N(51.6b^2 + 72.6b + 419) \text{ mil}^2. \end{aligned}$$

The percent increase or decrease in power consumption, device count and chip size over the single tap approach is also noted in Table 4. Going from a single tap to a two-tap filter would decrease power consumption by 46%, but would also increase chip size by 41%.

Chip yield must be considered when using a multi-tap approach. The low power filter is a substantial size chip even

when a single tap approach is used (Section 2.C.5). If extra multipliers and adders are to be added for a multi-tap approach, then fabrication yields would decrease as chip size increases. A multi-chip implementation may be used, but this approach requires inter-chip communication, and as shown in Section 2.C.4, significant power is consumed by output drivers handling off-chip capacitances. We believe that this power will not be compensated for by operating at a lower speed.

D. OTHER DESIGN CONSIDERATIONS

So far this report has assumed that the inputs are provided in digital form, and the input signal does not contain aliasing. This assumes that the original analog signal has been "conditioned"; that is, it has been sent through an automatic gain control (AGC) circuit and pre-filtered with a cut-off at about 4KHz to prevent aliasing. Furthermore, the conditioned signal has to be digitized. Whether the analog-to-digital (A/D) converter is to be included as part of the low power filter chip has to be considered. We have not examined the speed and power requirements of the A/D converter in this report, but it is an important part of the filter and needs further study. We believe, however, that the device count and power dissipation in an A/D converter would be small compared with the total device count and power consumption for the filter.

E. COST OF FABRICATION

A preliminary estimate of the cost of fabricating the low power filter chip using Hughes VHSIC CMOS/SOS technology has been made by our Industrial Electronics Group (IEG) at Carlsbad. This cost estimate is presented in Table 13, and includes processing 2 lots of wafers - almost a necessity for a chip of the filter's complexity. Errors in the first design will be eliminated in the second lot. A tentative schedule for fabrication of the filter chip over a 24 month period is shown in Figure 15.

Table 13. Cost of Fabrication for Filter Using the Hughes VHSIC CMOS/SOS Power Process

Design		\$
Design and Layout	300K	
Simulation	35K	
Generate CALMA tape, PG tape, mask set	<u>37K</u>	372K
Test (test hardware, test program, generate test vectors)		65K
Reiteration (redesign and new masks)		40K
Processing (2 lots, parametric measure and probe)		70K
Assembly and Test		25K
Project Engineer		50K
Program Management and Administration		<u>40K</u>
	Total	662K
	G&A	73K
	COM	10K

