UNIVERSITY OF ADELAIDE LIBRARY
18.7.96

Design of a Very High Speed Dynamic RAM in Gallium Arsenide for an ATM Switch.

Michael K. McGeever, B.E. (Hons.)

A Thesis submitted for the degree of

Master of Engineering Science

The Department of Electrical and Electronic Engineering,

The University of Adelaide,

South Australia

November 1995.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent for this copy of my thesis, when deposited in the University Library, to be made available for loan and photocopying.

DATE: 10/1/96

Table Of Contents

List of Figures

List of Tables

Published Papers

1. M. K. McGeever, K. Eshraghian and M. J. Liebelt, "A 14kb Three Transistor GaAs DRAM for an ATM Switch," *Proc. 13th Australian Microelectronics Conference*, pp. 169 - 174, July 1995.

2. J. F. Lopez, K. Eshraghian, M. K. McGeever, A. Nunez and R. Sarmiento, "Gallium Arsenide MESFET Memory Architectures," Proc. *IEEE International Workshop on Memory Technology*, Design and Testing, pp. 103 - 108, August 1995.

List of Abbreviations

As ..... **Ars**enic

ATM ..... **A**synchronous **T**ransfer **M**ode

BiCMOS ..... **Bi**polar **C**omplimentary **M**etal **O**xide **S**emiconductor

BISDN ..... **B**roadband **I**ntegrated **S**ervices **D**igital **N**etwork

CCITT ..... International Consultative Committee on Telephone and Telegraph

CMOS ..... Complimentary Metal Oxide Semiconductor

DCFL ..... **D**irect **C**oupled **F**ET **L**ogic

DSBFL ..... **D**ouble **S**uper **B**uffer **F**ET **L**ogic

ECL ..... **E**mitter **C**oupled **L**ogic

FET ..... **F**ield **E**ffect **T**ransistor

Ga ..... **Ga**llium

GaAs ..... **Ga**llium **Ars**enide

HEMT ..... **H**igh **E**lectron **M**obility **T**ransistor

IC ..... Integrated Circuit

ISDN ..... Integrated Services Digital Network

JFET .....Junction Field Effect Transistor

MESFET ..... MEtal Semiconductor Field Effect Transistor

MIM .....Metal Insulator Metal

MMIC. .....Monolithic Microwave Integrated Circuit

MOSFET ..... Metal Oxide Semiconductor Field Effect Transistor

PCML ..... Psuedo Current Mode Logic

SAGA..... Self Aligned Gate Array

SBFL ..... Super Buffer FET Logic

SDCFL.....Source follower Direct Coupled FET Logic

SFFL .....Source Follower FET Logic

Si .....Silicon

SPICE ..... Simulation Program for Integrated Circuit Estimation

VLSI..... Very Large Scale Integration

Acknowledgements

I would firstly like thank both of my supervisors, Professor Kamran Eshraghian (now at Edith Cowan University, Perth) and Mr. Michael Liebelt from the University of Adelaide for their support and guidance throughout the period of this research.

I would also like to thank Dr. Jens Jakobsen (from Jydsk Telefon, Denmark) and Mr. Eric Chu (University of Adelaide), who were also working on this project, for their advice and assistance.

Mr. Andrew Beaumont-Smith deserves special recognition for his many helpful discussions regarding GaAs design and the use of the software tools utilised in this work.

I would also like to acknowledge Mr. Tarmo Rhotla for is efforts in keeping the computers and software running dependably.

I would like to thank the staff and my many friends and colleagues at the university, especially Andrew Blanksby, Ali Moini, Shannon Morton, Sam Appleton, Joan Liu, Thong Nguyen, Said Al-Sarawi and Richard Beare.

Mike Patton and Pete Muller deserve thanks for their help and friendship.

Finally I would like to acknowledge the support of my parents and sister, and especially my closest friend, Libby Mapletoft.

Abstract

Broadband ISDN will be the foundation of the next generation of telecommunications networks, offering a single high bandwidth network capable of efficiently transmitting all forms of data, including voice, image, video and computer data.

Asynchronous Transfer Mode (ATM) is the protocol used to transmit data in Broadband ISDN networks. Traditionally data rates have been limited by transmission technology and not switching, however due to recent advantages in transmission technologies such as optical fibre, the opposite is now true. To support this new ATM network, switches capable of switching several gigabits of data per second must be designed.

Gallium Arsenide (GaAs) is an excellent material for the implementation of high speed ATM switches because of its high electron mobility and low parasitic capacitance, and superior speed/power performance over silicon.

One of the main requirements of an ATM switch is buffering. To maximise throughput, this buffering is ideally located on-chip. Due to the size of the ATM cell, large amounts of buffer space are required and the memory cell must therefore be compact and its power dissipation small while still

retaining acceptable performance. These properties are usually associated with Dynamic RAM, however due to it's leaky nature, most GaAs RAM designs are static in nature.

This thesis discusses the design of a Dynamic RAM in gallium arsenide for use as a buffer in an ATM switch. The causes of leakage are investigated and methods to overcome or compensate the leakage are devised, resulting in a memory cell with a large storage time, high speed and low power dissipation. Based upon the cell, a 14kbit (128 words $\times$ 112 bits) RAM array is designed and layed out in gallium arsenide. The RAM array is designed to operate over a -25 °C to +125 °C temperature range using process parameters which vary by up to 2 σ from typical.

UNIVERSITY OF ADELAIDE LIBRARY

Chapter 1: Introduction to Broadband ISDN and ATM

Traditional communications networks have evolved around voice, and the fact that the most scarce resource was bandwidth. With the advent of modern optical fibre, bandwidth is no longer of concern, with the bottleneck forming instead at the switch nodes. Combined with this is the desire to create a single all-purpose network, suitable not only for voice communications, but many other forms such as data, facsimile and digitised video and images. These services require high bandwidth digital transmission, hence the need for the Broadband Integrated Services Digital Network (BISDN).

1.1 Broadband ISDN

The Broadband Integrated Services Digital Network (BISDN) will be the foundation of the next generation of telecommunications networks, integrating voice, video and data onto a single, worldwide, public access network. The first transition from the current analogue telephone networks to BISDN has been implemented as a low bit rate (typically 64 kbit/s) service called Narrowband ISDN (NISDN). However, because of its limited band-

width it is unsuitable for video, images and other high bit rate services. Currently standard bit rates for Broadband ISDN are set at 150 Mbit/s, 622 Mbit/s and 2.4 Gbit/s, with higher bit rates becoming standard as switching technology improves.

NISDN is implemented using circuit switching, as is the current analog network. In circuit switching, time is segmented into fixed size recurring frames, with each frame being divided into a number of slots. A channel is identified by the position of the time slot it uses within the frame. A particular call is allocated a channel (hence a slot within the frame) and keeps that channel for the duration of the call [5]. Because of this fixed structure, the call has a fixed bandwidth available to it, and is hence most suitable for fixed bit-rate services. If a variable bit-rate service were to be transmitted using circuit switching, at times when its bit-rate is low it may not require all of the allocated slots resulting in an inefficient use of bandwidth. When its bit-rate is high, it may require more bandwidth than has been allocated resulting in long delays in the transmission of this extra data. For these reasons circuit switching is unsuitable for variable bit-rate services, and hence a different protocol is required for BISDN. The protocol, chosen by the CCITT Study Group XVIII Task Group on ISDN Broadband Aspects [1], is called Asynchronous Transfer Mode [2].

1.2 Asynchronous Transfer Mode

In Asynchronous Transfer Mode, time is again segmented into fixed size slots, however the slots are allocated to a particular service on demand, with any service having access to any particular slot, unlike the fixed allocation structure used in circuit switching. In this way variable bit-rate services are more efficiently transmitted. Because of the lack of fixed structure of ATM, a particular service cannot be identified by the position of its slot. Each packet of information being transmitted must therefore have its own identification. This information (along with other information which will be discussed later) is contained in a header which is appended to the data at the start of the cell. Therefore, as well as transmitting the data, a

header must also be transmitted resulting in a less efficient use of bandwidth than transmitting the data alone. In terms of efficiency it is therefore desirable to have the header comprising the smallest proportion of the cell length as possible which can be achieved by making the cells very big. However this introduces packetisation delay. The cell cannot be transmitted until each packet of data is full and the longer the packet is, the more time a cell will take to be filled. Packetisation delay is therefore minimised by minimising the size of a cell. Obviously a trade-off has to be made between bandwidth efficiency and delay. The ATM cell has been standardised at a length of 53 bytes, with 5 bytes of header information and a 48 byte data payload [3]. The general structure of an ATM cell is shown in Figure 1.1.

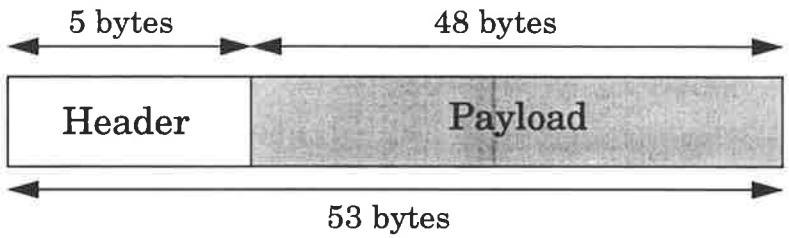


Figure 1.1. General structure of an ATM cell

1.2.1 ATM Cell Format

There are two different header formats defined in ATM, that at the User Network Interface (UNI) and at the Network Node Interface (NNI). The cell format at the Network Node Interface is shown in Figure 1.2 [6]. The header includes the following fields:

- VPI: Virtual Path Identifier (12 bits) - This field defines the Virtual Path which contains this cell's Virtual Channel. A Virtual Path is a collection of Virtual Channels.
- VCI: Virtual Channel Identifier (16 bits) - This field defines the Virtual Channel within the Virtual Path along which the cell will be routed.

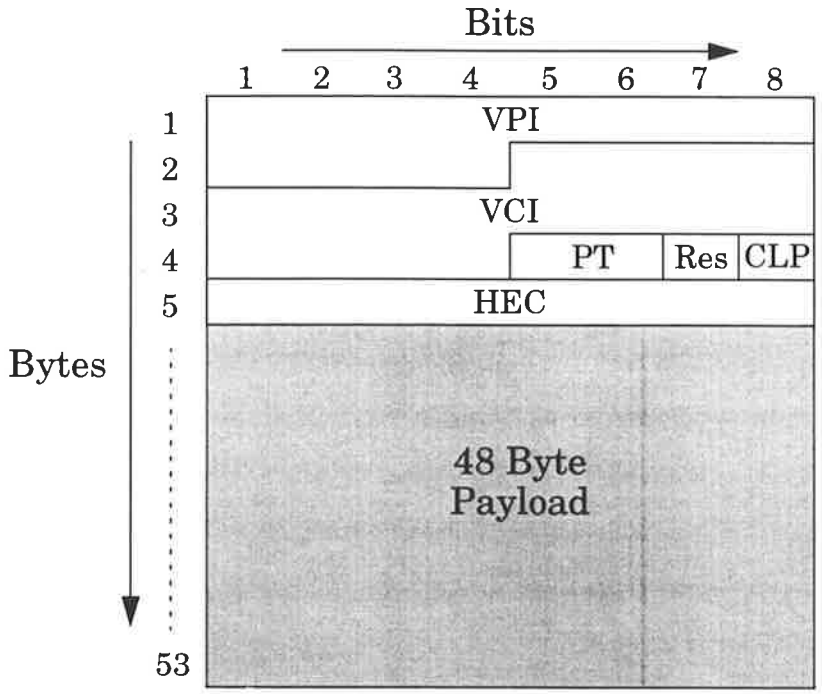


Figure 1.2. Format of the ATM cell at the Network Node Interface

- PT: Payload Type (2 bits) - This field defines the type of information contained in the payload. e.g. user, network control etc.
- Res - Reserved for future use (1 bit).
- CLP: Cell Loss Priority (1 bit) - A high priority cell will have preferential treatment over normal priority cells throughout the switch fabric to increase its probability of being successfully reaching its destination in the minimum time.
- HEC: Header Error Control (8 bits)- This field contains a Cyclic Redundancy Code (CRC) to detect errors within the header.

1.3 ATM Switching Requirements

The role of an N × N ATM switch is to accept cells at any of N inputs, and route them to any of N outputs. In the case that more than one of the input

cells require simultaneous access to an output, some form of buffering must be supplied to allow cells to be routed sequentially. Thus an ATM switch has two functions:

1. Routing

2. Buffering

The quantity and architecture of the buffering will determine the number of cells that will be lost in the switch. This is referred to as the cell loss rate. The cell loss rate should ideally be no more than that of optical fibre (the likely transmission medium), approximately 10^{-12} [7].

1.4 ATM Switches

A typical ATM switch block consists of an ATM switch, Input and Output Line units and some form of control as shown in Figure 1.3.

The Line Units are the interface between the switch architecture and the transmission medium. They must perform functions like header error correction, synchronisation and translate the data from the raw ATM form into that required by the switch, if necessary. This is further discussed in Section 1.5.1.

The Switch Control is used to monitor traffic within the switch and ensure that blocking throughout the switch is minimised. To do this, cells are routed through areas of minimum congestion.

1.4.1 Switch Architectures

There are three basic types of ATM switching architectures, named after the configuration of their buffering: Input buffered, Shared buffered and Output buffered [4]. The basic architecture of each is shown in Figure 1.4. Each of the particular architectures has advantages and disadvantages.

In summary:

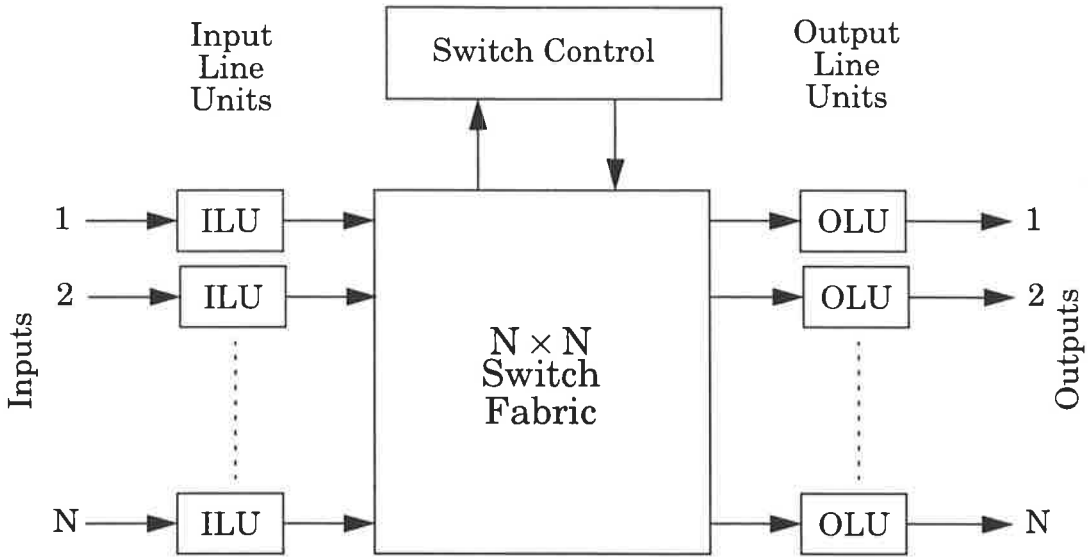
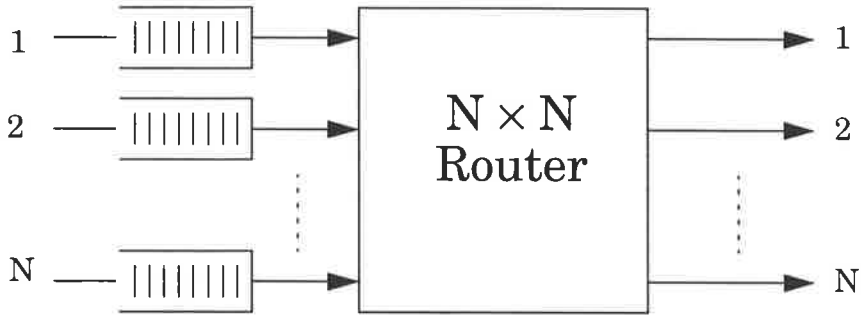


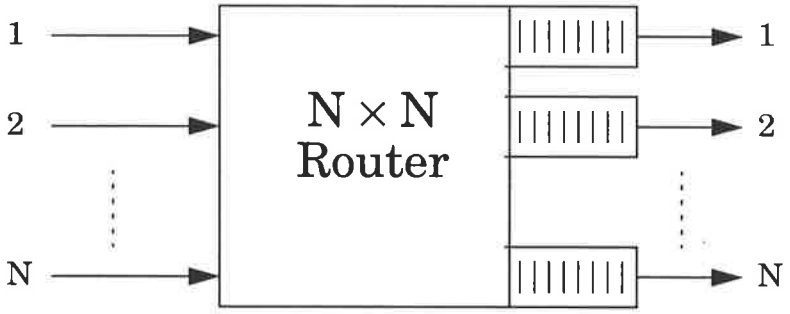
Figure 1.3. A typical ATM Switch Block

1. Input Buffered switches have FIFO buffers on each input. Incoming cells are placed into the buffers. If contention occurs, the cell at the front of the FIFO is held back. This leads to what is called Head Of Line (HOL) blocking. This architecture has the poorest performance of all three in terms of cell loss [4]. However, the buffer needs only to read from one input and write to one output and hence the required memory speed is much lower than other types at only 2B, where B is the input data rate.

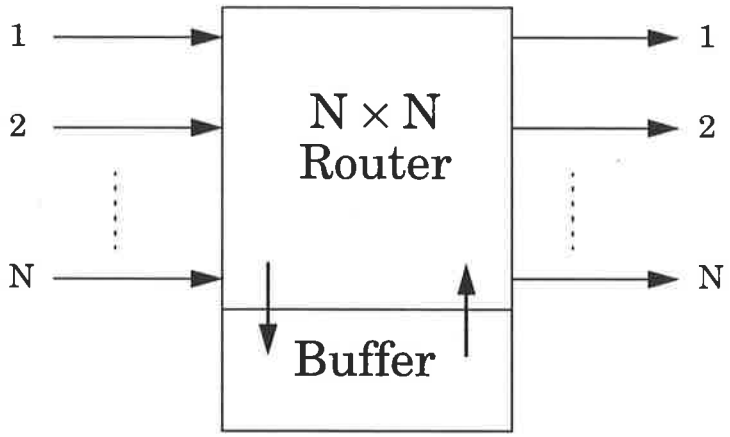
2. Output Buffered switches have FIFO buffers at each output of the switch. The problem of HOL blocking disappears, hence these switches have much better performance than input buffered switches, however this comes at the expense of increased memory speed. The buffer must read from each input and write to one output and hence the required memory speed is (N+1)B where N is the switch size and B is the data rate.



Input Buffered Switch



Output Buffered Switch



Shared Buffer Switch

Figure 1.4. ATM Switch Architectures

3. Shared Buffer switches offer a more efficient way of using memory than Output buffered switches, hence for the same cell loss rate, a smaller buffer size can be used, saving considerable chip area. However this comes at the expense of increased switch complexity and higher memory speed. The Buffer needs to read from N inputs and write to N outputs hence the required buffer speed is 2NB, where N is the switch size and B is the data rate.

Many ATM switches have been designed around these architectures [4][8][9].

1.5 2×2 ATM Switch

Due to the experimental nature of this research the chosen architecture was designed to be small and scalable. As such the basic switch is 2×2 in size, and has been segmented into three distinct chips, composing a multiplexer chip, a router chip and a buffer chip. Figure 1.5 shows the 2×2 switch. The two incoming ATM cell streams are multiplexed onto a single line of twice the bit-rate and passed into a router chip for routing. The router then interprets the header information and sends the cell to either of its outputs. The outputs from the router are then passed into a buffer chip which halves the bit-rate and thus the outputs from the buffer chip are at the same bit-rate as the inputs to the switch. This technique is called *Bit Rate Conversion* [10]. The switch architecture is of the output buffered type. The switch has been designed to operate with two input streams with bit-rates of 2.4 Gb/s, by parallelisation of the input data stream onto 16 lines. The architecture was proposed by Jakobsen [11]. The full specifications of the chip set can be found in [12].

1.5.1 Cell Format Within the Switch

To enable routing of the cell without modification to the original ATM cell, an additional header is appended to the cell before it enters the switch fabric by an Input Line Unit (ILU). This 3-byte header is shown in Figure 1.6. The header contains the following information:

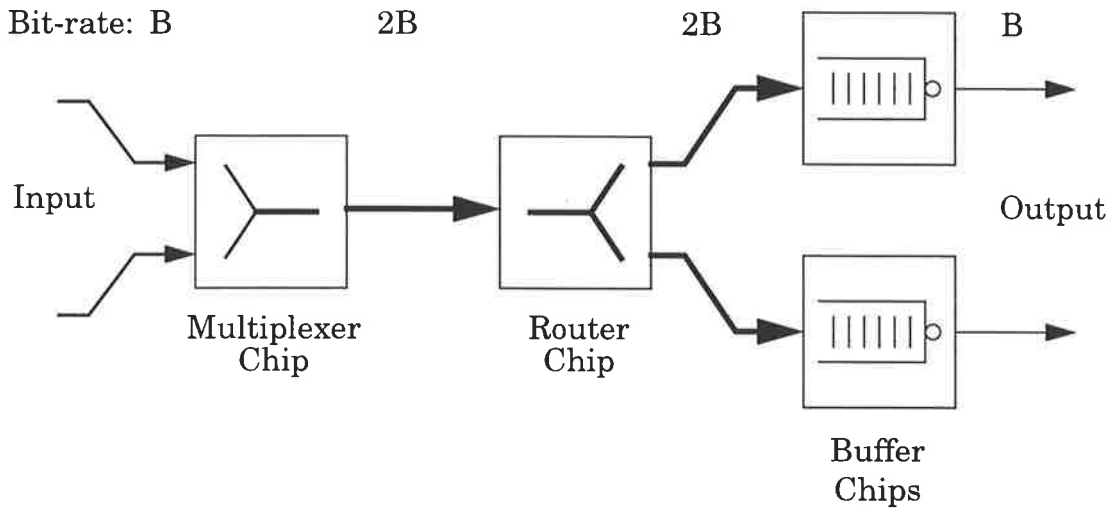


Figure 1.5. 2 × 2 Switch Architecture

- S - Start Bit: If set to 1 the cell is valid. If the entire header is set to 0 the cell is empty or idle and is ignored by the switch.
- P - Priority: This is the same as the Cell Loss Priority (CLP) as in Section 1.2.1.
- B - Broadcast: If set to 1 the cell is broadcast to both outputs of the router.
- A - Address Field: A 16 Bit Address field. The first bit is used by the current router and then shifted out to prepare for the next router. Thus a cell can pass through a maximum of 16 routers or switches before a new address field must be written.
- R - Reserved Field: 5 Bits reserved for future use.

Of course, this extra header is an additional overhead in the switch which reduces its efficiency and effective data throughput. An alternative way of routing is to use an on-chip look up table which determines from the VCI/VPI of the cell which output the cell should be routed to [15], but this adds significant extra complexity into the switch design. The bit-rate must also

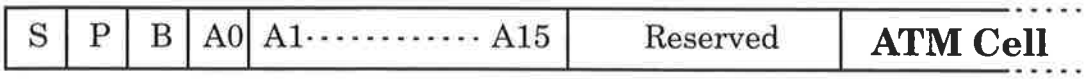


Figure 1.6. Extra Header inside the Switch Fabric

be increased proportionately to allow transmission of the extra information. For example the 2.4 Gb/s data-rate must be increased to 2.53 Gb/s.

The total size of each cell is therefore 48 bytes ATM payload + 5 bytes ATM header + 3 bytes extra Switch header = 56 bytes. Therefore the size of the 32 cell buffer is $32 \times 56 \times 8 = 14336$ bits = 14 kilobits.

1.5.2 Buffer Chip

The buffer chip receives one input stream and outputs one input stream at one half the input rate. Because the input stream is coming from the router which has two outputs, on average half of the input cells will be information bearing and the other half will be empty. However there will be fluctuations where more than half of the router traffic will be diverted to one buffer (equivalent to contention within the switch). These fluctuations must be absorbed by a buffer otherwise the excess cells will be lost. The larger the buffer, the fewer the number of cells that will be lost. The buffer size has been chosen to be 32 cells, according to [4], such that the switch will have a cell loss rate of 10^{-12} at an offered traffic load of 0.6. i.e. 60% of the ATM cells arriving at the switch contain information and have to be processed, while the remaining cells are empty and are discarded by the switch.

The buffer consists of Serial to Parallel converters, to convert the input data stream into 112 bit words (see Section 6.1), a 32 cell memory where the ATM cells are temporarily buffered, and Parallel to Serial converters to convert the output data stream back to the serial format of the input stream. The Input Controller, Buffer Manager and Output Controller con-

trol the operation of the Serial to Parallel Converters, Memory Array and Parallel to Serial converters respectively. A floor plan of the buffer chip is shown in Figure 1.7 [16].

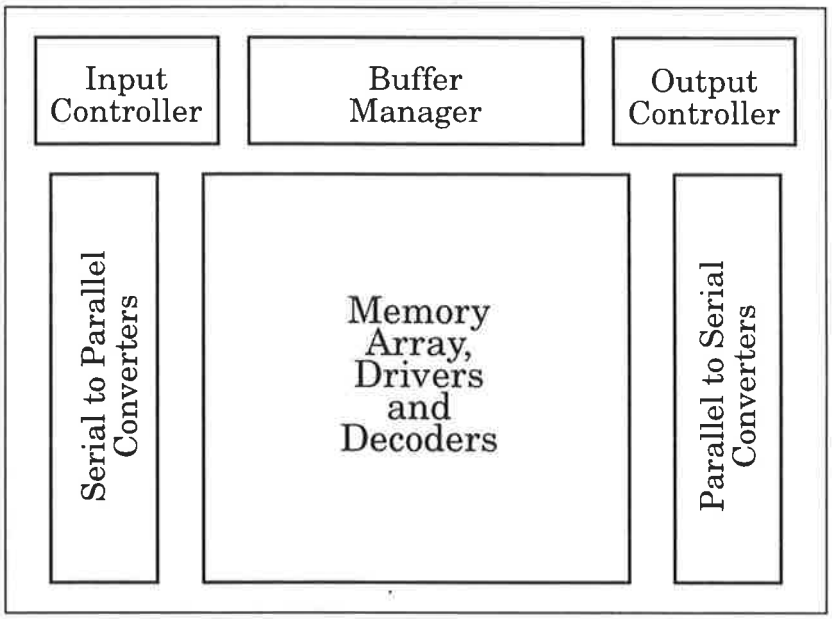


Figure 1.7. Floor plan of the Buffer Chip

The design and implementation of the 32 cell memory is the main concern of this thesis. The design and implementation of the control logic associated with the buffer chip was done by Eric Chu and can be found in [16].

1.5.2.1 Buffer Memory Performance Requirements

The buffer memory performance specifications are as follows:

1. Storage time:

   The data is being read from the buffer at 2.53 Gb/s. The maximum time a cell must be stored will occur when the cell is in the 32nd memory cell. Each cell is 56 bytes long (Section 1.5.1). Therefore the maximum require storage time is:

$$\frac{32 \times 56 \times 8}{2.53 \times 10^{9}} = 5.66\ \mu s$$

The buffer therefore has to have a storage time of at least 5.66 μs. Note that for slower bit rates, the storage time required will increase proportionately. For example at a data rate of 622 Mb/s (increased to 657 Mb/s), the storage time required is 21.8 μs.

2. Read and Write Cycle Times:

   The read and write cycle times will depend on the word size of the buffer (i.e. how much data is read/written to the buffer per operation). The array is dimensioned into 128 words each of 112 bits as described in Chapter 6. Therefore, as the data is being fed into the buffer at 5.06 Gb/s, a write operation must be performed every $112/5.06 \times 10^9 = 22.1$ ns. The data is read out of the buffer at 2.53 Gb/s, so a read operation is performed every 44.2 ns. Therefore the read and write cycle times must be less than 44.2 ns and 22.1 ns respectively.

3. The data is only read from the cell once, and so no advantage is obtained using a cell with a non-destructive read cycle.

4. Because there is no synchronisation between input and output circuits, the memory cell will ideally be dual ported, so that the array can be both read from and written to simultaneously [16]. This also maximises available read and write cycle times.

1.5.3 Multiplexer Chip

The multiplexer receives two input streams and multiplexes them onto a single output line at twice the bit-rate. In this way no cell loss occurs inside the multiplexer chip. The complete specifications for the multiplexer chip can be found in [12].

1.5.4 Router Chip

The router chip receives a single input stream from the multiplexer and interprets the first bit of the cell, then deciding to route the cell to either of the two outputs. Both output streams run at the same speed as the input

stream to ensure no cell loss occurs within the router chip. The full specifications of the router chip can be found in [12].

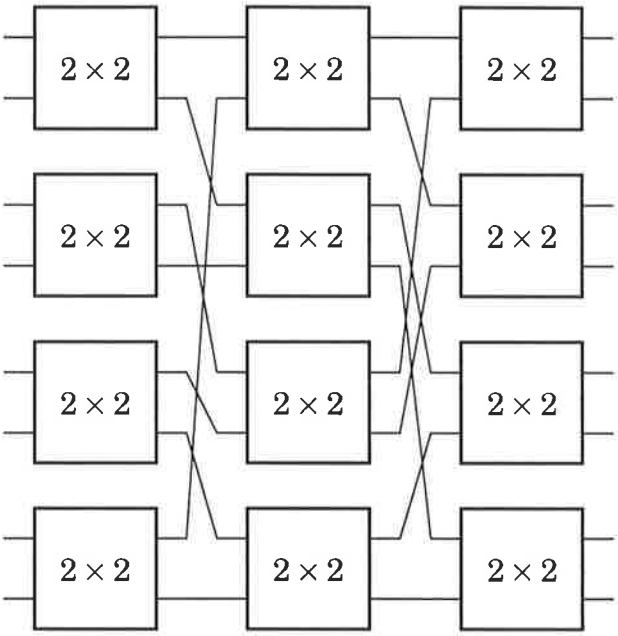
1.5.5 Larger Switch Networks

Once the two by two switch has been completed, larger switches can be constructed using multiple smaller switches. There are many different configurations for such switches, such as the Delta network [13] and the Benes network [14], both shown in Figure 1.8. The Delta network uses a smaller number of switches to implement the larger switch network, but only has one possible path from each input to each output. The Benes network, whilst using a larger number of smaller switches, has multiple paths from each input to output and therefore has superior performance as it is less easily blocked. In fact, the Benes network is the smallest network for which all permutations of interconnect are realisable [4].

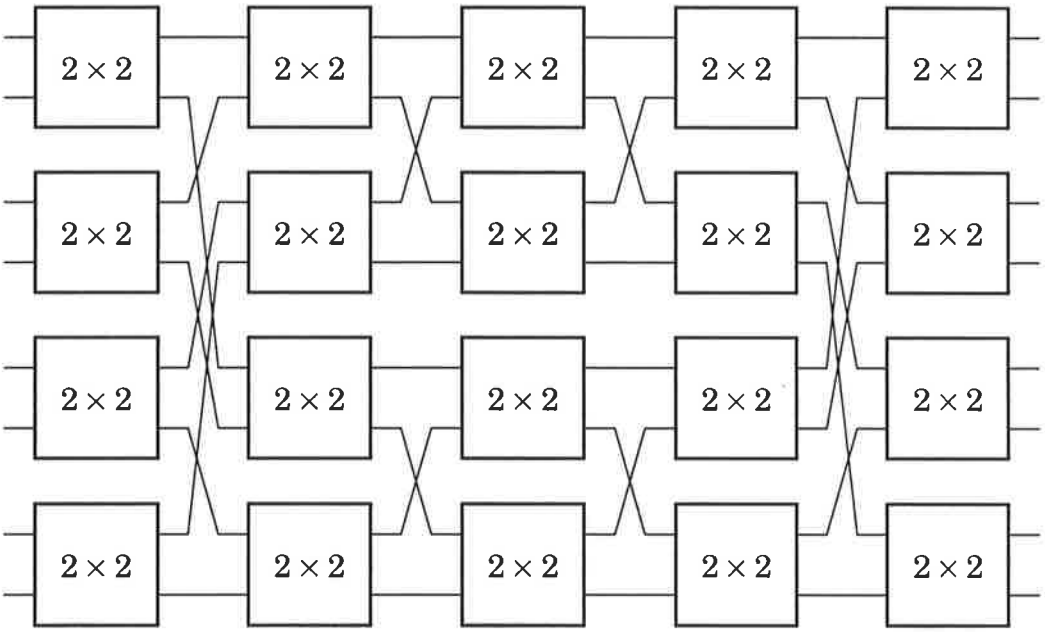
1.6 Gallium Arsenide Implementation

To maximise the bit-rate in the switch, it will be implemented in gallium arsenide (GaAs), a compound semiconductor with superior speed/power characteristics to silicon [19]. However GaAs is not without its disadvantages. In many instances it is harder to design than silicon, as currently it is a non-complimentary process (similar to nMOS), it is prone to leakage currents and other detrimental effects, and is also subject to a large variation in characteristics across the process. These properties make the design of a suitable memory buffer, with low power dissipation, high density, fast cycle times and long storage time a particularly challenging task.

The chip set is being implemented in the Vitesse H-GaAs II process, a 0.8 μm Self Aligned Gate Array (SAGA) process [26]. This process is specifically designed and optimised for high speed digital logic. However, as will be seen in the following chapters, the Dynamic RAM storage time may be improved in a process designed for analogue circuits with more emphasis on minimising leakage and process variation effects.



(b) 8 × 8 Delta network switch made using 2 × 2 switches



(b) 8 × 8 Benes network switch made using 2 × 2 switches

Figure 1.8. Constructing Larger Switches

The memory array designed in this thesis has been designed to operate over a temperature range of -25°C to +125°C, and over a process variation in parameters of 2σ slow to 2σ fast from typical parameters (For information on process spread please see Section 2.4.5). Because of this very wide operating range, many of the circuits have been over designed, particularly to function at 2σ slow. The result is an increased power consumption at typical operation, but also an increased yield in chips which will operate to specification.

Chapter 2: Gallium Arsenide VLSI

This chapter will provide an introduction to gallium arsenide, its properties and advantages/disadvantages over silicon, and the processes used to fabricate GaAs VLSI (Very Large Scale Integration) wafers. The operation and modelling of GaAs devices will then be discussed, including inaccuracies in the modelling process. A summary of GaAs logic families will then be presented, with a detailed discussion of those families used in the thesis.

2.1 Structure

Gallium arsenide (GaAs) is a compound semi-insulator material composed of the elements gallium (Chemical Symbol Ga, Group III of the periodic table) and arsenic (Symbol As, Group V). The atoms are arranged in a crystal lattice structure. To enable switching devices to be fabricated, it is necessary to introduce impurities into the GaAs substrate which enables the GaAs to become conductive. This process is referred to as doping. Once the material has been doped, it must be heated to a high temperature to allow the dopant atoms to become part of the crystal lattice structure, a process called annealing. The dopant usually used is silicon, and after annealing,

the silicon atoms take the place of the large Ga atoms, not the smaller As. The structure of gallium arsenide both before and after doping and annealing is shown in Figure 2.1 [17].

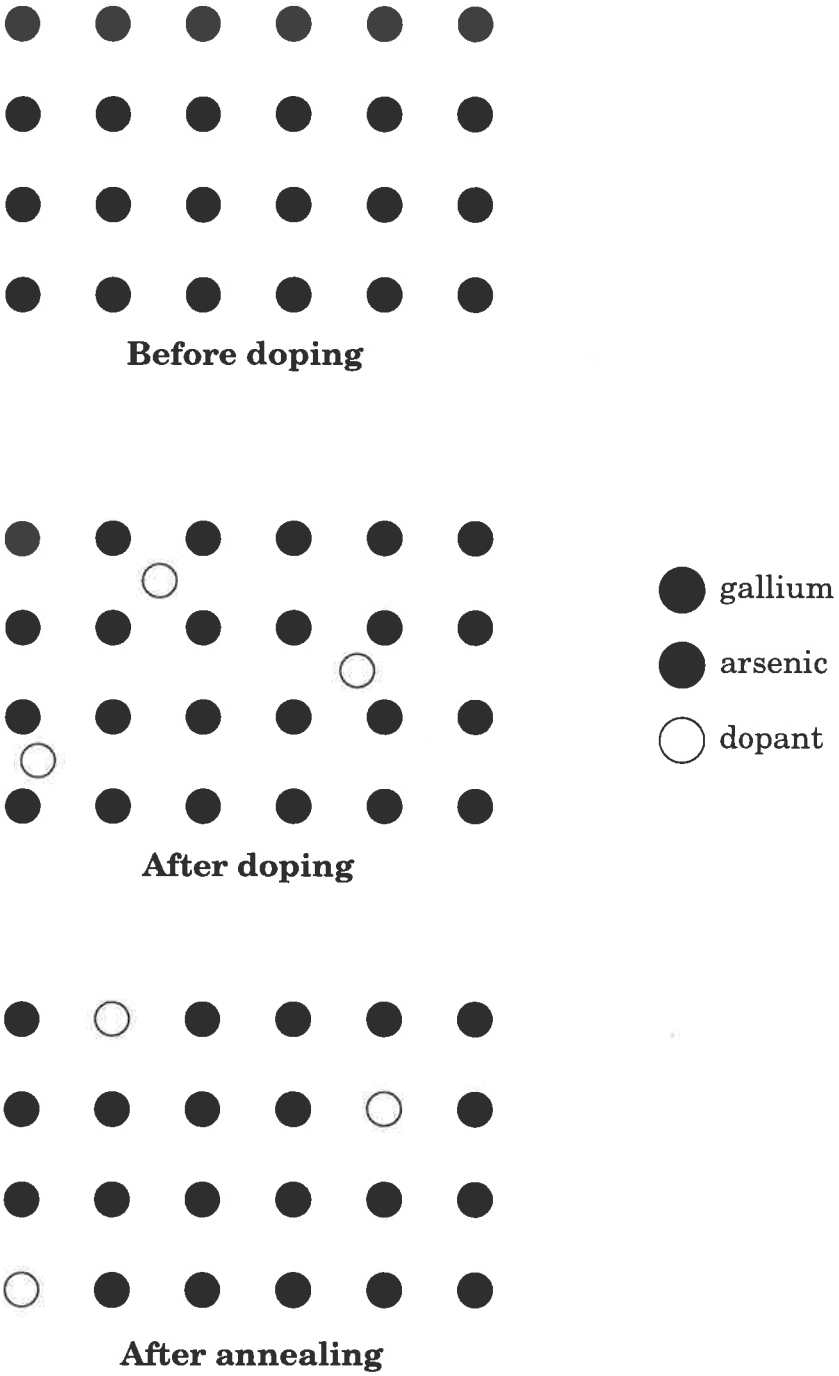


Figure 2.1. Crystal Lattice structure of GaAs.

2.1.1 Comparison with Silicon

Gallium Arsenide's properties make it suited to high speed micro-electronics and microwave applications. These properties include:

1. Gallium Arsenide's electron mobility, μ, and saturated drift velocity is higher than that of silicon [18], with the result that for a given Electric Field, E, the velocity, v, of electrons in GaAs will be higher, related by the expression:

$$v = \mu \cdot E \qquad \text{(EQ 2.1)}$$

Figure 2.2 shows the drift velocity of electrons in both GaAs and Silicon as a function of the Electric field [20], emphasising the superiority of GaAs. As the Electric Field is increased to very high levels (>> 10 kV/cm), the saturated drift velocities of GaAs and Silicon approach one another [19], and so the maximum speed advantage is only obtained at low values of Electric Field. Although dependant on a large number of factors, this speed increase is a factor of approximately 6 [18][20].

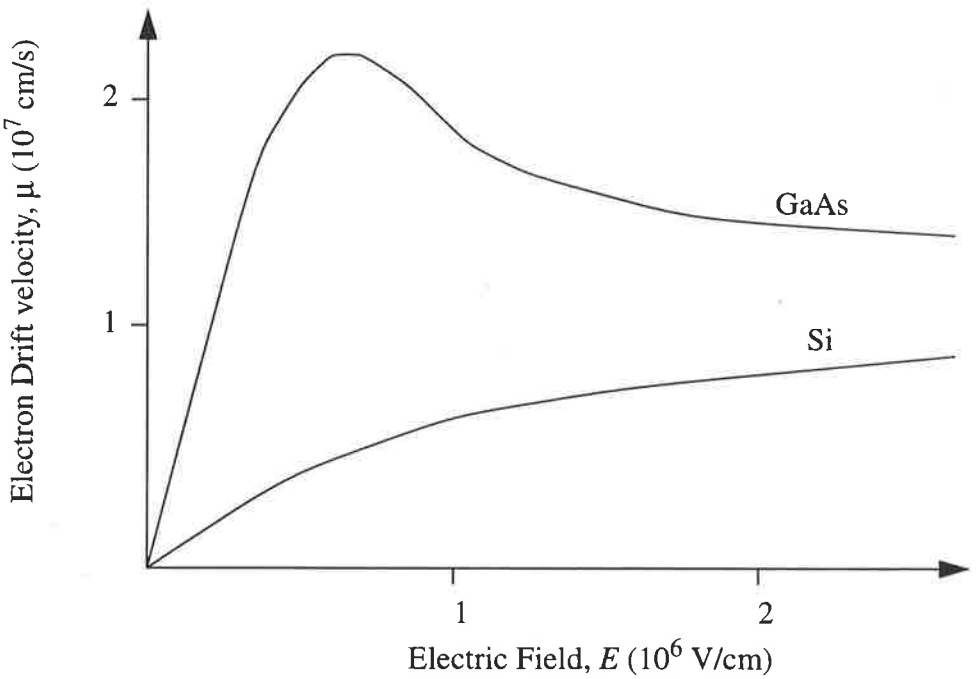


Figure 2.2. Drift Velocity of electrons in GaAs and silicon.

2. Gallium Arsenide's band gap is higher than that of Silicon. The band gap is the distance (in energy) between the valence and conduction bands. In large band gap materials, few electrons have enough thermal energy to jump from the valence band to the conduction band and hence the material will be an insulator. In conductors, the band gap is very low and many electrons have the energy to reach the conduction band.

   Gallium Arsenide's higher band gap therefore makes it a better insulator than silicon and its resistivity is consequently approximately 3 orders of magnitude higher. This provides for better insulation between devices without the pn junction isolation required in Silicon[20]. More importantly, however, the semi-insulating nature of the substrate reduces the magnitude of parasitic capacitances compared to the better-conducting silicon substrate, further increasing the relative speed advantage. However, in large digital circuits, often metal lines are closely packed and highly interconnected resulting in high coupling capacitances. This can result in capacitive loading in GaAs being of similar size to a silicon circuit [22].

   Figure 2.3 shows the Energy band diagrams of GaAs and Silicon [18][21].

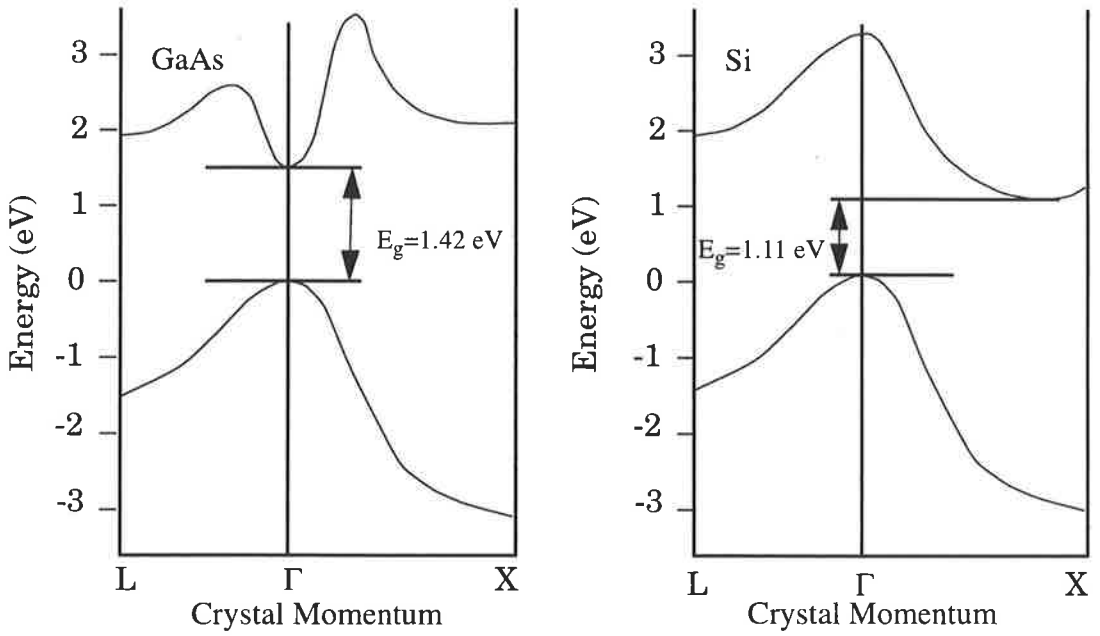


Figure 2.3. Energy Band Diagrams of Gallium Arsenide and Silicon.

3. Gallium Arsenide is a direct band gap semiconductor. A direct band gap material has the minimum of the conduction band occurring at the same crystal momentum as the maximum of the valence band [18], as demonstrated in Figure 2.3. The most likely transition between valence and conduction band will be the one that requires the least energy change, i.e. from the peak of the valence band to the minima of the conduction band. Thus in GaAs, these transitions require no momentum change and can be effected by emitting or absorbing a photon. This gives GaAs excellent optoelectronic properties, allowing integration of optical and electronic components onto a GaAs chip. This will be important in Ultra High Speed systems of the future [17], where fast optical interconnections will play in important role.

   Figure 2.3 also shows that silicon is not a direct band semiconductor, and thus electrons require both a change in energy and momentum to transfer from one energy band to the other. The momentum change implies some physical interaction with the lattice and is much less likely to occur.

4. Gallium Arsenide has a higher resistance to radiation damage (radiation hardness) than silicon. This makes GaAs an ideal technology for use in space where high radiation levels are present, giving higher reliability ICs and lighter spacecraft as the need for heavy shielding is reduced [23].

Although these advantages make GaAs appear a perfect semiconductor material, there are many problems and disadvantages associated with designing, fabricating and using it. While some of these are elaborated further on, a short list is presented here:

1. Gallium Arsenide material is generally more brittle and fragile than silicon, and requires very careful handling. It is also highly poisonous.

2. Gallium Arsenide is more susceptible to damage by Electro-Static Discharge (ESD).

3. Gallium Arsenide circuits are usually more susceptible to leakage effects.

4. Present Gallium Arsenide MESFET technology provides only n-type devices, meaning all logic families are non-complimentary - similar to silicon nMOS, although complimentary GaAs processes, offering both n-FET and p-FET devices are in the developmental stage [24][25].

5. Gallium Arsenide process technology is about 5 years behind present silicon technology. This is reflected in larger device sizes and less sophisticated modelling.

6. Gallium Arsenide fabrication costs are more expensive than an equivalent silicon process.

2.2 Technology

Gallium Arsenide technology currently provides two major technologies; MESFET (MEtal Semiconductor Field Effect transistor) and HEMT (High Electron Mobility Transistor), with the MESFET being most readily available. MESFET technology has also reached a point where Very Large Scale Integrated Circuits, with more than 1 000 000 transistors on a wafer are achievable [27]. HEMT technology is still at a lower level of integration, making it useful for small, very high speed circuits.

2.2.1 GaAs Schottky Barrier Diodes

The Schottky Barrier diode is the simplest of the GaAs devices. A Schottky barrier diode is a rectifying metal-semiconductor junction, named after W. Schottky [28]. In GaAs, such a diode consists of a layer of gate metal over n-doped substrate (active area), the metal becoming the anode connection and the semiconductor becoming the cathode. This forms a metal-semiconductor junction. The current flow mechanisms in the schottky diode are different from those in the *pn* junction diode (minority carrier devices) as schottky diodes are primarily majority carrier devices [19]. Many detailed discussions of schottky barrier diode operation are available [19][28][29].