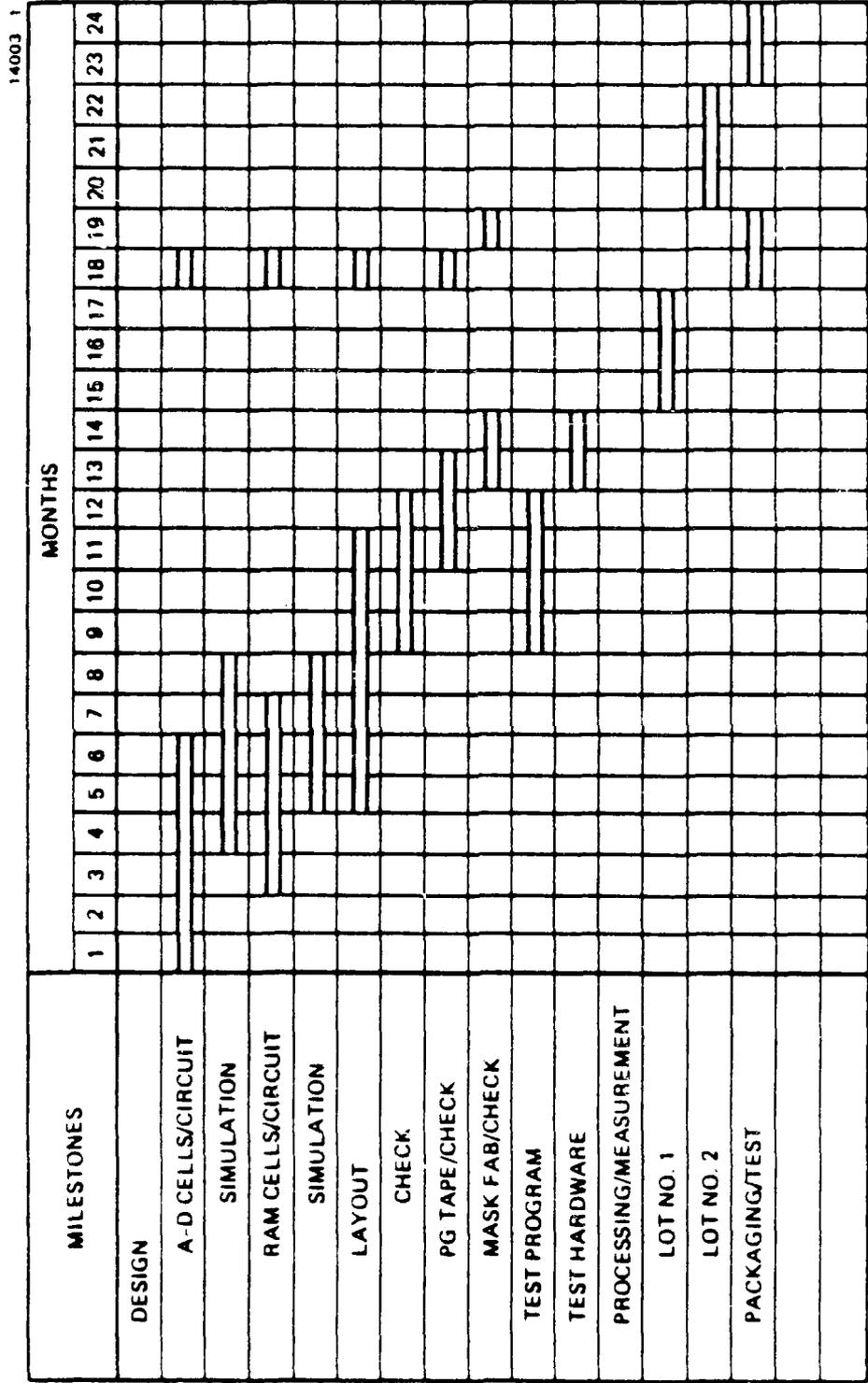


Figure 15. Tentative schedule for processing low power filter chip.

REFERENCES

1. J.G. Nash and G.R. Nudd, "Ultra-Low Power Digital Filter", NOSC Contract No. N66001-82-C-0504, Oct. 1982.
2. E. Sun, "Performance Comparison of MOS Transistors and Circuits Fabricated Using Bulk and SOS Technology", Hewlett-Packard Technical Report, May 1980.
3. R.C. Henderson, et al. "Develop Submicron Devices", HRL Technical Report, NOSC Contract No. N00123-79-C-0271, Oct. 1980.
4. D. Leong, "VHSIC Revision 7 Layout Rules", Hughes IDC, Feb. 1984.
5. Hughes VHSIC Cell Library, 1984.
6. Digest of Technical Papers, ISSCC Conference, Feb. 1984.
7. D.C. Mayer et al., "A short-channel CMOS/SOS Technology in Recrystallized 0.3 μm -Thick Silicon-on-Sapphire Films," HRL Technical Report, to be published in Electron Device Letters.
8. S. Taguchi et al., "Feasibility study of SOS VLSI: Capacitance Analysis in Downward Scaling and Improvement of Thin Films by a Solid-Phase Epitaxy," 1981 Symposium on VLSI Technology, Sept. 1981, pp. 92-93.

TOSHIBA 256 K CMOS STATIC RAM

ISSCC 84 / THURSDAY, FEBRUARY 23, 1984 / CONTINENTAL BALLROOMS 5-9 / THPM 15.1

[see page 340 for Figure 1.]