A potential barrier is formed at the junction of the metal and semiconductor. The potential barrier is often referred to as the built-in voltage, V_{BI}. At the equilibrium point of the metal and semiconductor, the built-in voltage is such that zero net current flows across the junction, with current flow due to the built in voltage from the channel to the metal being matched by thermionic emission of electrons in the opposite direction from anode to cathode [28]. This results in the highly n^- doped substrate being positively biased with respect to the metal gate, as in Figure 2.4. The current due to the built-in voltage is called the saturation current, I_S.

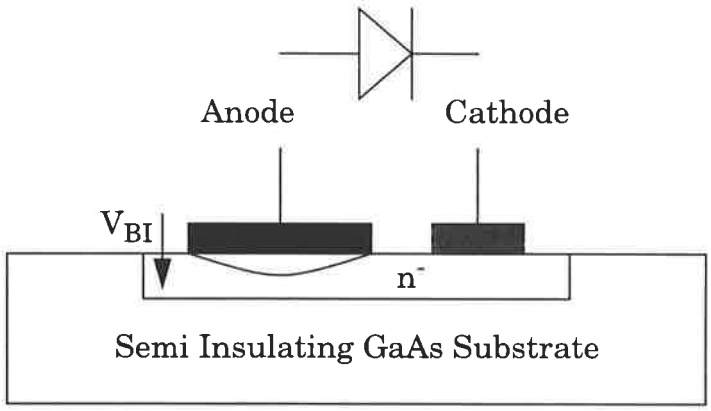


Figure 2.4. Schottky Barrier Diode.

If a forward bias is applied to the diode, the built-in voltage is reduced and electrons will begin to flow from the channel and into the metal. As the forward bias is increased, the current flowing will increase exponentially. The reverse current due to thermionic emission, with value $-I_S$ remains independent of the applied voltage [19]. Under the reverse-bias case, the built-in voltage is increased which effectively reduces the current flowing from channel to metal to zero, leaving a constant reverse bias current of electrons flowing from the metal into the channel. The current flowing in a schottky diode can be expressed as:

$$I = I_S \cdot \exp\left(\frac{q \cdot V}{n \cdot k \cdot T}\right) - I_S \qquad \text{(EQ 2.2)}$$

where I_S is the saturation current, V is the voltage applied to the diode terminals, n is the ideality factor and kT/q is the thermal voltage. The first term in Equation 2.2 is the forward current, the second being the reverse current. Referring to the above explanation of diode operation:

1. Under forward bias, as V increases, then I increases exponentially.

2. Under no external bias (V = 0), the forward and reverse currents are equal, resulting in zero net current flow.

3. Under reverse bias, as V becomes very negative, the first term (forward component of the current) becomes zero and the current becomes the reverse-bias saturation current.

The above analysis is for an ideal schottky diode, but there are several second order effects which are not taken into account in Equation 2.2, including: non-saturation and breakdown in reverse bias [33][34], edge effects [30], high level injection effects [31] and defects in the diode junction [32].

2.2.2 GaAs MESFETs

The GaAs MESFET is quite similar in structure to a silicon MOSFET. There are two varieties of MESFET; the enhancement mode MESFET (normally off) and the depletion mode MESFET (normally on).

The MESFET is formed by implanting n-type ions in the semi-insulating GaAs substrate, forming the conductive channel, and placing a layer of metal over the channel to form the gate. Drain and Source connections are then made at each end of the channel using an ohmic contact - a metal semiconductor junction which does not produce a schottky barrier junction. The gate metal and implanted channel form a schottky gate junction and due to re-combination of electrons and holes, a depletion region is formed on either side of this junction. On the channel side, this depletion region extends into the channel some distance, reducing the effective thickness of the implanted conductive channel. Depending on the impurity concentration and depth of the channel, the depletion region may extend down to the semi-insulating substrate, making conduction from drain to source impos-

sible. This condition is referred to as cut-off. By applying a voltage to the gate relative to the source, the depth of the depletion region can be altered. A positive voltage on the gate will attract electrons from the substrate into the channel, increasing their concentration and thus reducing the depth of the depletion region, thereby increasing the effective thickness of the conducting channel. Conversely, by applying a negative voltage to the gate, electrons in the channel are repelled, reducing their concentration and increasing the depth of the depletion region. The effective thickness of the conducting channel is reduced. The voltage at which conduction begins (i.e. channel thickness becomes non-zero) is called the threshold voltage.

An enhancement mode, or normally off transistor is one which under no external gate-source bias is cut off. A positive gate source voltage must be applied to turn on the transistor. In the Vitesse H-GaAs II process, the threshold voltage of an enhancement mode transistor is typically 0.227 volts [26].

A depletion mode, or normally on transistor is one which is conducting under no external gate-source voltage. A negative gate-source bias must be applied in order to switch the transistor off. Depletion mode transistors have a typical threshold voltage of -0.87 volts [26].

The operation of enhancement and depletion mode transistors is illustrated in Figure 2.5.

Figure 2.6 shows a typical MESFET characteristic. Three distinct areas of operation are visible:

1. Subthreshold: The gate-source voltage applied to the MESFET is below the threshold voltage. This region is characterised by very low levels of current flowing in the drain-source region as the transistor is in its off state. This region of transistor operation is currently poorly understood and modelled. However, current research indicates that the effects of the reverse bias schottky leakage currents dominate subthreshold leakage.

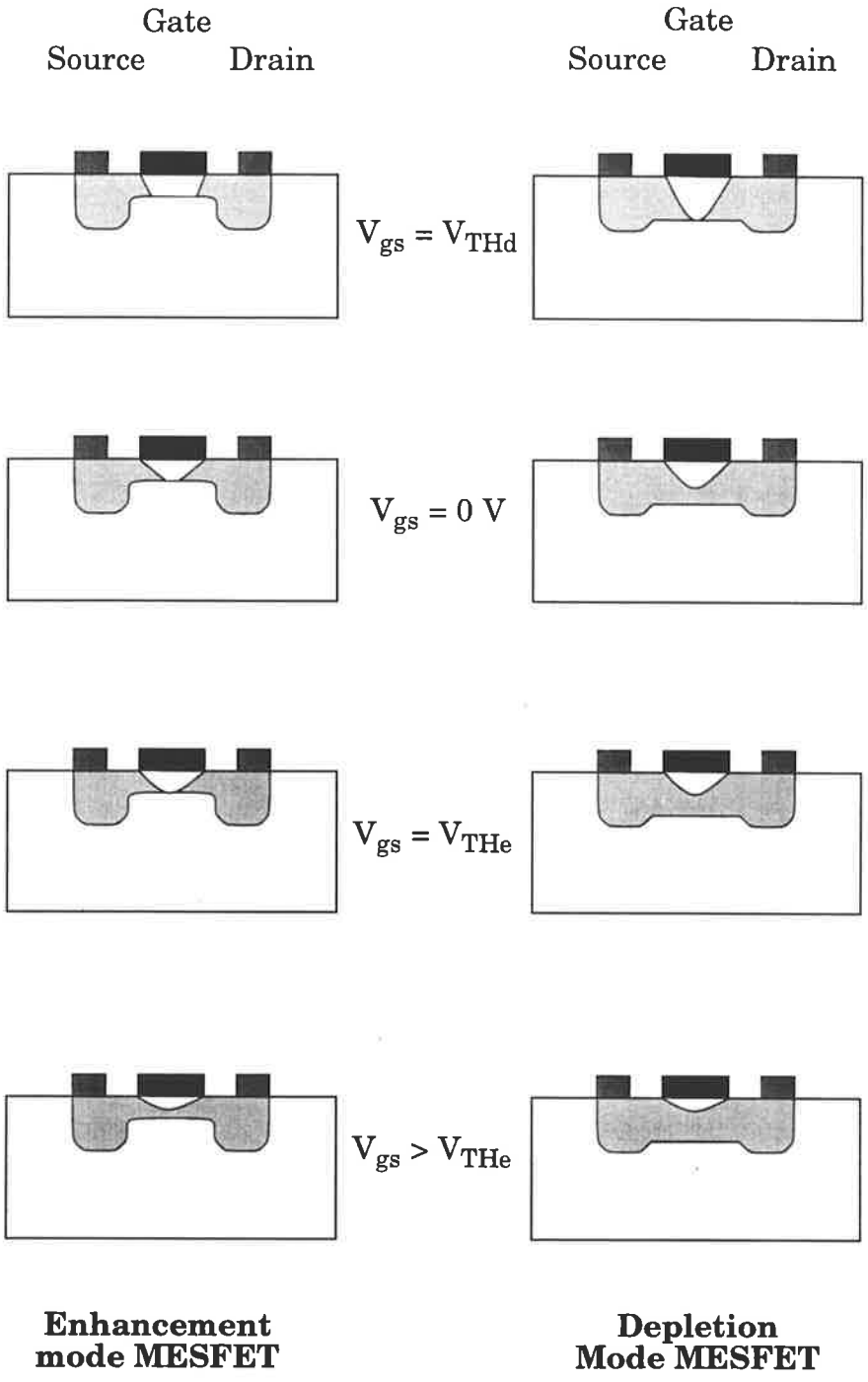


Figure 2.5. MESFET Operation.

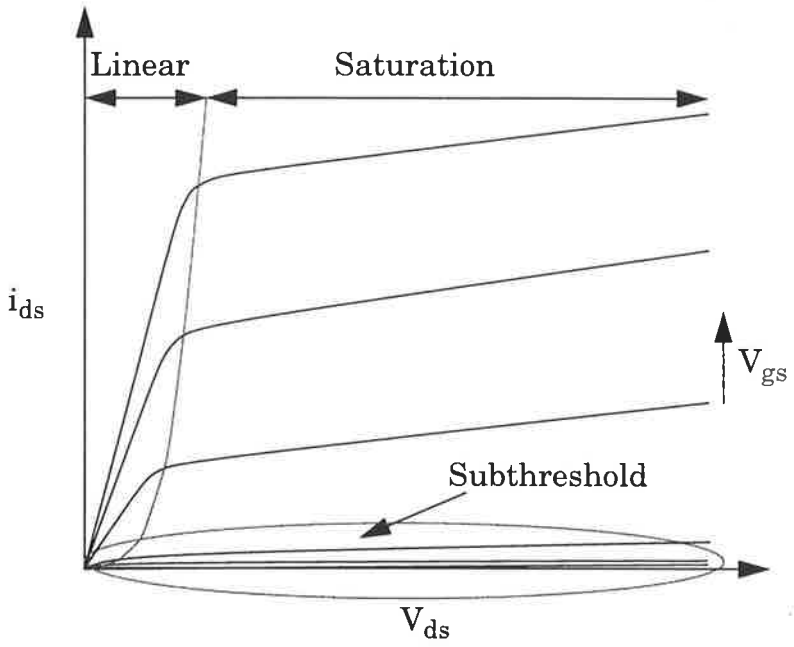


Figure 2.6. Operating regions of MESFET.

2. Linear: The transistors exhibit a linear region where drain-source current is approximately proportional to drain-source voltage when drain-source voltage is low.

3. Saturation: As drain-source voltage is increased, the drain-source current saturates and the current remains approximately constant. Some slight increase is observed due to the dependence of threshold voltage on drain-source voltage.

2.2.3 Fabrication

The Vitesse H-GaAs II process is a self-aligned gate additive implant MESFET process [26]. The self-aligned gate process defines the schottky gate first, and then positions the drain and source of the transistor relative to the gate, reducing alignment errors. All transistors are initially fabricated as enhancement mode, and a second mask layer is then used to add extra implant to those transistors which are required to be depletion mode, hence the term additive implant.

The basic masking steps for an E-MESFET are shown in Figure 2.7 [17]. Extra masking steps are required depending on the number of layers of metal used.

2.3 Layout

The software that was used to produce the VLSI layout is *Magic* [35]. This software uses a graphical environment to design the layout. The software includes a real-time error checking facility, so that any design rule errors are shown and can be fixed immediately. Once a layout is generated, a spice compatible netlist can be produced. Once the external stimulus has been added to the file, it can be input to *hspice*, the simulation program used for this work.

The layers used in the H-GaAs II process are shown in Figure 2.26.

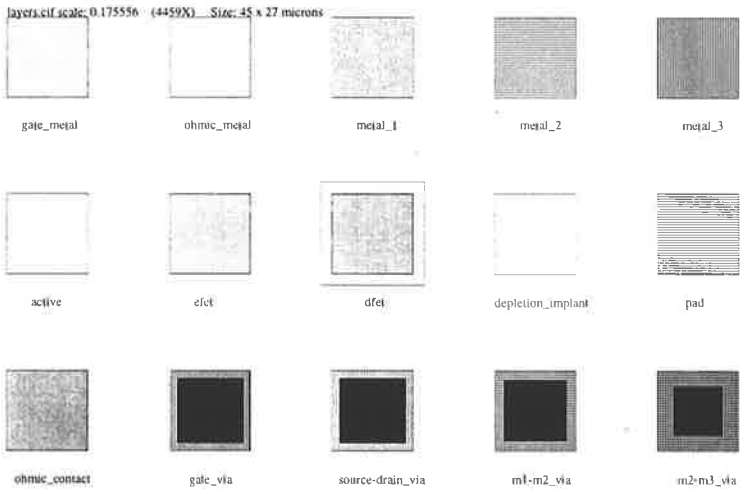


Figure 2.8. H-GaAs II layers.

2.4 Modelling and Simulation

All simulation in this thesis was done using *hspice*, an advanced SPICE compatible device level simulator from Meta Software [36]. Hspice uses equations to model all devices. These equations require specific parame-

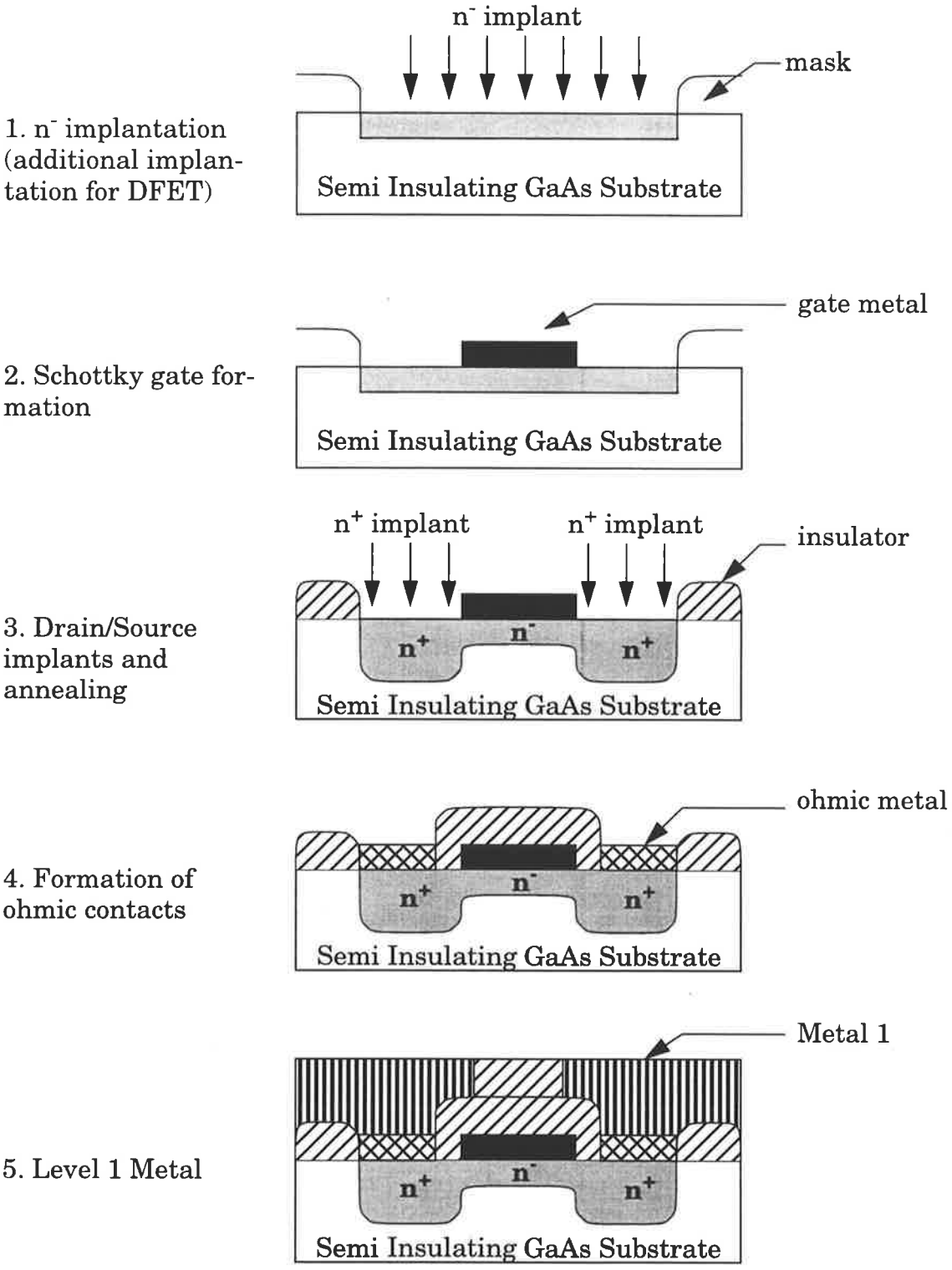


Figure 2.7. Basic Fabrication Steps.

ters to be input regarding the characteristics of the device. The parameters are supplied by the foundry, and are the result of curve-fitting on the measured characteristics of many devices. The parameters describing the process have a spread of values associated with them. See Section 2.4.5 for a more detailed description.

While the schottky diode is equivalent in functionality to a MESFET with drain and source shorted together, modelling of schottky diodes is not done using the MEFET parameters. Instead, a separate set of parameters used specifically for diode modelling is used, which provides a more accurate match to diode characteristics than the MESFET characteristics which are more concerned with modelling MESFET behaviour. The diode models are denoted by the prefix DIO in the parameter file, while the MESFET parameters have the prefix JFET [26].

2.4.1 Diode Characteristics

GaAs schottky diodes are modelled in *hspice* using the ideal diode equation as discussed above. To increase simulation speed, a reverse saturation voltage is defined, below which the diode current is assumed to be equal to the saturation current. The hspice diode equations are:

$$\begin{aligned} i_d &= i_{S_{eff}}(t) \cdot \left(e^{\frac{v_d}{N \cdot v_t}} - 1 \right) && v_d > -10 \cdot N \cdot v_t \\ i_d &= -i_{S_{eff}}(t) && v_d \leq -10 \cdot N \cdot v_t \end{aligned} \qquad \text{(EQ 2.3)}$$

The value of saturation voltage is typically low (of the order of several hundred millivolts). Table 2.1 shows the reverse bias saturation voltage for both enhancement and depletion mode schottky gate-source/drain diodes over various temperatures. Typical parameters are assumed.

However, as discussed in Section 2.2.1, the behaviour of GaAs schottky barrier diodes is not ideal, and this causes inaccuracies in the simulation result. In normal static logic, these inaccuracies are negligible, but because

dynamic memory operates in the subthreshold region with very small currents (in the order of picoamps), these errors are significant. In particular, the non-saturating nature of the reverse-bias characteristic means that the net leakage current cannot be accurately simulated and a generous safety factor must be allowed, as discussed in Section 5.5.

Table 2.1. Diode Reverse Saturation voltages versus Temperature for typical minimum sized D- and E-MESFETs.

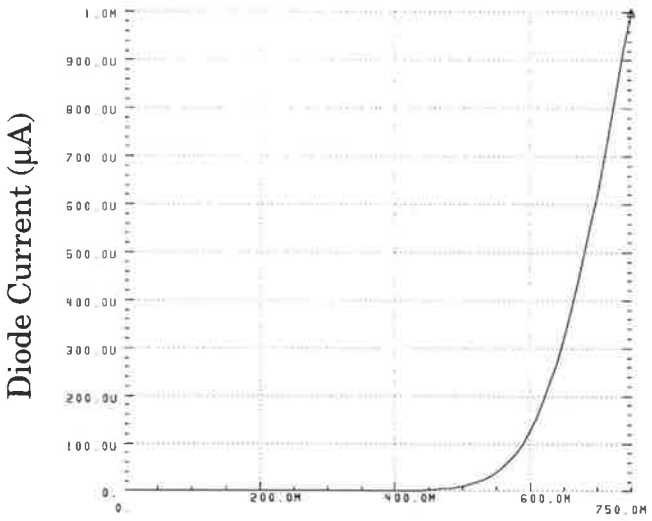
Saturation Voltage $V_{sat} = -10 \cdot N \cdot v_t$	Enhancement Mode	Depletion Mode
25 °C	-0.315v	-0.352v
75 °C	-0.368v	-0.412v
125 °C	-0.421v	-0.471v

Due to the difference in the channel thickness and doping concentration of enhancement mode and depletion mode transistors, the schottky diodes at the gates of these transistors will have different characteristics. Diode current characteristics of similar MESFETs will also vary over the wafer and process, as discussed in Section 2.4.5.

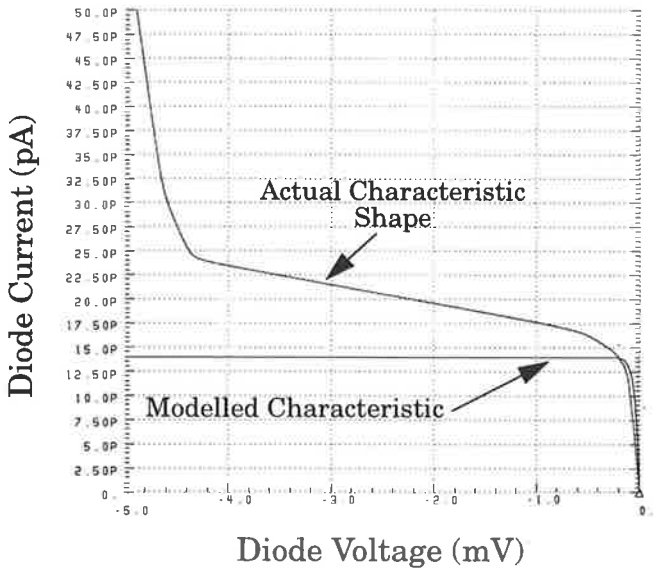
A *hspice* simulation of the forward and reverse characteristics of a typical enhancement mode schottky diode at 75 °C are shown in Figure 2.9. it can be clearly seen that significant forward conduction starts at a voltage of approximately 0.6 volts. On the reverse characteristic, the simulated response is shown, along with the general shape of the actual schottky barrier diode response, although this is not to scale and is shown only to demonstrate the deficiencies in the ideal *hspice* model.

2.4.2 MESFET Equivalent Circuit and Model

The MESFET equivalent circuit used in *hspice* [36] is shown in Figure 2.10. A full description of the *hspice* model can be found in [36]. The equivalent circuit is composed of drain and source resistances, gate-source and gate-drain schottky diodes, gate-source and gate-drain capacitances and a drain-source current source.



Forward Biassed



Reverse Biassed

Figure 2.9. 1.2μm x 10.0 μm schottky diode characteristics.

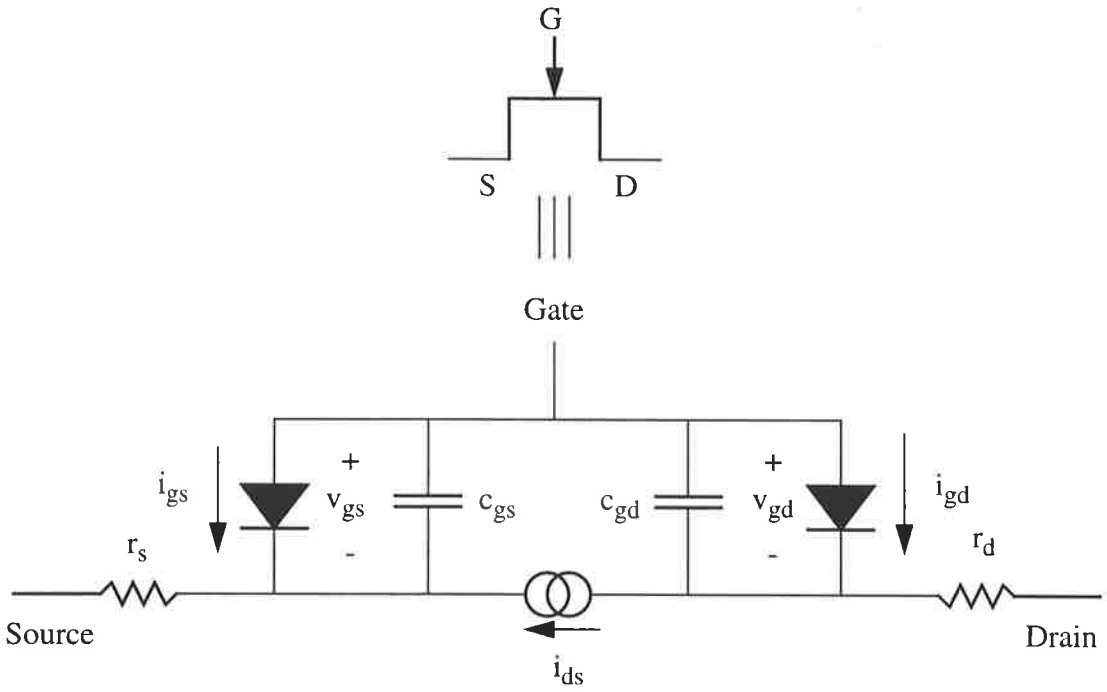


Figure 2.10. *Hspice* MESFET model.

This circuit is a simplified version of more complex equivalents [37][38]. *Hspice* offers three levels of simulation, each increasing in complexity. The level 1 model is the basic SPICE model, based upon the work done by Schichmann-Hodges [39]. The level 3 model is used in all simulations in this work and is significantly more advanced, being based on the Curtice model [40] with alterations by Statz [41].

The basic Curtice model MESFET equation [40] is:

$$i_{ds} = \beta \cdot \frac{W}{L} \cdot (v_{gs} - v_{TH})^2 \cdot (1 + \lambda \cdot v_{ds}) \cdot \tanh(\alpha \cdot v_{ds}) \qquad \text{(EQ 2.4)}$$

where β is the transconductance of the MESFET, W and L are the width and length of the gate respectively, λ is the proportionality constant between drain-source voltage and threshold voltage and α is the constant to account for current saturation at high drain-source voltages. The Curtice

Model does not include the subthreshold region and so a different model is used in this region, as discussed in Section 2.4.2.1.

A *hspice* simulation showing both enhancement and depletion mode characteristics is shown in Figure 2.11.

In Chapter 3, tests on some fabricated MESFETs used to verify the accuracy of the *hspice* modelling are described.

2.4.2.1 Subthreshold Model

The MESFET subthreshold model uses the same equations used in modelling the silicon MOSFET subthreshold region. This equation is extremely complex, and is similar to that presented in [42]. The subthreshold current model used in *hspice* is:

The subthreshold current expression is:

$$I_{sub} = \frac{I_{lim} \cdot I_{ex}}{I_{lim} + I_{ex}} \qquad \text{(EQ 2.5)}$$

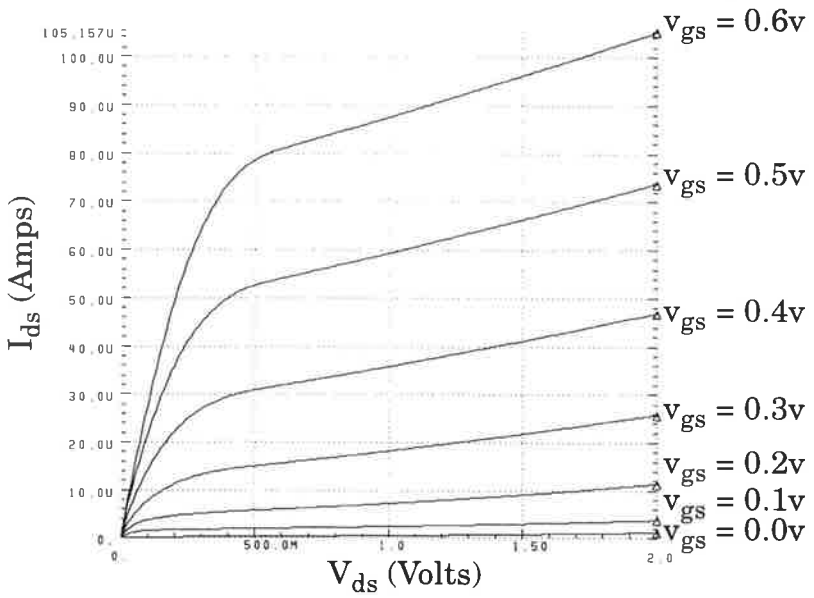
where:

$$I_{ex} = \beta_{eff} \cdot v_t^2 \cdot e^{1.8} \cdot e^{\frac{v_{gs} - V_{TH}}{xn \cdot v_t}} \cdot \left(1 - e^{-\frac{v_{ds}}{v_t}}\right)$$

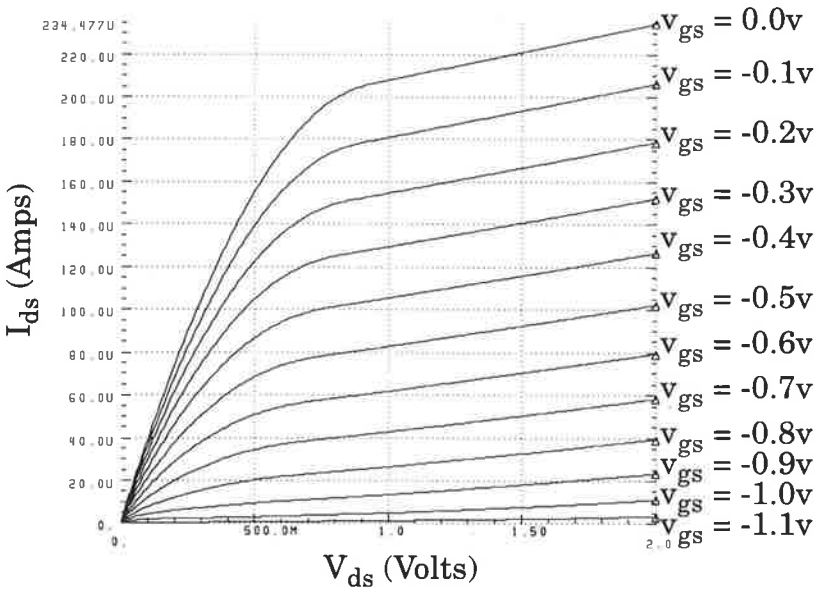
$$I_{lim} = 4.5 \cdot \beta_{eff} \cdot v_t^2$$

$$xn = NG + ND \cdot v_{ds}$$

NG and ND are the gate and drain subthreshold parameters, β_{eff} is the effective gain of the transistor, V_{TH} is the threshold voltage, and v_t is the thermal voltage (kT/q).



(a) Minimum size (1.2μm x 2.0μm) enhancement mode MESFET



(b) Minimum size (1.2μm x 2.0μm) depletion mode MESFET

Figure 2.11. MESFET Characteristics.

From Equation 2.5 it is obvious that as (v_{gs} - V_{TH}) becomes more negative, I_{sub} will approach zero, therefore $i_{sub} \rightarrow 0$ as $v_{gs} - V_{TH} \rightarrow -\infty$. To minimise subthreshold current it is necessary to make v_{gs} as negative as possible.

Although inaccuracies may exist within this model, it has been shown experimentally that provided the gate-source voltage is several hundred millivolts below the threshold voltage, the reverse-biassed schottky leakage currents will be much larger than the subthreshold currents, and thus the total drain current will be approximately equal to the reverse bias schottky leakage [43]. This is demonstrated in Figure 2.12.

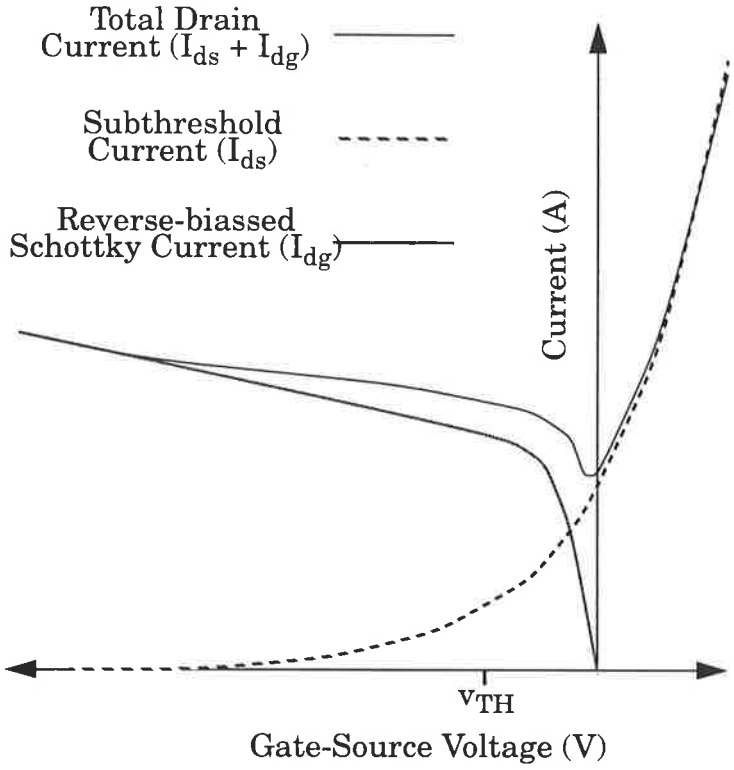


Figure 2.12. Subthreshold and Reverse-bias schottky currents.

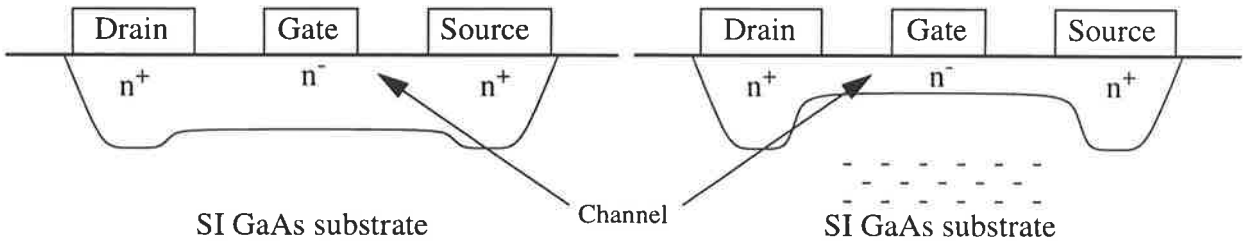
2.4.3 Backgating

Backgating and sidegating effects occur in Gallium Arsenide circuits if changes occur in the space-charge region of the channel-substrate junction.

Such changes can be caused by current flowing into the substrate, voltage being applied to the substrate or illumination of the substrate [44]. When electrons become trapped within the substrate, electrons in the channel will be repelled, increasing the thickness of the depletion region and lowering the effective channel thickness. This will reduce the drain-source current and can be modelled as an increase in the transistor threshold voltage. The transistors thus appear 'slower' and a performance reduction occurs. This problem is enhanced by the fact that the channel/junction capacitance is high and thus suitable for storing large amounts of charge. This effect is demonstrated in Figure 2.13. In hspice, the backgating effect is modelled as in increase in threshold voltage, using the equation:

$$V_{TH} = V_{TO} + \gamma_{ds} \cdot v_{ds} + K1 \cdot V_{BS} \qquad \text{(EQ 2.6)}$$

Where K1 is the backgating coefficient.



No Backgating/Sidegating

Sidegating present showing a reduction in channel depth

Figure 2.13. Cross-section of MESFET.

The backgating and sidegating phenomena and its adverse effects on circuit operation have been well researched [46] - [50]. Three distinct processes can be identified [51]:

1. Self-Backgating:

   Self backgating occurs when a portion of a MESFET's drain-source current flows into the substrate, due to the finite resistance of the semi-insulating substrate.

2. Backgating:

   Backgating is caused by a negative voltage (with respect to the source of the affected transistor) being applied to the substrate. This is usually done via a transistor drain/source being biassed negatively. This is analogous to the body effect encountered in CMOS circuits. However, due to the highly resistive nature of the substrate, its potential is not well known [19]. It will depend on the number and nature of transistors in the local area.

3. Sidegating:

   Sidegating occurs when negative charge flows from a sidegate node which is negatively biassed with respect to the source of the affected transistor, into the substrate and then into the channel of the affected transistor. Because the substrate is slightly p^+, this current is akin to collector-emitter current flowing in an NPN transistor, with the substrate as the base.

To account for the backgating/sidegating effect in simulation, it is recommended that the substrate node in the MESFET model be connected to 0.6 volts above the most negative supply [26]. Research has shown that due to the high resistivity of the substrate, backgating and sidegating are only localised phenomena, with affected MESFETs in a range of the order of 10 to 50 μm from the negatively biassed sidegate/backgate node [48][51]. This range is highly dependent on the magnitude of the negative bias. Several methods of reducing the effect of backgating have been investigated, including compensating the substrate material to minimise the channel/substrate capacitance [47], using a p-Type ring to isolate the MESFETs [52], use of a negatively biassed Schottky gate metal ring as isolation [53] and use of a buffer layer between channel and substrate [54].

2.4.4 Temperature

Temperature effects are built in to the *hspice* MESFET and diode models. The schottky diode current varies exponentially with temperature and this

is modelled as an exponential increase in the saturation current, i_S, with temperature.

MESFETs exhibit various changes with temperature:

1. A reduction in threshold voltage as temperature is increased, acting to increase the drain-source current.

2. A reduction in transconductance as temperature is increased, acting to reduce the drain-source current.

3. Variations in gate-source and gate-drain capacitances.

The overall effect of temperature on drain current is dependent on gate voltage: at low v_{gs}, the threshold voltage shift is dominant and hence drain current increases with temperature. At high v_{gs}, the opposite is true and a reduction in drain current is observed as temperature increases [19].

2.4.5 Process Variation

Due to variations during fabrication in process parameters such as channel implant dose, activation efficiency, built in voltage and substrate material, and short and long channel effects, supposedly identical MESFETs will exhibit large variations in characteristics over a wafer (local variation) or set of wafers (global variation). These variations are characterised by an average or typical value, and a standard variation from the average. To maximise yield, it is important that the design be tolerant to a wide variation in process parameters.

Process variation can occur in two forms [26]:

1. *Fast:* Under fast variation, the threshold voltage of the transistor is lower than that of the typical case. This results in the transistor turning on at a lower gate-source voltage and hence the circuit will operate faster than typical. The power dissipation is also increased.

2. *Slow:* When slow process variation occurs, the fabricated transistor has a higher threshold voltage than the typical case. The transistor hence requires a higher gate source voltage to turn on and the circuit will therefore operate at a slower speed. The power dissipation is proportionally reduced.

A diagram shown in Figure 2.14 explains these concepts more clearly. The example is given for an enhancement mode transistor. The waveform shown is an input to the gate of the MESFET. On the horizontal axis are shown the times when the input voltage is enough to turn the transistor on. It can be seen that the fast transistor is turned on quickest, followed by the typical and slow transistor.

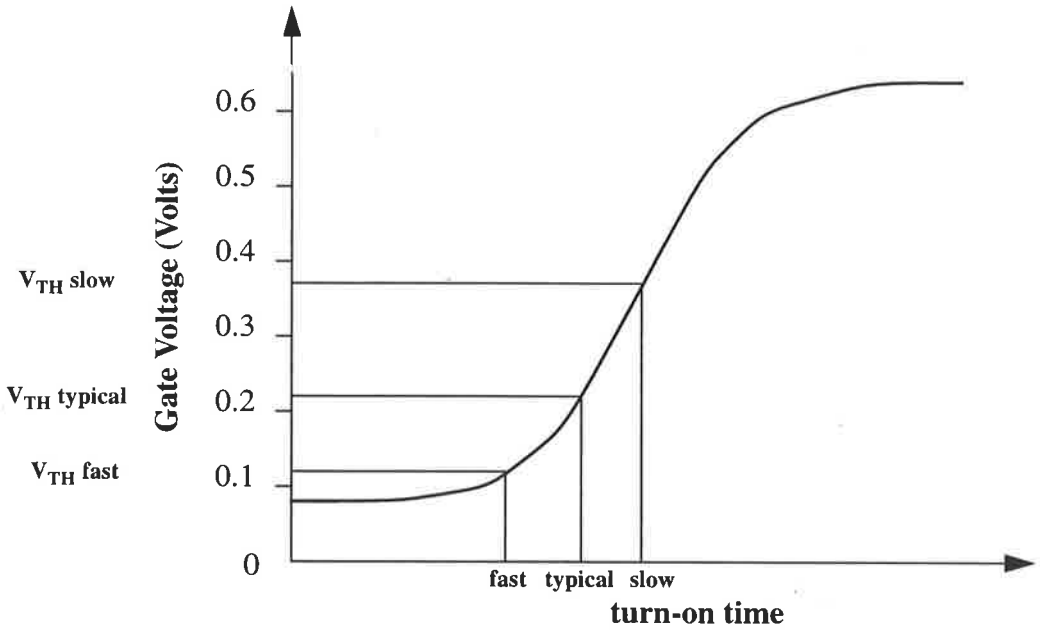


Figure 2.14. Demonstration of fast, typical and slow transistors.

In addition to the delay caused by the increase of threshold voltage, a second order effect further reduces operational speed as transistors become 'slower'. The maximum value of v_{gs} is limited by the schottky gate to approximately 0.6 volts. Therefore as the threshold voltage, v_{TH}, becomes higher, we can see from Equation 2.4 the drain source current will reduce with the square of the threshold voltage. Therefore as transistors become

slower the time taken to charge capacitive gates will increase in proportion to $(v_{gs} - v_{TH})^2$. This further compounds the delay, as not only does the transistor gate need to be charge to a higher value to be turned on, the charging itself is occurring at a slower speed as even when turned on, they are 'less turned on' than in the typical case.

Following a similar argument, fast transistors will sink/source more current than the typical case which further increases the speed advantage gained from the lower threshold voltage.

The process spread occurs in a normal distribution, centred around the typical parameters. Figure 2.15 shows a graphical representation of the distribution of threshold voltages. The bottom left corner is the fast corner of the process, with fast enhancement and depletion mode MESFETs. The top right corner is the slow corner of the process. As discussed above, the H-GaAs II process uses an additive implant technique, whereby depletion mode transistors are formed by re-implanting enhancement mode transistors hence increasing the channel concentration. If the original enhancement mode MESFET was slow, then the added implant is likely to make a slow depletion mode transistor and conversely for the fast case. Thus enhancement and depletion mode MESFETs will be formed along the diagonal connecting the slow and fast corners shown in Figure 2.15. Therefore simulation is done using parameters where both enhancement and depletion mode MESFETs are slow, typical or fast, and the cases of fast E-MESFETs, slow D-MESFETs and fast D-MESFETs, slow E-MESFETs are not simulated as fabrication of such transistors is unlikely.

2.5 Layout Methodology - Ring Notation

Layout methodology is an important part of any VLSI design. *Ring notation* [55] is a method of placing transistors which gives a high packing density, reduces the length of interconnects and provides for noise isolation.

It provides a simple intermediate symbolic stage between schematic and layout which can be drawn quickly by hand showing all transistors and

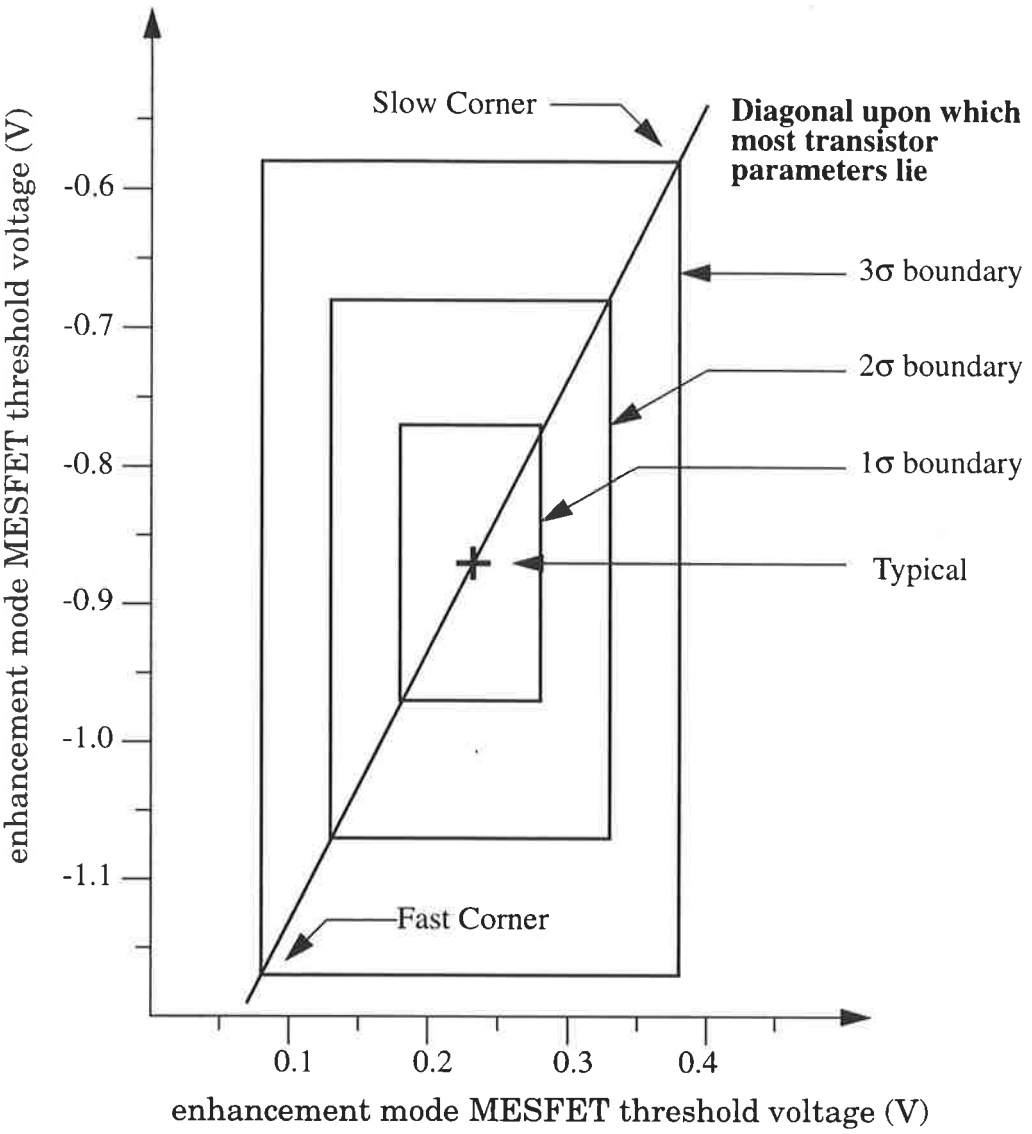
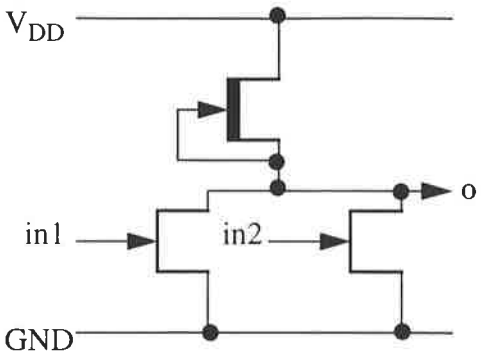


Figure 2.15. Graphical representation of process variation.

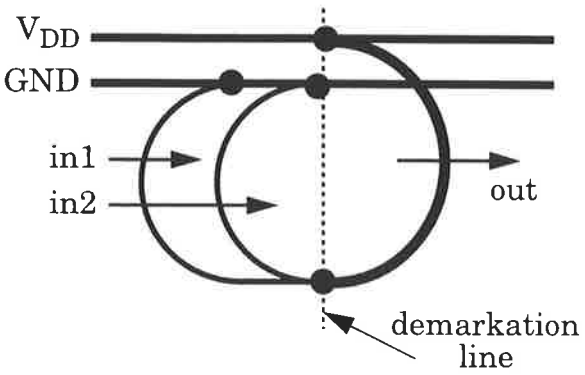
interconnects. This symbolic stage can then be mapped directly into a layout, and the result is a uniform layout, as opposed to using a haphazard approach and placing transistors 'at random'.

An example of a ring notation design is shown in Figure 2.16 for a DCFL NOR gate.

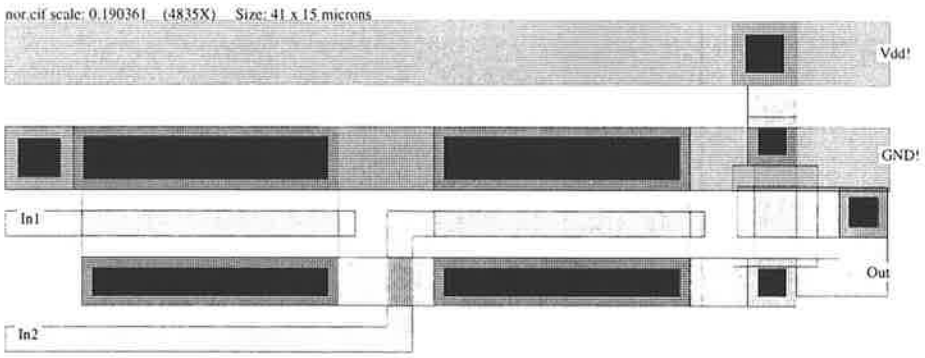
Most layouts in this thesis use the ring notation methodology, however in certain places where circuits were required to have minimal width (e.g. column drivers), ring notation was not practical.



DCFL NOR gate schematic



Ring Notation Symbolic Diagram



NOR gate layout

Figure 2.16. Ring Notation.

2.6 Metal Line Sizing

Because of the finite resistance and current carrying capacity of the metal layers, the sizing of both power supply and interconnect lines is critical to ensure correct, reliable operation of the circuit. The lines are dimensioned according to two criteria [26]:

1. Current carrying capacity:

   Each metal layer is capable of supplying a certain amount of current per unit width. If this limit is exceeded, the metal line is subject to degradation via metal migration or heating until eventual failure occurs [64].

   The current carrying capacity of a line can be expressed as:

$$I_{MAX} = I_{cl} \cdot (W - \Delta W) \quad \text{(EQ 2.7)}$$

Where I_{cl} is the Maximum Current Limit per unit width, W is the width of the line and ΔW is the process control factor to allow for inaccuracies which may occur in the line fabrication width.

2. Voltage drop:

Metal lines have a finite resistance proportional to their length, and so a voltage drop proportional to the current flowing and line length will occur. For supply lines, this difference should not exceed 5% of the supply voltage [26]. For a 2 volt supply, the allowable voltage drop is therefore 0.1 volts. The voltage drop across a line can be expressed, using Ohms law, as:

$$V_{DROP} = R_{line} \cdot I_{line} = \frac{L}{W} \cdot R_S \cdot I_{line} \quad \text{(EQ 2.8)}$$

Where L and W are the length and width of the line respectively and R_S is the sheet resistance of the layer in Ω/□.

Table 2.2 shows the Layer Resistances and Maximum Current limits of different metal layers for the H-GaAs II process [26].

Table 2.2. Layer Resistances and Maximum Current Limits.

Metal Layer	R_S (Ω/□)	Maximum Current Limit (mA/μm)			ΔW (μm)
		DC	AC	Peak	
Gate Metal	0.5 - 1.5	5.0	5.0	25.0	0.4
Ohmic (n+) Implant	190 - 230	1.0	1.0	2.0	0.0
Ohmic Metal	< 10.0	0.3	0.3	0.6	0.0
Metal 1	< 0.070	1.0	1.0	5.0	0.2
Metal 2	< 0.035	2.0	2.0	10.0	0.0
Metal 3	< 0.025	2.8	2.8	14.0	0.0
Metal 4	< 0.025	2.8	2.8	14.0	0.0

The metal width must therefore be chosen such that it satisfies both of the above conditions under worst case conditions, and the greater of the two line widths must be used.

2.6.1 Contact Sizing

The current limit of a contact is determined by the current handling ability of the two adjoining metals which form the contact, and the width of contact perpendicular to the primary direction of current flow. Therefore, in general, if the contact extends the full width of the line, and the line is adequately dimensioned with a sufficient safety factor, the contact will also be adequately dimensioned. Because of fixed overlap being required on contacts due to design rules, thin contacts will more susceptible to being below minimum width than wider contacts.

2.7 Logic Families

Due to the fact that present gallium arsenide processes use only n-type semiconductor for active channels, the logic families used in GaAs closely resemble those in nMOS, with no complimentary logic families currently in use.

The logic families can be divided into two types: Normally on and Normally off.

Early GaAs processes offered only Depletion mode MESFETs. Any logic constructed had to be made using only D-MESFETs. Depletion mode MESFETs are "normally on", that is when no gate voltage is applied, they conduct from drain to source, and thus logic families utilising only depletion devices are known as "Normally on".

As process technology improved, it became possible to fabricate both enhancement and depletion mode devices together. Enhancement mode devices have no conduction from drain to source when zero gate voltage is applied, and are thus referred to as "Normally off". Logic families utilising both enhancement and depletion mode MESFETs are called "Normally off" logic families. They require no level shifting diodes and only a single supply voltage and are thus simpler than Normally on families.

Although Normally on circuits are the most mature form of GaAs logic, Normally off families are often simpler, dissipate less power and are much more widely used, as most modern commercial GaAs processes offer both enhancement and depletion mode MESFETs.

Most GaAs logic families allow only NOR/OR (parallel pull-down) type operations, and not AND/NAND which require two or more pull-down transistors on series. The reason for this is due to problems with noise margin due to the increase in pull-down voltage. This can be compensated to allow operation of two input NAND gates, but higher fan-ins will have unsatisfactory performance [19].

Using DeMorgan's logic laws, any expression can be rearranged and expressed in terms of NOR/OR functions, so this limitation is acceptable.

2.7.1 Performance Measures

An important part of the evaluation of logic families is to compare their performance. In this section, some of the more common performance measurements are discussed.

2.7.1.1 Noise Margin

Noise margin gives a measure of the tolerance or susceptibility a circuit has to the influence of noise. Noise margin is particularly important in GaAs design due to the low voltage swing caused by the schottky barrier gate diodes.

There are several methods of measuring noise margin, but the two most common are referred to as *maximum square method* [73] and the *slope = -1 criterion* [19]. Both methods require the use of the characteristic (i.e. V_{in} vs. V_{out}) of the circuit being tested.

To measure the maximum square noise margin, a chain of three identical characteristics is required, with the middle inverter characteristic being used as the scale for the x-axis and the characteristic of the inverters on either side being plotted. Figure 2.17 shows the resulting output for a

series of DCFL inverters. The maximum square noise margin can be obtained by finding the dimension of the largest square that can fit in between the two curves.

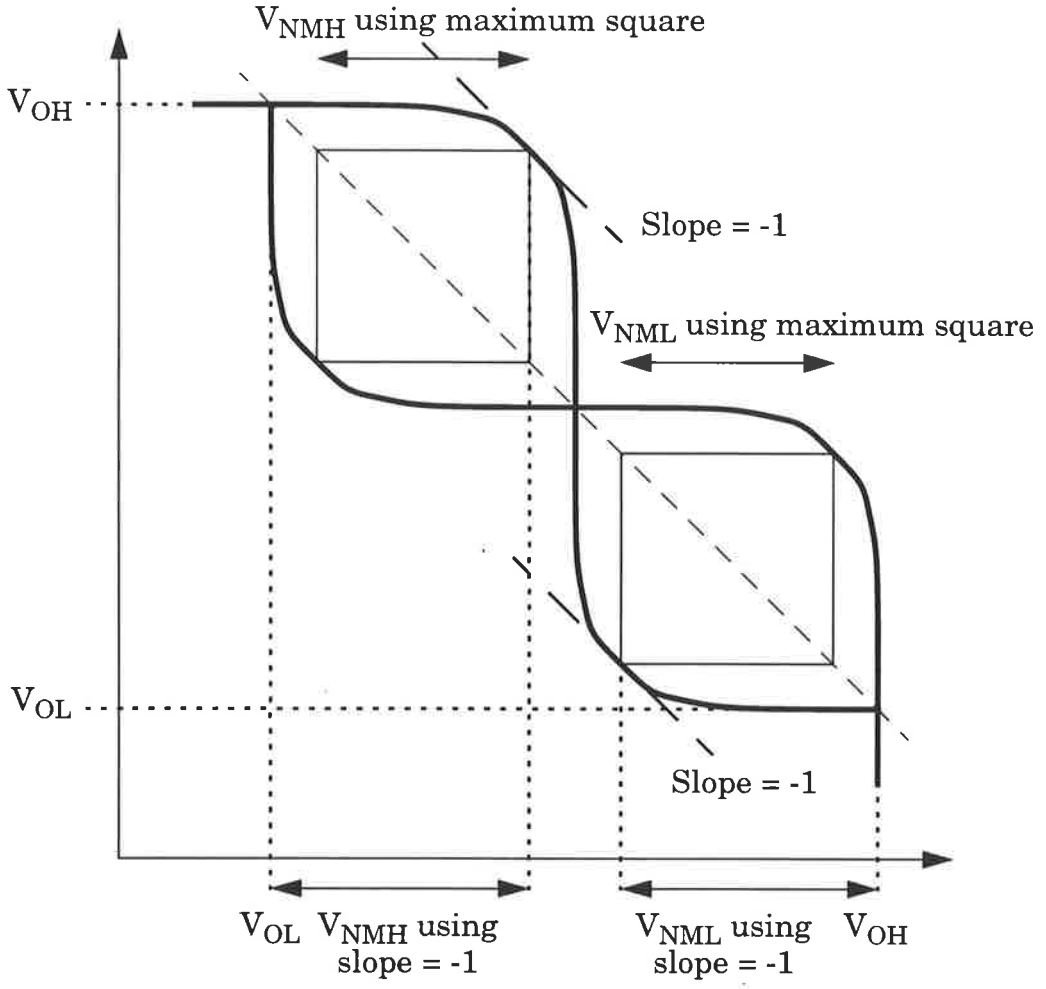


Figure 2.17. Noise Margin.

The noise margin according to the slope=-1 method is found simply using the point on the characteristics at which the slope is -1, again as shown in the figure.

While much debate has occurred regarding the validity and benefits of each method, the maximum square method gives a lower noise margin and can be regarded as a worst case when designing.

2.7.1.2 Fan-out

A circuit's ability to drive those following it can be characterised in terms of its fan-out performance. The fan-out assumes that the circuit being measured is only driving circuits of the same family and ratio as itself, and can be defined as the ratio of the total size of the circuits being driven to the size of the circuit being tested. Therefore, when driving identical circuits in parallel, the fan-out is simply equal to the number of circuits being driven. A circuit with the same pull-up to pull-down ratio, but with transistors three times the size represents a fan-out of 3, for example.

As the fan-out increases, the current required to drive the following transistors rises, resulting in a lower logic high output, and the capacitance also increases which causes extra delay. There are therefore both timing and noise margin considerations when determining the fan-out performance of circuits.

2.7.1.3 Fan-in

Fan-in measures the number of inputs to a circuit. A two-input NOR gate has a fan-in of 2, a four-input gate has a fan-in of 4. As the fan-in increases, capacitance at the input node is increased, causing increased delay when charging the node, and the amount of leakage through pull-down transistors when they are 'off' is also increased, resulting in a slower pull-up that is also reduced in voltage. As with fan-out, there are both noise margin and timing considerations.

Circuits which exhibit good fan-in performance are used to perform logic functions with a large number of variables.

2.7.2 Normally on Logic Families

Normally on logic families require extensive use of level shifting diodes to negatively shift logic levels to a point where they can turn off depletion mode transistors, and multiple supply lines. Normally on circuits are the most mature form of GaAs logic, and many such logic families have been developed, for example:

- Capacitively Coupled Domino Logic [59]
- Inverted Common Drain Logic [60]
- Schottky Diode FET Logic [61]
- Capacitor Diode FET Logic
- Source Coupled FET Logic [62][63]
- Buffered FET Logic [64]
- Unbuffered FET Logic
- Capacitively Coupled FET Logic [65]

No normally on logic families were used in the thesis and so no detailed discussion of their operation is included.

2.7.3 Normally off Logic Families

In the following section, the GaAs logic families used in the designs in this thesis are described, and the advantages/disadvantages of their operation are summarised. The simplest logic family, Direct Coupled FET Logic, is analysed in detail.

2.7.3.1 Direct Coupled FET Logic

Direct Coupled FET Logic (DCFL) [66] is the simplest GaAs logic family. It requires only a single supply rail and two transistors in its most simple form of an inverter. In the general NOR format, the number of transistors required for an n-input NOR gate is n+1. The DCFL inverter and two-input NOR gates are shown in Figure 2.18. A detailed explanation of DCFL operation follows, along with some analysis on the circuit to obtain appropriate transistor sizes for satisfactory operation.

The pull-up transistor, Q2, is a depletion mode MESFET, with gate and source shorted together (i.e. v_{gs} = 0 volts). Therefore this transistor is permanently on. The pull-down transistor, Q1, is an enhancement mode MESFET, the gate of which forms the input to the inverter. The inverter has two states of operation:

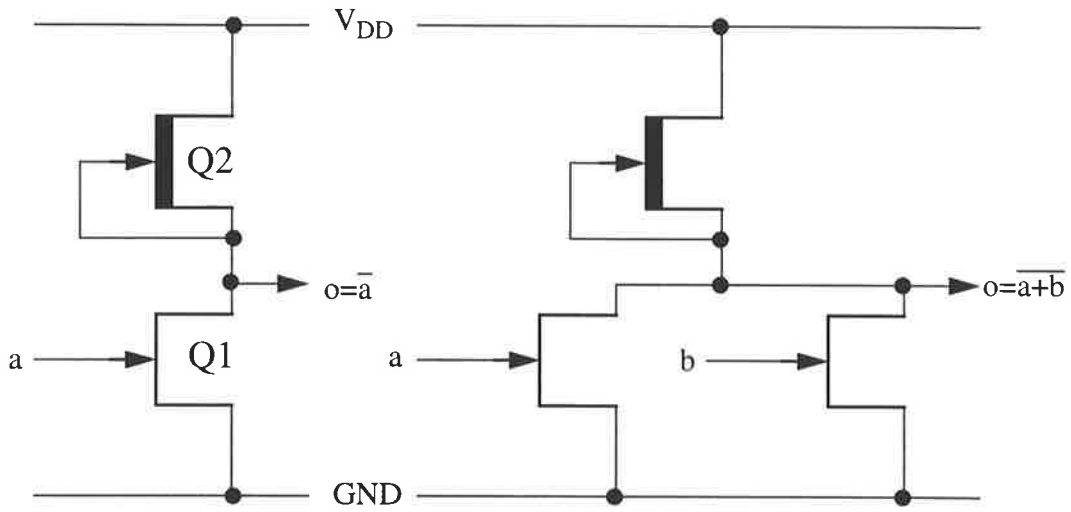


Figure 2.18. DCFL Inverter and 2-input NOR gate.

1. When the input, a, is low, Q1 is off and ideally no current will flow from its drain to source. Therefore the output will be tied to V_{DD} by the permanently on pull-up, Q2. If the output is unloaded, the resultant output voltage will be very close to V_{DD} as there will be very little current flowing through Q2 and hence the voltage drop across it will be small.

   In practical operation there will most often be a load circuit connected to the output of the inverter. In this case the output voltage will depend upon the current drawn by the load. The most common load will be the input of another DCFL circuit, and this is therefore equivalent to a forward biassed schottky diode between the output and ground, as shown in Figure 2.19.

   The output voltage will be that at which the current flowing through the transistor Q2 equals that through diode D1. i.e.

$$V_o \text{ such that: } i_{ds_{Q2}}\Big|_{v_{ds} = V_{DD} - V_o} = i_{D1}\Big|_{v = V_o} \qquad \text{(EQ 2.9)}$$

   Due to the exponential nature of the schottky diode current-voltage char-

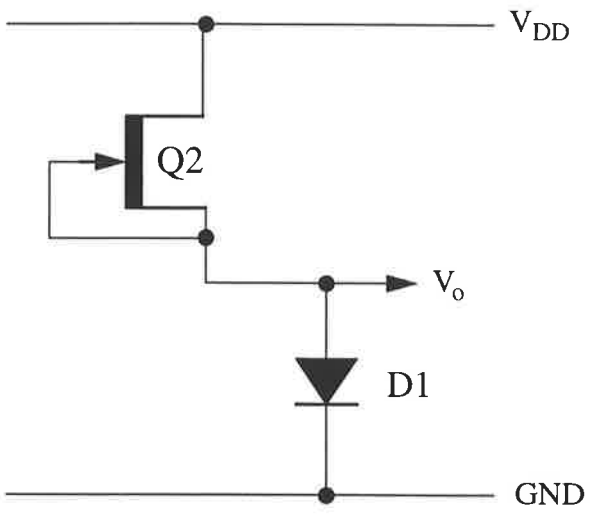


Figure 2.19. DCFL Output Stage Equivalent Circuit.

acteristic, this point occurs at approximately 0.6 volts, with variation of approximately ±0.1 volt due to temperature and process variation.

2. When the input, a, is high, transistor Q1 is on. This causes current to be drawn from V_{DD} through Q2 and Q1 to ground. This results in a lower output voltage than when Q1 was off. The maximum voltage at the gate of Q1 to turn it on is limited by the gate-source schottky barrier diode to approximately 0.6 volts. A gate voltage higher than this will result in significant forward conduction from input to output via the gate-drain schottky diode and increase the output voltage.

   For correct function of the inverter, we require that this output be interpreted as a logic low, and hence the output voltage when Q1 is on must be less than the threshold voltage of an E-MESFET. Because the pull-up transistor is permanently on, and permanently in saturation, it is sourcing an almost constant amount of current. Therefore, if we assume that when Q1 is on, the majority of this current flows through Q1 and very little through the output load schottky diode, the current flowing through Q1 will not vary as the dimensions of Q1 change. To provide an adequate

noise margin, an output low voltage in the region of 0.1 volts should be achieved. This will give the drain-source voltage of Q1 as $v_{dsQ1} = V_o \approx 0.1$ volts and therefore Q1 will be in the linear region. Again, Q2 will be in saturation as $v_{dsQ2} = V_{DD} - V_o \approx 1.9$ volts.

From Equation 2.4, we see that $i_{ds} \propto \beta\, v_{ds}$ and as $\beta = W/L$, $i_{ds} \propto W/L\, v_{ds}$. Therefore, if i_{ds} remains constant, increasing the transistor width to length ratio will reduce the drain-source voltage. By equating expressions for drain-source current in the pull-up and pull-down MESFETs, it is possible to evaluate the value of pull-up to pull-down ratio required to give the output of $V_o = 0.1$ volts. The following analysis refers to the quantities shown in Figure 2.20.

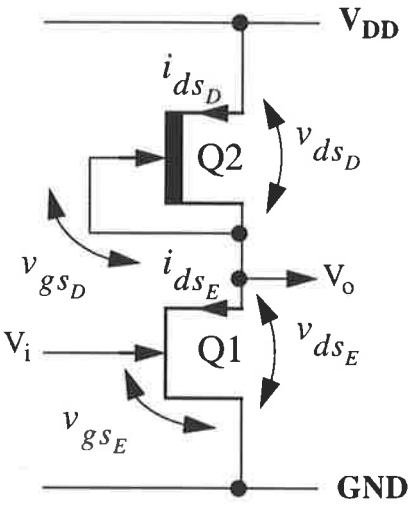


Figure 2.20. Definitions of DCFL inverter voltages and currents.

From Equation 2.4, the current in the pull down MESFET is:

$$i_{ds_E} = \beta_E \cdot \frac{W_E}{L_E} \cdot (v_{gs_E} - v_{TH_E})^2 \cdot (1 + \lambda_E \cdot v_{ds_E}) \cdot \tanh(\alpha_E \cdot V_{DS}) \qquad \text{(EQ 2.10)}$$

Assuming that no current flows to the output (v_o), the two currents must be equal (Kirchoffs current law), and substituting:

$$v_{gs_E} = v_i \qquad v_{ds_E} = v_o \qquad v_{gs_D} = 0\,\text{V} \qquad v_{ds_D} = V_{DD} - V_o$$

and rearranging gives:

$$\frac{\frac{W_E}{L_E}}{\frac{W_D}{L_D}} = \frac{(v_{TH_D})^2 \cdot (1 + \lambda_D \cdot (V_{DD} - v_o)) \cdot \tanh(\alpha_D \cdot (V_{DD} - v_o)) \cdot \beta_D}{(v_i - v_{TH_E})^2 \cdot (1 + \lambda_E \cdot v_o) \cdot \tanh(\alpha_E \cdot v_o) \cdot \beta_E}$$

As discussed above, in the high input, low output state, v_i = 0.6 volts and v_o = 0.1 volts. Also, at typical parameters, 25 °C [26]:

$$v_{TH_E} = 0.227 \text{ V}, \; v_{TH_D} = -0.870 \text{ V},$$
$$\lambda_E = 0.072, \; \lambda_D = 0.037$$
$$\alpha_E = 6.53, \; \alpha_D = 3.91$$
$$\beta_E = 3.02 \times 10^{-4}, \; \beta_D = 2.65 \times 10^{-4}$$

and substituting these parameters and voltages gives a pull-up to pull-down ratio of:

$$\frac{Z_{p.u}}{Z_{p.d.}} = \frac{\frac{W_E}{L_E}}{\frac{W_D}{L_D}} = \frac{\frac{L_D}{W_D}}{\frac{L_E}{W_E}} = \frac{(0.870)^2 \cdot (1 + 0.037 \cdot (2 - 0.1)) \cdot \tanh(3.91 \cdot (2 - 0.1)) \cdot 2.65 \times 10^{-4}}{(0.6 - 0.227)^2 \cdot (1 + 0.072 \cdot 0.1) \cdot \tanh(6.53 \cdot 0.1) \cdot 3.02 \times 10^{-4}} = 8.85$$

We can thus see that a pull-up to pull-down ratio of at least 8.85 (i.e. the pull up transistor must be 8.85 times *more* resistive than the pull-down) is needed for the inverter to pull-down to 0.1 volts at 25 °C under typical conditions. Any increase in this ratio will reduce the pull-down voltage further.

As process parameters and temperature vary, the required pull-up to pull-down ratio also changes. The Vitesse design manual [26] recommends using a ratio of 14, however hspice simulations carried out by the author revealed that a ratio of 12 was satisfactory for function over a temperature range of -25 to +125 °C and a process variation of 2σ slow to 2σ fast. All DCFL logic designed in this thesis uses a pull-up to pull-down ratio of 12 unless stated otherwise.

Once the ratio has been chosen, the physical sizes of the transistors must be decided. This involves trading off speed and drive capability for area and power dissipation. An inverter with large transistors will be fast and able to drive a large load, but it will also consume more power and occupy more area.

There have been many attempts at optimising the transistor sizes of DCFL circuits [56][57][58]. The minimum possible size will be determined by the process limitations. In [56], the authors conclude that the DCFL inverter power delay product is minimised when the width of the pull-down enhancement mode transistor, W_E, is in the region of 8μm to 12μm (assuming minimum gate length of 1.2μm). If W_E is made smaller, the inverter delay becomes excessive, and as W_E is increased, power dissipation becomes larger. The size of the depletion mode pull-up transistor is then determined by the pull-up transistors. In some parts of the memory (address decoders, sense amplifiers etc.), memory performance was sacrificed to enable a lower power dissipation, in which case a smaller pull-down transistor was used. Some examples of DCFL inverter designs are shown in Figure 2.21.

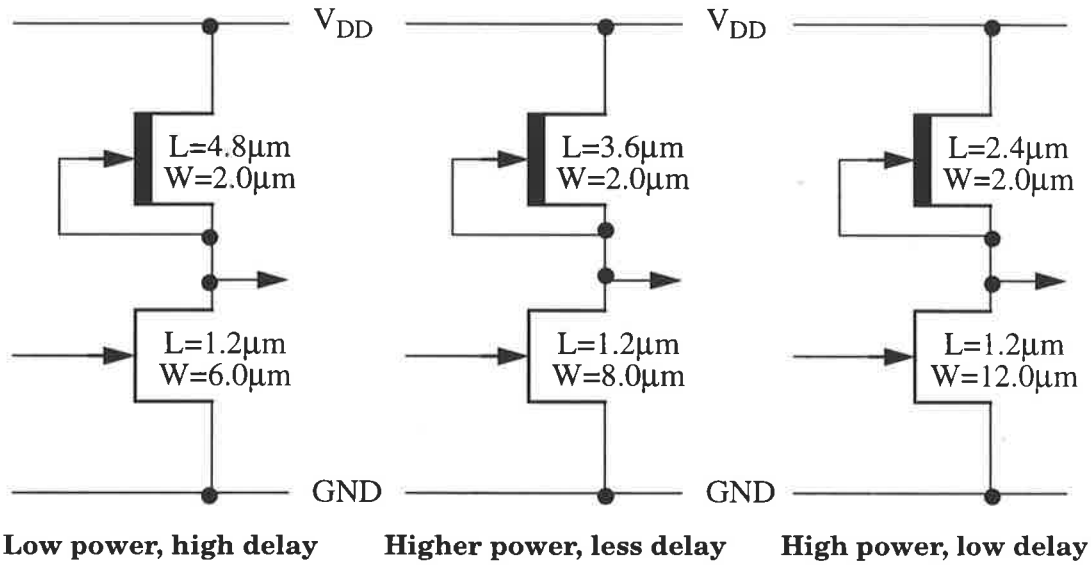


Figure 2.21. Various DCFL Inverters.

A *hspice* simulation showing a DCFL inverter transfer characteristics is shown in Figure 2.22. The result shows the output from two inverters at 25 °C and typical parameters: the first inverter has the pull-up to pull-down ratio previously derived of 8.85 and the second has a ratio of 12 as used in the thesis and shown Figure 2.21b. It can be seen that the inverter with a p.u./p.d. ratio of 8.85 pulled down to approximately 100 mV as designed, and that the second inverter pulled down further, giving it a better low noise margin. The increase in output voltage due to forward conduction with an input voltage larger than 0.6 volts is clearly seen. This is caused by the forward current passing through the gate-source parasitic resitance, r_s, as shown in Figure 2.10.

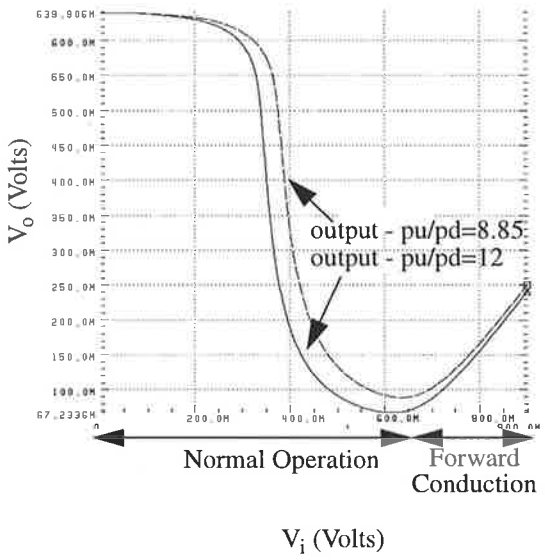


Figure 2.22. Transfer characteristic of typical DCFL at 25 °C.

2.7.3.2 Source-follower Direct Coupled FET Logic

Source-follower Direct Coupled FET Logic (SDCFL) [55] is similar in structure to DCFL but has a source follower output stage. A source follower is an inverted DCFL inverter, as shown in Figure 2.23 and so called because the value at the output (or source of the E-FET) will directly follow the value at the input (the gate of the E-FET). The source follower structure is therefore non-inverting. The source follower utilises a permanently on depletion mode MESFET as its pull-down, with the pull-up as an enhancement mode MESFET.

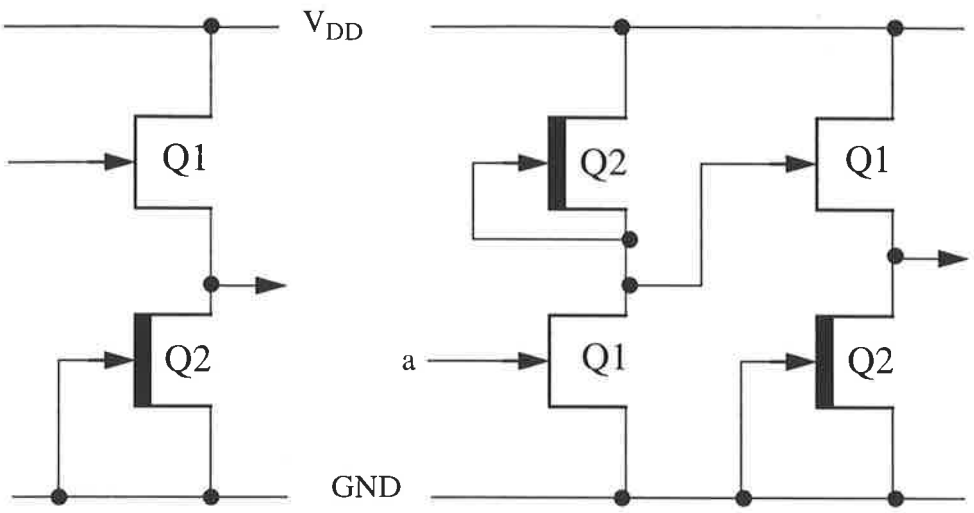


Figure 2.23. (a) Source Follower; (b) SDCFL inverter.

The advantages of SDCFL over DCFL are that the Source follower output stage has a better driving capability than the DCFL inverter, and that the output low voltage is not determined by a pull-up to pull-down ratio - when Q1 in Figure 2.23a is off, the output low voltage is 0 volts. This results in a better low noise margin. However, due to the presence of the source follower, power dissipation is significantly increased.

2.7.3.3 Source Follower FET Logic

Source Follower FET Logic (SFFL) [57] is similar in structure to SDCFL, except that the source follower is used as an input stage that feeds into a DCFL output stage. An SFFL inverter and two-input NOR gate are shown in Figure 2.23. The source follower input stage presents a different load to the DCFL inverter. There is a schottky diode and a depletion mode transistor in series between the input and ground. As a result the output voltage of the previous stage is not limited by the schottky diode to 0.6 volts, but a higher value (approximately 1.0 volts). This results in an increased logic high noise margin. The source follower also allows a bigger fan-in.

The major disadvantage of SFFL is the high leakage current in the forward biassed schottky gate diodes of the inputs (Q1 and Q2 in Figure 2.23b).

This problem is compounded with a high fan-in. SFFL also has an increased power dissipation over DCFL.

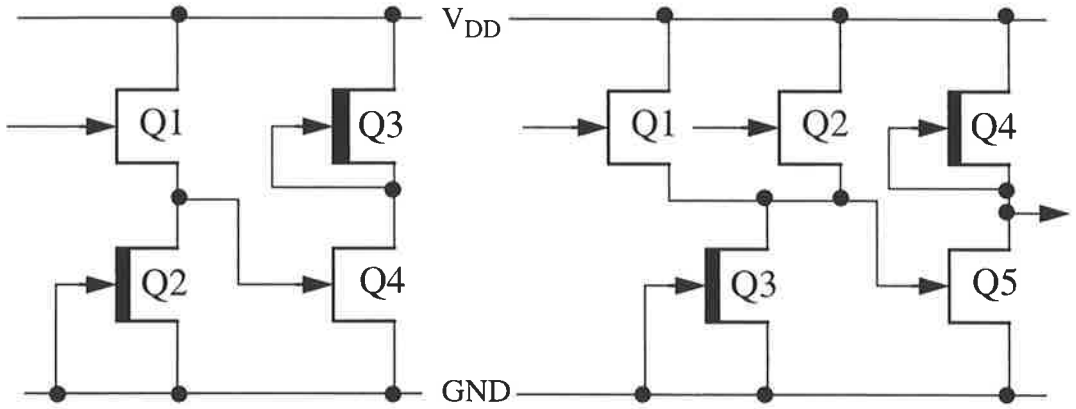


Figure 2.24. (a) SFFL Inverter; (b) SFFL Two input NOR gate

2.7.3.4 Super Buffer FET Logic

Super Buffer FET Logic (SBFL) [19] is a logic family designed specifically for driving large fan-outs and capacitive loads. Its output stage is different from those discussed, in that it is complimentary in operation. The SBFL inverter and two-input NOR gate is shown in Figure 2.25. In Figure 2.25a, it can be seen that the pull-up output transistor, Q4 is connected as a source follower to the output of the preceeding DCFL stage, and the pull-down output transistor, Q3 has its gate connected to the input of the preceeding DCFL inverter stage. The operation is as follows:

Assuming the input, a, is low, then Q1 and Q3 are both off. Q2 is permanently on and so the gate of Q4 is high and therefore Q4 will also be turned on allowing the output to be charged high. As input a is brought high, Q1 and Q3 are turned on, causing the output to be discharged via Q3. The output of the DCFL stage feeding into the gate of Q4 is also bought low via Q1, turning off Q4. The output is therefore low. However, due to the delay in Q4 turning off due to the DCFL stage, Q3 and Q4 are simultaneously on, causing a low impedance path from V_{DD} to GND, resulting in a large current spike. If input a is then taken low again, Q1

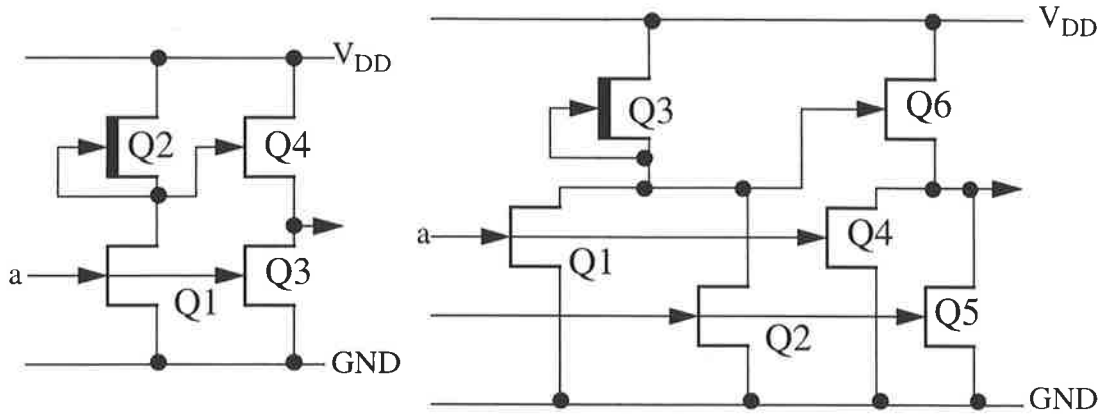


Figure 2.25. (a) SBFL Inverter; (b) SBFL two input NOR gate.

and Q3 will be switched off. The gate of Q4 will then be charged via Q2, and hence Q4 will be switched on allowing the output to be charged high via Q4. Hence there is no current spike associated with the negative input transition (positive output transition), only the positive input transition (negative output transition).

SBFL offers excellent high and low logic levels due to the complimentary nature of the output stage, however its power dissipation is high. The large current spike during the negative output transition can induce noise into the supply line, with unwelcome consequences on the operation of digital logic.

2.7.3.5 Ultra Buffer FET Logic

Ultra-Buffer FET logic was reported in [56] as a means of maintaining the high driving capacity of SBFL, but minimising the supply noise injection by preventing the direct path to ground during positive input transition. It achieves this using a feedback mechanism from the output to the input, via a 2 input DCFL NOR gate instead of the DCFL inverter used in SBFL. The feedback is used to turn off the pull-up transistor once the output is charged high. The output high is then held using a permanently on deple-

tion mode MESFET. A schematic of a UBFL inverter is shown in Figure 2.26.

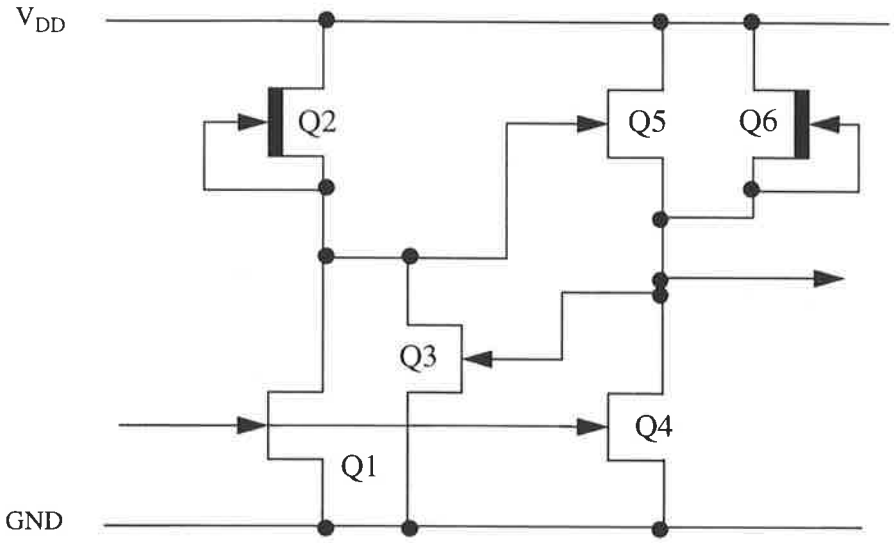


Figure 2.26. UBFL inverter.

The operation of UBFL is as follows:

> When the input, a, is high, Q1 and Q4 are both on. The output of the DCFL NOR stage is therefore low so Q5 is off and the output is low because of the DCFL inverter effect between Q6 and Q4, and Q3 is therefore also off. As the input goes low, Q1 and Q4 are turned off, and as the output is low, Q3 is still off. Therefore, the output of the NOR gate will rise and turn on Q5 which will charge the output high. However, as the output becomes high, Q3 is turned on, lowering the voltage of the NOR gate and turning Q5 off. Therefore in the output high state, both Q4 and Q5 are off, with the output high value being held by Q6. When the input goes high again, Q1 and Q4 are turned on, but Q5 is already off and therefore no direct current path to ground occurs as it did with SBFL. The output is discharged via Q4.

While the supply line noise is reduced using this method, UBFL has several disadvantages caused by the modifications. Due to the NOR gate and pull-up depletion mode MESFET, average static power dissipation is signif-

icantly increased. However, the feedback into the DCFL NOR gate also limits the output high voltage of the UBFL driver to approximately 0.6 volts, making it useful for driving only DCFL circuits.

2.7.3.6 Double Super Buffer FET Logic

Double Super Buffer FET Logic was devised by the author to obtain a driver with the speed of SBFL but a lower power dissipation. The main cause of static dissipation in SBFL is the DCFL stage which drives the output transistors. To reduce the power dissipation, it is necessary to reduce the size of this inverter, and this reduces the load driving capacity and speed. A solution is to use an intermediate super buffer stage between the DCFL stage and the output stage, as shown in Figure 2.26. This allows a smaller inverter to charge the intermediate stage as a super buffer, and then the large output stage is driven by the intermediate stage. This allows a reduction in power dissipation, although small increases in delay will be caused by the extra stage. DSBFL also has a large input capacitance due to the gates of Q1, Q3 and Q5. As with SBFL, the problem of the direct DC path to ground during a positive input transition is also present.

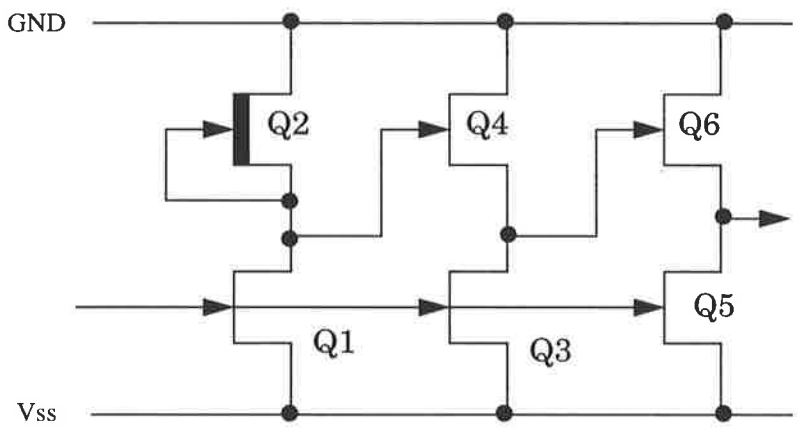


Figure 2.27. DSBFL inverter

2.7.3.7 Other Logic Families

There are many other normally off logic families which have not been used in this thesis, due to their complexity, poor performance or other disadvantages. Such logic families include:

- Capacitively Coupled FET Logic [65]
- Pseudo Current Mode Logic [68]
- Two-phase Dynamic FET Logic [69][70]
- Differential Pass Transistor Logic [71]
- FET FET Logic [72]

2.7.4 Power Supply

The value of power supply chosen is a trade-off between several factors. A high supply voltage will give faster operation and better noise margin, but at the expense of increased power dissipation. If the supply voltage is too large, breakdown effects in the MESFETs and diodes will occur. With these factors in mind, a supply voltage of V_{DD} = 2 volts was used throughout this work.

2.7.5 Summary and discussion

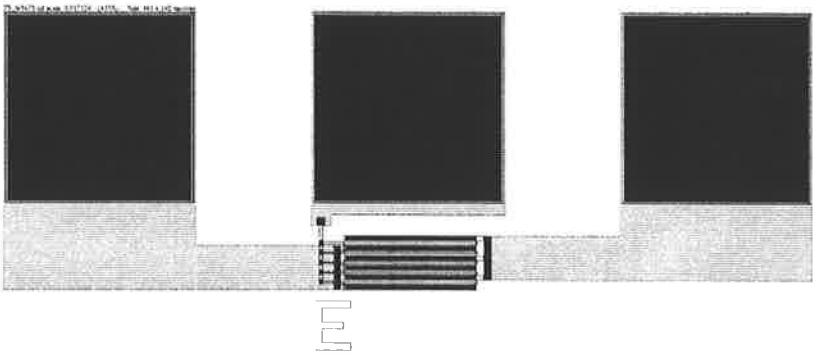
This chapter has provided an introduction to the chemistry and capabilities of gallium arsenide, outlining its advantages over silicon. It is these advantages, particularly the speed advantage, which make it suitable for the implementation of a high speed ATM switch. The modelling and simulation of gallium arsenide circuits was then discussed, followed by an analysis of GaAs logic families suitable for the implementation of the memory. In the following chapter, the testing of some fabricated gallium arsenide transistors is presented.

Chapter 3: MESFET Testing

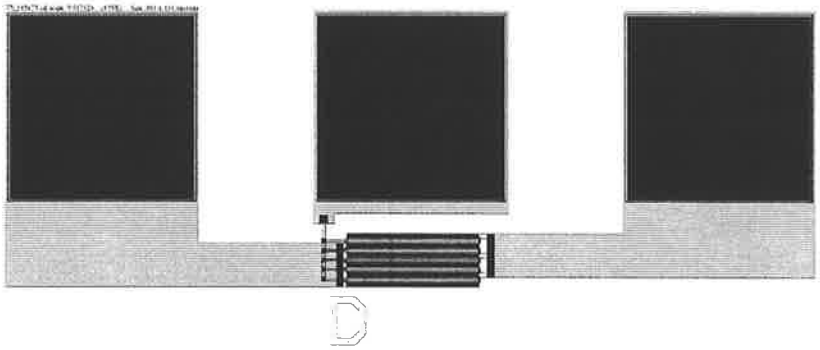
This chapter presents results obtained from the testing of two gallium arsenide MESFET devices fabricated on a test chip. More detailed information can be found in [76].

3.1 Devices Under Test

Two MESFETs, an enhancement mode and a depletion mode, were fabricated in the H-GaAs II process on a chip for testing. Once the transistor operating characteristics were obtained, they were compared with simulation results in an attempt to verify the accuracy of the *hspice* models used in this thesis. The layouts of the enhancement mode and depletion mode MESFETs are shown in Figure 3.1. The MESFETs are layed out as five FETs of width 74.8 μm in parallel in a 'fingered' structure, giving an effective width of 374 μm. This is an exceptionally wide MESFET, made so to afford some protection against current damage by the testing equipment. There are no pads with current protection etc. so as not to distort or influence the transistor characteristics. The MESFET gates, sources and drains are bonded directly to the chip pins.



Enhancement Mode MESFET



Depletion Mode MESFET

Figure 3.1. Layout of fingered MESFETs.

3.2 Test Equipment and Setup

To obtain the transistor characteristics, a transistor curve tracer was used. This curve tracer offered only an analog output displayed on a CRT (similar to an oscilloscope) and thus to obtain numerical data for comparison with simulation results it was necessary to photograph the output and manually read points from the photograph. Not only was this a laborious task, but it is obvious that some errors will have inevitably been introduced at this stage. This process was not helped by the lack of a scale illumination control on the transistor tracer, resulting in scale lines which are very hard to discern in the photographs. The tester has a scan frequency of approximately 1 kHz.

To obtain the characteristics, the drain-source voltage was swept over a 0 - 4 volts range. The gate source voltage was varied over a range suitable for the testing of the particular MESFET: -1.2 volts to 0.4 volts for the depletion mode MESFET (V_{TH} = -0.9 volts nominally) and 0 to 0.7 volts for the enhancement mode MESFET (V_{TH} = 0.225 volts nominally).

Before testing the MESFETs, it was necessary to connect a capacitor between the drain and source terminals of each FET, as shown in Figure 3.2, to prevent oscillations between the drain and source occurring. These oscillations are caused by the parasitic capacitances of the transistor and the inductance of the leads forming a resonating circuit, and the transistor acting as an amplifier. The capacitance has to be significantly larger than the magnitude of the parasitic capacitances, and the value used was 0.1 μF.

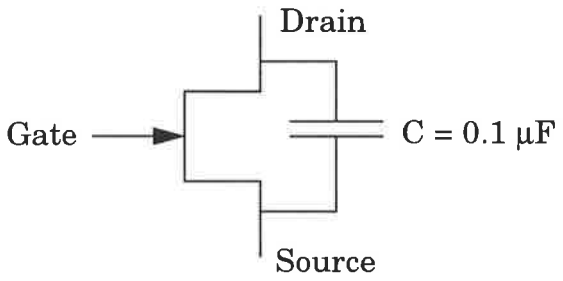


Figure 3.2. Capacitance used for MESFET testing.

3.3 Test Results

3.3.1 Enhancement Mode MESFET

For the enhancement mode MESFET, the transistor tracer was set up to provide six 0.1 volt steps of gate voltage starting at 0.2 volts, i.e. V_{gs} = 0.2, 0.3, 0.4, 0.5, 0.6 and 0.7 volts.

A photograph of the transistor tracer screen is shown in Figure 3.3. It can be seen that the characteristic curves are generally of the shape that we

would expect, however a large amount of hysteresis appears to be present. Similar hysteresis has been noted in [74]. The modulation of the drain voltage over a 0 to 4 volt range is also changing the voltage and charge underneath the substrate and thus the hysteresis is occurring due to back-gating effects as discussed in Section 2.4.3. Because of the high resistivity of the semi-insulating GaAs substrate, the charge has a long transient delay time (1 ms - 1s) [75], which is similar to the scan frequency of the transistor tracer. Hence the hysteresis effects in the curves.