SESSION XV: STATIC RAMs

Chairman: Richard Pashley
Intel Corp.
Santa Clara, CA

THPM 15.1: A 46ns 256K CMOS RAM

Mitsuo Isobe, Junichi Matsunaga, Takayasu Sakurai, Takayuki Ohtani,
Kazuhiko Sawada, Hiroshi Nozawa, Tetsuya Izuka, Susumu Kohyama
Toshiba Semiconductor Device Engineering Laboratory
Kawasaki, Japan

DEMANDS FOR HIGH DENSITY, high speed and low power dissipation are increasing with recent static RAMs. Actually, as storage capacity on a chip increases, reduction in delay time and power consumption within the core area, i.e., word line, bit line and sensing circuit delays, becomes vitally important to meet the requirements. This paper will describe an asynchronous 256Kb CMOS static RAM, which utilizes a double word line technique^{1,2}, automatic power down function, and internally clocked circuitry.

The chip microphotograph and the typical characteristics are shown in Figure 1 and Table 1. The row decoder is placed on the left side of the memory arrays. The cell size is 11 x 11.2µm, and the chip measures 6.68 x 9.86mm, which fits into a standard 28pin DIP. The RAM offers typically 46ns access time, 100µW operating power, and 30µW standby power.

For achieving high packing density and improved performance with this sub-micron channel length VLSI memory, a plasma F₂ based double polysilicon CMOS technology has been developed³, which includes a narrow field isolation and 0.2µm type transistors for 5V operation. A 1.2µm ground rule was employed for highest density regions. Table 2 lists the design rules and device parameters, in comparison with the preceding generation device⁴. Double level Al structure is characterized by a unique low temperature interlayer formation and plasma etch technique, combined with a hillock suppressed metal layer deposition.

The block diagram of the RAM is shown in Figure 2. Address translation detector circuits are employed to generate column select signals, chip activating pulses and an automatic power down pulse. The chip activating pulses are used to equalize bit lines and data lines, and also to activate word line drivers and sense amplifiers. The automatic power down pulse is generated after the read-out operation. The pulse deselects column drivers and sense amplifiers, so that no dc current flows in the circuit in standby current. Therefore, the active power dissipation is reduced at low operating frequencies, as shown in Figure 3.

To reduce both the word line delay and the active power dissipation, a double word line structure is introduced. Row lines consist of aluminum main word lines which select one of 512 rows, and polysilicon section word lines which activate one of 16 sections. The section word line is activated by the main word line and a column select signal. Since only 32 memory cells connected to one section word line are accessed in a cycle, column current flows only in a selected section. In addition, an RC time delay of each section word line is reduced to 1/256 compared with conventional arrangements. Therefore, the total word line delay is reduced to 8.5ns from 30ns as is the case of conventional 4 block word line configurations. The circuit design is realized by utilizing double aluminum structure.

A schematic diagram of a memory cell and peripheral circuits is illustrated in Figure 4. A two-stage current mirror type CMOS sense amplifier is used to achieve high speed read operation. The first stage amplifies a small signal from one of the four bit line pairs. The second stage amplifies the first stage output signal to a large swing level.

The bit lines and the first sense amplifier output are equalized by the chip activating pulse before the read operation. To improve fabrication yield, a redundancy circuit is employed without any speed degradation.

The oscillograph of the address input and data output signal waveforms at V_{DD} = 5V with 100pF load capacitance is shown in Figure 5, which indicates a 46ns address access time.

Acknowledgments

The authors wish to thank Y. Nishi and Y. Uchida for their continuous support and encouragement, and H. Naneko and K. Sato for their great efforts in processing and evaluations.

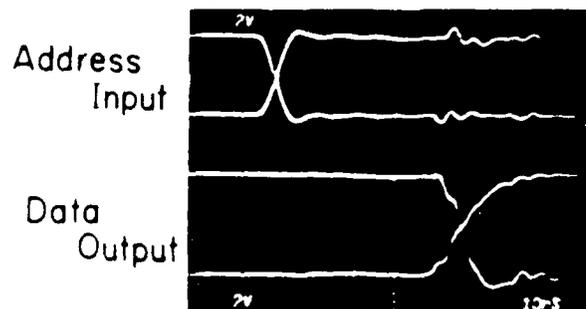


FIGURE 5—Oscilloscope of address input and data output waveforms.