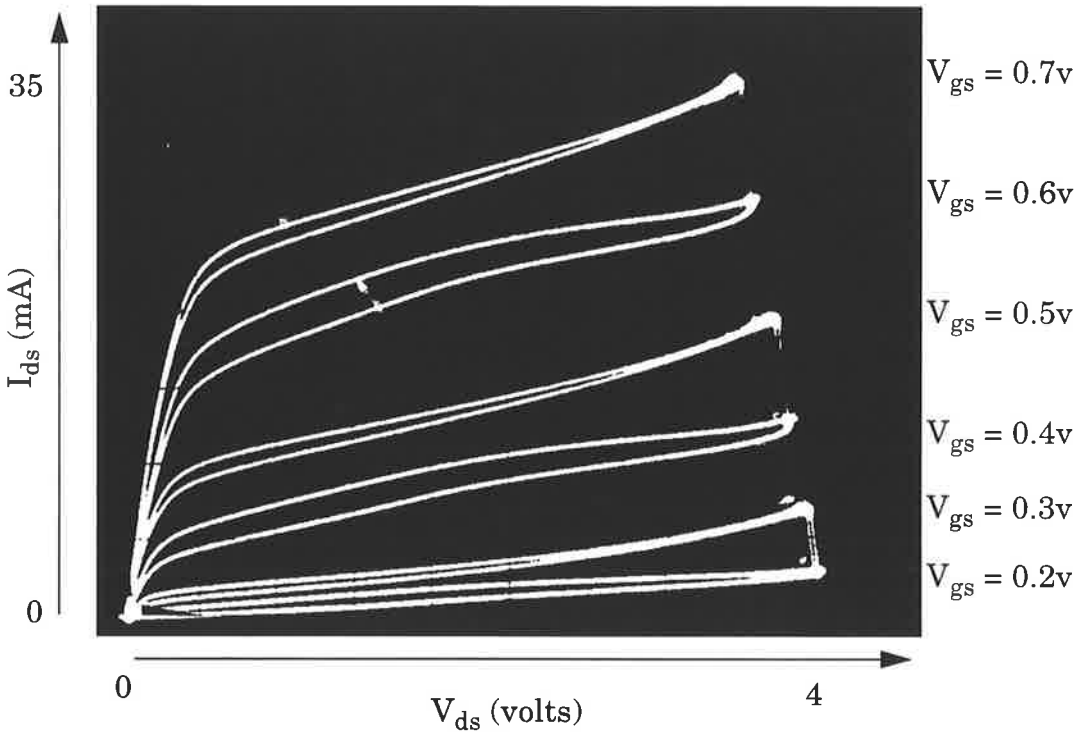


Figure 3.3. Photograph of the E-MESFET characteristics.

The hysteresis is not modelled, so for comparison with simulation, the mid-point of each curve was used as the characteristic.

In finding the appropriate simulation characteristics, two unknowns must be solved: temperature and process variation.

1. The gate junction temperature will be above ambient room temperature due to dissipation within the transistor, but due to the size of the chip and only moderate dissipation occurring within the transistor, the heating effect should not be excessive. The temperature of the gate junction was estimated to be in the region of 30 °C - 60 °C.

2. The degree of process variation experienced by the chip is unknown, and hence a range of simulations over different process variations and the temperature range indicated above were executed, and the best match to the measured characteristics sought. Simulations were carried out at the following values of process variation: 3σ slow, 2s slow, 1σ slow, 0.5σ slow, 0.25σ slow, typical, 0.25σ fast, 0.5σ fast, 1σ fast, 2σ fast and 3σ fast. It was found that the best simulation match occurred at a temperature of 50 °C with typical parameters.

The results of the *hspice* simulation superimposed with the measured E-MESFET characteristics is shown in Figure 3.4. It can be seen that a reasonable agreement was obtained between measurement and simulated parameters, with the maximum error being approximately 12%.

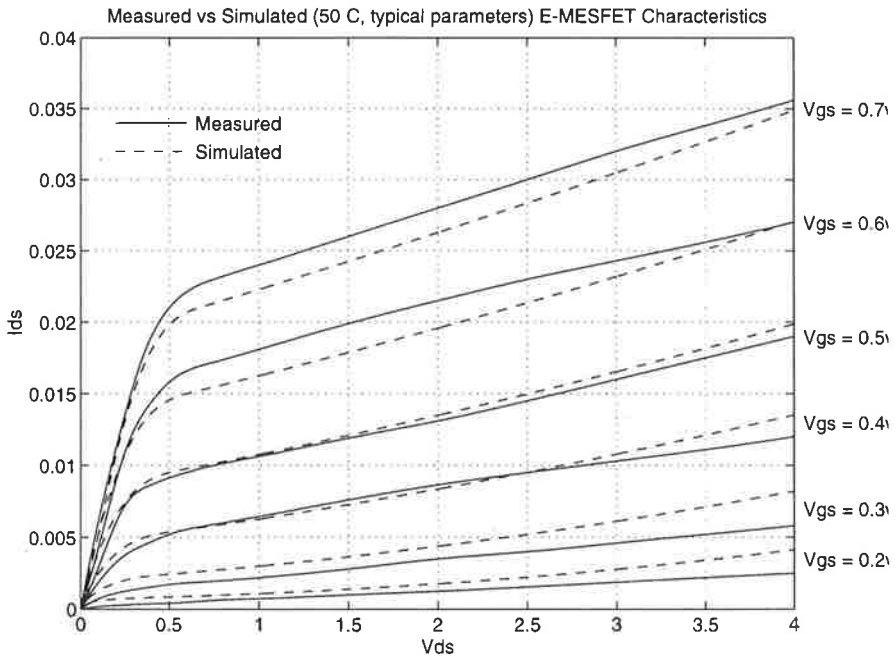


Figure 3.4. Simulated and Measured E-MESFET characteristics.

3.3.2 Depletion Mode MESFET

To test the depletion mode MESFET, the tracer was set up to give 8 steps at 0.2 volts per step, beginning at -1.2 volts. This gives the characteristics for V_{gs} = -1.2, -1.0, -0.8, -0.6, -0.4, -0.2, 0 and 0.2 volts. A photograph of the transistor tracer output for the depletion mode transistor is shown in Figure 3.5. Again, significant hysteresis in the transistor curves is evident.

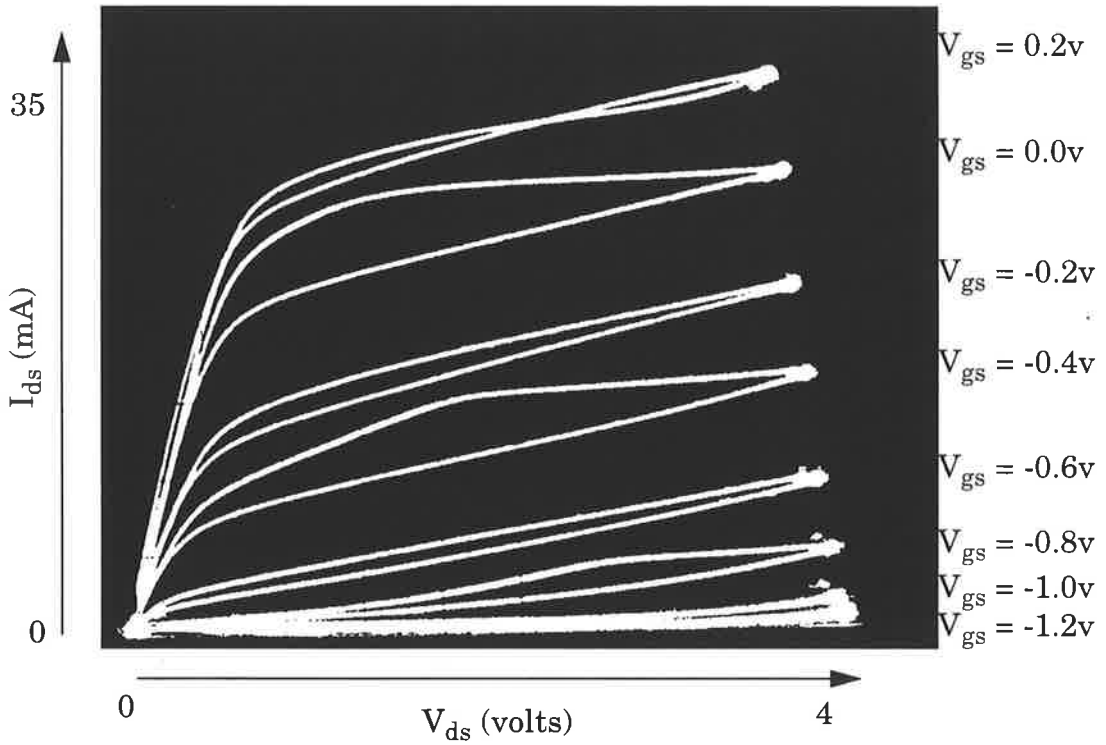


Figure 3.5. Photograph of the D-MESFET characteristics.

Hspice simulations were again done and the best match was found to occur again at a temperature of 50 °C, but using 1σ slow parameters. The results of the hspice simulation superimposed with the measured characteristics obtained from the photograph is shown in Figure 3.6. Once again, a reasonable correlation between measurements and simulation was obtained, with a maximum error of about 10%.

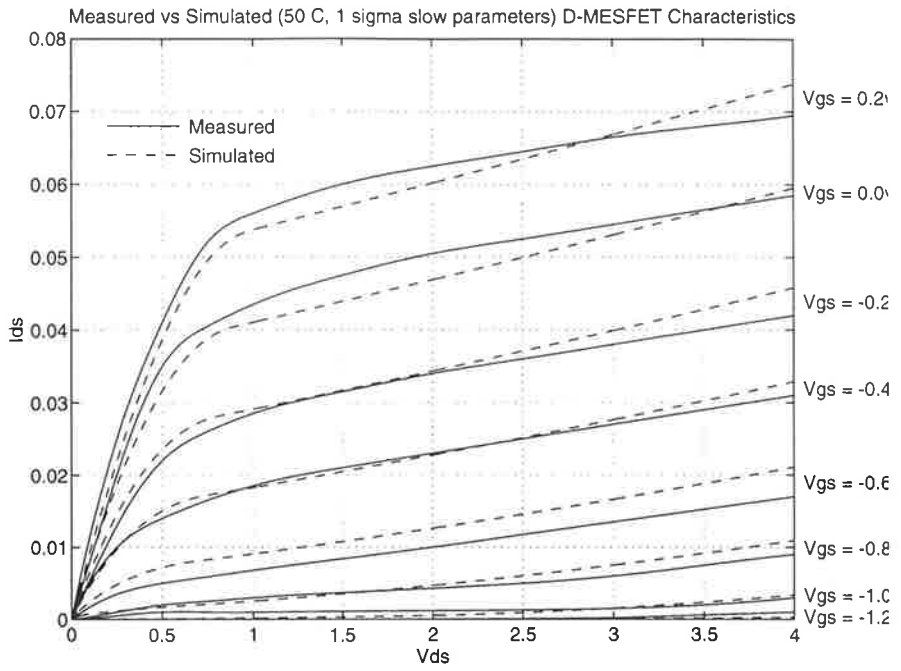


Figure 3.6. Simulated and Measured D-MESFET characteristics.

3.4 Discussion

It can be seen that in both cases, the model provided a reasonable degree of accuracy in approximating the measured transistor characteristics. At low values of gate-source voltage, the model appears to over-estimate the measured characteristics, while at high gate-source voltages the measured characteristics are higher than those simulated. The closest correlation appears at middle values of gate-source voltage.

There are some factors which may account for these discrepancies:

- The model parameters were determined using 10 μm wide transistors. Therefore, some non-linear scaling effects in these very wide transistors may be occurring which would not be accounted for by the model.
- The simulation does not model any geometrical considerations of the of the fingered MESFET structure, and treats the fingered MESFET simply as 5 MESFETs in parallel.

- The models were determined using Vitesse's H-GaAs II foundry. The chips were actually fabricated at Thomsons' (TCS) foundry which has licensed the process from Vitesse. Although the foundries should produce very similar results, there will undoubtedly be some variation in performance between the two. Unfortunately, Thomson's have not characterised their process and supply actual models derived from their foundry.

The results of the tests performed do, however, indicate reasonable validity of the supplied models under a limited range of test conditions.

With both transistors exhibiting different process variation, the importance of simulating designs over a wide range of process variation is evident.

Chapter 4: Memory Cells in Gallium Arsenide

This chapter will first provide a brief overview of basic MOS memory cells, and the performance of each cell when implemented in gallium arsenide. Each memory cell's characteristics are then analysed and the cell which is considered to be best suited to the desired application is determined.

4.1 Memory cells

The operation of all VLSI memory cells relies on the storage of charge at a capacitance, either the gate of a transistor or a capacitor. Because of its highly leaky nature (due mostly to the presence of the schottky barrier diode at the transistor gate), the storage of charge in gallium arsenide is more difficult than silicon. Thus, a direct translation from silicon to GaAs may produce a memory cell that is much less efficient than its silicon counterpart, or worse, one that barely functions at all. The following sub-sections show the evolution of the MOS memory cells from the 6 transistor static cell, as in [77] and give their direct GaAs equivalents. The merits of the cell's GaAs implementation and suitability based on the project requirements detailed in Section 1.5.2.1 will then be discussed. Note that no in-

depth simulation (such as cycle times, etc.) was done on the cells, but simple storage time estimation and performance estimating simulations were executed.

4.2 Six Transistor static cell

A typical nMOS 6 transistor memory cell [77] is shown in Figure 4.1a, using the classic cross-coupled inverter structure, whereby each inverter is constantly maintaining the voltage stored at the others' gate, hence this cell is a Static RAM. This can be converted to an equivalent MESFET structure of cross-coupled DCFL inverters as shown in Figure 4.1b. Depletion type MESFETs are used as the pass transistors (see Section 5.3)

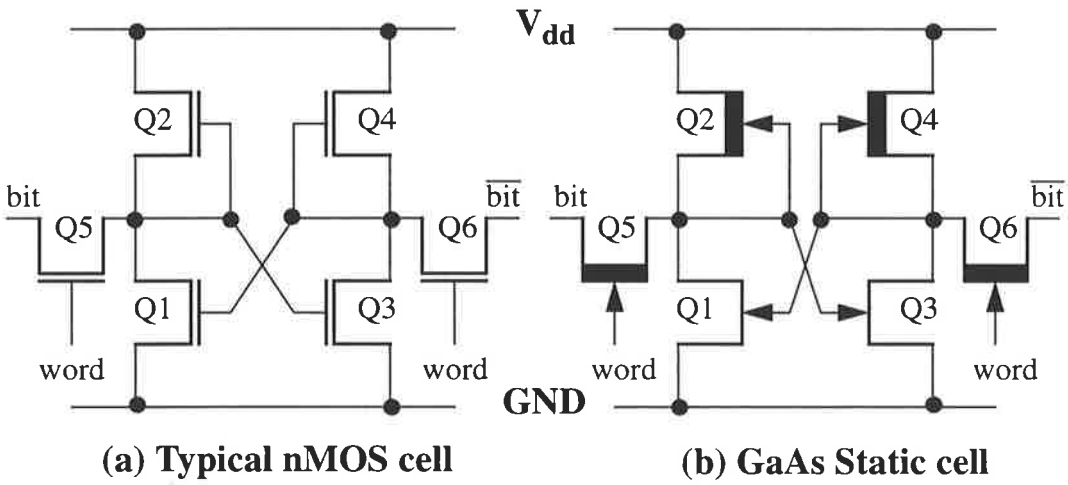


Figure 4.1. Six Transistor static memory cells

To write to the memory cell, the data and $\overline{data}$ are placed on the bit and $\overline{bit}$ lines respectively. The word line is then enabled, transferring charge onto the gates of Q1 and Q3. To read the cell, both bit and $\overline{bit}$ lines are first precharged to a certain voltage, either the high voltage level, midway between the high and low voltage [78], or ground. The word signal is then enabled, and one of the bit lines is then charged and the other discharged relative to

the precharge value. This difference is then fed into a differential amplifier and sensed accordingly. The operation of the cell in GaAs is virtually identical.

There have been many reports of GaAs RAM being implemented in this form [79] - [81].

There are several disadvantages with using the static approach. Firstly, the cell has a large stand-by power because of the continuous dissipation by both inverters being permanently on. At first glance, it may appear that only one inverter is on, but the supply voltage is usually set such that both pull-up D-MESFETs remain in saturation (in the range 1.5v - 2v),and because of the 0.6v schottky diode limited output high, both D-MESFETs are continuously conducting. This results in a very little variation in current drawn from the supply rail which minimises noise, but up to 200μW continuous dissipation for each cell. Secondly, the GaAs cell is typically large in area, because of the large number of transistors required and the need for a suitable pull-up/pull-down ratio in the DCFL inverters. Thirdly, the cell is essentially a one port structure, and can only be written to or read from at a particular time.

It is possible to significantly reduce the stand-by power dissipation of the static GaAs cells by reducing the value of the V_{DD} supply to approximately the value of the schottky diode clamp voltage (0.6v). By limiting the supply voltage to the logic high level, the dissipation in the on inverter is substantially reduced ($P \propto V^2$), while the dissipation in the other inverter becomes virtually negligible. Cell dissipations as low as 5 μW have been reported [83]. However, in such cases the noise margin is also reduced causing problematic operation, particularly as the temperature is increased [80][84].

Due to the static nature of these cells, problems with leakage currents are not as important in determining memory performance as they are in the dynamic cells which are discussed in detail in future sections.

4.3 Six Transistor dynamic cell

In order to reduce the power dissipation, (remembering that NMOS also has static power dissipation), the permanently-on inverter pull-up loads of Figure 4.1 can be replaced with switched loads that can be dynamically turned on and off as shown in Figure 4.2.

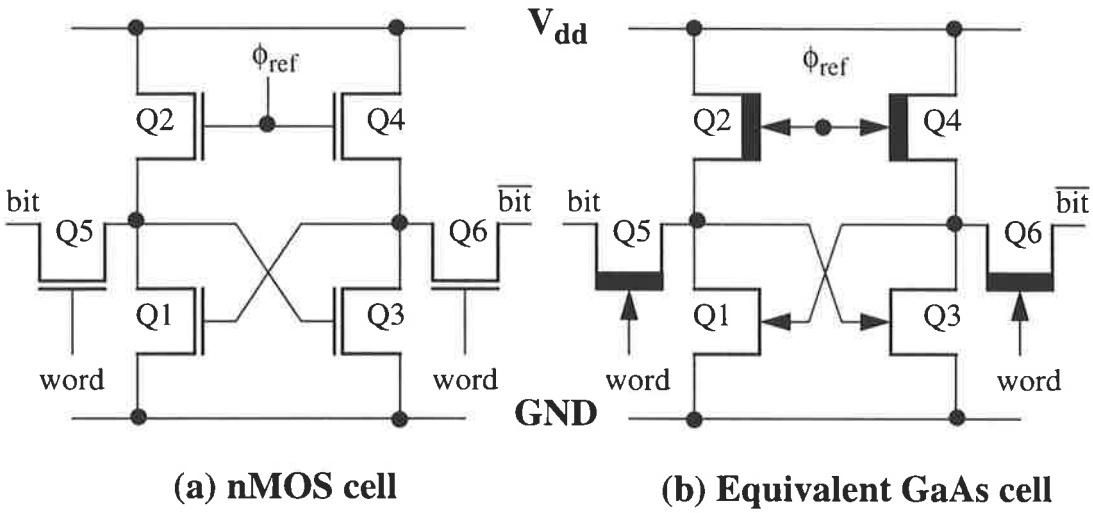


Figure 4.2. Six Transistor dynamic memory cells.

Before the advent of more advanced dynamic cells, the MOS form of this cell was used quite widely [78] [85]. The operation of the cell is similar to that of the six transistor static cell, except that instead of constant static inverters constantly refreshing each gate charge, the inverters are turned on by a refresh clock, ϕ_{ref}, only when necessary to prevent storage degradation by refreshing the charge stored on the gates of Q1 and Q3. Because of the low leakage levels in MOS, refreshing needs to be done only in the order of milliseconds. The advantage of such a cell is a very simple refresh strategy, in that instead of having to read the cell and then write it back, a single clock signal can be used.

A GaAs implementation of the circuit, however, is not as effective due to leakage factors which are orders of magnitude higher than those present in MOS circuits, via the forward biassed gate-source schottky diodes and sub-

threshold currents present in transistors Q1 and Q3. The large subthreshold leakage is exaggerated by the fact that the low level DCFL voltage is approximately 100mV, and that even though this is substantially below the E-MESFET threshold voltage, a large subthreshold current is still present to discharge the node storing the high level. Simulations indicate that this current is in fact larger than the schottky diode leakage current. Figure 4.2 shows the dominant leakage modes present in such a cell.

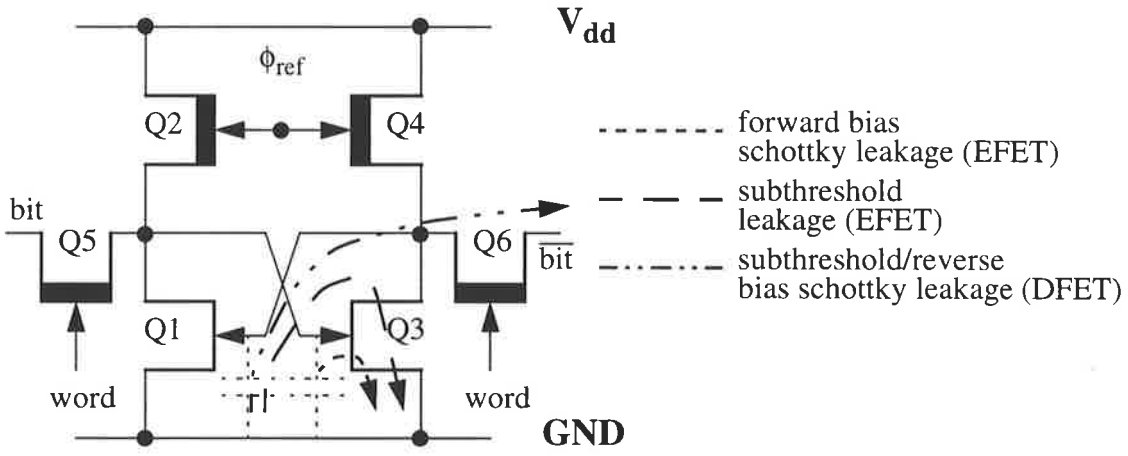


Figure 4.3. Leakage modes in 6 transistor dynamic GaAs DRAM cell.

It can be seen that three forms of leakage are present:

1. Forward biassed schottky diode leakage via gate-source diode of Q1 and Q3.

2. Subthreshold leakage via drain-source channel of Q1 and Q3 (Typically, $v_{gs} - v_{TH} \approx -100$ mV)

3. Subthreshold and reverse-biassed schottky diode leakage via pass transistors Q5 and Q6 (Typically, $v_{gs} - v_{TH} \approx -1$ V)

As the leakage currents are always draining charge from the gates of Q1 and Q3, the limit affecting storage time will be determined by the decay rate of charge stored on the logic high gate. Leakage mode 2 is dominant over 1, because of the fact that the forward biassed schottky diode leakage

is strongly dependant on gate source voltage. i.e. as the high gate is discharge via modes 1 and 2, the schottky diode causing leakage mode 1 is forward biassed less, resulting in less leakage via mode 1. Leakage mode 3 is negligible because of the small (in comparison) magnitudes of both the reverse-biassed schottky diode leakage and subthreshold current when the depletion mode transistor is turned off by a large negative voltage.

Simulations show that the storage node is discharged in a time of the order of 1 ns. This time can be increased somewhat by reducing leakage mode 1. One method is to use pass transistors to isolate the gates from the drains using pass transistors, and turning them on only during refresh. This will increase storage time to around 10ns, but requires more complex control, and the noise margin is reduced considerably because of the voltage drop across the pass transistors. Future sections will show that far better performance can be obtained from a GaAs DRAM, so this form is not used.

4.4 Four Transistor dynamic cell

The next step in the development of DRAM was the removal of the two load transistors to leave the 4 transistor DRAM cell [85], as shown in Figure 4.4.

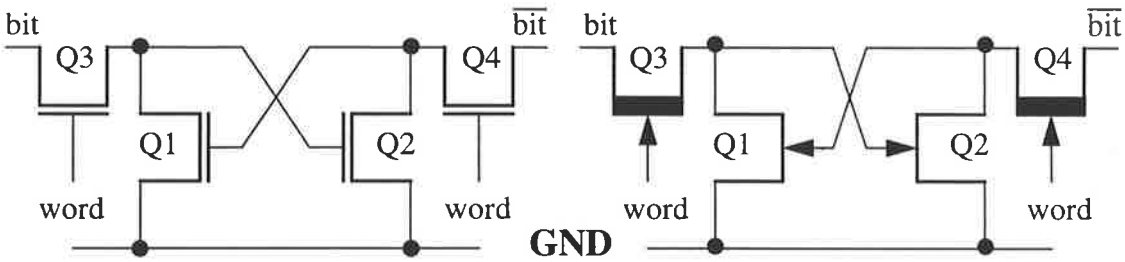


Figure 4.4. Four Transistor dynamic memory cells.

The operation of this cell is similar to the Six Transistor dynamic cell, except this cell has no internal refresh capability. A refresh must therefore be done by reading and then writing back the contents of the cell. The

advantage is that the cell is smaller in area. The equivalent GaAs cell, shown in Figure 4.4b, is subject to the same leakage modes discussed in the previous section and so once again, only a very short storage time can be achieved.

4.5 Three Transistor dynamic cell

The feedback path of the 4 transistor cell can be removed to leave a 3 transistor cell [86] [87] shown in Figure 4.6. There is now a single storage node at the gate of Q1. This introduces a number of benefits in cell operation.

The enhancement mode MESFET subthreshold leakage current (the most prominent form of leakage in the above sections) has been eliminated from the remaining storage node, leaving forward biassed schottky gate leakage and subthreshold and reverse-biassed schottky leakage via the pass transistors.

Because of the removal of the feedback system, only a single transistor is required for the write process. The second transistor may be used to facili-

tate reading of the cell, making the cell a dual port memory with separate read and write bit lines and read and write word enable lines.

Writing of the cell is accomplished by placing the data to be written on the write bit line and asserting write enable. Reading of the cell is done by pre-charging the $\overline{\text{read bit}}$ line high and then asserting the read enable line. If the data stored on the gate of Q1 was high, the $\overline{\text{read bit}}$ line will be discharged. Otherwise it will remain high. Hence the cell is inverting.

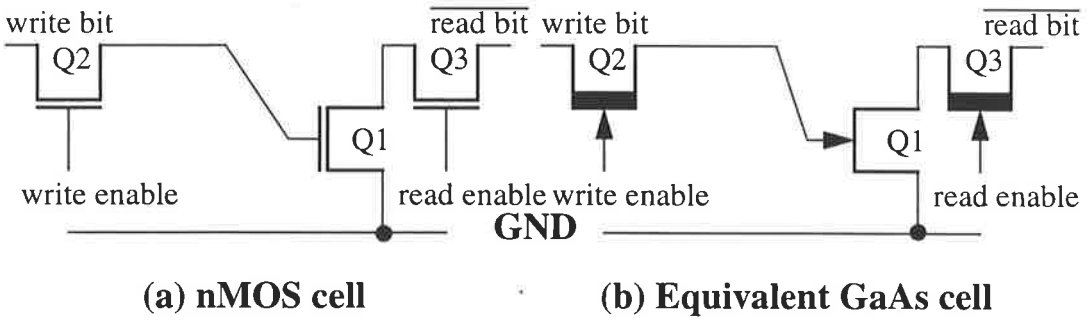


Figure 4.6. Three Transistor memory cells.

Again, however, storage time is limited to the order of nanoseconds because of the high forward biassed schottky diode leakage from the gate of Q1. To remove the forward biassed diode leakage, the structure of the cell can be inverted [88], resulting in the cell shown in Figure 4.6. Because the lack of feedback transistor has removed subthreshold leakage from the gate of Q1, the success of the inverting operation on this cell is much greater than that of the 4 transistor memory. The only leakage affecting the cell charge storage is via reverse-biassed schottky gate diodes on Q1 and Q2 and small subthreshold leakage via Q2's drain-source. Because reverse biassed schottky leakage is several orders of magnitude lower then forward biassed leakage, a corresponding increase in storage time can be expected. The leakage is also arranged such that some is acting to charge the storage node and some is acting to discharge the storage node, with both being of similar magnitude, further reducing the *net* leakage to/from the storage

node. The leakage present in the cell is explained and analysed in detail in Chapter 5.

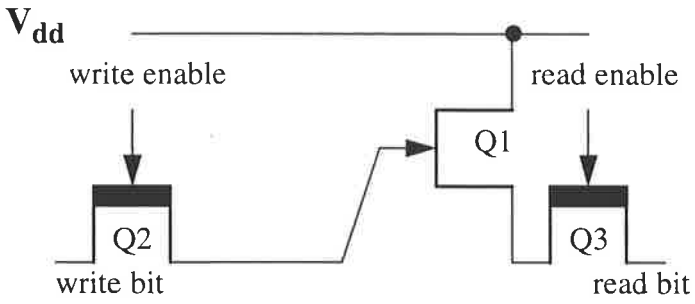


Figure 4.7. Inverted 3 transistor GaAs memory cell.

To write to the memory cell, transistor Q2 is turned on via write enable so the data at Write Bit can be transferred to the storage node on the gate of Q1. If Write Bit is high, charge will be transferred onto Q1's gate capacitance, storing a logic high, while if Write Bit is low, any charge on the gate of Q1 will be discharged through Q2, resulting in a logic low being stored in the memory cell.

Reading the memory cell is controlled by pass transistor Q3. Before reading, the Read Bit bus is precharged to logic low, as there is no pull-down facility within the memory cell. Read Enable is then set high to perform the read operation. If there is a logic low stored, transistor Q1 will be off and the Read Bit bus will remain low. If a logic high has been stored, Q1 will be turned on and the Read Bit bus will be charged via Q1 and Q3. When a logic high is read, the gate source diode of Q1 is forward biassed, resulting in significant forward conduction from the storage node. This discharges the storage node. Although the node will not be completely discharged, data integrity cannot be guaranteed after the first read. The cell thus has a *destructive* read cycle.

4.6 Single transistor dynamic cell

The final step in DRAM evolution was the removal of the read and storage transistors and the use of a capacitor as the charge storage element. The single transistor dynamic RAM cell [89] is the result, shown in Figure 4.8. This cell is the most commonly used MOS DRAM cell, because of its small area and low transistor count, and therefore high density [90]- [92]. There has also been a large amount of research into improving cell structure and capacitance formation [93], with fabrication processes being specifically designed for the manufacture of DRAMs only. The cell also has the disadvantages of a fully destructive read and needs very sensitive sense amplifiers because of the small storage capacitance/bit-line capacitance ratio. The cell is also a single port structure.

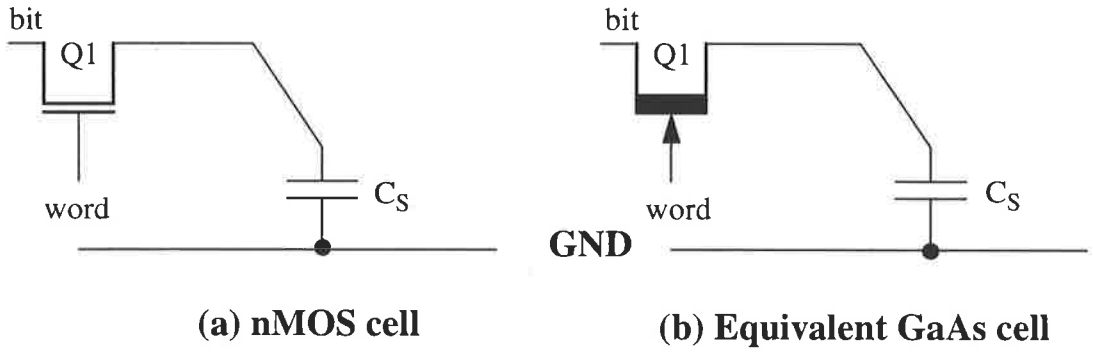
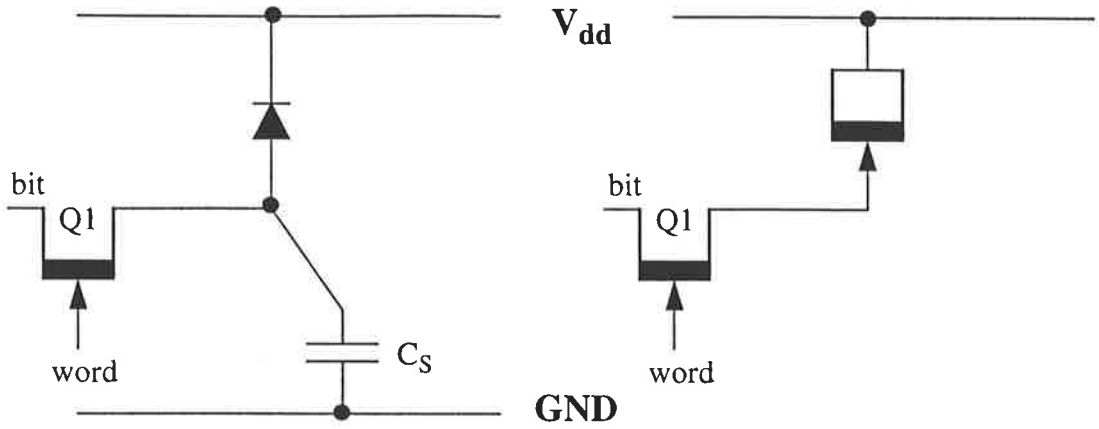


Figure 4.8. Single Transistor memory cells.

This GaAs implementation of this cell suffers from charge degradation via subthreshold and reverse-bias schottky leakage in transistor Q1. One possible way to increase storage time is to add an extra transistor to provide a facility to add charge into the storage capacitor at approximately the same rate at which it is being taken out. As discussed in Section 2.4.2.1, the reverse-bias schottky leakage dominates subthreshold leakage, and so a possible solution is shown in Figure 4.9a. Depending on storage time constraints, the cell can be reduced to a 2 transistor configuration using a MESFET gate capacitance for storage as shown in Figure 4.9b.



(a) Leakage compensated GaAs Single Transistor cell **(b) Compensated cell using a Gate as a storage element**

Figure 4.9. GaAs Single Transistor DRAM cells.

An implementation of the GaAs cell in Figure 4.8b has been reported in [94] and shows storage times in the order of 2.5 ms at room temperature, although the process used was optimised for cell performance. Simulations show that the cells shown in Figure 4.9 have superior storage performance over that in Figure 4.8b, but due to limitations discussed in Section 2.2.1, the accuracy of these simulations may be limited.

4.7 Discussion

RAM cell designs suitable for a GaAs buffer have been discussed. Due to the nature of the buffer, a small high density, low power, dual port dynamic RAM cell is ideal. The three transistor cell, discussed in Section 4.5, was considered to have the best characteristics in these respects and hence was chosen, analysed and implemented. This is described in the following Chapters of the thesis.

The single transistor cell is not suitable for this buffer due to it's single port nature.

The static RAM cell is also unsuitable for this application due to its high power consumption and the large area required for each cell.

Chapter 5: Three Transistor Dynamic Memory Cell Optimisation

The operation of the three transistor memory cell chosen in the previous chapter is analysed in detail. From this analysis the leakage modes are determined, and the total leakage minimised so that the storage time of the cell is maximised.

5.1 Pass Transistors

The Three Transistor dynamic memory cell schematic is shown in Figure 5.1. The reading and writing of the cell is controlled using pass transistors. Obviously either enhancement or depletion mode MESFETs may be used for implementing these. The pass transistor should be chosen so that the maximum voltage is passed from drain to source when the transistor is turned on. For the purposes of this analysis the voltage swing at the source is defined as:

$$V_{s_{min}} \leq V_s \leq V_{s_{max}} \qquad \text{(EQ 5.1)}$$

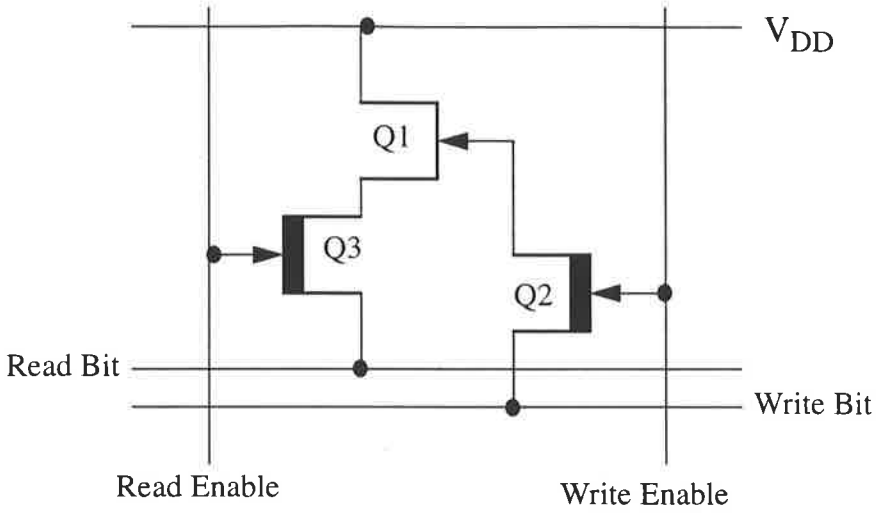


Figure 5.1. Three Transistor GaAs dynamic memory cell.

The condition for a transistor to be in its on state can be expressed as:

$$V_{gs} = V_g - V_s \geq V_{TH} \qquad \text{(EQ 5.2)}$$

Therefore, when the transistor is on, the source voltage is:

$$\begin{aligned} V_s &\leq V_g - V_{TH} \\ \therefore V_{s_{max}} &= V_{g_{max}} - V_{TH} \end{aligned} \qquad \text{(EQ 5.3)}$$

But, the forward biassed schottky diode at the gate of the transistor limits the gate source voltage to the schottky clamp voltage, V_{SC}:

$$V_{gs} = V_g - V_s \leq V_{SC} \qquad \text{(EQ 5.4)}$$

Therefore the maximum voltage at the gate to turn the transistor on is:

$$V_g \leq V_s + V_{SC} \qquad \text{(EQ 5.5)}$$

The maximum gate voltage that can be applied will therefore be higher for V_{Smax} than for V_{Smin}, so we must set V_S to V_{Smin} so that Equation 5.5 is

valid for all V_S. Therefore, the maximum gate voltage which can be applied to turn the transistor on is:

$$V_{g_{max}} = V_{s_{min}} + V_{SC} \quad \text{(EQ 5.6)}$$

Substituting Equation 5.6 into Equation 5.3 gives:

$$V_{s_{max}} = V_{g_{max}} - V_t = V_{S_{min}} + V_{SC} - V_t \quad \text{(EQ 5.7)}$$

Rearranging, the maximum available voltage swing at the source of the pass transistor is:

$$V_{s_{max}} - V_{s_{min}} = V_{SC} - V_t \quad \text{(EQ 5.8)}$$

Therefore the maximum voltage swing at the source can be increased by increasing either the schottky diode clamp voltage or decreasing the transistor threshold voltage. However, because the schottky diode clamp voltage is fixed (at approximately 0.6 volts), to maximise the source voltage swing it is necessary to minimise the transistor threshold voltage.

The typical D-MESFET has a threshold voltage of about -0.871 volts, compared with the typical E-MESFET threshold of 0.227 volts [26]. Using D-MESFETs as pass transistors thus offers approximately 1.1 volts of extra voltage swing. For this reason, depletion mode MESFETs are superior pass transistors. The major disadvantage in using depletion mode pass transistors is that a negative voltage must be used to turn them off, increasing design complexity.

5.2 Voltage Levels in the DRAM cell

5.2.1 Write Bit Line Voltages

To effectively store a low level, Q1 must be off when the cell is being read (i.e. Q3 is on) and as the Read Bit bus is pre-charged to 0 volts, the source

voltage of Q2 will also be approximately 0 volts. To turn off a transistor, we require:

$$V_{gs} < V_{TH} \quad \text{(EQ 5.9)}$$

We therefore require a logic low stored on the gate of Q1 to be less than the threshold voltage of an enhancement mode MESFET, nominally 0.227 volts.

In the case of successfully storing a high level, when the cell is read the Read Bit bus must be charged to a logic high level, typically 0.6 volts. For transistor Q1 to be on in this case, we see from Equation 5.2 that the gate voltage must be at least a threshold voltage higher than this high level, and the further above this value it is the more transistor Q1 will be turned on. As we want to maximise the rate of charge of the Read Bit bus, the logic high level on the gate should be maximised.

From the above arguments, 0 volts for a low level and 2 volts for a high level will be ideal. However, due to logic family limitations in the design of the Write Bit driver bus (Section 6.2.6), the maximum logic high voltage available is approximately 1.5 volts.

5.2.2 Pass Transistor Turn Off Voltages

In turning off the pass transistors, we seek to minimise the currents flowing in them, thereby reducing leakage currents. As discussed in Section 2.4.2.1, while the accuracy of the subthreshold model is not known, provided that the gate-source voltage is several hundred millivolts below the threshold voltage, reverse-biassed schottky leakage currents will be dominant. Therefore to minimise the effect of the subthreshold leakage currents in the pass transistors, we need only ensure that the gate-source voltage of the transistors in their off state is sufficiently low for this to occur.

However, as threshold voltage varies with temperature and process variation, the point at which the subthreshold current becomes negligible will

also change. The worst case will occur when the transistor threshold voltage is at its lowest (i.e. the transistor will be fastest). In this case, the gate-source voltage required to minimise the subthreshold current will be at its maximum negative value. Care must be taken to ensure that the turn off voltage is not too low, as this will result in increases in reverse bias schottky leakage, and eventually failure if the breakdown voltage is reached, as discussed in Section 2.4.1. Nominal minimum size depletion mode MESFET threshold voltages are shown in Table 5.1.

Table 5.1. Nominal minimum size D-MESFET threshold voltages for different process variations (25 °C).

D-MESFET	2σ Slow	1σ Slow	Typical	1σ Fast	2σ Fast
V_{TH} *(volts)*	-0.677	-0.774	-0.87	-0.969	-1.067

As temperature increases, threshold voltage is reduced (Section 2.4.4) and therefore at 125 °C, threshold voltages will be significantly lower than those shown in Table 5.1. By this reasoning, a turn off voltage of -2 volts was decided upon. This value was verified using simulation.

Figure 5.2 shows the result of an *hspice* simulation of the subthreshold currents in a minimum sized D-MESFET. The graph is a plot of drain-source current at a drain source voltage of 2 volts and temperature of 75 °C versus gate-source voltage. It can be clearly seen that the gate-source voltage required to cut off the subthreshold currents becomes more negative as the transistor threshold voltage reduces (i.e. becomes *more* negative), and that a cut-off voltage of -2 volts is adequate.

Table 5.2 presents a list of required turn-off voltages for a range of process variations and temperatures. The increase of cut off voltage with temperature can be clearly seen. It can be seen that using a -2 volt cut off voltage will provide enough swing to turn the transistors very close to fully off over the required operating range.

5.2.3 Pass Transistor Turn On Voltages

The higher the turn on levels for the pass transistors are set, the greater the voltage swing that will be passed across the pass transistors

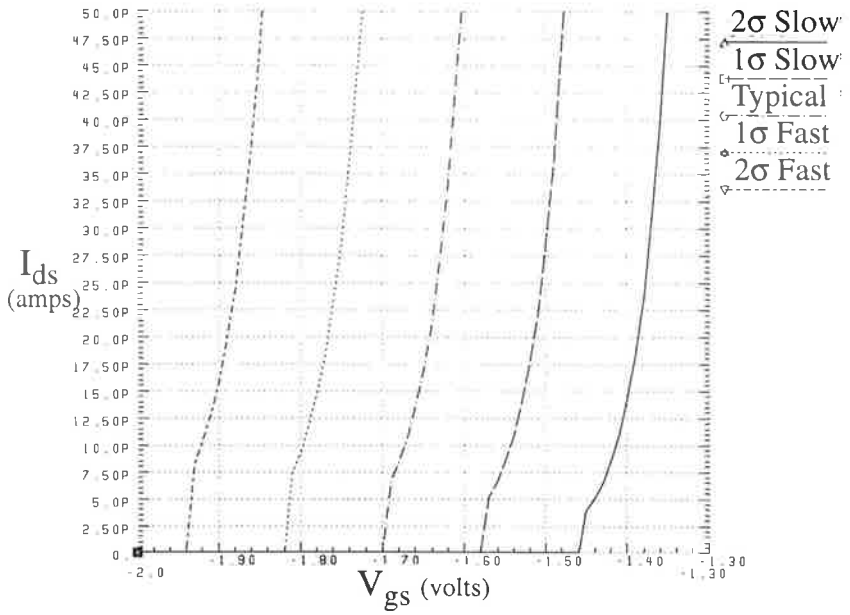


Figure 5.2. Subthreshold currents in a minimum sized D-MESFET.

Table 5.2. Voltages required to turn off D-MESFET subthreshold currents for various temperatures and process variations

Temperature	2σ Slow	1σ Slow	Typical	1σ Fast	2σ Fast
-25 °C	-1.20	-1.32	-1.44	-1.55	-1.66
75 °C	-1.46	-1.58	-1.70	-1.82	-1.94
125 °C	-1.58	-1.71	-1.84	-1.96	-2.08

(Equation 5.8). It is therefore advantageous to make the levels as high as possible. However if the level is set too high, forward biassed schottky leakage into the storage node or Read Bit line may degrade operation.

To turn on the pass transistors, a voltage greater than their threshold voltage (nominally -0.87 volts) is required. Without introducing another power supply voltage into the design, two possibilities are evident: a 0 volt turn on signal can be generated by using GND as the positive rail, or by using V_{DD} as the positive rail, a higher turn on signal (which will be schottky diode limited at the gates of the pass transistors to approximately 0.6 volts) can be produced.

At a gate-source voltage of 0 volts they will be turned on, and because the gate-source voltage is 0 volts, there will be no forward biassed schottky leakage. At the higher turn on voltage, a higher voltage swing will occur, but significant forward conduction of current through the schottky diode may effect the logic low level in particular.

5.2.3.1 Write Enable Transistor On Voltage.

When the Write Enable on signal is set to 0 volts, satisfactory operation occurs, and the maximum logic high value which is stored is approximately 1 volt. When a logic low is stored, no forward conduction occurs and so a logic low value of 0 volts occurs.

A higher value of stored logic high voltage will result in a better read from the cell as Q1 will be turned on more. To increase the value of stored logic high voltage, the Write Enable on signal can be increased in voltage. If the output of the Write Enable signal driver is instead tied to the high voltage rail, the Enable signal on voltage will be increased.

When storing a logic low value, the Write Enable signal will be constrained to 0.6 volts by the gate-source schottky diode of the Write Enable transistor, Q2. This voltage will result in significant forward conduction through into the storage node. Initially, it was thought this process would be damaging to the cell operation by storing a higher value of logic low. However, this is not the case as illustrated in Figure 5.3. Assume that when storing a logic low, 0 volts is present on the Write Bit line. When the Write Enable signal is applied, it increases to the point at which schottky clamping will occur - approximately 0.6 volts. Current then flows into the storage node via the gate, increasing the voltage at the storage node above 0 volts. Eventually, the Write Enable signal is turned off, and the voltage level of the signal falls. As it falls towards 0 volts, the transistor is still on (as the depletion mode threshold voltage is $\approx$ -0.8 volts), but the forward conduction ceases. Further, because the Storage node (> 0 volts) is now significantly higher than the Write Bit line ($\approx$ 0 volts), the Storage node is discharged through the Write transistor back to 0 volts.