[Right]

FIGURE 2—Block diagram of the 256Kb CMOS RAM

¹Sakurai, T. et al., "Methods for Analyzing a Word Line Delay and Their Applications," The Symposium of Solid State Device of IEE of Japan, SSD82-72, p. 15-21, Oct. 1982.

²Matsumoto, M. et al., "A 64Kb 5V CMOS RAM with Double Word Line Structure," ISSCC DIGEST OF TECHNICAL PAPERS, p. 58-59, Feb. 1983.

³Matsunaga, J. et al., "1.2µm Process Design for CMOS Devices," Electrochem. Soc. Meeting, San Francisco, Extended Abstracts 93-1, p. 566, May 1983.

⁴Kohyama, S. et al., "A 64Kb CMOS RAM," ISSCC DIGEST OF TECHNICAL PAPERS, p. 258-259, Feb. 1982.

PARAMETERS	64K-CMOS RAM	256K-CMOS RAM
PROCESS	DOUBLE-LEVEL POLY-Si SINGLE-LEVEL Al	DOUBLE-LEVEL POLY-Si DOUBLE-LEVEL Al
GATE LENGTH (NMOS)	2.0 μm	1.2 μm
PMOS)	1.2 μm	5 μm
GATE OXIDE THICKNESS	450 Å	250 Å
JUNCTION DEPTH (N ⁺)	0.25 μm	0.20 μm
(P ⁺)	0.5 μm	0.35 μm
POLY-Si WIDTH/SPACING)	2 μm/2 μm	1.2 μm/1.2 μm
1st Al (WIDTH/SPACING)	2 μm/2 μm	1.2 μm/1.6 μm
1st CONTACT HOLE	2 μm x 2 μm	1.2 μm x 1.2 μm
2nd Al (WIDTH/SPACING)	—	2.0 μm/2.0 μm
2nd CONTACT HOLE	—	2.0 μm x 2.0 μm

TABLE 2

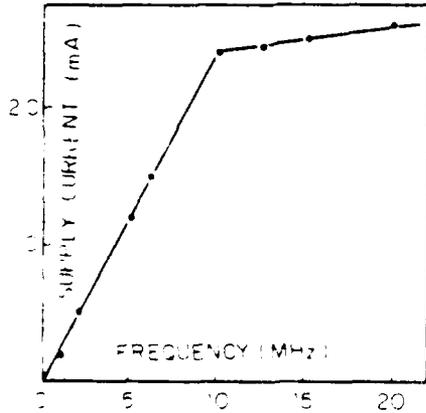
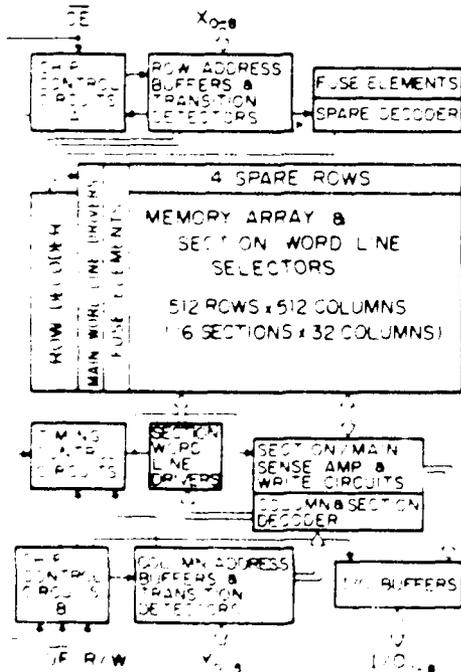


FIGURE 3



OPERATION	FULLY ASYNCHRONOUS (ADDRESS ACTIVATED CLOCKED OPERATION) AUTO POWER-DOWN FUNCTION
ORGANIZATION	32K WORDS x 8BIT
REDUNDANCY	4 SPARE ROWS
CHIP SIZE	6.68 x 8.86 mm
CELL SIZE	11 x 13.5 μm
I/O INTERFACE	TTL COMPATIBLE
ADDRESS ACCESS TIME	46 ns
ACTIVE POWER	10 mW (1MHz)
STANDBY POWER	30 μW
PACKAGE	STANDARD 28-PIN DIP

TABLE 1—Typical characteristics of the 256Kb CMOS RAM.

[Left: Top]

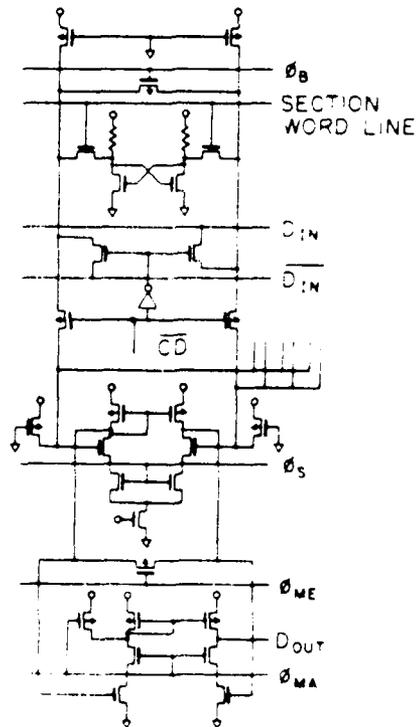
TABLE 2—Design rules and device parameters of the 256Kb CMOS RAM.

[Left]

FIGURE 3—Supply current versus operating frequencies.

[Below]

FIGURE 4—Schematic of memory cell and peripheral circuits.



HITACHI 64 K CMOS STATIC RAM

ISSCC 84 / THURSDAY, FEBRUARY 23, 1984 / CONTINENTAL BALLROOMS 5-9 / 3:45 P.M.

[See page 343 for Figure 3.]

SESSION XV: STATIC RAMs

THPM 15.5: A 20ns 64K CMOS SRAM

Osamu Minato, Toshiak Masuhara, Toshio Sasaki, Yoshio Sakai, Tetsuya Hayashida
Hitachi, Ltd.
Tokyo, Japan

IN RECENT YEARS, several circuit techniques have been combined with scaling to realize MOS static RAMs, having a speed comparable to bipolar RAMs^{1,2}. It has also been possible to realize bipolar RAMs with densities comparable to MOS devices³. This paper will report on a 64K x 1b CMOS static RAM with a 20ns typical address access time and 70mW active power dissipation.

The RAM performance has been achieved by the development of a pulsed word-line (PWL) technique and double P-well bipolar CMOS (BCMOS) circuitry. Also, gate length of MOS transistors has been scaled down to 1.3 μ m.

The pulsed word-line technique is illustrated in Figure 1. The schematic of the part of the RAM controlled by clocks is shown in Figure 2. The RAM circuitry is activated by the internal clock formed by detecting all address, \overline{WF} and \overline{CS} transistors. The basic clock for the RAM circuitry operation is XD which controls word lines through the X-decoder and sense amplifiers. For precharge and equilibration of data lines, precharge clock PC, generated by XD and equivalent to XD, is used. When the basic clock XD becomes high, a selected word line becomes high because the X-decoder is activated. At the same time, the sense amplifiers are activated by XD. At this moment, since the PC clock is high, the precharge transistor works as the usual data line load. Consequently, small differential data of the selected cell appear on the data lines. These data of the cell are transferred to the sense amplifier which consists of two stages of two single-ended active load differential amplifiers, and then to the output buffer. This operation is completely static, and all operations begin after the data lines are equilibrated. Then precharge clock PC goes low, cutting all data line loads. From this transition no dc current is consumed by the cell. After the signal is transferred to the output buffer, the data is latched to the output buffer by the DL clock. Further, this clock pulls down all word lines to the low state. At this moment, data lines are immediately precharged by the PC clock to prepare for subsequent data read. This circuit technique

differs from existing data line equilibration techniques⁴ and latched column techniques⁵ in that the word line is kept high when a specific address is read. The former suffers from large current through memory cells, although fast access is achieved. The latter achieves low power by latching the signal at the column by pulling one of the bit lines to a completely low state resulting in slow access time and large bit line recovery time. Pulsed word-line (PWL) techniques make it possible to reduce current through the transmission gate of the cell by the PC clock and to obtain fast access time by static operation during which both XD and PC are high. Also the PWL technique reduces the data line recovery time.