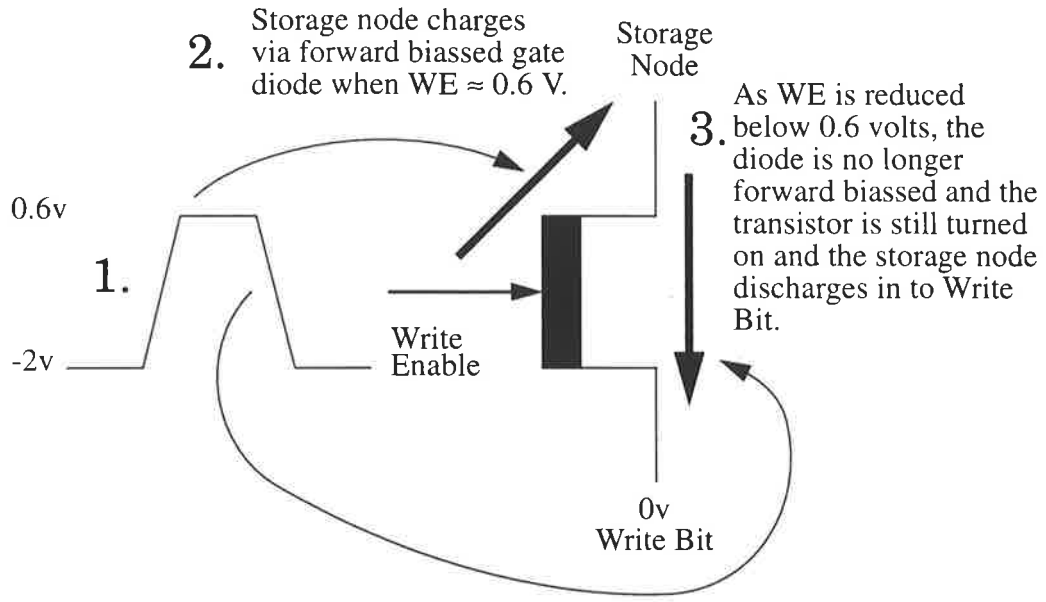


Figure 5.3. Why a high WE voltage does not increase low stored voltage.

Therefore, a Write Enable voltage high level of 0.6 volts is used as it maximises the high level stored in the memory, but does not increase the low level stored. Using this level, simulations will show a logic high value of approximately 1.6 volts on the storage node.

5.2.3.2 Read Enable Transistor On Voltage Level

The Read Enable level can only be set to a maximum of 0 volts. Increasing this voltage results in significant forward conduction into the Read Bit node. This problem is exacerbated by the very high gain sense amplifier which must be used to extract the low output high voltage from the cell. Simulations indicate that a logic high when reading can be as low as 400 mV, while using a 0.6 volt Read Enable signal can result in an output of a logic low as high as 300 mV. It can be seen that this situation is unacceptable due to the noise margin problems it produces. Therefore the Read Enable level is set to a maximum of 0 volts.

5.2.4 Summary

The control voltage levels used in the memory cell are summarised in the Table 5.3:

Table 5.3. Control Voltages for optimum memory cell performance.

Signal	Low level	High level
Read Enable	-2 volts	0 volts
Write Enable	-2 volts	0.6 volts
Write Bit	0 volts	2 volts

The shaded cells in the above table represent the voltages which are applied to the memory cell under steady state conditions, i.e. when no reading or writing is occurring, and a bit is being stored in the memory cell.

5.3 Leakage Model

To maximise the storage time of the memory, a full understanding of the ways in which current is leaking into/out of the storage node must be obtained. The MESFET equivalent circuit model, Figure 2.10, can be substituted into the Dynamic RAM cell (Figure 5.1), to obtain the equivalent circuit of the memory cell. The resulting circuit is shown in Figure 5.4.

The currents which act on the storage and internal nodes are shown in Figure 5.4, and defined as:

i_{gde} - Current flowing from V_{DD} to the *Storage* node via the E-MESFET (Q1) gate-drain schottky diode.

i_{dse} - Current flowing from V_{DD} to the *Internal* node via the E-MESFET (Q1) drain-source current source.

i_{gse} - Current flowing from the *Storage* node to the *Internal* node via the E-MESFET (Q1) gate-source schottky diode.

i_{gdww} - Current flowing from the *Storage* node to *Write Enable* via the Write Enable D-MESFET (Q2) gate-drain schottky diode.

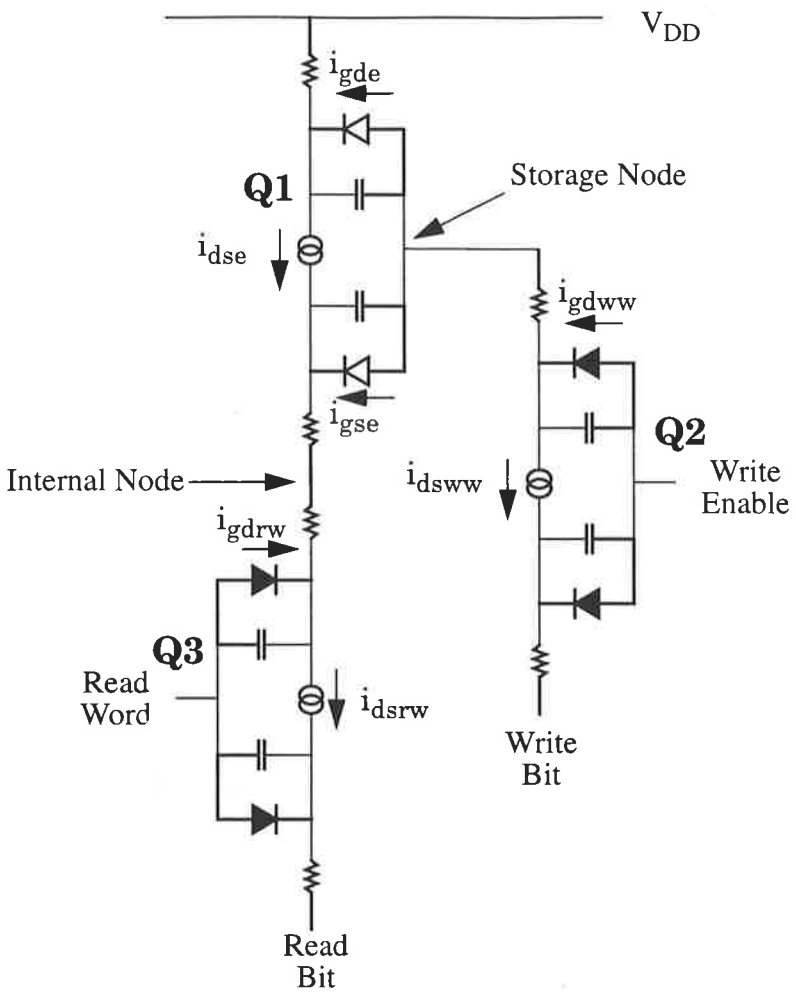


Figure 5.4. Memory Cell equivalent circuit.

i_{dsww} - Current flowing from the *Storage* node to *Write Bit* via the Write Enable D-MESFET (Q2) drain-source current source.

i_{gdrw} - Current flowing from the *Internal* node to *Read Enable* via the Read Enable D-MESFET (Q3) gate-drain schottky diode.

i_{dsrw} - Current flowing from the *Internal* node to *Read Bit* via the drain-source current source of the Read Enable D-MESFET (Q3).

To simplify analysis of the leakage currents acting on the storage node, it is assumed that the memory cell is in steady state. i.e. that no reading or writing is occurring in the cell. The Write (Q2) and Read (Q3) Word tran-

sistors are therefore turned off by -2 volts as discussed in Section 5.2.2. We can therefore assume that their subthreshold currents are negligible compared to the reverse-bias schottky leakage and remove them from the circuit. By assuming the resistances are negligibly small compared to the off resistance of the transistors, and ignoring the effect of the capacitors (steady state), the leakage model can be somewhat simplified to that shown in Figure 5.5.

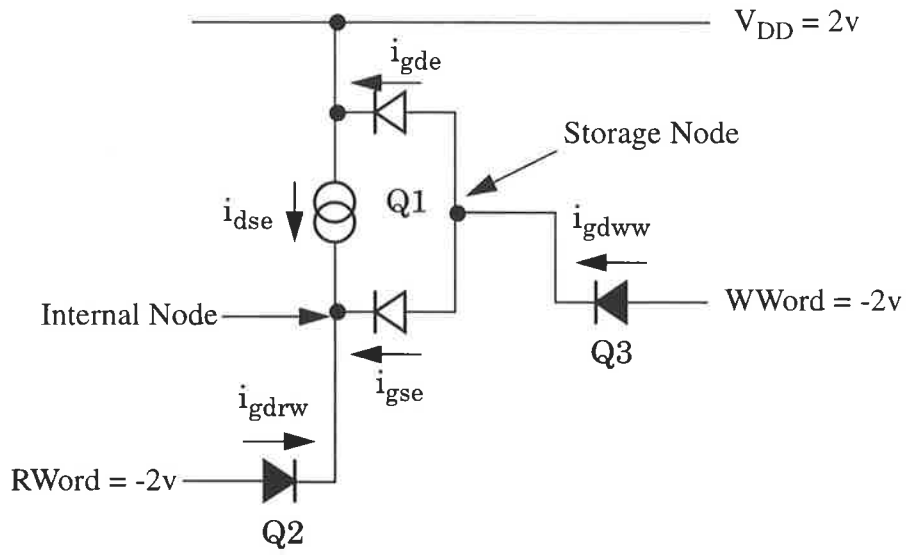


Figure 5.5. Simplified Leakage Model of Dynamic RAM Cell.

By assuming that the voltages at the storage and internal nodes are somewhere between +2 volts and -2 volts, we can summarise the leakage effects in Figure 5.5 as:

- Reverse bias schottky leakage, i_{gde}, will charge the storage node from V_{DD}.
- Reverse bias schottky leakage, i_{gdww}, will drain current from the storage node into the Write Word line.
- There will be some current flow, i_{gse}, either charging or discharging the storage node via the internal node, depending on the relative voltages of those two nodes.

The relative voltage between the storage node and the internal node must therefore be found. Summing the currents at the internal node gives:

$$i_{dse} + i_{gse} + i_{gdrw} = 0 \qquad \text{(EQ 5.10)}$$

Rearranging gives:

$$i_{dse} = -i_{gse} - i_{gdrw} \qquad \text{(EQ 5.11)}$$

We know that the drain current in Q1, i_{dse}, cannot be negative as current will always flow out of the supply, V_{DD}, and into the circuit. Substituting $i_{dse} \geq 0$ into Equation 5.11 gives:

$$-i_{gse} - i_{gdrw} \geq 0 \qquad \text{(EQ 5.12)}$$

Rearranging gives:

$$i_{gse} \leq -i_{gdrw} \qquad \text{(EQ 5.13)}$$

We assumed before that the internal node voltage was higher than -2 volts, and therefore i_{gdrw} will be a reverse bias schottky diode current. Hence, current i_{gse} can be one of three possibilities:

1. If $v_{gse} > 0$ volts, a very small forward bias schottky current (lower in magnitude than the reverse-bias schottky current i_{gdrw}), flowing from the storage node into the internal node.
2. If $v_{gse} = 0$ volts, $i_{gse} = 0$.
3. If $v_{gse} < 0$ volts, a reverse bias schottky current flowing from the internal node to the storage node.

If we assume that case 1 is true, then the magnitude of i_{dse} will be smaller than the magnitude of the reverse-bias schottky current i_{gdrw} (from Equation 5.11). Such a small drain source current requires that the transistor is heavily into the subthreshold region. From [43] and discussions in Section 2.4.2.1, we know that if the drain source current is less than the

reverse-bias schottky leakage, then the gate source voltage must be several hundred millivolts below the threshold voltage, such that the transistor is well into the subthreshold region. We would expect that a typical E-MESFET with a nominal threshold voltage of 0.227 volts would require a gate source voltage of the order of -0.5 volts. This was verified in comprehensive simulation over all operating conditions. Therefore for this small drain-source current we require $v_{gse} < 0$ volts, but the first assumption is that $v_{gse} > 0$ volts. Therefore case 1 cannot be true.

If we assume case 2 is true, then the magnitude of i_{dse} will be equal to the magnitude of reverse-bias schottky current i_{gdrw}. A similar argument to case 1 can be followed, with the result that if $v_{gse} = 0$ volts, the transistor will not be far enough into the subthreshold region to have such a small drain current and that therefore the actual value of v_{gse} must be lower resulting in it being non zero and the initial assumption cannot be possible.

Now, assuming that case 3 is true, then from Equation 5.11 we know that the magnitude of i_{dse} will be the sum of two reverse-bias schottky diode currents, i_{gse} and i_{gsrw}. Once again this is a very low drain-source current which will require the transistor to be heavily into the subthreshold region. An enhancement mode transistor will require a negative gate-source voltage, therefore $v_{gse} < 0$, which is the original assumption we made. Therefore case 3 is the only case which can exist under steady state conditions. Once again this was verified by simulation over the full range of operating conditions. Therefore i_{gse} is a reverse-bias schottky diode current flowing from the internal node into the storage node. i.e. charging the storage node. We can then further simplify the leakage mode to its simplest format, shown in Figure 5.6.

The magnitude of the current i_{gse} charging the Storage node will depend on the value of voltage difference between the storage and internal nodes, v_{gse}. The exact value of this voltage can be found by substituting in to Equation 5.10, as shown below in Equation 5.14.

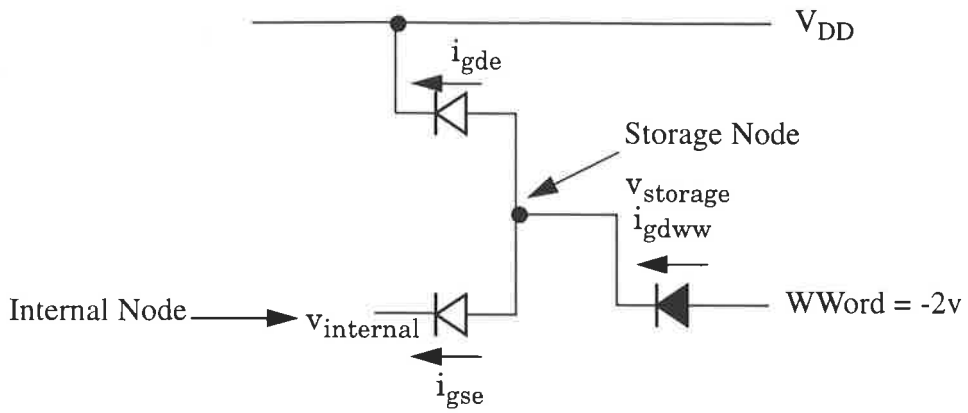


Figure 5.6. Most simplified Leakage Model of the Dynamic RAM Cell

$$\left.i_{dse}\right|_{v_{gs}=v_{Storage}-v_{Internal},\, v_{ds}=V_{dd}-V_{Internal}} + \left.i_{gse}\right|_{v_{gs}=v_{Storage}-v_{Internal}} + \left.i_{gdrw}\right|_{v_{gd}=v_{Internal}-V_{WE}} = 0 \qquad \text{(EQ 5.14)}$$

Solving this equation by hand is extremely complex, however it can be solved by *hspice* simulation, using $v_{Internal}$ as the variable and evaluating the points at which the curve crosses the x-axis. As an example, a simulation of Equation 5.14 for fixed $v_{Storage}$ = 0 volts (storing logic low) using minimum sized transistors at 125 °C and typical parameters is shown in Figure 5.7. It can be seen from this example that where the curves cross the 0 current axis, $v_{Internal}$ ranges from 420 mV to 480 mV at the equilibrium point for different process spreads, and since $v_{Storage}$ is set to 0v, this corresponds to a v_{gse} in the range of -420 mV to -480 mV.

By performing a similar simulation over the whole range of operating conditions and values of voltage at the storage node, it can be shown that v_{gse} is always negative.

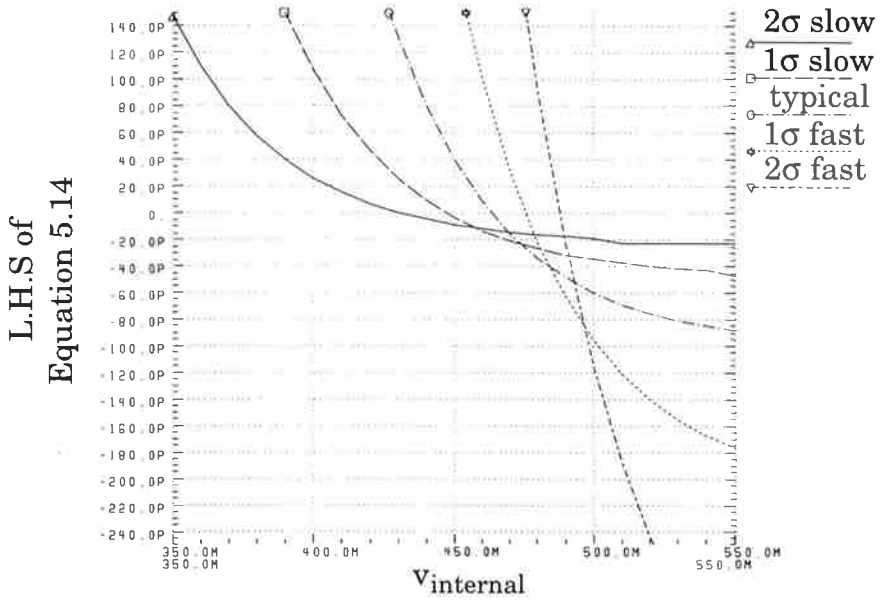


Figure 5.7. Memory equilibrium point at $V_{storage}$ = 0, 125 °C, 2σ variation.

Table 5.4 shows the values of $v_{storage}$ - $v_{internal}$ at equilibrium points over a range of temperature and process variations at $V_{Storage}$ = 0V, that typical of a logic low level being stored. A similar result is obtained as the voltage stored on the storage gate is increased from 0v up to that stored for a logic high (typically around 1.2 volts). Values of $v_{storage}$ - $v_{internal}$ at equilibrium points over a range of temperature and process variations at $V_{Storage}$ = 1.2V, that typical of a logic high level being stored, are shown in Table 5.5.

Further, by comparing the values of the reverse-bias voltage v_{gse} shown in Table 5.4 and Table 5.5 with values of reverse saturation voltage such as those shown in Table 2.1, it is evident that the equilibrium point is always such that i_{gse} will be in, or very close to, reverse saturation.

We can therefore validly assume that all three diodes in Figure 5.6 are in reverse saturation. By determining the total current flowing into/out of the

Table 5.4. Table of $V_{Storage}$ - $V_{Internal}$ under various operational conditions for $V_{Storage}$ = 0V.

Temperature	2σ Slow	1σ Slow	Typical	1σ Fast	2σ Fast
25 °C	-270 mV	-310 mV	-360 mV	-400 mV	-439 mV
75 °C	-387 mV	-425 mV	-470 mV	-485 mV	-500 mV
125 °C	-430 mV	-447 mV	-463 mV	-476 mV	-488 mV

Table 5.5. Table of $V_{Storage}$ - $V_{Internal}$ under various operational conditions for $V_{Storage}$ = 1.2V.

Temperature	2σ Slow	1σ Slow	Typical	1σ Fast	2σ Fast
25 °C	-140 mV	-190 mV	-230 mV	-269mV	-309 mV
75 °C	-256 mV	-298 mV	-312 mV	-324 mV	-345 mV
125 °C	-269 mV	-288 mV	-307 mV	-322 mV	-336 mV

Storage node, an estimate of the cell storage time can be made. The transistor sizes can then be optimised to minimise the current and hence maximise the storage time.

5.4 Leakage Minimisation

To simplify analysis, a re-labelled version of the fully reduced leakage model is shown in Figure 5.8.

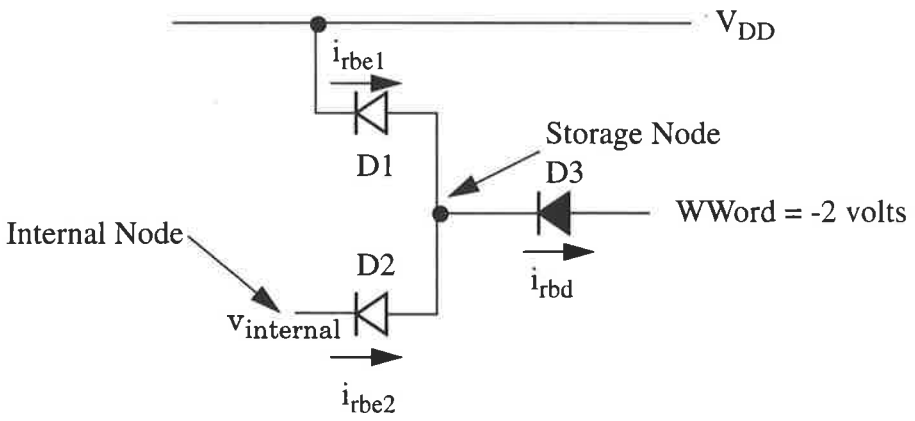


Figure 5.8. Simplest Leakage Model of the Dynamic RAM Cell.

By summing currents at the storage node, we gain the total leakage current, i_{leak}:

$$i_{leak} = i_{rbe1} + i_{rbe2} - i_{rbd} \qquad \text{(EQ 5.15)}$$

Thus the current through the two reversed biassed E-MESFET (Q1) diodes, D1 and D2, acts to increase the charge stored on the gate of Q1, while the current through the reversed biassed gate-drain diode of D-MESFET Q3 acts to reduce the charge stored at the gate of Q2. The expression in Equation 5.15 can be plotted against temperature and process spreads to find its maximum value, which will correspond to the minimum storage time.

Three important points can be stated about the leakage, due to the fact it is caused by the schottky diodes, as discussed in Chapter 2:

1. It is directly proportional to diode area. Therefore, all transistors are set to be minimum size as an initial attempt to minimise leakage.

2. It increases exponentially with temperature, and therefore as temperature increases the net leakage current affecting the storage node will also increase. Power dissipated by the chip should therefore be minimised.

3. It increases as process spread becomes faster. Therefore leakage will be a bigger problem when transistors are 2σ fast.

A simulation result of Equation 5.15 is shown in Figure 5.9, using all minimum sized transistors over a temperature range of 0 to 125 °C for process variations ranging from 2σ slow to 2σ fast. It can be seen that as temperature and process speed increases, a large net current is flowing into the storage node, due to the current imbalance between the two E-FET diodes and single D-FET diode. By increasing the width of the depletion mode Write Enable transistor Q2, the current leaking out of the Storage node will increase and reduce the net current flowing into the Storage node. If the width is increased too much, eventually a large net current will flow out of

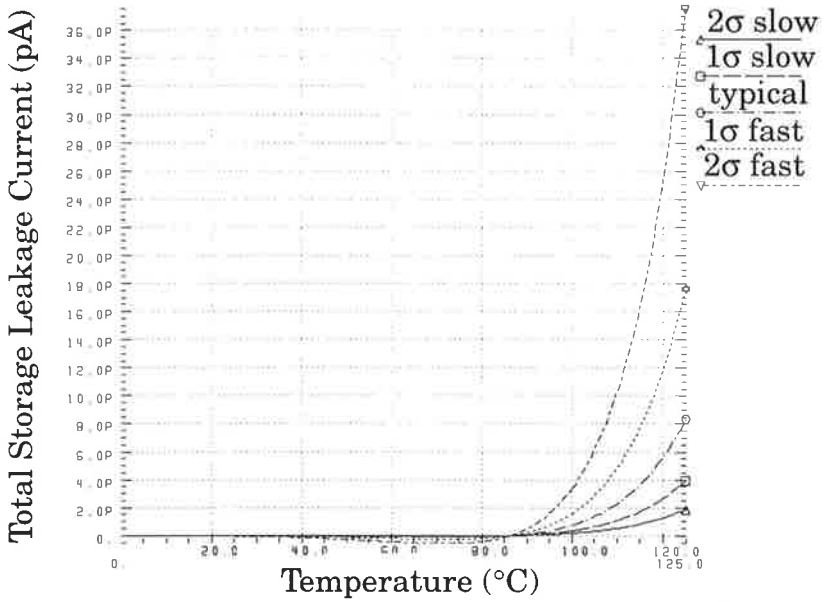


Figure 5.9. Simulation of Memory leakage current.

the storage node. The appropriate balance must be found such that the maximum leakage current is minimised. A table comparing maximum leakage currents for differing widths of transistor Q2 is shown in Table 5.6. It can be seen from Table 5.6 that a width of 2.2μm and 2.3 μm offer similar maximum magnitude of leakage current. However, a width of 2.3 μm was chosen because leakage over the operating range is always negative, meaning the storage node will always be discharged. This provides better performance, due to the fact the when storing a logic low, the node can only be charged by about 200 mV before it will be at the E-MESFET threshold voltage and at risk of providing a false reading. A logic high level can be discharged by significantly more before reading falsely. A simulation of the leakage with a 2.3 μm wide transistor for Q2 is shown Figure 5.10. The worst case net leakage current is -8.12 pA.

Table 5.6. Table of maximum leakage currents from Storage node for different widths of D-MESFET Q2 over a 0 - 125 °C temperature range

Q2 Width (μm)	Maximum Positive Leakage	Maximum Negative Leakage
2.0	38.6 pA	-581.7 fA
2.1	25.5 pA	-1.4 pA
2.2	9.6 pA	-3.5 pA
2.3	--	-8.12 pA
2.4	--	-18.22 pA
2.5	--	-39.65 pA

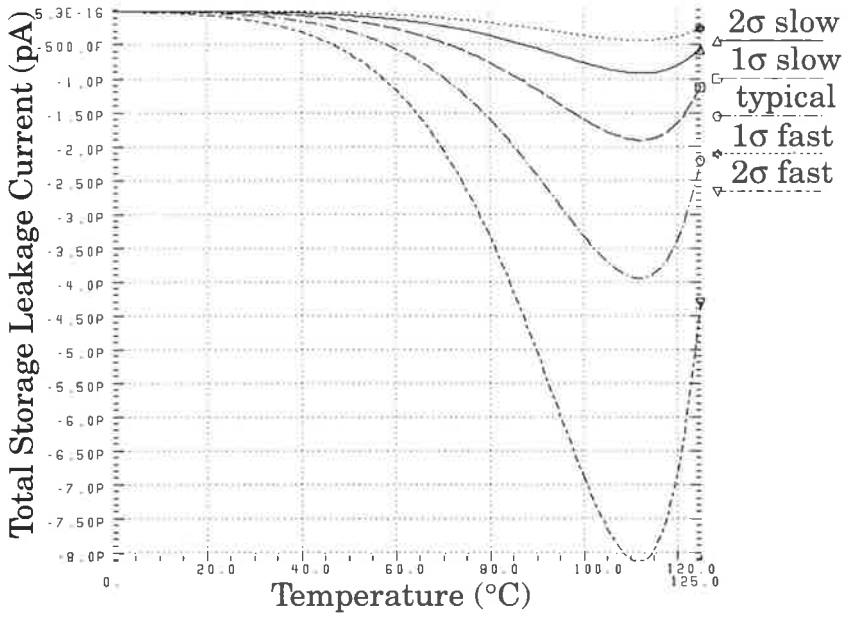


Figure 5.10. Memory leakage current for Q2 = 2.3 μm, T = 0 to 125 °C

5.5 Storage Time Calculation

Now that the maximum leakage currents are known, the storage time can be determined. We know that at a particular temperature and process variation, the leakage current will be approximately constant as the diodes are in reverse saturation. The voltage current relationship for a capacitance is:

$$\Delta Q = I \cdot \Delta t = C \cdot \Delta V \qquad \text{(EQ 5.16)}$$

Therefore:

$$\Delta t = \frac{C_{store} \cdot \Delta V}{I_{leak}} \qquad \text{(EQ 5.17)}$$

Where: C_{store} = Total Capacitance at storage node, ΔV = Allowable change in storage node voltage, I_{leak} = Leakage current into storage node and Δt = storage time of memory cell.

The critical case determining the maximum ΔV will be for the high level case, as we have assumed that net leakage is always negative over the operating conditions for the transistor sizes shown. A typical high storage voltage is around 1.2 volts, and being conservative, a 300 mV drop in this will still give a true read output. The maximum leakage current, from Table 5.6, will be 8.12 pA.

The storage time required for the memory is 5.6 μs (from Section 1.5.2.1). The capacitance at the storage node is that of a gate, typically of the order of 5 fF. This gives a storage time of approximately 180 μs. This offers a safety factor of roughly 30, however due to uncertainty in the accuracy of the reverse-bias diode models (discussed in Section 2.4.1), the safety factor allowed should be made as large as is feasibly possible. Storage time can be increased by increasing the capacitance at the storage node (Equation 5.17). A capacitor may be fabricated on chip for each cell and placed in parallel with the gate capacitance, as shown in Figure 5.11. The capacitor will require a large amount of area, and hence there will be a trade-off between cell storage time and area. It was found that a fabricated capacitance of 50 fF provided a reasonable trade off with area, giving a total storage capacitance of approximately 55 fF, and a *worst case* simulated storage time of approximately 2ms. The equates to a worst case safety factor of roughly 300. Under more favourable operating conditions, the storage time and safety factor will be many more times higher, as the leakage current will be significantly reduced. For example, at 75 °C and

UNIVERSITY OF ADELAIDE

typical parameters, a simulated storage time of about 30 ms is obtained, giving a safety factor of approximately 5000.

5.6 Read Enable Transistor Sizing

The above analysis also shows that the read transistor, Q3, has no effect on the steady state storage time of the memory cell. Therefore its size can be adjusted to optimise the read process. At first glance, it would seem appropriate to increase this width and speed up the read cycle. However, this also has the effect of reducing the value of the output high voltage, due to the increased pull-down of the wider E-MESFET (an identical effect to that seen in the DCFL inverter). Therefore, increasing the width of Q2 will cause the cell to reach a lower output high voltage, faster. Conversely, increasing the L/W ratio of Q3 will cause the cell to reach a higher output high voltage, but at a slower speed. For these reasons, the size of Q3 is left at its minimum size as this represents an effective trade-off between the two cases.

5.7 Optimised Memory Cell and Layout

A schematic of the final memory cell, optimised for our requirements, is shown in Figure 5.11.

The layout of the RAM cell must be as compact as possible. Any saving in area in the cell will result in a large saving in the 14k array. A large capacitance will take a substantial portion of the area, and so must be designed to be as small as possible. The capacitor must be layed out as a parallel plate, as no special capacitor facilities are provided with the H-GaAs II process. A table showing the values of interlayer capacitance between various layers in the H-GaAs II process is shown in Table 5.7 [26]. C_{pp} represents the parallel plate capacitance between two layers of metal, while C_f represents the fringing capacitance component as illustrated in Figure 5.12.

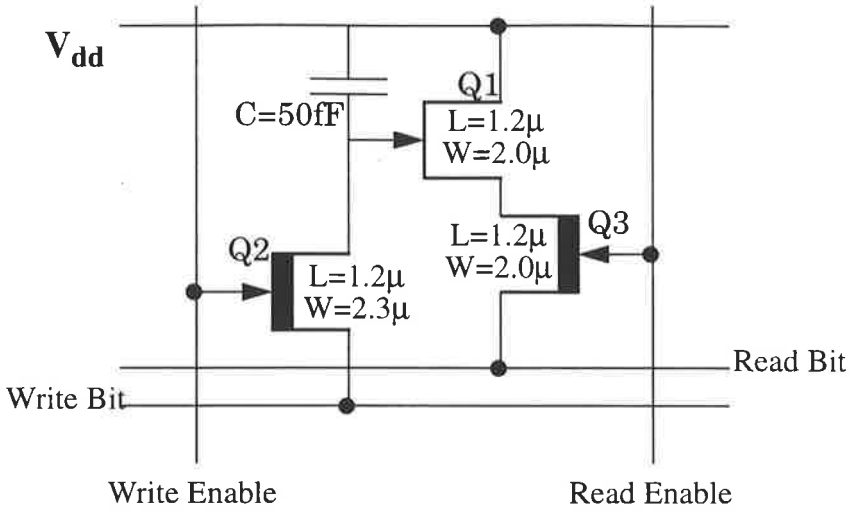


Figure 5.11. Optimised GaAs Dynamic RAM Cell

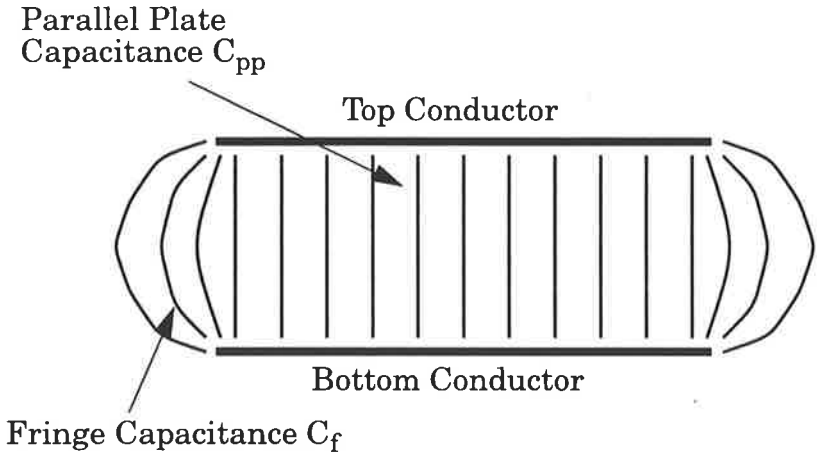


Figure 5.12. Parallel Plate and Fringe Capacitances

Using the largest capacitance metals will obviously provide the most efficient use of area when producing the additional capacitor. The largest interlayer capacitances occur between metal 1 and gate metal and metal 1 and ohmic metal. However, it should be noted that gate metal also has a substantial capacitance to the substrate. Therefore by using gate metal as the storage node plate and metal 1 as the V_{DD} plate, a large capacitance

Table 5.7. Interlayer Parallel Plate and Fringe Capacitances

Top Layer	Bottom Layer	C_{pp} (fF/μm²)			C_f (fF/μm)
		minimum	nominal	maximum	nominal
Metal 4	Metal 3	0.0278	0.0293	0.0308	0.080
	Metal 2	0.0138	0.0146	0.0153	0.026
	Metal 1	0.0105	0.0110	0.0116	0.008
	Gate Metal	0.0093	0.0096	0.0100	0
	Ohmic Metal	0.0094	0.0100	0.0105	0
	Substrate	0.0083	0.0088	0.0092	0
Metal 3	Metal 2	0.048	0.051	0.054	0.048
	Metal 1	0.032	0.033	0.035	0.035
	Gate Metal	0.026	0.028	0.029	0.030
	Ohmic Metal	0.026	0.028	0.029	0.030
	Substrate	0.020	0.022	0.023	0.035
Metal 2	Metal 1	0.069	0.073	0.076	0.049
	Gate Metal	0.047	0.050	0.053	0.045
	Ohmic Metal	0.047	0.050	0.053	0.045
	Substrate	0.030	0.032	0.035	0.042
Metal 1	Gate Metal	0.121	0.127	0.134	0.051
	Ohmic Metal	0.121	0.127	0.134	0.051
	Substrate	0.048	0.052	0.057	0.044
Gate Metal	Substrate	0.074	0.076	0.079	0.045

will be made between the storage node and V_{DD} and the storage node and ground. Using ohmic metal as the storage node plate and metal 1 as the V_{DD} plate will only give the Vdd to storage node capacitance.

The layout is shown in Figure 5.13. The cell was designed to be as compact as possible with the large capacitance required. This meant laying out the capacitor as efficiently as possible, hence the square nature of the design. The cell dimensions measure 25.5 μm square, giving approximately 1538 cells/mm^2. The storage capacitor now comprises a total of 35 femptofarads capacitance between V_{DD} and the storage node, and another 20 femptofarads of parasitic capacitance between the storage node and the substrate (effectively ground). This gives an effective nominal storage capacitance of 50 fF.

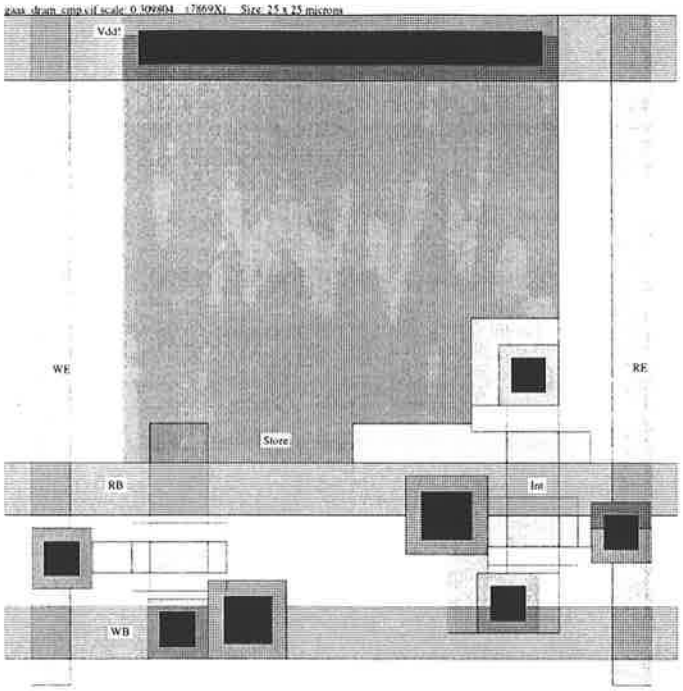


Figure 5.13. H-GaAs II Layout of DRAM cell

5.8 Memory Cell Performance

In all memory simulations, the Read Bit bus is driving a typical DCFL inverter load (resulting in Read Bit discharging through the forward biassed gate diode once the read cycle has finished, as seen in the simulations), and a D-MESFET is used to pre-discharge the bus.

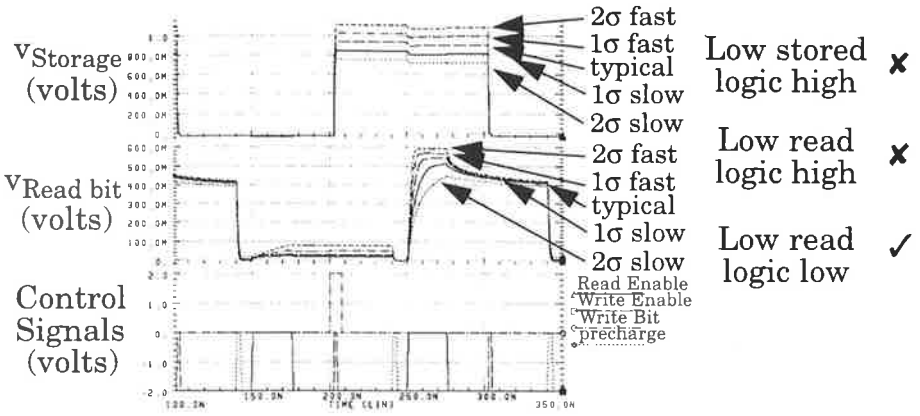
To demonstrate the performance of the memory cell with the chosen voltage levels decided in Section 5.2, Figure 5.14 shows three simulation results. The first, Figure 5.14a shows a simulation of the memory cell at 75 °C, using 2σ slow to 2σ fast parameters and 0 volt control signal on the Read and Write Enable transistors to read and write the memory. The lower storage and output high voltages can be clearly seen.

Figure 5.14b shows another simulation under the same conditions, this time using a 0.6 volt signal to enable the Read and Write transistors. It can be seen that significantly more current flows from the Read Enable line

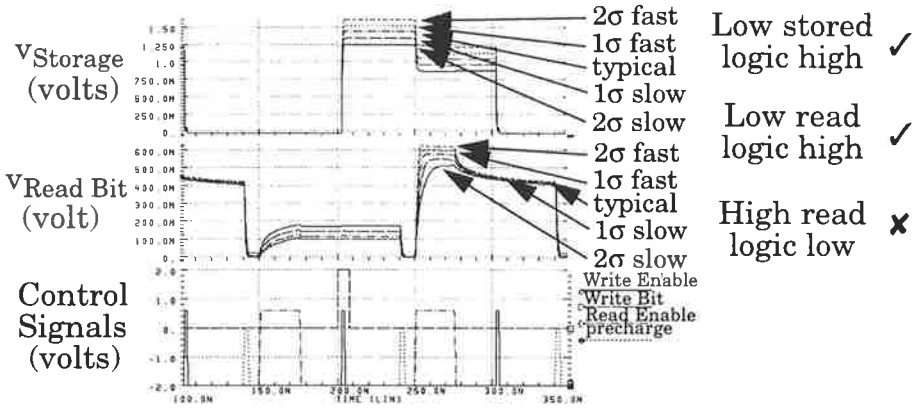
onto the Read Bit bus via the forward biassed diode when a logic low is being read, although the stored logic high level, and consequently read logic high level and read access time is significantly increased, relative to Figure 5.14a. This higher logic low level is exacerbated at higher temperatures.

In Figure 5.14c, a simulation with a 0.6 volt signal to write the data to memory and a 0 volt signal to read from memory is shown. The same high voltage stored in Figure 5.14b is evident, and the forward conduction is significantly reduced, which results in better performance than Figure 5.14a and Figure 5.14b. The logic levels chosen above are therefore justified.

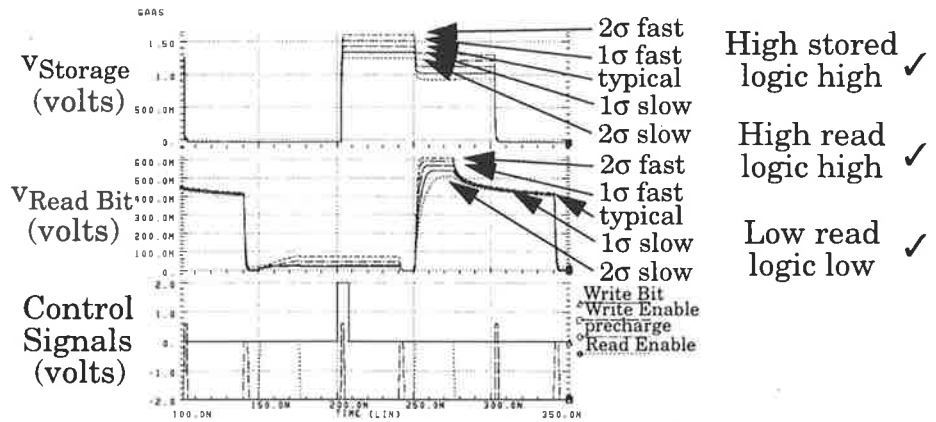
Simulations were also done to ensure that the storage time requirements of the memory cell were being met. A simulation showing long term storage time at 75 °C over a 2σ slow to 2σ fast process variation is shown in Figure 5.15. It can be seen that under such conditions the memory storage time is exceeding 2.5 ms. At lower temperatures, the safety factor is even larger, while as temperature increases it reduces exponentially, although at no temperature below 125 °C will the safety factor be less than 80. As it is unlikely temperatures will reach this point, the storage time should be adequate. Note the storage time simulation differs from the predicted storage time due to the use of the MESFET model when simulating the memory cell. As discussed earlier, this model does not accurately model diode behaviour. Unfortunately, the Diode model does not model the MESFET operation at all and thus to simulate the memory cell the MESFET model must be used, at the expense of not accurately modelling storage time. The MESFET model gives a much larger reverse-bias leakage current than the diode model and so leakage currents in the simulation are greater than when using the diode model. This result suggests that the cell will be able to store for longer periods of time than indicated by simulation.



(a) Memory Cell simulation at 75 °C, 2σ slow to 2σ fast parameters using -2v to 0v control signals



(b) Memory Cell simulation at 75 °C, 2σ slow to 2σ fast parameters using -2v to 0.6v control signals



(c) Memory Cell simulation at 75 °C, 2σ slow to 2σ fast parameters using -2v to 0.6v on WE signal, -2v to 0v on RE signal

Figure 5.14. Memory Cell Simulations.

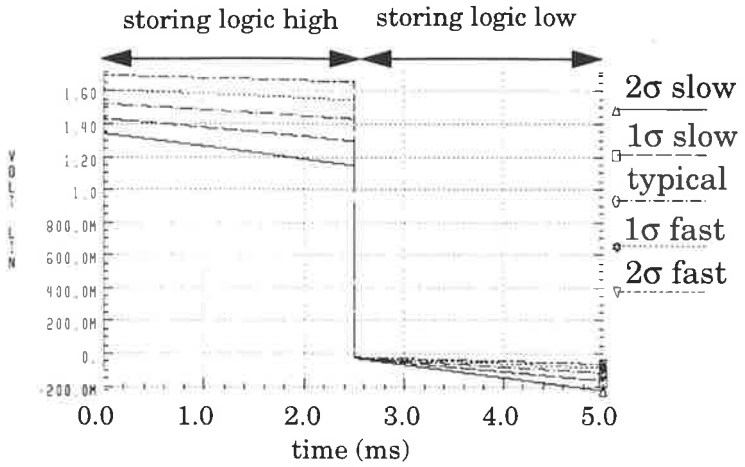


Figure 5.15. Long term storage time simulation at 75 °C, 2σ variation.

5.9 Conclusion

We have now obtained a suitable memory cell for the implementation of the ATM buffer. The cell has been optimised so that its storage time is maximised over a wide range of process and temperature variation. The value of storage capacitance required, allowing for significant inaccuracies in the *hspice* models, has been determined. The optimum memory cell has been layed out and simulated to verify performance. The design of the 14 kilobit memory array, based on this cell, is presented in the next chapter.

Chapter 6: Dynamic RAM Array Design

This chapter discusses the design of the full 14 kbit Dynamic RAM array, along with all associated word and bit drivers, level shifters and sense amplifiers.

6.1 DRAM Array

The layout of the DRAM was designed to be as close to square as possible, and hence the array is partitioned into 128 words each of 112 bits, giving the core memory dimensions of 2.856 mm by 3.264 mm. A floor plan of the memory array is shown in Figure 6.1. Seven bit address lines are required to decode all 128 addresses.

To enable simultaneous reading and writing of the memory, two sets of word drivers are provided to decode the read and write addresses being sent from the output and input controllers respectively.

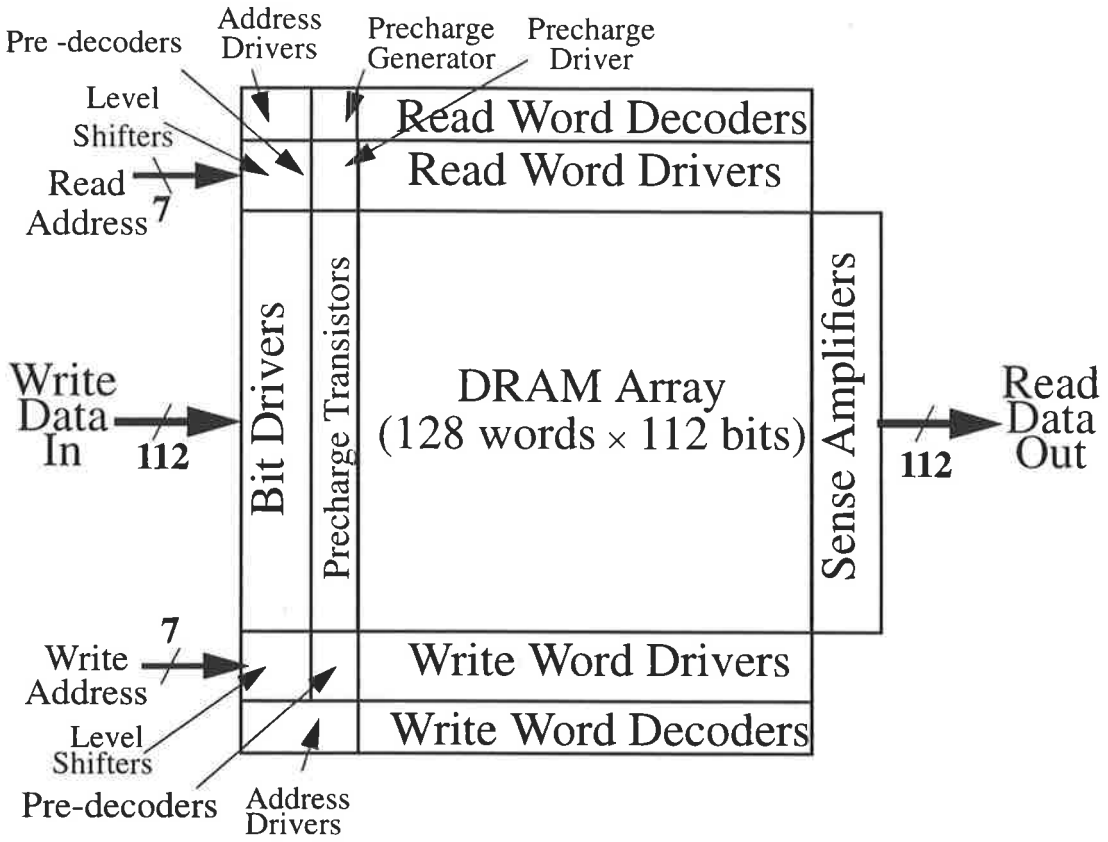


Figure 6.1. Detailed floor-plan of the DRAM

6.1.1 Voltage Levels in Memory Control Logic

As discussed in Section 5.2, the control voltages for the DRAM cell use -2 volts as the low supply rail and lower limit of the voltage levels. The controller logic has been designed to supply signals which use ground (0 volts) as the low supply rail. This means that at some point there must be some level shifting of the voltage of signals from a 0 volt low supply logic to -2 volt low supply logic. Two options present themselves, namely:

- Decode the address lines and the read/write enable signal using 0 to +2 volt logic, and level shift the control signal down before entering the word drivers.
- Level shift address lines and the read/write enable signal down before decoding and then perform decoding using -2 to 0 volt logic. No further level shifting is required, as the word drivers can then be driven directly.

Clearly, the second option is superior, as it requires only one level shifter for each of the address/enable signals - a total of 16 level shifters (7 address signals + 1 enable signal for read and write). The first option requires one level shifter for each word line - a total of 256 level shifters. This will make a significant addition to the total chip dissipation and area.

For this reason, all address decoding is done using -2 to 0 volt powered logic.

6.1.2 Address Decoding Strategy

The classical address decoding equation for writing to a word in the memory array is given by Equation 6.1:

$$Select = A_0 \cdot A_1 \cdot A_2 \cdot A_3 \cdot A_4 \cdot A_5 \cdot A_6 \cdot WE \qquad \text{(EQ 6.1)}$$

When the *Select* line is enabled, the data on the write bus will be written into the appropriate memory word. Only one memory word may be *Select*ed at a time. By mapping the inputs of the AND gate representing Equation 6.1 to the address lines and their complements, all words of the memory can be accessed. The *Select* line is only enabled when the Write Enable signal, WE is high. When Write Enable is low, all *Select* lines are low, and no memory cells are being written to.

Because of the inability to produce AND functions in GaAs, it is necessary to convert Equation 6.1 into a form which can be implemented using NOR gates. This can be done using deMorgans law:

$$Select = \overline{\overline{A_0} + \overline{A_1} + \overline{A_2} + \overline{A_3} + \overline{A_4} + \overline{A_5} + \overline{A_6} + \overline{WE}} \qquad \text{(EQ 6.2)}$$

An 8-input NOR gate will therefore be required to implement the decoding function. Because of the poor fan-in performance of GaAs, the 8-input NOR gate will have to be partitioned into combinations of smaller fan-in NOR gates. There are several methods of partitioning the NOR gate, and the optimum method is chosen based upon delay and the number of lines which must be driven to the decoders. If the 8 input NOR gate is used, it is

necessary to drive each address line and its complement and the Write Enable signal, $\overline{WE}$, to the decoder. The number of lines which have to be driven to the decoders is 2n + 1, where n is the number of address lines. In our case we have 7 address lines and therefore the number of lines which must be driven to the decoders is 15.

The 8-input NOR gate can be partitioned into two 4-input NOR gates feeding into a 2-input NOR gate. Therefore our gate count is increased 3 fold - resulting in a large increase in area and dissipation.

Consider however an alternative rearrangement of Equation 6.1:

$$Select = \overline{\overline{A_0 \cdot A_1} + \overline{A_2 \cdot A_3} + \overline{A_4 \cdot A_5} + \overline{A_6 \cdot WE}} \qquad \text{(EQ 6.3)}$$

This can be further altered to NOR/OR form as:

$$Select = \overline{\left(\overline{A_0} + \overline{A_1}\right) + \left(\overline{A_2} + \overline{A_3}\right) + \left(\overline{A_4} + \overline{A_5}\right) + \left(\overline{A_6} + \overline{WE}\right)} \qquad \text{(EQ 6.4)}$$

It can be seen that in this form only a single 4-input NOR gate is required at each decoder, if some decoding is done prior to delivering the signals to the decoder via 2 input NOR gates. The number of lines necessary to be driven to the decoders is now $3 \times 2^2 + 2 = 14$ lines. i.e. 2^2 combinations for each of the first three terms of Equation 6.4, and two for the last (A_6 and $\overline{A}_6$). A total saving of 512 gates and one driver is made over the chip, saving significant amounts of power, area and delay.

Further partitioning of Equation 6.1 is not beneficial, as:

$$Select = \overline{\left(\overline{A_0} + \overline{A_1} + \overline{A_2}\right) + \left(\overline{A_3} + \overline{A_4} + \overline{A_5}\right) + \left(\overline{A_6} + \overline{WE}\right)} \qquad \text{(EQ 6.5)}$$

Here we require $2 \times 2^3 + 2 = 14$ driving lines, but only a 3-input NOR gate as each decoder. There is little difference in dissipated power between 3-input and 4-input NOR gates and driver complexity and number will be higher. Grouping the address lines into still larger groups, i.e. 4, 5, 6, 7 and 8 dramatically increases the number of drivers required (24 for 2 groups of

4, 36 for groups of 5 up to 128 for groups of 7 and 8). We can thus conclude that Equation 6.4 represents the optimum form in terms of partitioning.

However, under simulation it was found that this method produces extra delay in the output write enable signal, due to pre-processing the input write signal. As a compromise between delay and power dissipation/area, some pre-processing can be done on the address lines, while the write signal is passed directly into decoder. The equation can be expressed as:

$$Select = \overline{\left(\left(\overline{A_0} + \overline{A_1}\right) + \left(\overline{A_2} + \overline{A_3}\right) + \left(\overline{A_4} + \overline{A_5}\right) + \overline{A_6}\right) + \overline{WE}} \qquad \text{(EQ 6.6)}$$

This can be realised as in Figure 6.2.

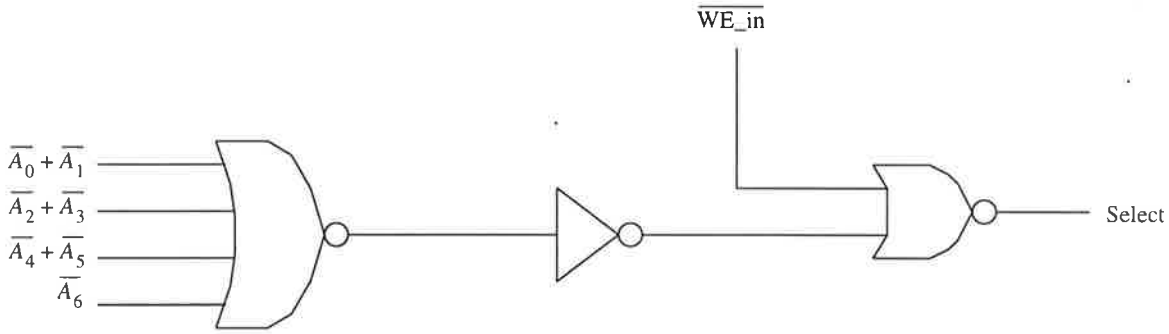


Figure 6.2. Circuit diagram of the address decoder

It is also necessary to design a floor plan for the address decoding scheme as there is a large number of signals involved which must be appropriately routed. A floor plan of the Write Address decoder section is shown in Figure 6.3. The Read Address decoder section has an identical floor plan, except it also includes provision for the pre-charge drivers and decoders.

6.2 Design of Functional Blocks

The functional blocks which make up the DRAM and their design are now discussed. The following blocks are designed:

- Input Level Shifters/Buffers

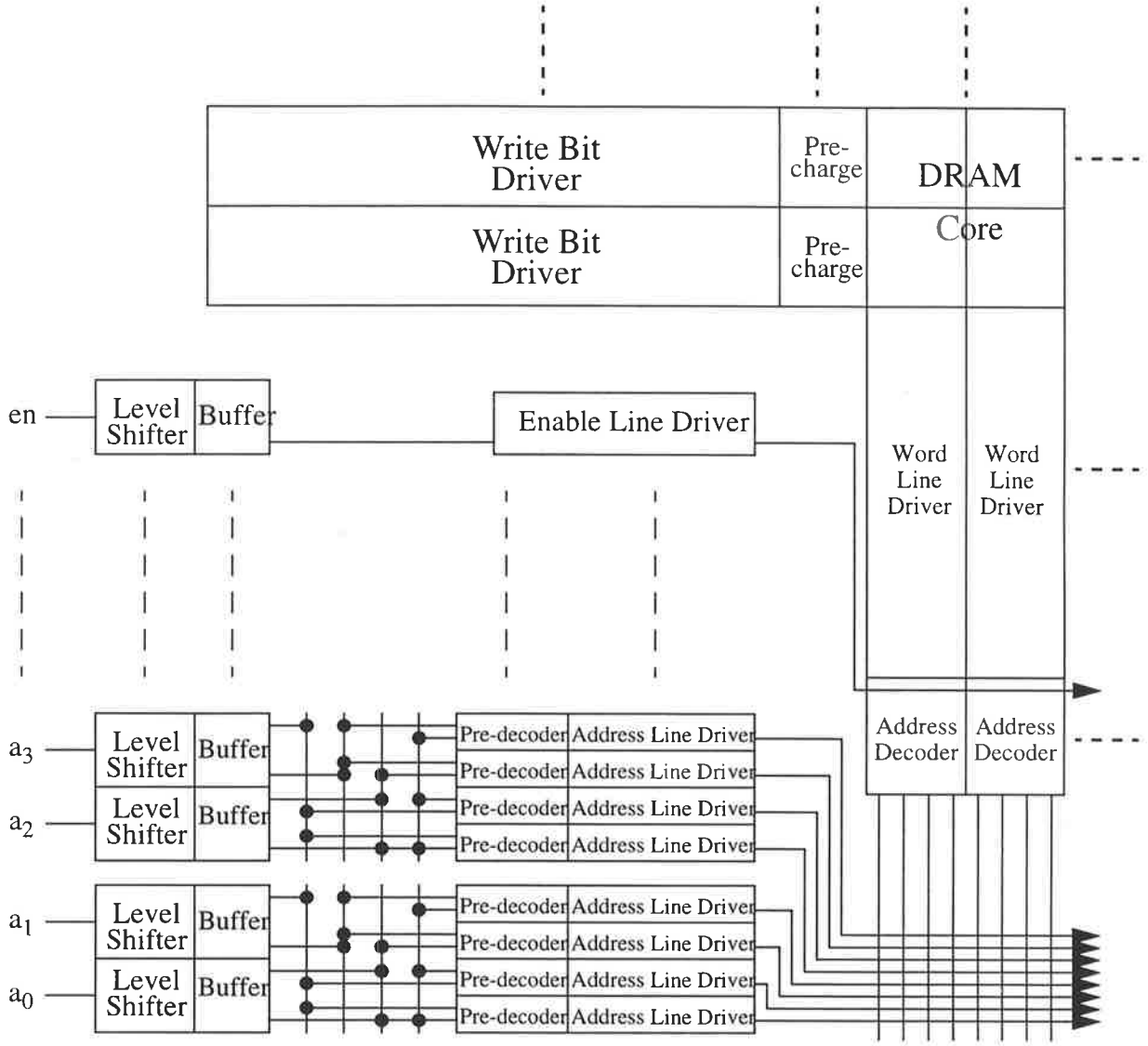


Figure 6.3. Floorplan of the DRAM write address decoding section

- Address Pre-Decoders and Decoders
- Read and Write Enable Line Drivers
- Address Line Drivers
- Write Bit Line Driver
- Read and Write Word Driver
- Pre-charge Circuit, Decoder and Driver
- Sense Amplifier

As we saw in Chapter 5, DRAM storage time reduces exponentially with increased temperature. Therefore, when designing the functional blocks,

the primary concern of the author was to keep power dissipation as low as possible while still giving adequate performance. This is a difficult task, mainly due to the need for large numbers of large capacity, high speed drivers (i.e. Read and Write Word drivers, of which 128 of each are needed, and which must drive a load of 112 transistors + about 1pF of capacitance each). Due to their number, a small amount of dissipation saved per driver will result in a large overall saving, while a high dissipation driver has its effects multiplied to result in an enormous dissipation when used in the completed array.

6.2.1 Level Shifter/Buffer Design

Because all address decoding is done using a -2 volt lower supply, and circuits external to the DRAM operate on a 0 volt lower supply, level shifters are required to convert the incoming address and enable signals. The buffer stage is necessary to provide the level-shifted output and its complement to the pre-decoder stages, and is easily done using two DCFL inverters with a β ratio of 12, arranged as shown in Figure 6.4. The first inverter is made slightly larger than the second as in addition to the pre-decoders, which will amount to a small connecting capacitance and a fan-out of approximately two, it must drive the second inverter. There will be a slight amount of delay between the output and its complement, but this it will be very small compared to the total read/write cycle time and will not create problems.

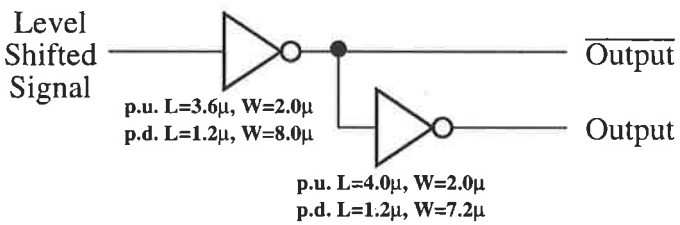


Figure 6.4. Buffer following the level shifter.

Level shifting in GaAs is done using constantly forward biassed schottky barrier diodes, which will give in the region of 0.5 to 0.7 volts of voltage drop, depending on process and temperature. To provide larger voltage drops, several diodes in series can be used. To level shift down by 2 volts, four level shifting diodes must be used. A schematic of a typical level shifter is shown in Figure 6.5. It has a small DCFL input stage, followed by a larger DCFL stage driving the level shifting diodes and the active pull-down D-MESFET load. The level shifted output is taken from across the load. A simulation showing the performance of the level shifter is shown in Figure 6.6. The level shifter is driving the buffer shown in Figure 6.4 as its load, and was done at 75 °C and 125 °C using 2σ slow to 2σ fast process parameters.

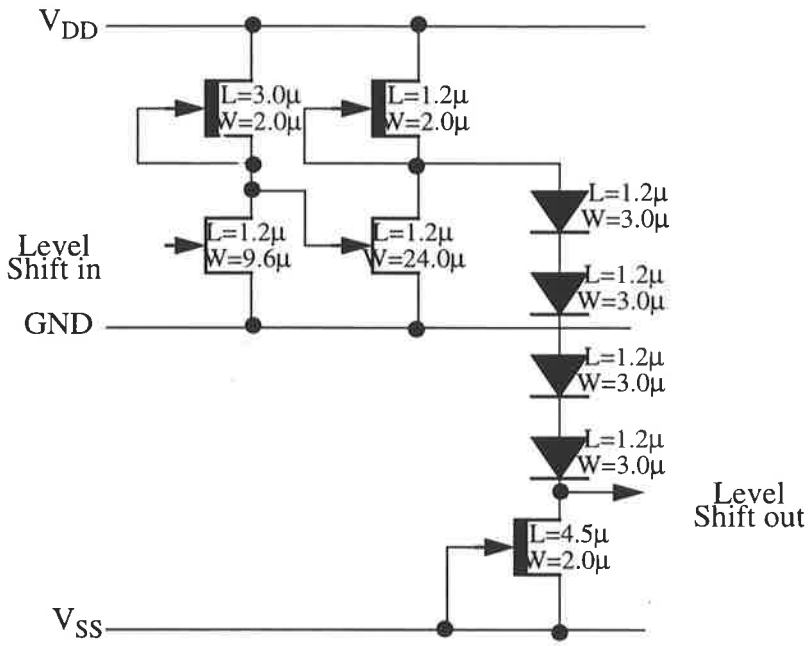


Figure 6.5. Typical level shifter design.

Both the pull-down and pull-up performance of the level shifter can be improved by using a source-follower type structure as shown in Figure 6.7. The pull-up performance is improved because of the better driving power of the wide enhancement mode transistor, while the lower pull-down voltage at the output is due to the fact that the pull-up enhancement mode transistor is completely turned off when the circuit is outputting a logic low. The

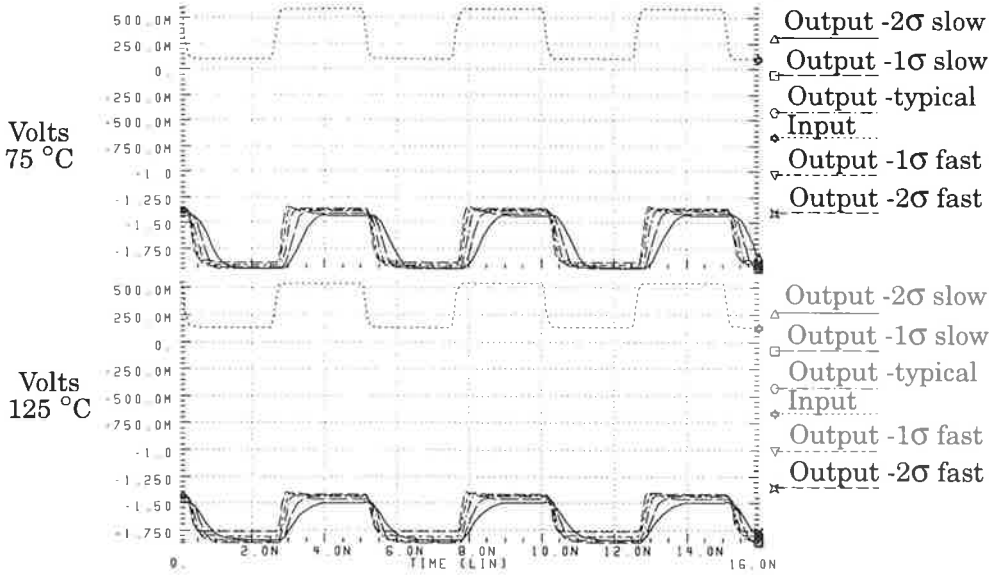


Figure 6.6. Simulation of the typical level shifter design.

greatest benefit is the lower pull-down voltage, which enables the following buffer section to function with less leakage over large process and temperature variation.

Because the DCFL stages are not directly driving the diodes, they can be reduced in size and the average power dissipation of the circuit is therefore reduced. There is no notable increase in delay. A simulation of this level shifter, again driving the DCFL buffer, is shown in Figure 6.8. The simulation was carried out at 75 °C and 125 °C using 2σ slow to 2σ fast process variation models. As an illustration of the superior performance of the second level shifter, a comparison is shown in Figure 6.9, at typical parameters at 75 °C and 125 °C. The total current drawn from the supply by both circuits at 125 °C, typical parameters, is also shown. It can be seen that as well as superior performance, the source follower type level shifter has a slight increase in maximum power dissipation, but a lower average power dissipation of approximately 520 μW, compared to 550 mW for the typical level shifter design.

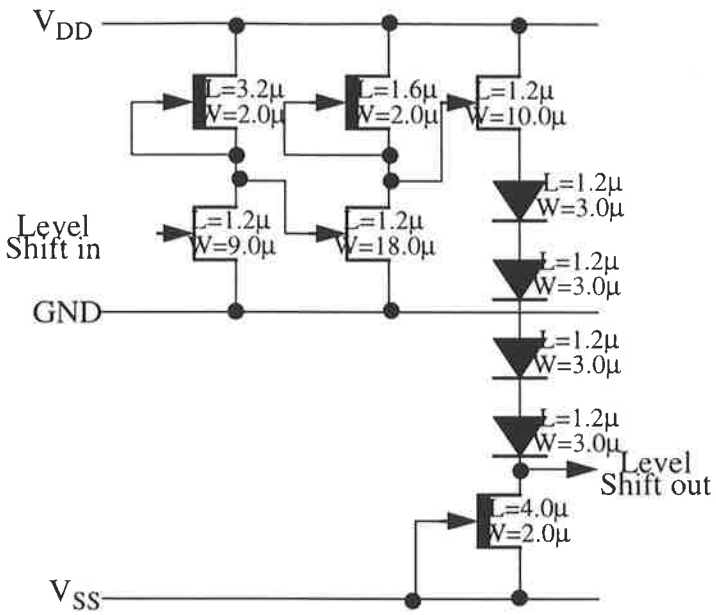


Figure 6.7. Source follower type level shifter.

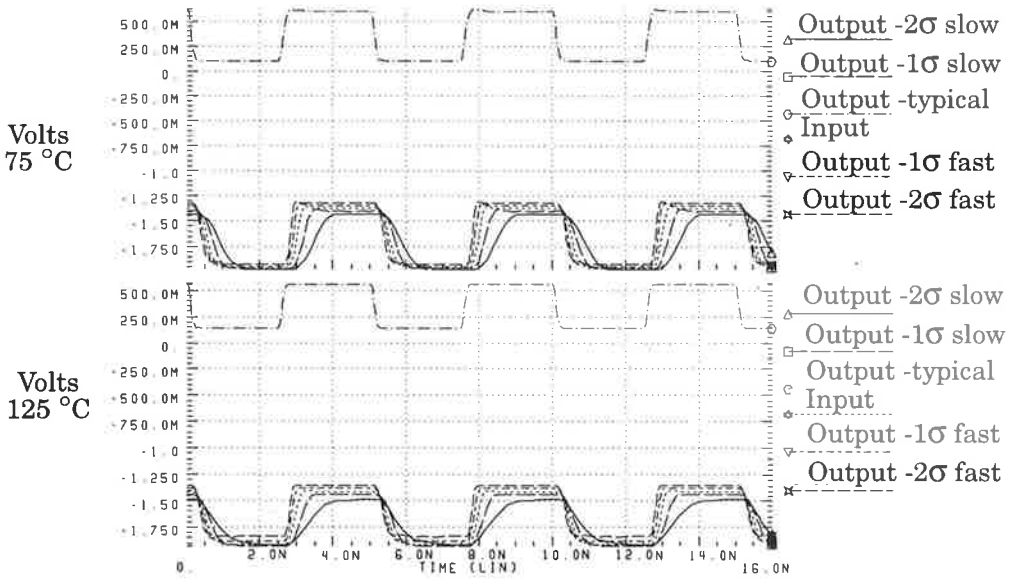


Figure 6.8. Simulation of the Source follower type level shifter design.

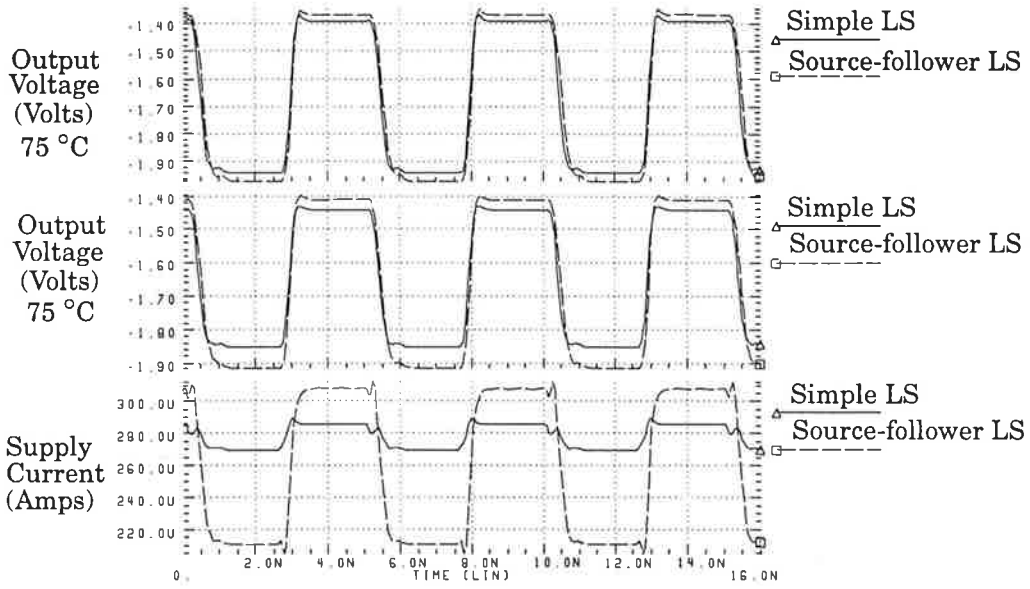


Figure 6.9. Comparison simulation of level shifters.

The same level shifter/buffer is used for the enable signal as the level shifted enable signal must also be buffered before being put into the enable driver.

The level shifter and buffer were layed out using the *ring-notation* methodology, and designed to complement the pre-decoder and address line driver layouts. The *ring notation* style produces a very wide layout, as can be seen in Figure 6.10. The level shifter and buffer sections are shown. The buffer outputs directly to the pre-decoders.

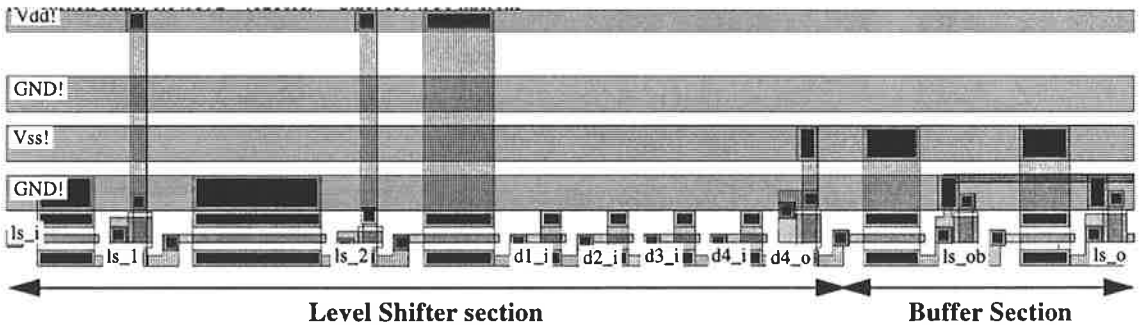


Figure 6.10. Layout of the level shifter and buffer.

6.2.2 Address Decoder Design

Straight DCFL logic was used to implement the decoder function shown in Figure 6.2, due to the fact that DCFL has adequate performance at a fan-in of 4, and it provides the lowest power dissipation. To improve fan-in performance on the 4-input NOR gate, leakage through the pull-down E-MESFETs while they are in the off state is minimised, by feeding the inputs using a driver capable of supplying a good logic low ($\approx$ 0 volts) rather than a DCFL level ($\approx$ 100 mV). To minimise the power dissipation of the 4 input nor gate and inverter, the pull-up/pull-down ratio of 12 was realised using 5.0μm × 2.0μm pull-up D-MESFETs and 1.2μm × 5.8μm pull-down E-MESFETs. To improve the speed of the two input nor gates, the transistor size was increased to 4.4μm × 2.0μm for the D-MESFET and 1.2μm × 6.4μm for the E-MESFETs. A schematic of the decoder is shown in Figure 6.11.

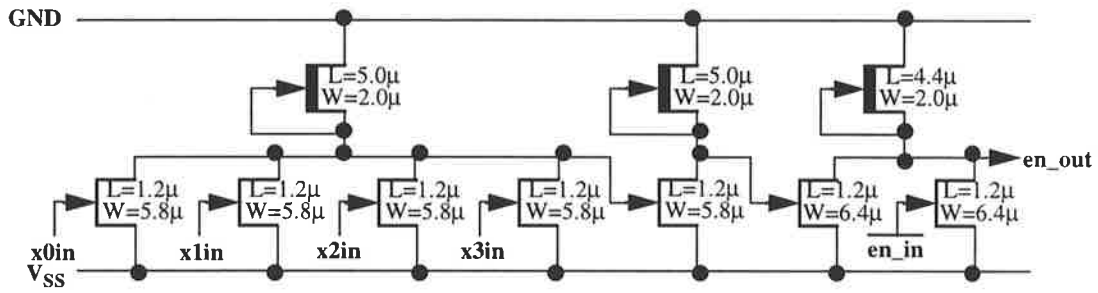


Figure 6.11. Schematic of the address decoder.

Figure 6.12 shows the results of a simulation on the address decoder performed at 75 °C and 125 °C using 2σ slow to 2σ fast process parameters. It can be seen that when all decoded address lines are low and $\overline{\text{en\_in}}$ goes low, the output enable signal goes high as required. The output high swing is approximately 1.2 volts as it is feeding into the Read and Write Word drivers, which both have SFFL input stages. The decoder has an average power dissipation of approximately 450 μW when in the output low state (at 125 °C, typical parameters), giving the total dissipation for the 256 decoders at approximately 115 mW. Because only one cell can be written to

and read from simultaneously, only 2 of the 256 decoders will be in the output high state.

The decoder must be of the same horizontal pitch as the memory cell, and therefore this requires a thin layout only 25.5μm in width, as shown in Figure 6.13. Because of its tendency to create layouts which have a large width, the general *ring notation* methodology cannot be followed.

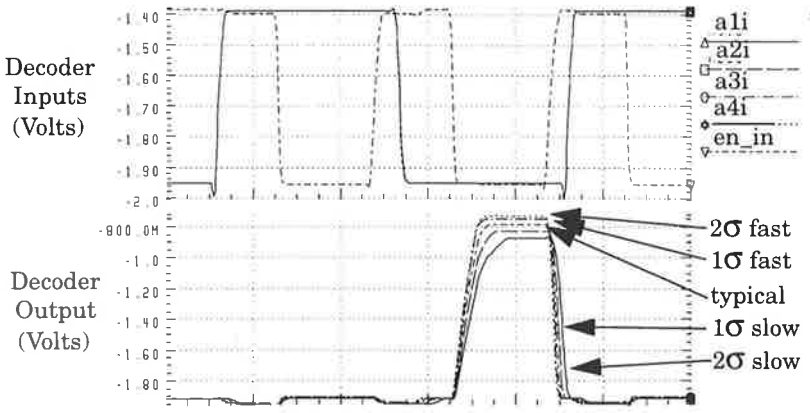


Figure 6.12. Address decoder driving a Write Word driver simulation.

6.2.3 Pre-Decoder Design

The Pre-Decoder design is very simple. It must perform an OR of the level shifted address inputs, as discussed in Section 6.1.2, and then drive the address line drivers, discussed in Section 6.2.4, requiring only a small DCFL stage. By making the Address Line driver, (designed in Section 6.2.5) inverting, the Pre-decoder can be made inverting, and hence only a single 2-input NOR gate is required. The schematic is shown in Figure 6.14, using a pull-up to pull-down ratio of 12. The pre-decoder has a peak dissipation of about 540 μW and an average dissipation of 320 μW.

Because there is an uneven number of address lines, a_6 is not paired with another line and so does not require pre-decoding. However, the same level-shifting and address line drivers will be used and for this reason a

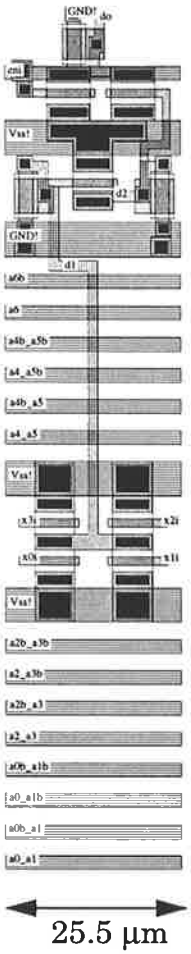


Figure 6.13. Layout of the Address Decoder.

simple, small DCFL inverter will be used in place of the pre-decoder, as shown in Figure 6.15.

The pre-decoder was layed out to fit the floor-plan shown in Figure 6.3, using the *ring-notation* methodology, and designed to complement the level shifter/buffer and address line driver designs. The layout of the pre-decoder is shown in Figure 6.16. The layout of the inverter is very similar.

6.2.4 Enable Line Driver Design

The enable line driver must drive the enable transistor and line into the array of 128 word decoders. Each decoder represents a maximum load of about 6 fF capacitance plus a transistor 5.8 µm wide. Therefore the maxi-

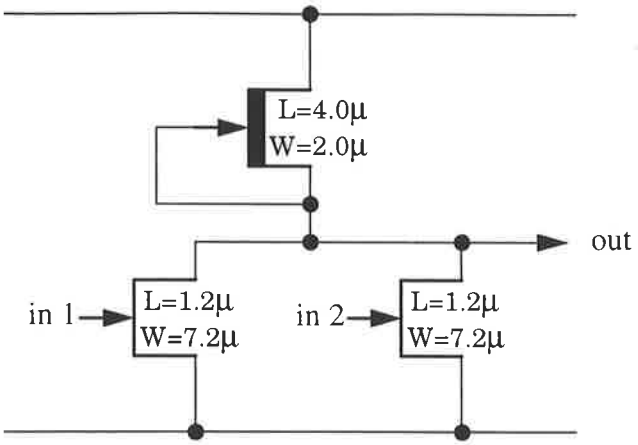


Figure 6.14. Schematic of the address pre-decoder.

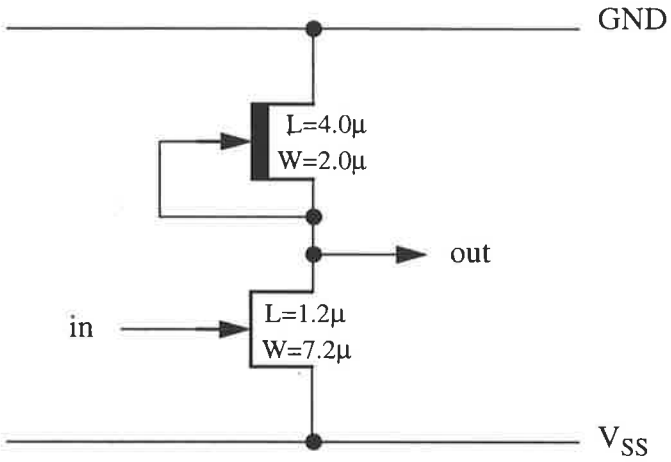


Figure 6.15. Replacement for pre-decoder for line a_6.

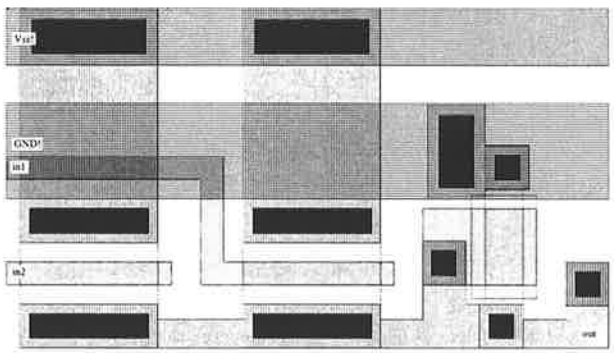


Figure 6.16. Layout of the pre-decoder.

mum equivalent load to be driven is 768 fF plus a transistor 742.4 μm wide. The level required for the decoders is -2 volts to -1.4 volts (DCFL inputs with -2v as lower supply). This therefore allows investigation of all logic families suitable for driving large loads, namely SBFL, UBFL, SBFL, DSBFL and SDCFL. As discussed in Section 6.2.2, to improve the rise time of the decoder it is necessary to drive the lines to a lower logic low level than typical DCFL, because of its high fan-in. After investigation the speed, power dissipation and logic levels of the above drivers, it was found that DSBFL provided superior operation. This process is discussed in Appendix A. The schematic of the Enable Line Driver is shown in Figure 6.17.

The drivers were layed out again using *ring-notation*, and to fit the floor plan and layouts of the pre-decoders and level shifters, as these three circuits form a sequential block. The layout of the driver is shown in Figure 6.18.

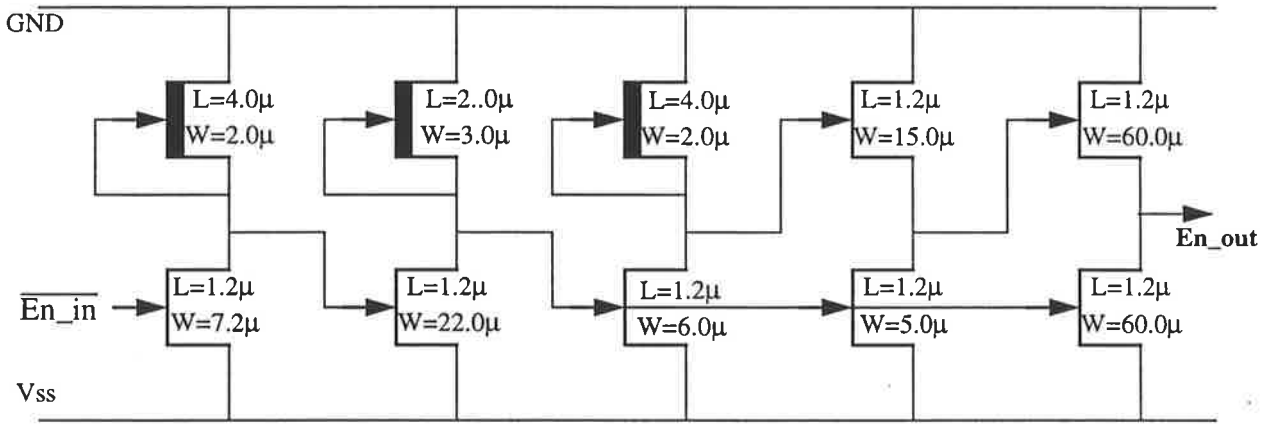


Figure 6.17. Schematic of the Enable line Driver

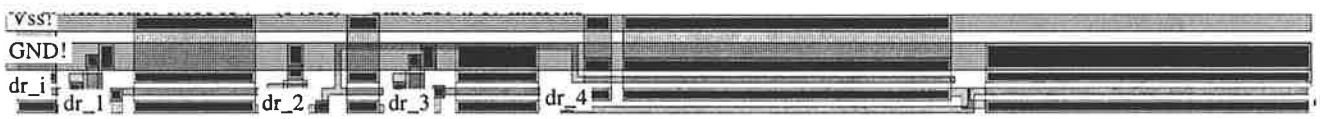


Figure 6.18. Layout of the Enable Line Driver

6.2.5 Address Line Driver Design

The address line driver has a similar properties and functionality to the enable line driver discussed above, except the active load it must drive is substantially reduced because of the fact that only one in every four address decoders is driven by each line. A smaller size driver must be used, or the address line high voltage would be over driven, reducing the decoder low level noise-margin via forward biassed schottky diode conduction. The capacitive load is slightly increased because of the slightly higher line capacitance of 8 fF per decoder. Therefore the total load to be driven is:

- A capacitance of approximately 1000 fF
- An enhancement MESFET of dimension L=1.2μ, W=180μ, configured as the pull-down of an inverter of β ratio 12.

A similar investigation similar to the previous section was carried out and again it was found that the DSBFL driver resulted in the best performance. A schematic of the driver is shown in Figure 6.19.

The Address Line driver layout is shown in Figure 6.20.

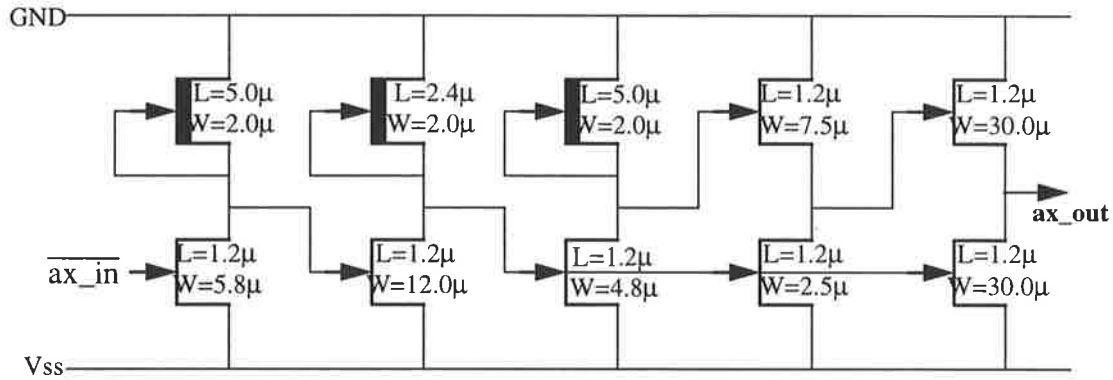


Figure 6.19. Schematic of the Address Line driver

6.2.6 Write Bit Line Driver Design

The line drivers must supply a signal on the Write Bit bus. To maximise the value of the high value stored on the storage node, the output from the

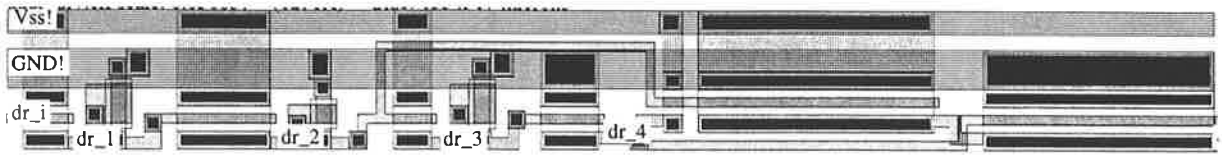


Figure 6.20. Layout of the address line driver

Write Bit line driver would ideally be as high as possible (Section 5.2). Because of the output level limitations of SBFL, UBFL and SDCFL, a DCFL circuit will give the best output. However, due to the large load presented by 128 Write Bits of memory cells in parallel, a DCFL output stage has the following disadvantages:

- A large load will require a large D-FET pull-up, and to maintain the pull-up to pull down ratio, a very large E-MESFET will be needed, resulting in a very large driving stage with high power dissipation
- To minimise the low level written, the output low voltage on the Write Bit line will ideally be close to 0 volts, lower than a typical low level DCFL signal. To achieve this the pull-up/pull-down ratio must be increased, exacerbating the previous problem.

Hence, a DCFL driver was not used. Instead a Superbuffer driver (SBFL) was designed. It is a much more efficient driver, and although the output high voltage is approximately 0.5 volts lower than the DCFL output, it was necessary to accept this because of the benefits in drive speed, rise and fall times, power dissipation, and reduced logic low level. A schematic of the Write Bit line driver is shown in Figure 6.21. The driver is non-inverting. A simulation of the driver is shown in Figure 6.22, carried out at 75°C and 125 °C, using a $\pm 2\sigma$ process variation, and driving 128 memory cells in series. The average power consumption is 700 μW, (at typical parameters, 75 °C). This gives a typical average power consumption for the 112 line drivers of 78 mW.

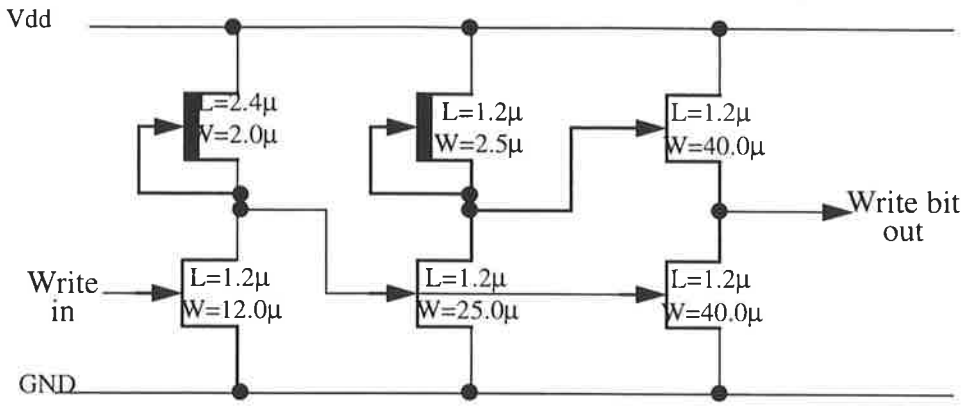


Figure 6.21. Write Bit line driver.

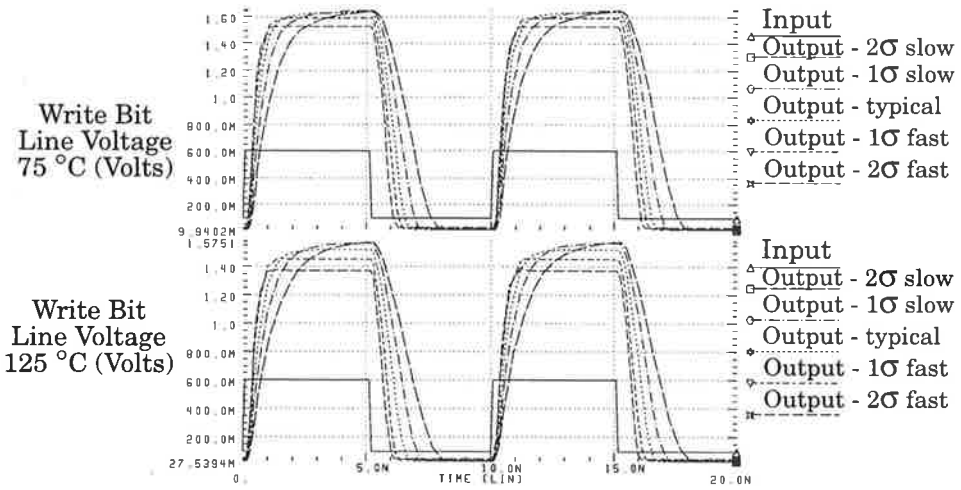


Figure 6.22. Simulation of Write Bit Line Driver.

The layout of the Write Bit line driver must be compatible with the memory array, and thus have a vertical pitch of 25.5 µm maximum. The layout of the Write Bit line driver is shown in Figure 6.23. The dimensions of the cell are 259.2 µm by 25.5 µm. The cell includes vertical Vdd and GND connections on metal 3.

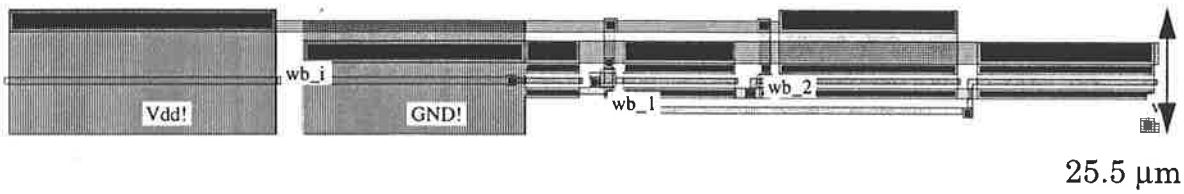


Figure 6.23. Layout of the Write Bit line driver.

6.2.7 Write Word Driver Design

The Write Word driver must be designed such that it provides a -2 volt low signal and unconstrained high output signal on the Write Enable line of a word of 112 memory cells in parallel. This represents an equivalent capacitance of approximately 1 pF, a very large capacitive load in GaAs VLSI design. There will be 128 Write Word drivers in the DRAM array. When designing the driver, it should be noted that only one driver will be in the output high position at any time, two cells can not be written simultaneously. The other 127 drivers will be in the -2 volt position. The ideal driver would therefore dissipate low power when in the output low state, while the power dissipated in the high output state and in switching between states is unimportant (within reason). Unfortunately, due to the nature of GaAs MESFET transistors, the only way of achieving such a large output voltage is using an inverting logic class, such as DCFL and SBFL. These classes rely on having a large pull-down transistor, and so when in the output low state, there is a continuous path from Vdd to ground via the pull-up D-MESFET and pull-down E-MESFET, resulting in large power consumption in the low state. In the output high state, the E-MESFET is turned off, and thus output power is reduced. A source-follower type non-inverting structure cannot be used, as the output voltage is constrained to be less than the input voltage applied to the gate of the pull-up E-MESFET. A source follower would also be unable to supply a sufficiently high output voltage because of the large, permanently on, D-MESFET pull-down transistor required.