The XD, PC clock generators and output buffer use a bipolar CMOS (BCMOS) configuration to assure fast risetime. The bipolar transistor is formed in a thin P-well to realize high V_T . Thus, this technology utilizes double P-wells, one for NMOS transistors and the other for high performance bipolar transistors. Risetime capability of the bipolar device is 0.013ns/V · pF. This is three times greater than that of bipolar devices formed in the usual P-well.

Third generation CMOS (Hi-CMOSIII) technology has been developed. Used are N- and P-channel MOS transistors having 1.3 μ m typical gate length and 1.3 μ m design rule. Basically, this technology is a 70 percent reduction in size, both horizontally and vertically, of the original Hi-CMOSII technology which utilizes a 2 μ m design rule.

The memory cell is a cross-coupled four-NMOS flipflop with high resistance loads. The cell is 128 μ m² (3 μ m x 16 μ m).

A photomicrograph of the chip is shown in Figure 3. The die measures 3.16mm x 6.0mm. To achieve fast access time, the RAM is organized so that the array is split into four planes of 64 columns x 256 rows. Corresponding to these arrays, four sense amplifiers are laid out and power switched according to address A14 and A15.

A typical 20ns address access time was achieved with a 70mW nominal power dissipation at 1MHz cycle time. Chip select access time is 22ns. The RAM output waveforms for a typical 30pF load capacitance are shown in Figure 4. Supply current versus operating frequencies are shown in Figure 5. Active power dissipation at low operating frequencies is reduced by the aid of the pulsed word line (PWL) techniques. Typical features of the RAM are summarized in Table 1. This RAM has realized a speed comparable to bipolar 16K and III 64K RAMs, even though it consumes much less power.

Acknowledgments

The authors wish to thank M. Kubo, S. Asai, S. Yoneyama, Y. Kosa and T. Yasui for their guidance, and N. Hashimoto, Y. Nagai, S. Yamamoto for the device fabrication.

¹Conner, K.J., et al., "A 5ns 4K x 1 NMOS Static RAM", ISSCC DIGEST OF TECHNICAL PAPERS, p. 194-195, Feb. 1983.

²Minato, O., "Hi-CMOSII 4K Static RAM", ISSCC DIGEST OF TECHNICAL PAPERS, p. 14-15, Feb. 1981.

³Aiedmann, J.K., et al., "A 25ns 8K x 8b Static MTL 321 RAM", ISSCC DIGEST OF TECHNICAL PAPERS, p. 110-111, Feb. 1983.

⁴Yamamoto, K., et al., "A 64K x 1b NMOS Static RAM", ISSCC DIGEST OF TECHNICAL PAPERS, p. 66-67, Feb. 1983.

⁵Kishino, S., et al., "A 64K CMOS RAM", ISSCC DIGEST OF TECHNICAL PAPERS, p. 258-259, Feb. 1982.