The nature of the logic levels also requires the inverting stage to be connected between V_{DD} (2v) and V_{SS} (-2v). The power consumption of this inverter in the output low state will be particularly high, due to an effective supply voltage of 4 volts. Thus, as power consumption is a primary concern, the size of this DCFL stage must be minimised. The speed of the driver has also been compromised to minimise the power requirement. To minimise the size of the DCFL stage, a super-buffer (SBFL) inverter is used as the driver.

During an output high state, the voltage across the pull-down MESFET will be at least 3 volts. Driving the super buffer via a smaller DCFL stage will provide approximately 100 mV for the low level on the gate of the E-MESFET. These two conditions, combined with high temperatures and/or fast process variations, results in a significant leakage current through the pull-down MESFET, adversely affecting pull-up time and voltage level. Therefore, the super buffer must be driven using logic capable of pulling down very low (lower than the typical DCFL level of about 100 mV above the low supply). The obvious choice would be to use a source-follower class, such as SDCFL. However a source follower in this case will use most power when the output of the driver is in its logic low state. For this reason, a large β (28) DCFL inverter was used, as its power dissipation all occurs when the driver is in its high output state. This inverter achieves an output low approximately 40 mV above the lowest supply. A source follower is used to drive the DCFL inverter to prevent its pull-up time being affected by leakage through the large E-MESFET pull-down. The source follower again uses most power when the driver is in its high output state. There is no need to clamp the output of the Write Word driver. A schematic of the Write Word driver is shown in Figure 6.24.

The Write Word driver has been optimised for minimum power while still maintaining adequate speed, and functions over a process range of 2σ slow to 2σ fast up to a temperature of 125 °C. Simulations of the Write Word driver at 75 °C and 125 °C over a 2σ process variation are shown in

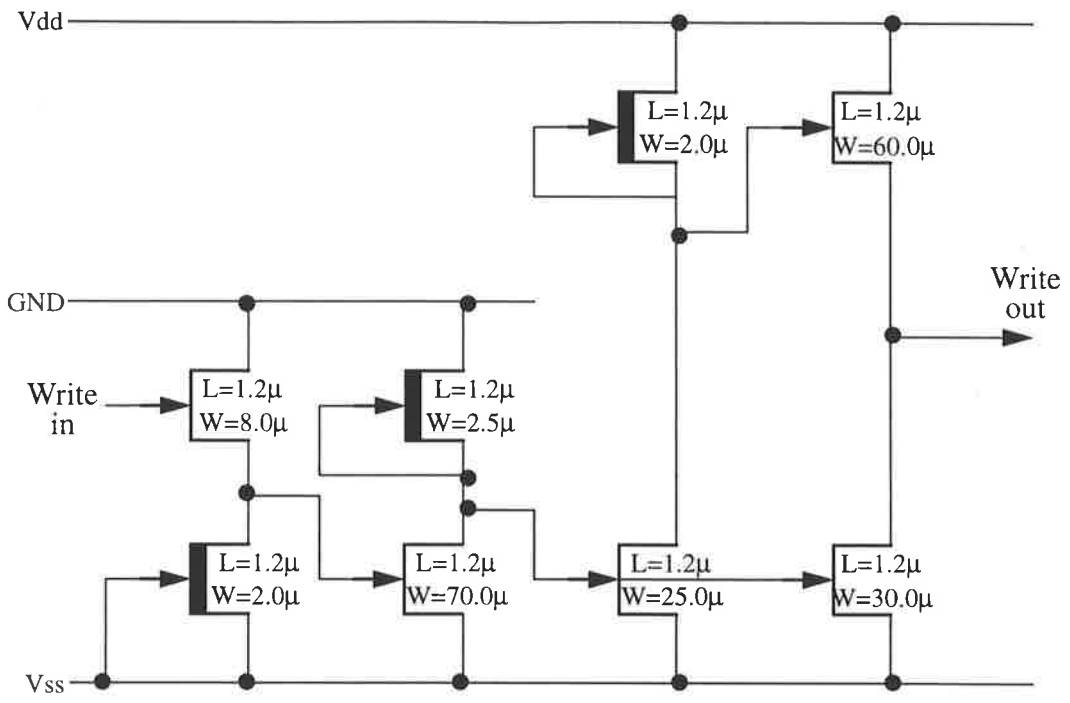


Figure 6.24. Schematic of the Write Word driver.

Figure 6.25. Also shown is the typical instantaneous power dissipation at 125 °C.

Typically, the driver dissipates 2.6 mW at 125 °C in its output low state. This gives a total of 332.8 mW of dissipation for the 128 drivers. As can be seen from Figure 6.25, the power dissipation is during a low level output is lower than that during switching and high level output.

The layout of the Write Word driver must match the horizontal pitch width of the memory cell, 25.5 μm. Once again, due to width constraints, the general *ring notation* methodology was not followed. The layout was designed such that the 25.5 μm pitch width was achieved with minimum height. The resulting layout is shown in Figure 6.26, and has dimensions of 25.5 μm × 159.6 μm. The very wide MESFETs (> 25 μm) had to be layed out in a parallel fingered format.

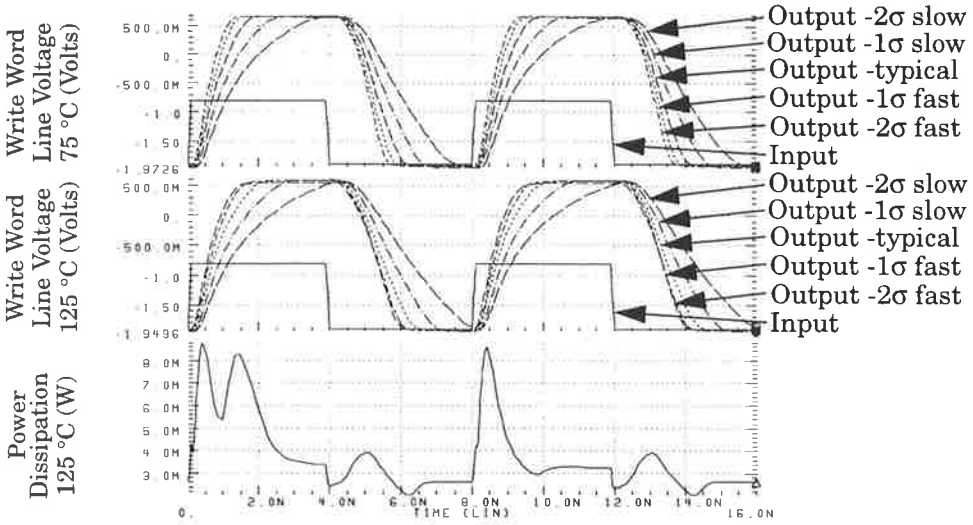


Figure 6.25. Simulation of the Write Word Driver.

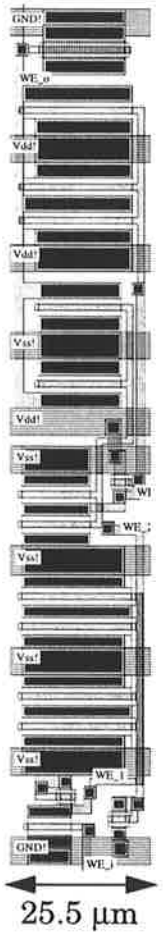


Figure 6.26. Layout of the Write Word Driver

6.2.8 Read Word Driver Design

The Read Word driver functionality is very similar to that of the Write Word driver, the only differences being that the output level of the Read Word line driver is limited to a maximum of 0 volts to prevent corruption of the read logic low signal by current flowing from the Read Word line to the Read Bit line via the forward biassed gate diode of the read transistor (Section 5.2.3.2). Also, because of the slower requirement of the read cycle, the speed of the driver is not as critical. Only one cell will be read at a time, and so again at any instant only one Read Word driver will be in the logic high position at any one time.

Because a lower output voltage than that of the Write Word driver is required, a technique can be used to further reduce the size of the V_{DD} - V_{SS} inverter, thereby reducing power dissipation. The large high output transistor of the super buffer can be driven by a super buffer itself, instead of directly from the inverter. Because of the gate diode on the high output E-MESFET, we know that the logic high output of a super buffer is approximately a diode drop lower than the logic high input. Therefore the voltage drop between the output of the driver and the high level of the inverter will be about 2 diode drops. Now the high level of the output is 0 volts, requiring the inverter to pull up to about 1.2 volts (given that a diode drop is approximately 0.6 volts). This is quite easily done as the pull-up D-MESFET remains in saturation. However, attempting the same method in the previous driver (Write Word line) was unsuccessful because the output high voltage is a diode drop itself, requiring the DCFL inverter to pull up to about 1.8 volts. The rise time will be slow because the pull-up D-MESFET is not in saturation at this voltage, and extensive leakage through the pull-down E-MESFET will increase rise time.

A schematic of the Read Word driver is shown below in Figure 6.27. The double super buffer driving stage previously discussed can be clearly seen. Once again the circuit was designed to perform down to 2σ slow process variation and at temperatures of up to 125 °C. By optimising the transistor sizes via extensive simulation, the typical driver power dissipation at

125 °C under a low output level was brought down to only 900 mW without significantly compromising the performance of the driver. This gives a total dissipation for the 128 Read Word Line drivers of 115.2 mW.

A simulation of the Read Word driver at 75 °C and 125 °C using 2σ slow to 2σ fast process variation parameters is shown in Figure 6.28. It can be seen that the driver has worst case rise and fall times of approximately 5 ns under 2σ slow process variation and 125 °C temperature. The simulation result also shows the power dissipated at 125 °C and typical process parameters. It can be seen that the power is a minimum when the driver is outputting a low value of -2 volts. This power is approximately 900 μW.

Again the pitch width of the Read Word driver must be matched to that of the memory cell, i.e. 25.5 μm. The layout was designed with this in mind, and the vertical dimension minimised. Because of the narrow nature of the layout requirement, *ring notation* was not followed. The cell is shown in Figure 6.29 and its dimensions are 109.0 μm tall × 25.5 μm wide. Because of the slower nature of the driver, significant area gains were made over the Write Word driver.

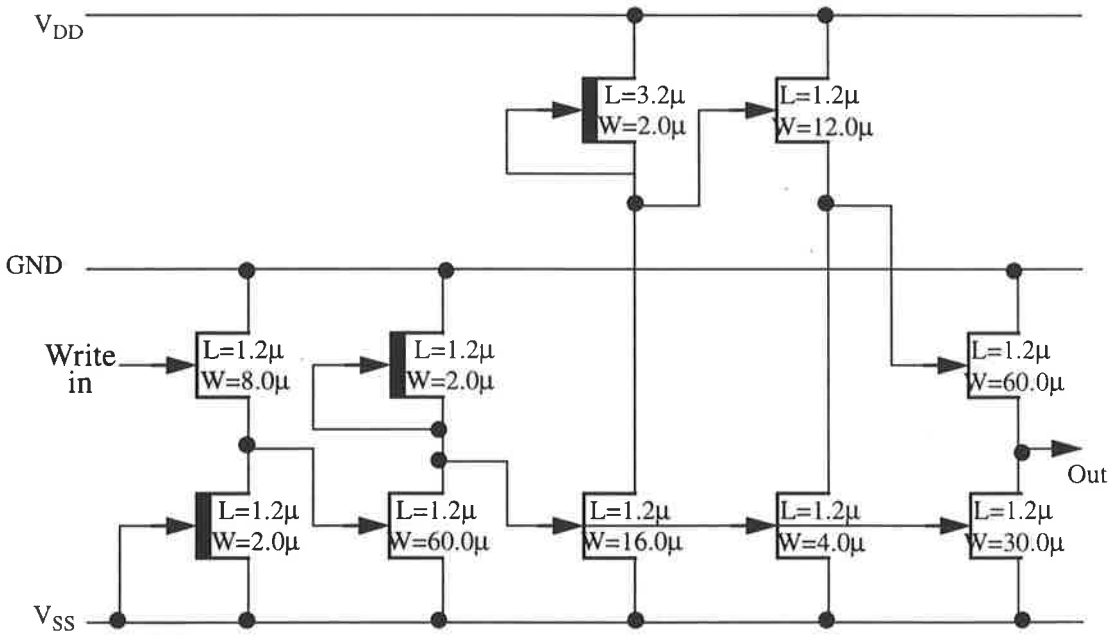


Figure 6.27. Schematic of the Read Word driver.

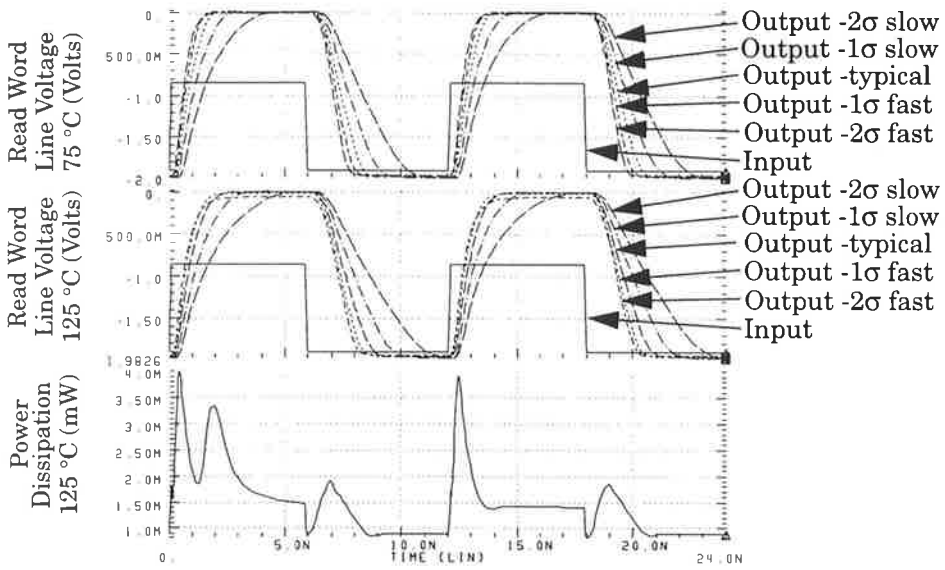


Figure 6.28. Simulation of the Read Word Driver.

6.2.9 Precharge Circuit Design

The function of the precharger is to discharge the Read Bit line before a memory cell is read. The simplest way of accomplishing this is using a transistor, with drain connected to the Read Bit line, source connected to ground and turning it on and off with a control signal to its gate. The Read Bit line consists of a capacitance of approximately 600 femtofarads. The design requires a trade-off between discharge time and transistor size. A wide transistor will discharge the line more quicker than a thinner transistor, but at the expense of area and a bigger driver. Both enhancement and depletion mode MESFETs were simulated for this role, and their advantages and disadvantages are summarised in Table 6.1.

Due to the disadvantages in using an enhancement mode transistor, particularly the inability to discharge completely to 0 volts, it was decided to use a depletion mode MESFET. It was found by simulation that a D-MESFET with a width of 6.0 μm was able to completely discharge the Read Bit bus in under 3 ns under all conditions from 2σ slow to 2σ fast process variation and at temperatures up to 125 °C. As a comparison between the two, Figure 6.30 shows the results of a simulation at 125 °C and both 2σ slow

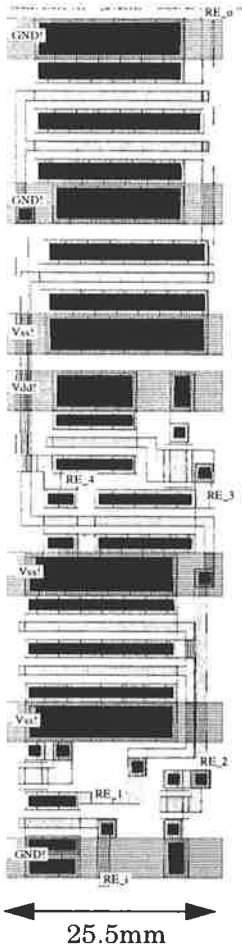


Figure 6.29. Layout of the Read Word Driver.

Table 6.1. Comparison of E-MESFET and D-MESFET as prechargers.

Transistor	E-MESFET	D-MESFET
Advantages	Control Signals easy to generate (0 - 0.6 volts)	Discharge Time faster
	Lower voltage swing - faster	Smaller FET required
		Discharges completely to 0 volts
Disadvantages	Discharge time slower	Control Signals harder to generate (-2.0 - 0 volts)
	Larger FET required	Larger voltage swing - slower
	Does not discharge completely to 0 V	

and typical parameters of 6.0 μm wide E- and D-MESFETs discharging the Read Bit bus. It can be clearly seen that the D-MESFET discharges the bus much quicker, and that the resulting low level is better than that achieved by the E-MESFET. The higher discharged value of the Read Bit bus when using the E-MESFET is due to forward conduction of the gate-drain diode. The inverse exponential nature of the discharge curve is caused by the reduction of the transistor current drain-source voltage with the discharging of the Read Bit bus.

Although the precharge transistor is only 6.0 μm wide, the cell layout is significantly wider so as to pitch match with the precharge driver. A driver similar to that of the Read Word driver, providing a -2 volt to 0 volts swing, will be used. The driver must be modified to account for the increased capacitive load, and the this is done by widening the driver transistors, and hence the driver is 50 μm wide. Hence the width of the precharge cell is 50 μm. The layout of the precharge circuit is shown in Figure 6.31. Its vertical dimension is 25.5 μm to pitch match the memory cell.

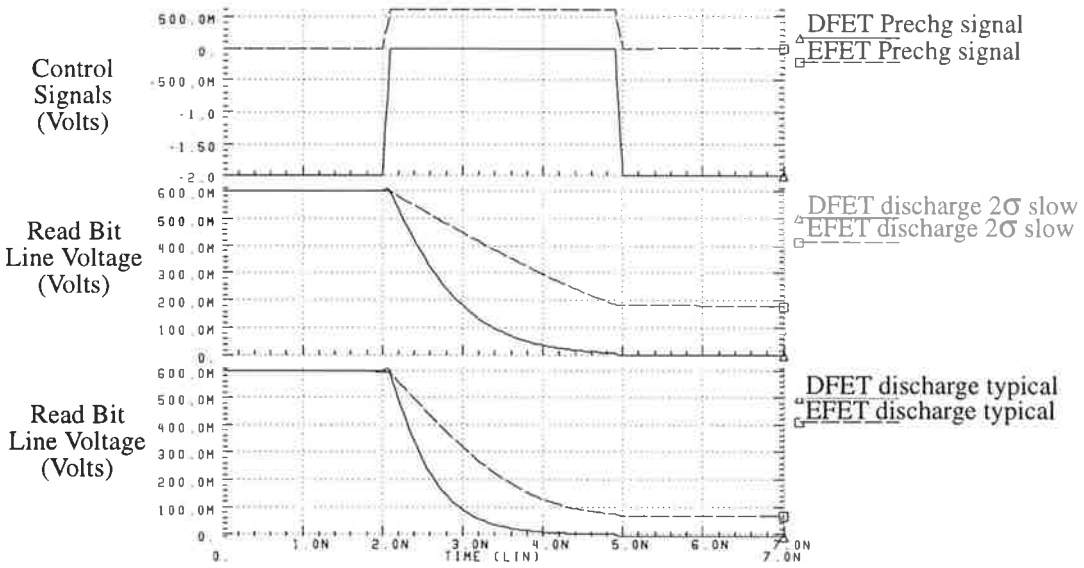


Figure 6.30. E- and D-MESFETs as Precharging transistors.

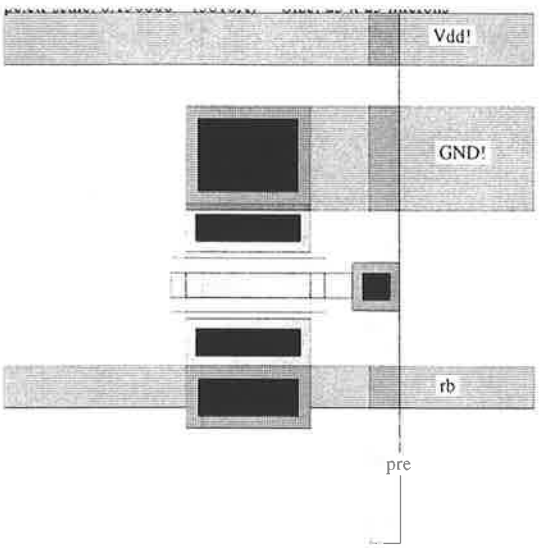


Figure 6.31. Layout of the Precharge Circuit.

6.2.10 Precharge Driver

The precharge driver must supply a -2.0 volt to 0 volt signal to the 112 precharge transistors, and is thus very similar in requirement to the Read Word driver. For this reason, a double super buffer structure (DSBFL) can be used. Because of the significantly larger width of the precharge transistors, the capacitive load to be driven is somewhat larger, approximately 1500 fF. The driver was also designed to be faster than the Read Word driver, so as to minimise the time between memory reads. Once again, the driver was designed to function over a 2σ process variation at temperatures up to 125 °C. Because of the load and speed requirements of the driver, its size is larger than that of the Read Word driver. The driver also utilises a SDCFL input stage instead of SFFL, to reduce the output high power dissipation and increase the output low power dissipation because of its more efficient design, and because the low state dissipation is not critical, as it was with the Read Word driver. A schematic of the Precharge driver is shown in Figure 6.32. Simulation results for the driver with a load of 112 precharge cells are shown in Figure 6.33. The simulations were done at 75 °C and 125 °C using 2σ slow to 2σ fast process variation models. It can be seen that the worst case rise and fall times are approximately 3 ns,

occurring when the process variation is 2σ slow and the temperature is 125 °C.

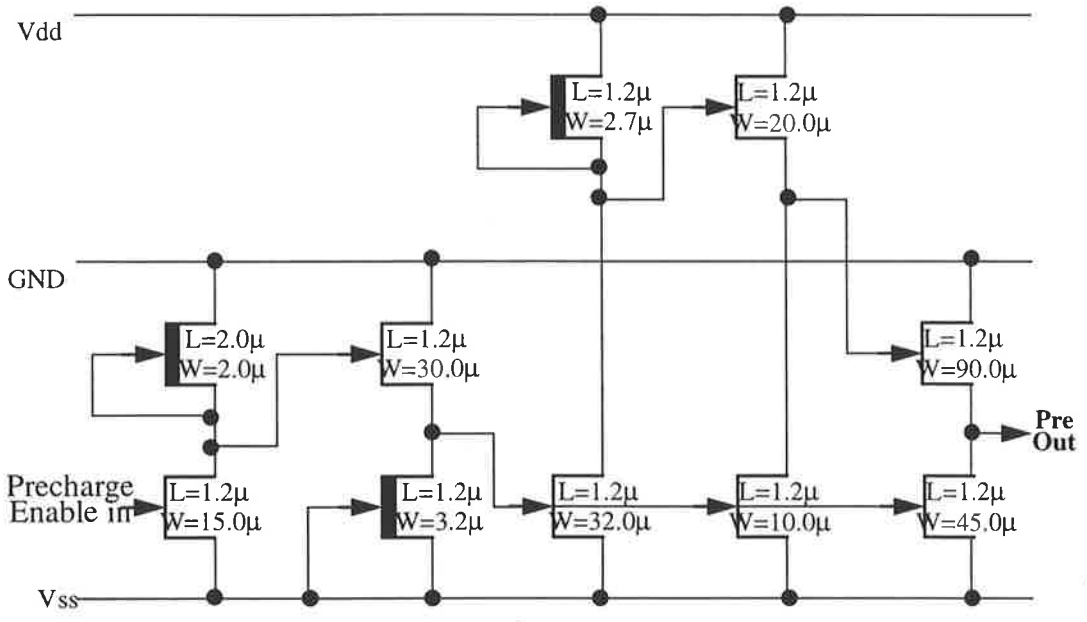


Figure 6.32. Schematic of the Precharge driver.

The layout of the Precharge driver was designed to use the supply rails provided from the Read Word drivers. To lay out the wide driving transistors, the cell width had to be extended beyond that of the Read Word driver. Further parallelisation of the drivers was not possible due to space restrictions between the supply rails. The resulting cell was layed out as in Figure 6.34. It has dimensions of 50µm wide by 109.0 µm high.

6.2.11 Precharge Decoder

We know from the discussion of memory cell operation that the Read Bit lines must be discharged to zero immediately before every read operation. The precharging will be done by supplying an enable signal to the precharge driver. This signal can be applied separately, or to simplify operation of the DRAM block, the precharge enable signal can be generated

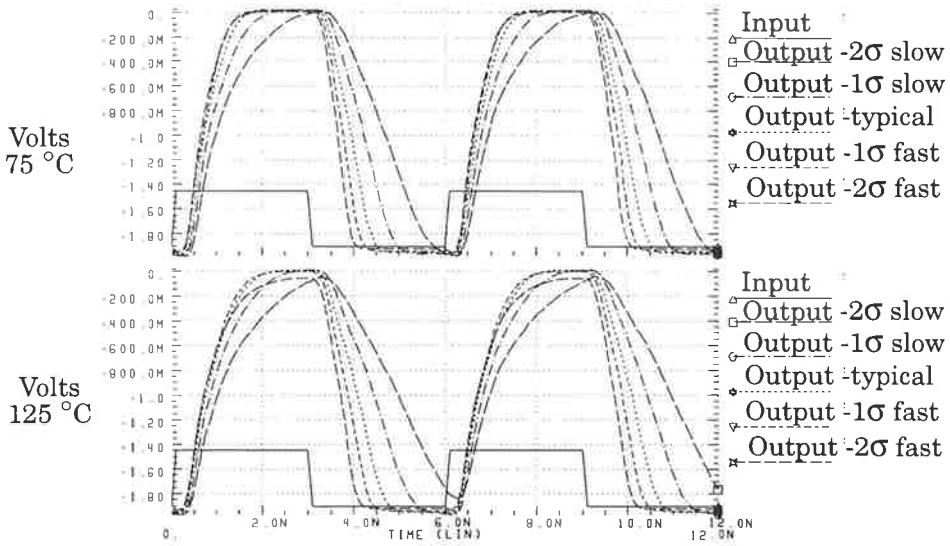


Figure 6.33. Simulation of Precharge driver.

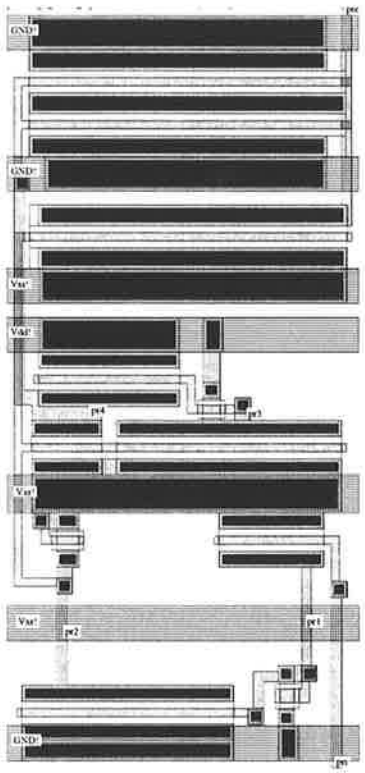


Figure 6.34. Layout of the Precharge driver.

automatically using the fact that when the memory is not being read, the precharge signal is on, and as soon as a cell is read, the precharge is switched off. Ideally, the precharge will switch off just before the Read Enable is sent to the cell, and turned on just after the Read Enable signal has returned low. This can be accomplished by NORing the enable signal input directly from the level shifters with a delayed version of the Read Enable signal from the Enable line driver which is sent to the decoders, as illustrated in Figure 6.36. A series of 9 inverters is used to generate a delay of approximately 1 nanosecond. A schematic of the precharge decoder is shown in Figure 6.35. To demonstrate the operation of the precharge system, a simulation is shown in Figure 6.37 showing the Read Enable signal sent to the memory cells and Precharge signal sent to the Precharge circuits at 75 °C and typical parameters. It can be seen that precharging is turned off before the read operation is started and turned on only once the read operation has finished.

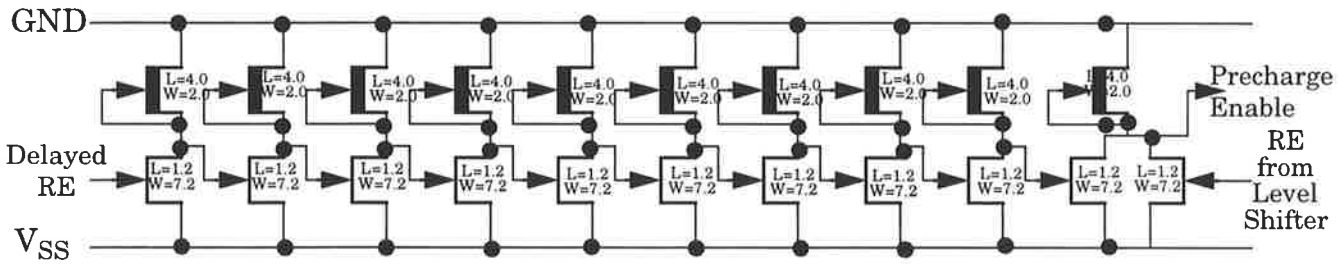


Figure 6.35. Schematic of the Precharge decoder.

6.2.12 Sense Amplifier Design

The sense amplifier must perform the opposite function to that of the Write Bit driver. It must accept the data coming from the Read Bit lines and convert it to a logic signal compatible with the circuitry outside of the memory array. Simulations such as those shown in Section 5.8 indicate that the read logic high level at higher temperatures and slow process variation can become as low as 400 mV. The problem in recognising this level is compounded by the increase in threshold voltage of MESFET devices at such process variation. The read logic low level could also be as high as 150 mV

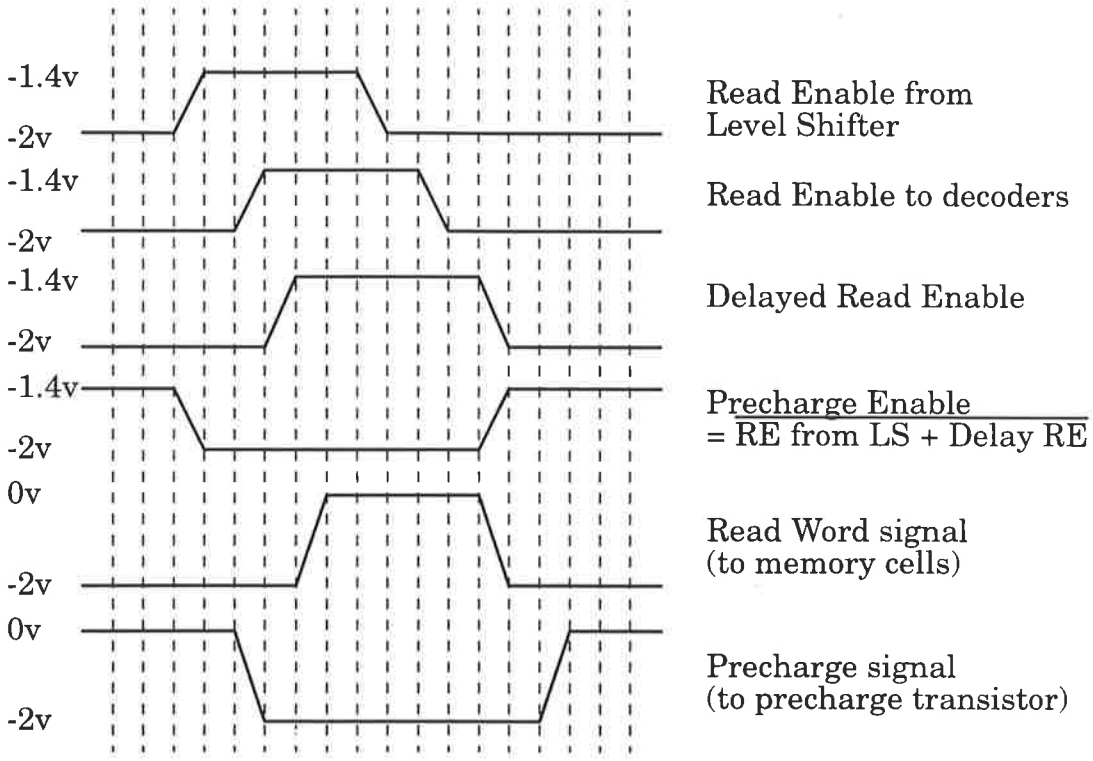


Figure 6.36. Generation of the Precharge enable.

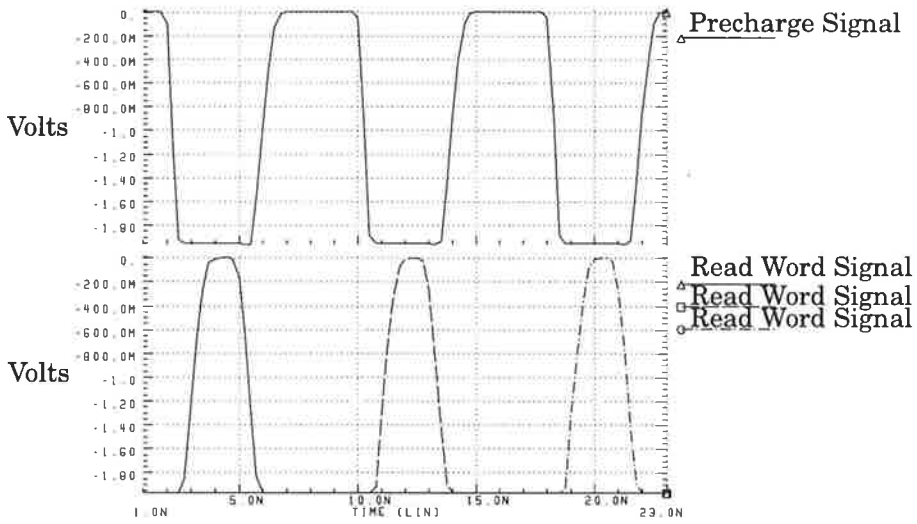


Figure 6.37. Read Word and Precharge signals at 75 °C, typical parameters.

under worst case conditions. This gives an input dynamic range of only about 250 mV, from which a signal of DCFL levels (100mV to 600 mV) in swing must be obtained. The ideal solution would, perhaps, be a differential amplifier with high gain acting as a comparator. However, this introduces the problem of adding another supply voltage of 300 mV to act as a reference voltage.

For this reason, the sense amplifier was simply designed as an inverter with high β (gain). After experimentation and simulation, a β ratio of 20 was found to provide good results. Because of the large time constant of the memory cell read cycle, the speed performance of the sense amplifier is not critical, and so a low power design was chosen. A second inverter is placed after the sense amplifier inverter to buffer the read signal. A schematic of the sense amplifier is shown in Figure 6.38. Figure 6.39 shows a simulation of the sense amplifier at 75 °C and 125 °C over a 2σ slow to 2σ fast process variation, using an input swing of 150 mV to 400 mV. In this simulation, the sense amplifier is driving a small DCFL inverter as a dummy load. It can be seen that the logic levels are correctly resolved. The sense amplifier has, under typical parameters, an average power dissipation of 360 µW at 75 °C, giving a total average of 40.3 mW for the 112 sense amplifiers.

The layout of the sense amplifier is shown in Figure 6.40. It also has a vertical pitch of 25.5 µm and the dimensions of the cell are 70.0 µm by 25.5 µm.

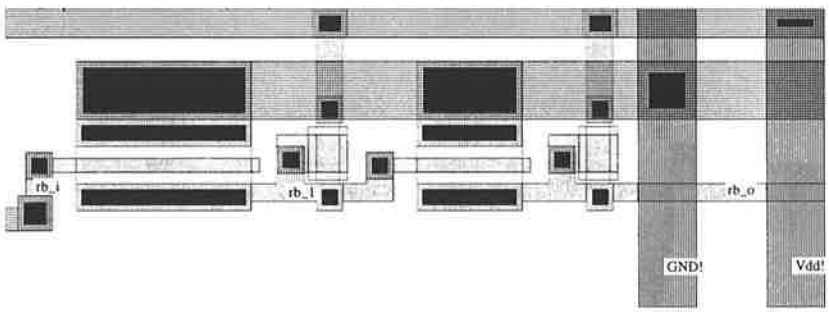


Figure 6.40. Sense Amplifier Layout.

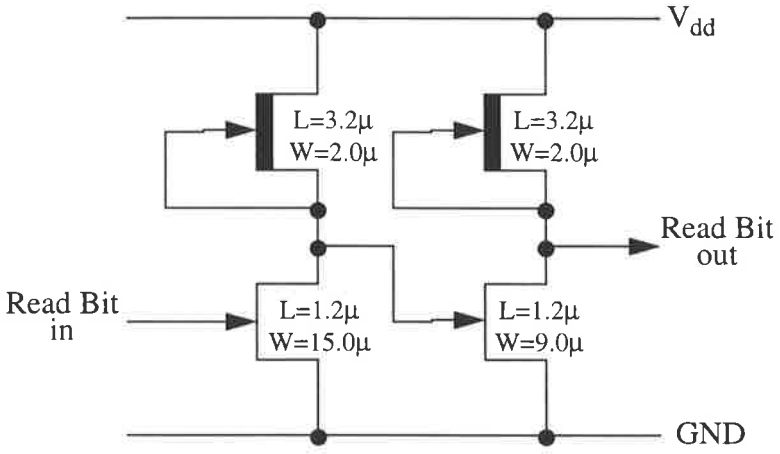


Figure 6.38. Schematic of the sense amplifier.

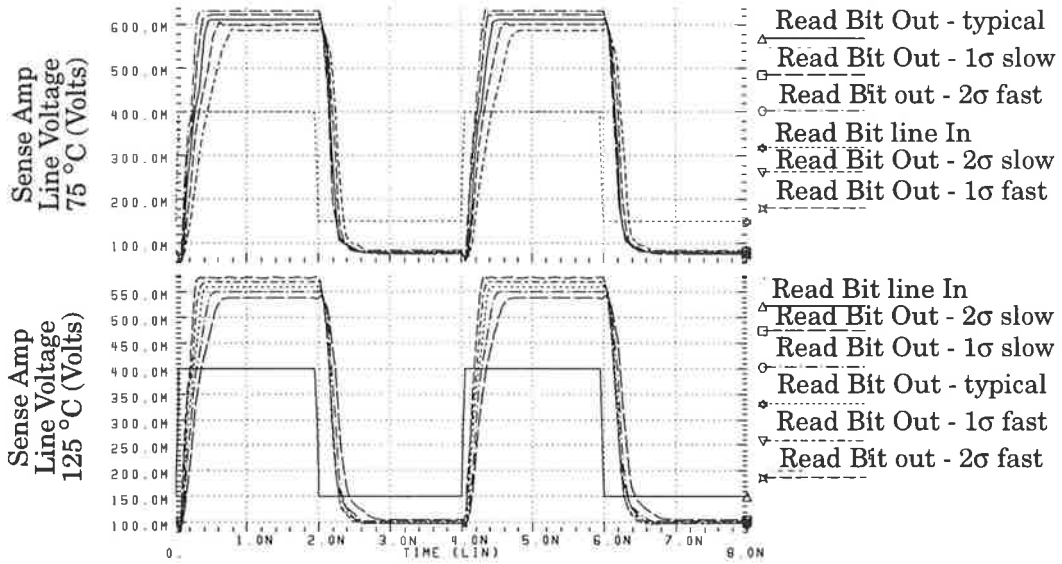


Figure 6.39. Simulation of the sense amplifier.

6.3 DRAM Array Layout

As all components of the Dynamic RAM have been designed and layed out, the full Dynamic RAM array must now be layed out using these components. The array is layed out to the floor plan discussed in Section 6.1, and shown in Figure 6.1.

6.3.1 Backgating Strategy

The DRAM design is potentially susceptible to backgating and sidegating performance degradation, due to the presence of circuits operating on a -2 to 0 volt supply range, which will be biassed negatively with respect to those operating on a 0 to 2 volt supply. Because of the low magnitude of backgate bias (-2 volts) the backgating effect will not be large. The modular nature of the memory array, also allows easy implementation of a schottky guard ring as described in [53] around the -2 volt circuits to minimise their effect on the 0 to 2 volt circuits. We can thus effectively separate the two groups of circuits and assume their effect on each other is minimal. Therefore, the 0 to 2 volt circuits, comprising Write Bit drivers, memory cells, precharge cells and Read Bit drivers were simulated with the substrate node connected to 0.6 volts. The -2 to 0 volt circuits, comprising Read and Write Word drivers, decoders, address line drivers, pre-decoders, level shifters and the precharge decoder and driver were all designed and simulated with the substrate node of each MESFET connected to -1.4 volts. In addition, these circuits were surrounded by a schottky metal ring of width 10μm to ensure backgating effect on the other circuits was minimised. This ring must be connected to a negative potential, estimated to be in the range of -5 to -10 volts from [53], to ensure proper circuit operation.

Because of the large number of connections to GND in the 0 to 2 volt section, and to -2 volts in the other section, we can be reasonably sure that the substrate in each section will be at an approximately uniform potential. Thus with this strategy, the performance of the circuit should closely match simulations.

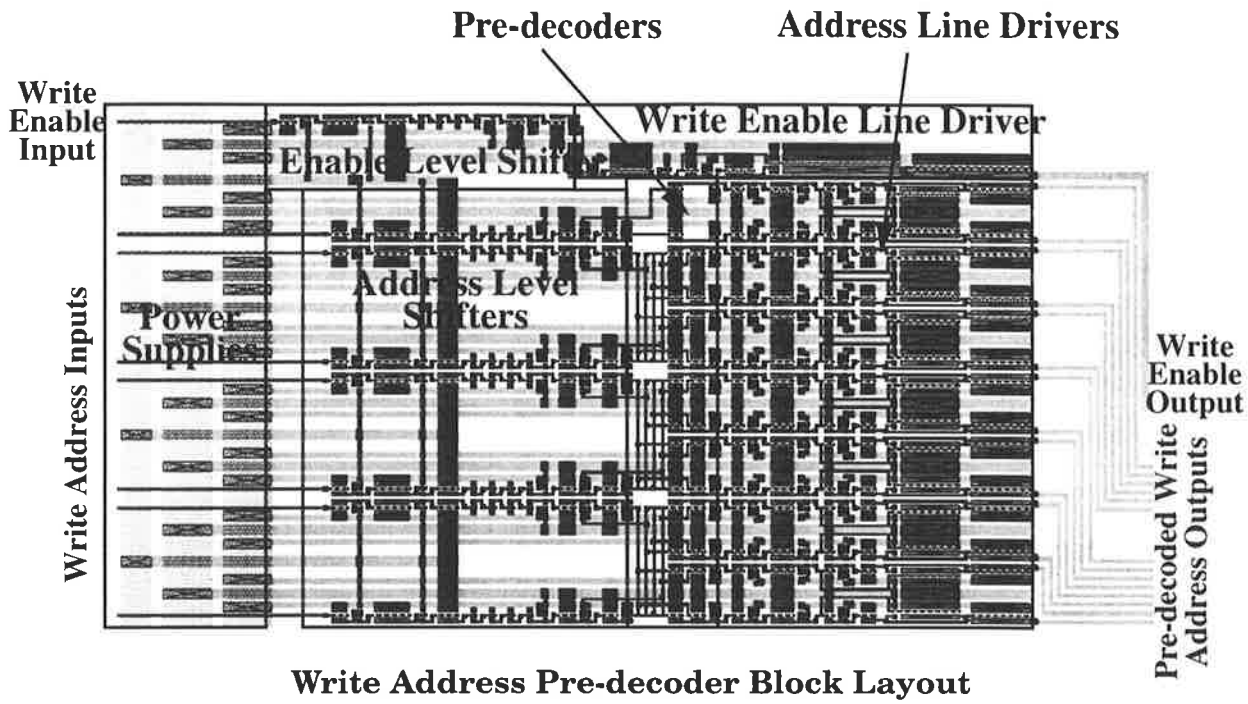
6.3.2 Pre-decoder/Driver block

The Pre-decoder driver block takes the 7 address and one enable input, does some pre-decoding on them and then drives these onto the address and enable lines which are then used by the address decoders. The module thus consists of several level shifters, pre-decoders and line drivers. The components were designed to fit with each other. To save space, the cells are paired so that they share a common supply bus. The entire pre-decoder/ driver block consists of 7 level shifters, 14 pre-decoders and 14 address line drivers for the address inputs, and a level shifter and enable line driver for the enable lines. The read block also has the precharge decoder. Layouts of the write and read pre-decoder blocks are shown in Figure 6.41, and they have transistor counts of 302 and 323 respectively.

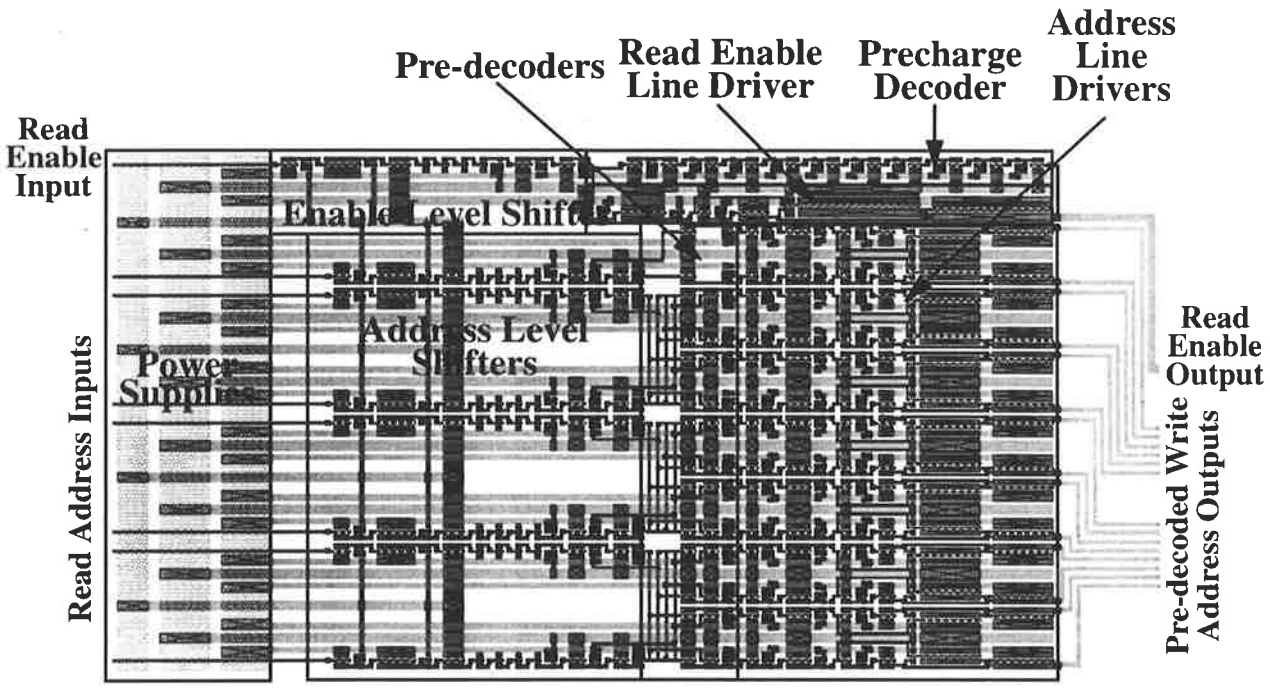
6.3.3 Read and Write Decoder/Driver blocks

The read and write decoder/driver blocks each consist of 128 individual decoders and drivers. The decoders are each connected to a unique combination of pre-decoded address lines, as discussed in Section 6.1.2, so that only one decoder is enabled for any particular combination of address inputs. The decoders then feed into their respective read or Write Word drivers which drive the memory cells.