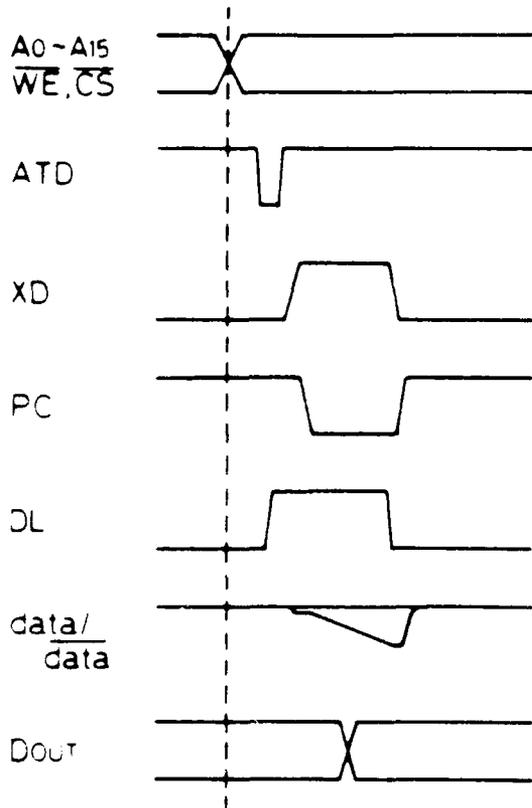


FIGURE 1—Timing diagram.

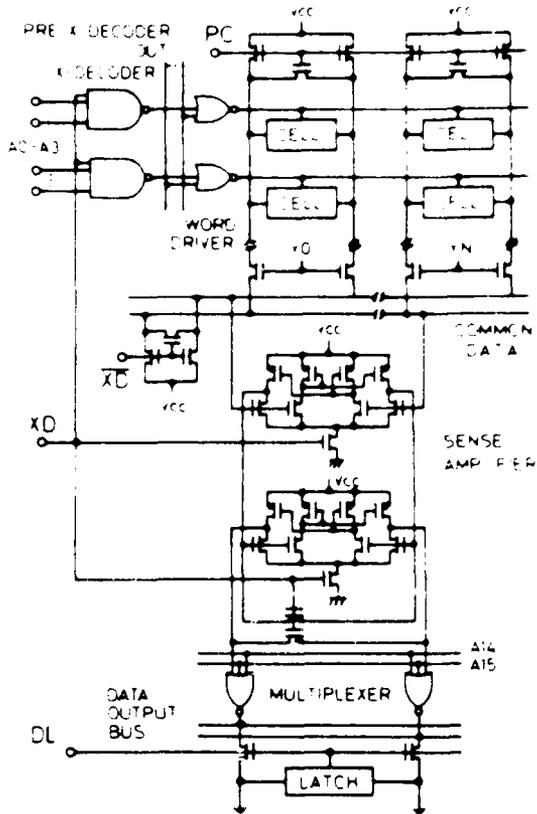


FIGURE 2—Circuit schematic of the RAM.

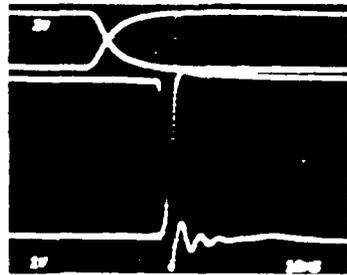


FIGURE 4—RAM output waveforms.

ADDRESS INPUT

DATA OUTPUT

Organization	64K word by 1 bit
Process	3µm Hi-CMOS III
Gate length	3µm for P- N-channel MOS transistor
Supply voltage	5 volts
Interface	72 compatible
Address access time	20 ns
Chip select access time	22 ns
Active power dissipation	70 mW (f _{clk} = 1 MHz)
Standby power dissipation	21 mW
Die size	315 µm x 60 µm
Pin size	80 µm x 16.0 µm

TABLE 1—Typical 64K SRAM features

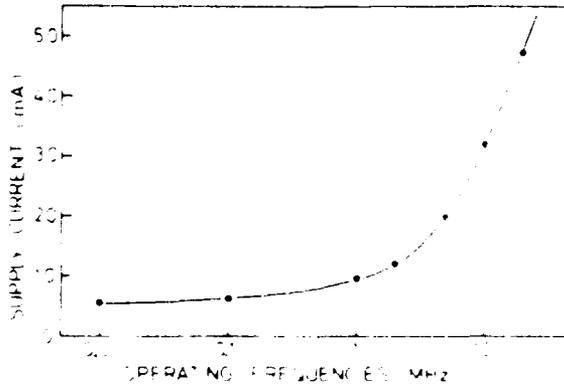


FIGURE 5—Supply current versus operating frequencies

**END
DATE
FILMED**

8-12-87