A section of layout of both the Read and Write Word decoder/driver array is shown in Figure 6.42. Before being integrated into the DRAM, the pre-decoder/decoder/drivers were simulated to ensure that decoding was being performed correctly. This was done by incrementing the address inputs a_0 to a_6 from 0 to 127, providing an enable signal for each address and ensuring that the correct driver 'fired'. The read and write decoder/driver blocks have transistor counts of 2816 and 3200 respectively. The precharge driver has a transistor count of 123.

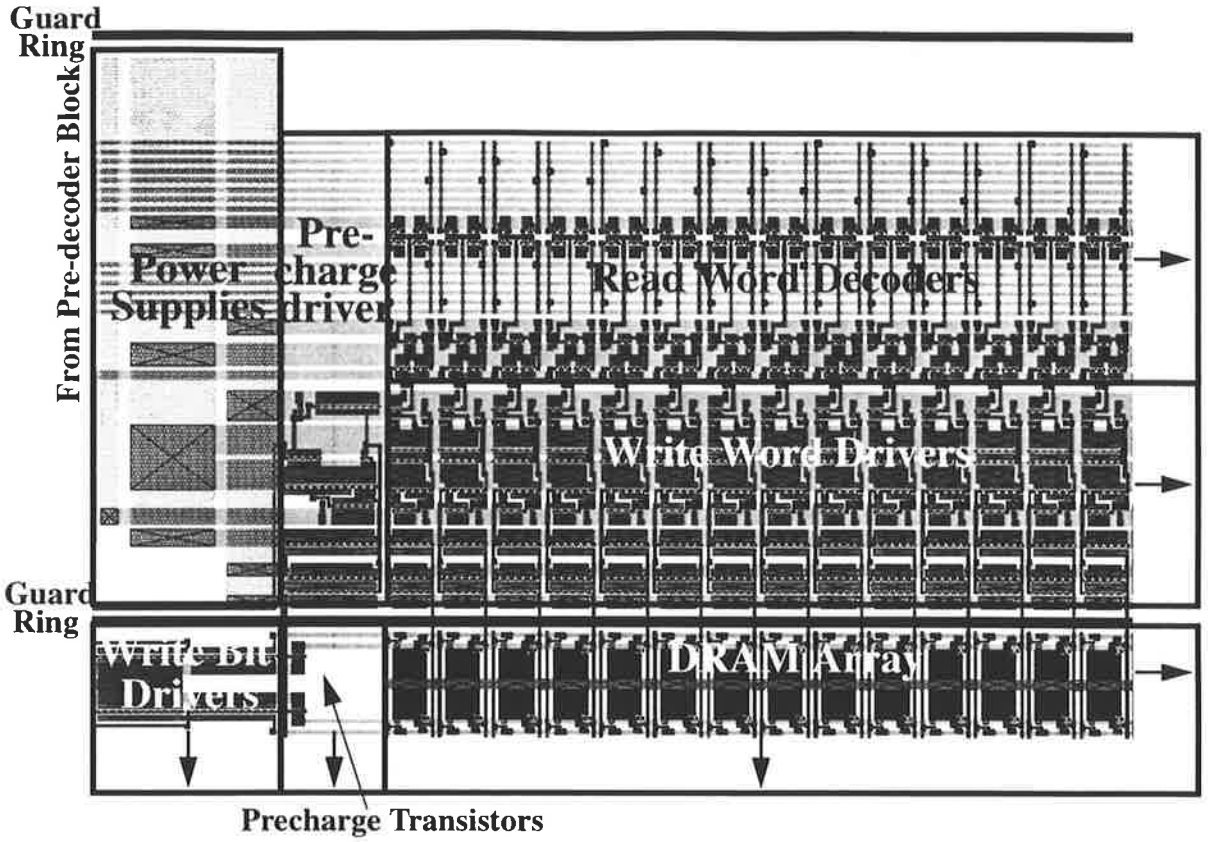


Write Address Pre-decoder Block Layout

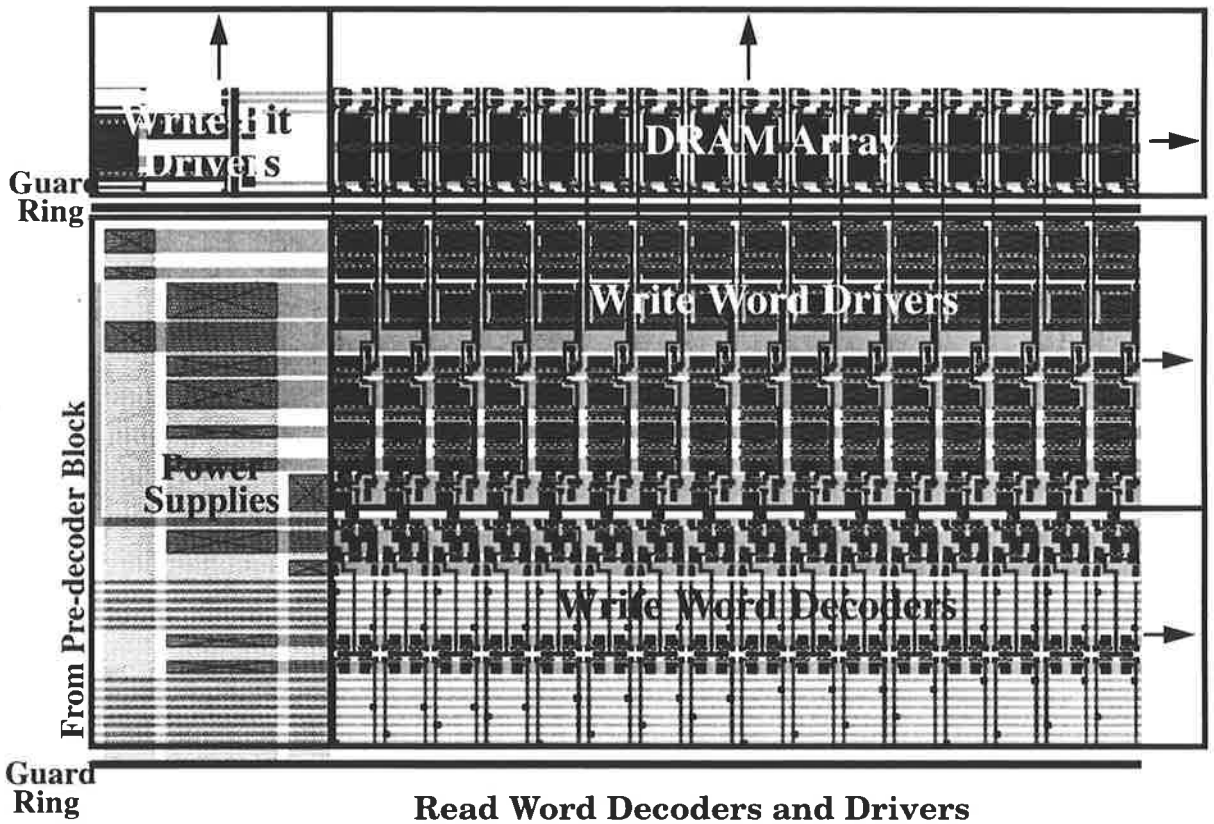


Read Address Pre-decoder Block Layout

Figure 6.41. Read and Write block layouts



Write Word Decoders and Drivers



Read Word Decoders and Drivers

Figure 6.42. Read and Write Decoders and Drivers

6.3.4 Write Bit Driver, DRAM and Sense Amplifier block

The Write bit driver section consists of 112 Write Bit drivers, taking 112 parallel lines of data and converting it into the format suitable for writing into the DRAM cells. It is then passed on to the Write Bit lines of the DRAM array, where it can be read in to any of 128 columns of memory cells by activating a particular column's Write Enable line. Any of the columns of RAM can then be read out onto 128 parallel Read Bit lines which are then passed through a sense amplifier to convert the voltage levels into a suitable form for interfacing with the rest of the buffer chip components. This is demonstrated by the layout section shown in Figure 6.43. The Write Bit driver and sense amplifier blocks have transistor counts of 672 and 448, while the 14kbit DRAM array uses 43008 transistors.

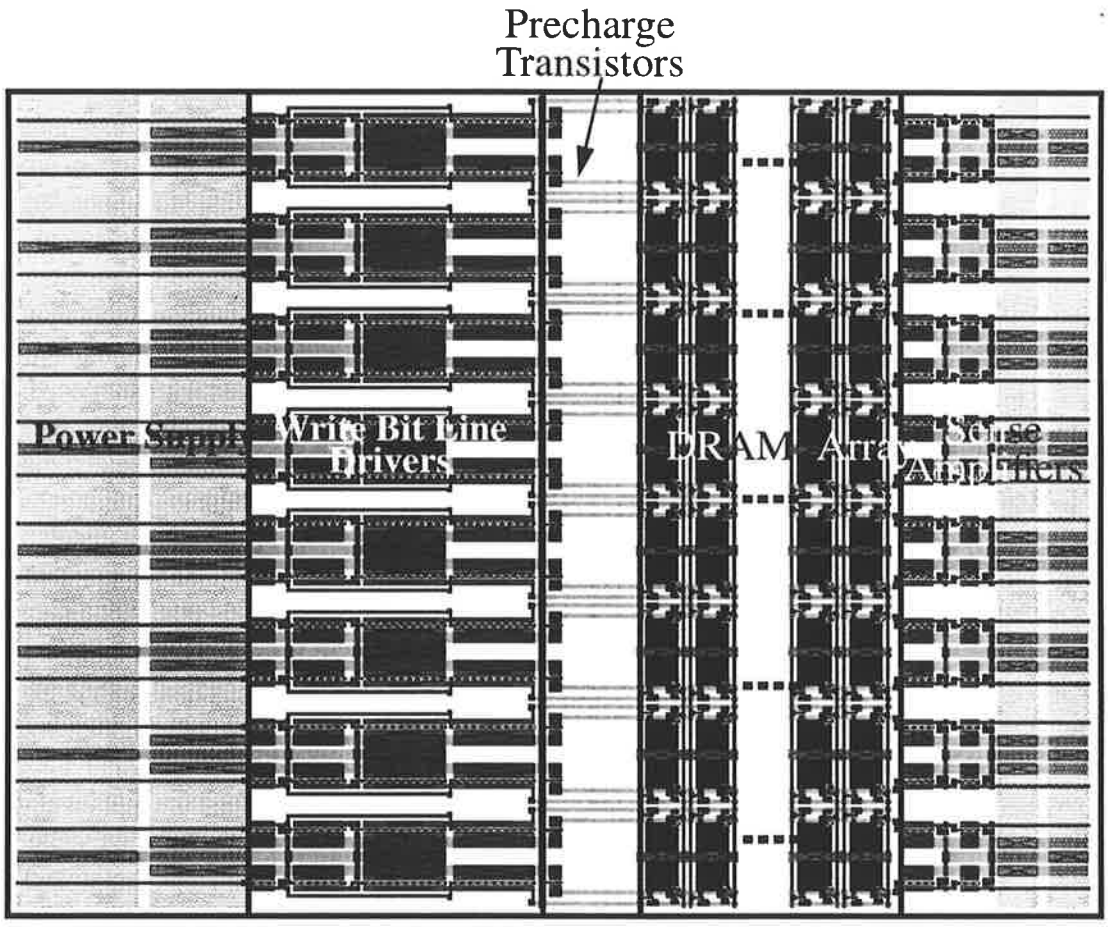


Figure 6.43. Write Bit line drivers, DRAM Array and Sense Amplifiers.

6.3.5 Full 14k DRAM Array

The full DRAM layout is shown in Figure 6.44. It has a transistor count of 50892. The total area occupied by the RAM is approximately 12.838 mm^2, giving a density of 3971 transistors/mm^2. This is of comparable density to other chips fabricated in the H-GaAs II process, even with such a large area taken up with RAM capacitors.

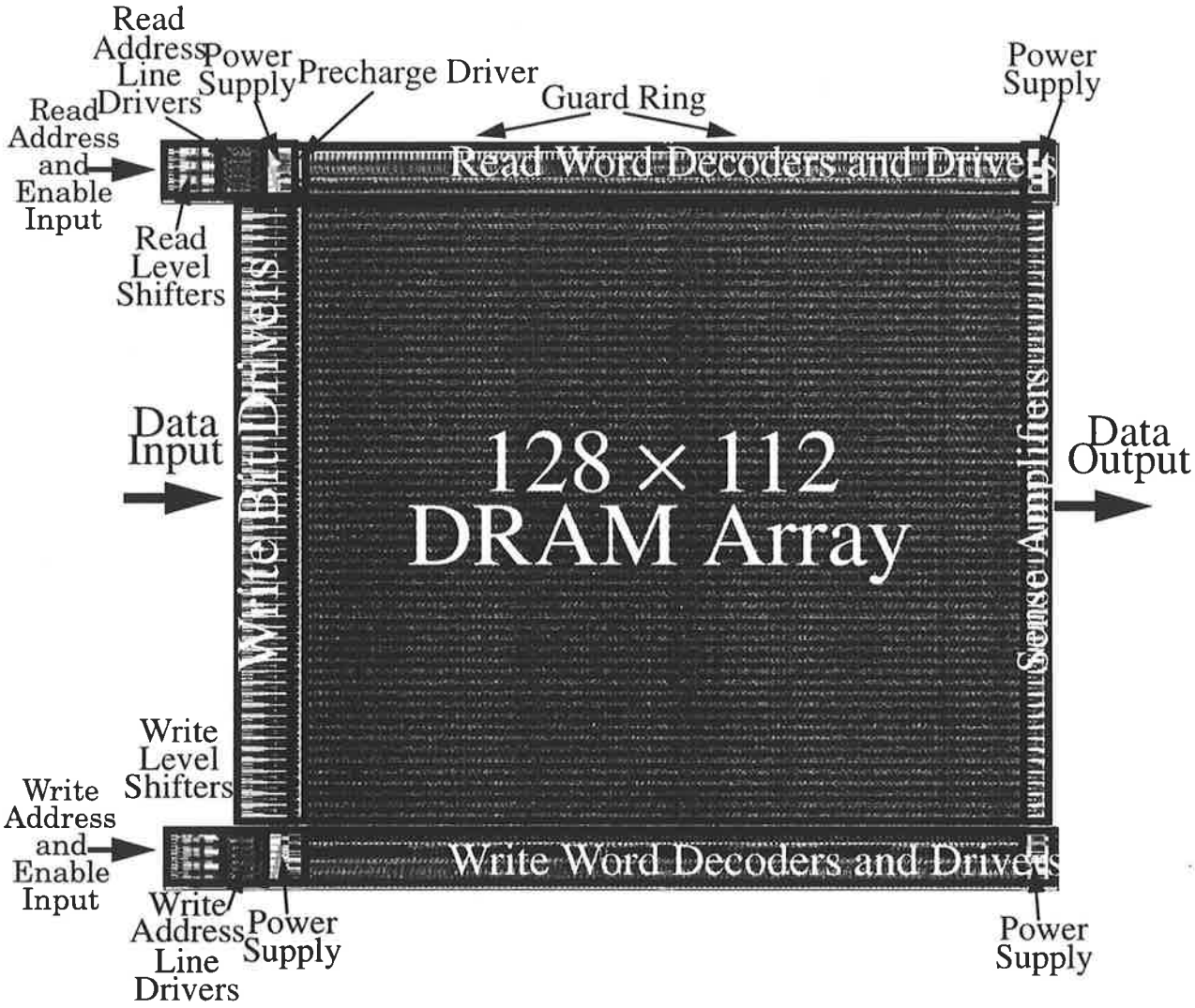


Figure 6.44. Full DRAM Array Layout

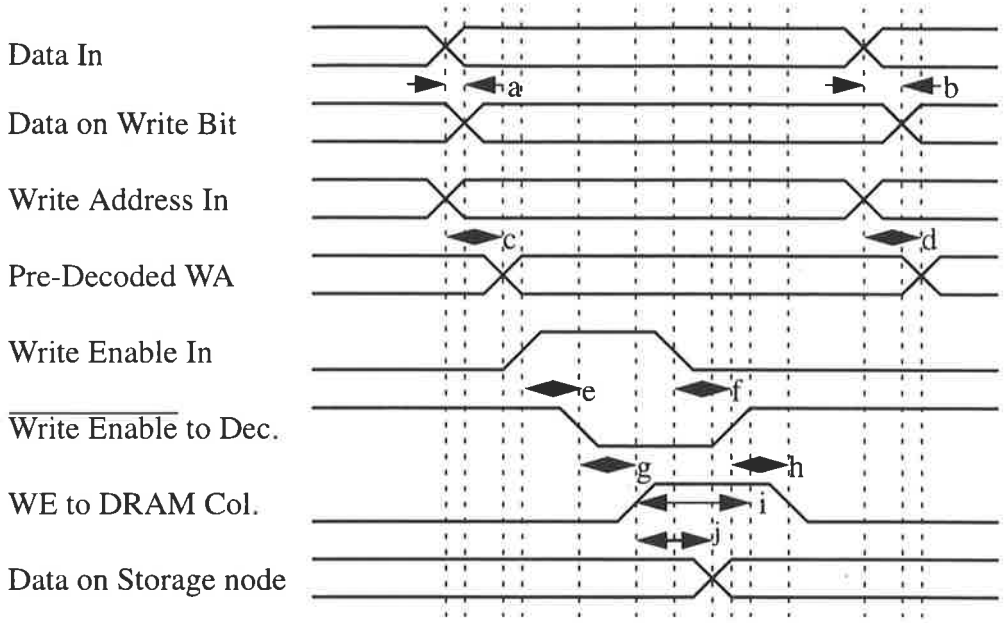
Simulations were carried out on the DRAM array over various temperatures and process variations to obtain typical, best and worst performance measures with respect to delay and read and write cycle times.

The DRAM array was fully functional over a 2σ slow to 2σ fast process variation, and over the designed operating temperature range of -25 °C to +125 °C. A timing diagram for the array is shown in Figure 6.45. Table 6.2 shows the times for the timing diagram at typical, slowest and fastest operation. It was found that worst case timing performance was obtained under conditions of 2σ slow at -25 °C, while the fastest operation is achieved at 125 °C using 2σ fast transistors. This is the reverse of the case for storage performance.

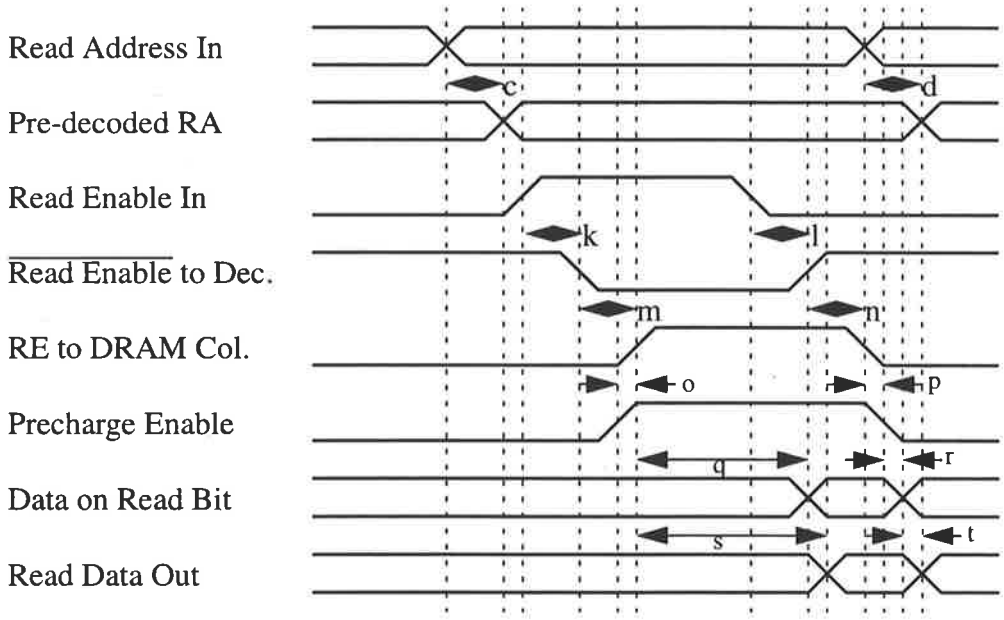
Table 6.2. Times for the timing diagram (see Figure 6.45).

Time	2σ slow, -25 °C	typical, 75 °C	2σ fast, 125 °C
a	820 ps	510 ps	410 ps
b	1.6 ns	870 ps	540 ps
c	4.2 ns	1.7 ns	1.2 ns
d	4.2 ns	1.7 ns	1.2 ns
e	3.0 ns	1.4 ns	1.0 ns
f	3.2 ns	1.4 ns	900 ps
g	1.9 ns	1.1 ns	900 ps
h	2.1 ns	1.6 ns	1.5 ns
i	2.2 ns	1.5 ns	1.3 ns
j	950 ps	500 ps	400 ps
k	3.2 ns	1.4 ns	910 ps
l	3.1 ns	1.4 ns	1.5 ns
m	2.3 ns	1.1 ns	1.3 ns
n	2.65 ns	1.7 ns	1.5 ns
o	1.25 ns	420 ps	400 ps
p	790 ps	120 ps	90 ps
q	3.7 ns	2.0 ns	1.4 ns
r	1.3 ns	620 ps	540 ps
s	8.5 ns	2.5 ns	1.5 ns
t	17 ps	280 ps	300 ps

Using the timing diagram and table of times (Figure 6.45 and Table 6.2 respectively), the worst case delay and cycle times for successful memory operation can be evaluated for the memory specifications. For the write cycle, under worst case conditions, the maximum delay from the time the data and write address are placed onto the inputs to the time the data is written is 7 ns. The read cycle delay, from when the read enable signal is asserted to when valid data is output is a maximum of 14 ns under worst case conditions. Under typical (75 °C, typical parameters) conditions, the



DRAM Write Cycle



DRAM Read Cycle

Figure 6.45. Timing diagrams for Read and Write Cycles.

delays are much smaller, with a write cycle delay of approximately 3.5 ns and a read cycle delay of about 6 ns.

A simulation under typical conditions is shown in Figure 6.46. For the sake of simplicity, the simulation result shows data rows 0 and 1 being written to DRAM columns 0 and 1 and then read out. The data word written to column 0 is '10', and the word written to column 1 is '01'. Outputs show that the read data from each column corresponds to that written, verifying correct operation of the Dynamic RAM array.

Power dissipation was also evaluated under the three operating conditions shown in Table 6.2. The typical average total power dissipation at 75 °C, typical parameters is 708 mW. The minimum power dissipation occurs under the same conditions as worst case performance at -25 °C, 2σ slow parameters, with a dissipation of 348 mW. The maximum dissipation is 1.294 W, occurring at 125 °C, 2σ fast parameters. The typical power dissipation can be broken down as shown in Table 6.3.

Table 6.3. Breakdown of power dissipation under typical conditions.

DRAM Section	Power Dissipation
Main DRAM Core	50 μW
Write Bit Drivers	78 mW
Sense Amplifiers	40 mW
Write Word Drivers	330 mW
Read Word Drivers	115 mW
Address Decoders	115 mW
Level Shifters, Pre-decoders and Address Line drivers	30 mW
Total	**726 mW**

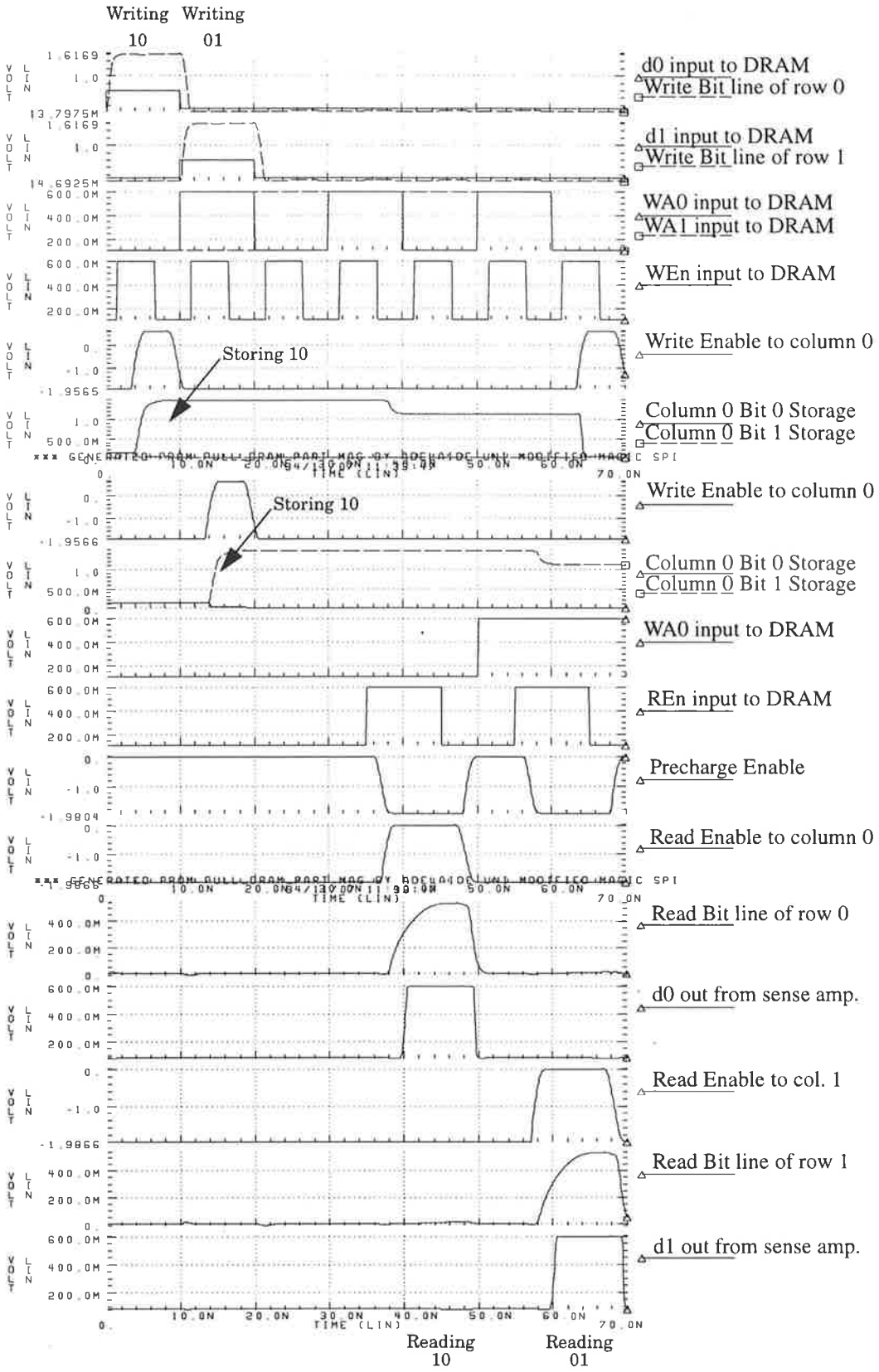


Figure 6.46. Entire DRAM block simulation of full Read and Write Cycle.

Chapter 7: Conclusion

7.1 Summary

Chapter 1 provided an introduction to the purpose of this thesis, that is the design of a Dynamic RAM suitable for use as a buffer in an ATM switch. It also discussed the motivation for the construction of the switch, and its use within a larger network.

Chapter 2 gave a detailed overview of the physical properties of gallium arsenide, along with a comparison of advantages and disadvantages over silicon. The basic MESFET device structure, operation and modelling was then discussed, along with second order effects which influence MESFET performance. A detailed examination of GaAs logic families was provided.

Chapter 3 discussed the testing of fabricated gallium arsenide MESFETs carried out by the author, in an attempt to validate the accuracy of the hspice MESFET model parameters supplied by the foundry. It was concluded that the models provided a reasonably accurate estimate of observed MESFET characteristics.

Chapter 4 provided a review of silicon memory cell design and discussed the benefits or disadvantages each different cell has in a direct translation

into a gallium arsenide implementation. Suggestions were then made on possible methods of improving the performance in GaAs. Based upon this, the memory cell which best supplies the required performance and functionality was chosen. The cell chosen was the inverted Three Transistor dynamic memory cell.

Chapter 5 gave a detailed analysis of the operation and methods of leakage of the Three Transistor memory cell, and using this analysis the performance of the cell was optimised such that the storage time is maximised over a wide range of temperature and operating conditions. The memory cell was then designed such that it satisfied storage time requirements.

Chapter 6 utilised the optimised cell obtained in Chapter 5 to construct the full 14338 bit (14kbit) memory array in a 128 word by 112 bit configuration. Included in this chapter was the detailed design and optimisation of all address decoders, word and bit drivers, sense amplifiers, level shifters and buffers required for the DRAM array. All circuits were designed to minimise power dissipation while still satisfying performance requirements, so that the operating temperature of the chip is lowered, minimising the effect of leakage currents on the memory cell storage time.

7.2 Discussion

In this thesis, the three transistor Dynamic RAM cell in gallium arsenide is investigated in detail and optimised to provide maximum storage time. Based upon this analysis, a 14kbit Dynamic RAM has been designed, layed out and simulated. The *worst case* performance of the Dynamic RAM gives a read cycle time of 14 ns and a write cycle of 7 ns, which satisfies the performance requirement by more than double. As a result, further parallelisation of the switch architecture into 16 parallel lines of 622 Mbit/s, giving a total input data rate of 9.6 Gb/s is possible, if other components of the switch can also be suitably modified.

The storage time requirement of 5.6 μs has also been met with a large safety factor included, to account for possible inaccuracies in the *hspice*

modelling. The worst case minimum storage time was simulated to be approximately 2 ms, offering a minimum safety factor of 300 times the required storage time. Under typical operating conditions the safety factor is much larger, with a storage time of about 30 ms being simulated.

The memory designed in this work is not just suitable for an ATM application. Indeed it would be ideal for any gallium arsenide design which requires a low power, high density, short term storage (or long term if a refreshing strategy is used).

While few suitable performance comparisons are available, in [88] a smaller memory array using a non-optimised three transistor GaAs cell was designed that had a storage time of above 1 ms under *typical* operation. It can be seen that through optimising the leakage currents, a significant increase in storage time has been obtained. In [94], a single transistor GaAs dynamic memory cell in a process optimised for RAM design was reported to have a room temperature storage time of 2.4 ms. Simulations indicate that our design has a *worst case* storage time of a similar value, at a much higher temperature and in a process optimised for high speed digital circuits.

The Dynamic RAM has yet to be fabricated and tested. Testing is necessary to verify the operation and performance of the memory cell. If simulation proves to be unreliable, the subthreshold and schottky diode models or parameters should be altered, based on the testing results, to provide increased simulation accuracy. Once simulation results are verified to be accurate, it may be possible to reduce or remove the additional capacitance, resulting in a faster, more compact memory cell.

References

[1] CCITT Draft Recommendation I.150, "B-ISDN ATM aspects," Geneva, Switzerland, January 1990.

[2] Thomas, A., et. al., "Asynchronous Time Division Technique: An experimental Panket Network Integrating Videocommunication," *Proc. ISS '84*, session 32, May 1984.

[3] CCITT Draft Recommenation I.361, "B-ISDN ATM Layer Specification," Geneva, Switzerland, June 1990.

[4] Tobagi, F. A., "Fast Packet Switch Architectures For Broadband Integrated Services Digital Networks," *Proc. IEEE*, Vol. 78, No. 1, pp. 133 - 166, January 1990.

[5] Minzer, S. E., "Broadband ISDN and Asynchronous Transfer Mode (ATM)," *IEEE Communications Magazine*, pp. 17 - 24, September 1989.

[6] Kawarasaki, M., "B-ISDN Architecture and Protocol," *IEEE J. Selected Areas in Communications*, Vol. 9, No. 9, pp. 1405 - 1415, December 1991.

[7] De Prycker, M., "Asynchronous Transfer Mode: Solution for Broadband ISDN," Ellis Horwood Ltd., 1991.

[8] Fischer, W, et. al., "A scalable ATM Switching System Architecture," *IEEE J. Selected Areas in Communications*, Vol. 9, No. 8, pp. 1299 - 1307, October 1991.

[9] Kozaki, T., et. al., "32 × 32 Shared Buffer Type ATM Switch VLSI's for B-ISDN's," *IEEE J. Selected Areas in Communications*, Vol. 9, No. 8, pp. 1239 - 1247, October 1991.

[10] Jakobsen, J., "Buffered Benes Networks with Bit Rate Conversion," *Proc. Australian Broadband Switching and Services Symposium 1993*, Vol. 2, pp. 363 - 370, July 1993.

[11] Jakobsen, J., "Systems Description (SYB) for High Bandwidth Packet Switch," Jydsk Telefon Internal Technical Report, Version 2, August 1993.

[12] Jakobsen, J., "Systems Specifications (SKS) for High Bandwidth Packet Switch," Jydsk Telefon Internal Technical Report, Version 2, April 1993.

[13] Patel, J. H., "Performance of Processor-Memory Interconnections for Multiprocessors," *IEEE Trans. Computers*, Vol. C-30, No. 10, pp. 771 - 780, October 1981.

[14] Benes, V. E., "Optimal Rearrangeable Multistage Connecting Networks," *Bell Systems Technical Journal*, Vol. 43, No. 7, pp. 1641 - 1656, July 1964.

[15] Katevenis, M., "Weighted Round-Robin Cell Multiplexing in a General-Purpose ATM Switch Chip," *IEEE J. Selected Areas in Communications*, Vol. 9, No. 8, pp. 1265 - 1279, October 1991.

[16] Chu, E., "Characterisation & Optimisation of Computational Functional Blocks for ATM Switches in GaAs MESFET Technology," Mas-

ters Thesis, Centre for Gallium Arsenide VLSI Technology, University of Adelaide, May 1994.

[17] Eshraghian, K., "Fundamentals of Gallium Arsenide VLSI Systems," to be published, Silicon Systems Engineering, McGraw-Hill, 1994.

[18] Williams, Ralph, "Modern GaAs Processing Methods," Artech House, Inc., 1990.

[19] Long, S. I., et. al., "Gallium Arsenide Digital Integrated Circuit Design," McGraw-Hill, Inc., 1990.

[20] Morgan, D. V., "Gallium Arsenide - a new generation of integrated circuits," *IEE Review*, September, 1988.

[21] Penzkofer, A., et. al., "Two-photon absorbtion and emission dynamics of bulk GaAs," *Optical and Quantum Electronics*, Vol. 21, pp. 283 - 306, 1989.

[22] Yuan, H. T., Lin, Y-T., and Chiang, S-Y., "Properties of Interconnection on Silicon, Sapphire and Semi-Insulating GaAs," *IEEE Trans. Electron Devices*, Vol. ED - 29, pp. 639 - 644, April 1982.

[23] Eshraghian, K., "Imaging System for Aerospace Vehicle," Centre for GaAs VLSI Technology Internal Report, University of Adelaide, February, 1992.

[24] Abrokwah, J. K., et. al., "A Manufacturable Complementary GaAs Process," *Proc. 1993 IEEE GaAs IC Symposium*, pp. 127 - 130, 1993.

[25] Bernhardt, B., et. al., "Complimentary GaAs (CGaAs): A High Performance BiCMOS Alternative," *Proc. 1995 GaAs IC Symposium*, 1995.

[26] Vitesse Semiconductor Corporation, "Foundry Design Manual Version 5.0," March 1991.

[27] Vitesse Semiconductor Corporation, "Foundry Design Manual Version 6.0," May 1993.

[28] Tyagi, M. S., "Physics of Schottky Barrier Junctions," *Metal-Semiconductor Schottky Barrier Junctions and their Applications*, Edited by B. L. Sharma, Plenum Press, New York, 1984.

[29] Sze, S. M., "Physics of Semiconductor Devices," 2nd Edition, Wiley Publications New York, 1981.

[30] Willis, A. J., "Edge Effects in Schottky Barrier Diodes," *Solid-State Electronics*, Vol. 33, No. 5, pp. 531 - 536, 1990.

[31] Ng, W. T., "Schottky Barrier Diode Characteristics Under High Level Injection," *Solid-State Electronics*, Vol. 33, No. 1, pp. 39 - 46, 1990.

[32] Meyer, E., "Defect generation by Schottky contacts on n-GaAs," *J. Vac. Sci. Technology*, Vol. 7, No. 3, June 1989.

[33] Andrews, J. M., "Reverse Current-Voltage Characteristics of Metal-Silicide Schottky Diodes," *Solid-State Electronics*, Vol. 13, pp. 1011 - 1023, 1970.

[34] Levine, J. D., "Power Law Reverse Current-Voltage Characteristic in Schottky Barriers," *Solid-State Electronics*, Vol. 17, pp. 1083 - 1086, 1974.

[35] Mayo, N., et. al., "1990 DECWRL/Livermore Magic Release," Western Research Laboratory, California USA, 1991.

[36] Meta Software, "Hspice Users Manual," 1992.

[37] George, P., et. al., "GaAs MESFET Model for Circuit Simulation," *Int. J. Electronics*, Vol. 66, No. 3, pp. 379 - 397, 1989.

[38] Goyal, R., et. al., "GaAs MESFET Model for Precision Analog IC Design," V*LSI Systems Design*, pp. 52 - 55, May 4, 1987.

[39] Schichmann, H., et. al., "Modeling and simulation of insulated gate field-effect transistor switching circuits," *IEEE J. Solid-State Circuits*, Vol. SC-3, pp. 285 - 289, September 1968.

[40] Curtice, W. R., "A MESFET Model for Use in the Design of GaAs Integrated Circuits," *IEEE Trans. Microwave Theory and Techniques*, Vol. MTT-28, No. 5, pp. 448 - 456, May 1980.

[41] Statz, H., et. al., "GaAs FET Device and Circuit Simulation in SPICE," *IEEE Trans. Electron Devices*, Vol. ED-34, pp. 160 - 169, February 1987.

[42] Troutman, R. R., "Subthreshold Design Considerations for Insulated Gate Field Effect Transistors," *IEEE J. Solid-State Circuits*, Vol. SC-9, pp. 55 - 60, April 1974.

[43] Liang, C. L., et. al., "A diffusion model of subthreshold current for GaAs MESFETs," *IEEE J. Solid State Electronics*, Vol. 34, p. 131 - 138, 1991.

[44] Larson, L. E., "An Improved GaAs MESFET Equivalent Circuit Model for Analog Integrated Circuit Applications," *IEEE J. Solid-State Circuits*, Vol. SC-22, pp. 567 - 574, 1987.

[45] Long, S. I., et. al., "Gallium Arsenide Digital Integrated Circuit Design," McGraw Hill, 1990.

[46] Kocot, C., "Backgating in GaAs MESFETs," *IEEE Trans. Electron Devices*, Vol. ED-29, No.7, pp. 1059 - 1064, July 1982.

[47] Birrittella, M. S., et. al., "The Effect of Backgating on the Design and Performance of GaAs Digital Integrated Circuits," *IEEE Trans. Electron Devices*, Vol. ED-29, No. 7, pp. 1135 - 1142, July 1982.

[48] Liu, Y., et. al., "Sidegating Effect of GaAs MESFETs and Leakage Current in a Semi-Insulating GaAs Substrate," *IEEE Electron Device Lett.*, Vol. 11, No. 11, pp. 505 -507, Nov. 1990.

[49] Fu, S., et. al, "Backgate-Induced Characteristics of Ion-Implanted GaAs MESFETs," *IEEE Trans. Electron Devices*, Vol. ED-34, p. 1245, 1987.

[50] Rao, H. V., et. al., "Sidegating Characteristics as a Function of Distance and Dose for Implanted and Metallic Sidegates on Proton-Isolated GaAs," *IEEE Trans. Electron Devices*, Vol. 35, No. 9, pp. 1558 - 1560, Sept. 1988.

[51] Kwok, H. L., "Modelling of the sidegating and the backgating effects in GaAs MESFETs", *IEEE Proc.*, Vol. 137, Pt. G, No. 6, Dec. 1990.

[52] Blum, A. S., et. al., "Use of a Surrounding p-Type Ring to Decrease Backgate Biasing in GaAs MESFETs," *IEEE Electron Device Lett.*, Vol. EDL-6, No. 2, pp. 97 - 99, Feb. 1985.

[53] Lee, C. P., et. al., "Shielding of Backgating Effect in GaAs Integrated Circuits," *IEEE Electron Device Lett.*, Vol. EDL-6, No. 4, April 1985.

[54] Chen, C., et. al., "Reduction of Sidegating in GaAs Analog and Digital Circuits Using a New Buffer Layer," *IEEE Trans. Electron Devices*, Vol. 36, No. 9, pp. 1546 - 1556, Sept. 1989.

[55] Eshraghian, K., "Design Methodology and Layout Style for Very High Speed Circuits and Subsystems," Internal Report GAAS-92-4, Centre for GaAs VLSI Technology, University of Adelaide, South Australia, January 1992.

[56] Chu, E., et. al., "Comparison of GaAs Static logic Families," *Proc. 11th Norchip Conference*, pp. 93 - 98, 1993.

[57] Eshraghian, K., et. al., "Gallium Arsenide Design Methodology & Performance Estimates for Very High Speed Circuits Using Normally-Off Classes of Logic," *Proc. 12th Australian Microelectronics Conference*, pp. 227 - 232, October 1993.

[58] Eshraghian, K., et. al., "Speed-Area-Power Optimisation for DCFL and SDCFL Classes of Logic Using Ring Notation," *Proc. 32nd Microprocessing and Microprogramming Conference*, North Holland, pp. 75 -82, 1991.

[59] Hoe, D. H., et. al., "Dynamic GaAs Capacitively Coupled Domino Logic (CCDL)," *IEEE J. of Solid-State Circuits*, vol. 26, pp. 844 - 849, June 1991.

[60] Abdel-Mortaleb, I., et. al., "GaAs Inverted Common Drain Logic (ICDL) and its Performance Compared with other GaAs Logic Families," *Solid-State Electronics*, Vol. 30, pp. 403 - 414, 1987.

[61] Long, S. I., et. al., "MSI High Speed, Low Power, GaAs Integrated Circuits using Schottky Diode FET Logic," *IEEE Trans. Microwave Theory and Techniques*, Vol. MTT-28, No. 5, p. 466, May 1980.

[62] Vu, T., et. al., "GaAs Source-Coupled FET Logic," Proc. IEEE International *Symposium on Circuits and Systems*, Vol. 3, pp. 1129 - 1130, May 1986.

[63] Katsy, S., et. al., "A Source Coupled FET Logic - A New Current Mode Approach to GaAs Logic," *IEEE Trans.Electron Devices*, Vol. ED-32, No. 6, pp. 1114 - 1118, 1985.

[64] Pucknell, D. A., et. al., "Basic VLSI Design: Systems and Circuits," Second Edition, Silicon Systems Engineering Series, Prentice Hall Australia, 1988.

[65] Mellor, P. J. T., et. al., "Capacitor-Coupled Logic using GaAs Depletion mode FETs," *Electron Letters*, Vol. 16, p. 749, September 1980.

[66] Peczalski, A., et. al., "Design Analysis of GaAs Direct Coupled Field Effect Transistor Logic," *IEEE Trans. Computer-Aided Design*, Vol. CAD-5, pp. 266 - 273, 1986.

[67] Fulkerson, D. E., "Feedback FET Logic: A Robust, High-speed, Low-power GaAs Logic Family," *IEEE J. Solid-State Circuits*, Vol. 26, pp. 70 - 74, 1991.

[68] Duncan, R. A., et. al., "Gallium Arsenide Pseudo-Current-Mode Logic," *Electronics Letters*, Vol. 26, No. 25, December 1990.

[69] Nary, K. R., et. al., "GaAs two-phase Dynamic FET logic: A Low power logic family for GaAs VLSI," *IEEE J. Solid State Circuits*, Vol. 27, pp. 1364 - 1371, October 1992.

[70] Lassen, P. S., et. al., "Ultralow-Power GaAs MESFET MSI Circuits Using Two-Phase Dynamic FET Logic," *IEEE J. of Solid-State Circuits*, Vol. 28, No. 10, October 1993.

[71] Pasternak, J. H., et. al., "GaAs MESFET Differential Pass Transistor Logic," *Electronics Letters*, Vol. 26, No. 19, September 1990.

[72] Larue, G., et. al., "FET FET Logic: A High Performance, High Noise Margin E/D Logic Family," *GaAs IC Symposium*, pp. 223 - 226, 1990.

[73] Hill, C. F., "Noise Margin and Noise Immunity in Logic Circuits," *Microelectronics*, Vol. 1, pp. 16 - 22, April 1968.

[74] Bayruns, R., et. al., "Design Issues Facing Analog Circuit Design Using GaAs MESFET Technology," Proc. *IEEE International Symposium on Circuits and Systems*, 1987.

[75] Hsu, M. S., "Aspects of Designing a High Speed Analog to Digital Converter," Master Of Engineering Science Degree Thesis, Dept. of Electrical and Electronic Engineering, University of Adelaide, September 1992.

[76] McGeever, M. K., et. al., "Testing the ATM test chip," Internal Report GAAS-94-1, Centre for GaAs VLSI Technology, University of Adelaide, February 1994.

[77] Stein, K. U., et. al., "Storage Array and Sense/Refresh Circuit for Single Transistor Memory Cells," *IEEE J. Solid-State Circuits*, Vol. SC-7, No. 5, pp. 336 - 340, October 1992.

[78] Lu, N.C., et. al., "Half-V_{DD} Bit-Line Sensing Scheme in CMOS DRAM's," *IEEE J. Solid-State Circuits*, Vol. SC-19, No. 4, pp. 451 - 454, August 1984.

[79] Makino, H., et. al., "A 7ns/850mW GaAs 4Kb SRAM Fully Operative at 75 °C," *Proc. IEEE GaAs IC Symposium*, pp. 71 - 74, November 1988.

[80] Hirayama, M., et. al., "A GaAs 16-Kbit Static RAM Using Dislocation-Free Crystal," *IEEE Trans. Electron Devices*, Vol. ED-33, No .1, pp. 104 - 110, January 1986.

[81] Bert, G., et. al., "High-Speed GaAs Static Random-Access Memory," *IEEE Trans. Electron Devices*, Vol. ED-29, No. 7, pp. 1110 - 1115, July 1982.

[82] Gabillard, B., et. al., "A GaAs 1K SRAM with 2ns Cycle time," *1987 ISSCC Digest of Technical Papers*, pp. 136 - 137, February 1987.

[83] Drangeid, K., et. al., "A Memory-Cell Array With Normally Off-Type Schottky-Barrier FET's," *IEEE. J. Solid-State Circuits*, Vol. SC-7, No. 4, pp. 277 - 282, August 1972.

[84] Tanino, N., et. al., "A 2.5nx/200mW GaAs 4Kb SRAM," *Proc. IEEE GaAs IC Symposium*, pp. 101 - 104, 1986.

[85] Lund, D., et. al., "Design of a Megabit Semiconductor Memory System," *Proc. AFIPS Fall Joint Computer Conference*, pp. 53 - 62, December 1970.

[86] Regitz, W. M., et. al., "A Three-Transistor Cell, 1024-bit, 500 ns MOS RAM," *ISSCC Dig. Tech. Paper*, pp. 42 - 43, February 1970.

[87] Terman, L. M., "MOSFET Memory Circuits," *Proc. IEEE*, Vol. 59 No. 7, pp. 1044 - 1058, July 1971.

[88] Jakobsen, J., "GaAs Dynamic Memory for Full Custom VLSI," *Proc. 11th Norchip Conference*, pp. 99 - 106, Trondheim, November 1993.

[89] Dennard, R. H., "Field-Effect Transistor Memory," U. S. Patent 3 387 286, June 1968.

[90] Yamada, T., et. al., "A 64-Mb DRAM with meshed power line," *IEEE J. Solid State Circuits*, Vol. 26, No. 11, pp. 1519 - 1529, November 1991.

[91] Kirihata, T., et. al., "A 14-ns 4-Mb CMOS DRAM with 300-mW active power," *IEEE J. Solid State Circuits*, Vol. 27, No. 9, pp. 1222 - 1228, September 1992.

[92] Hidaka, H., et. al., "A 34-ns 16-Mb DRAM with controllable voltage down-converter," *IEEE J. Solid State Circuits*, Vol. 27, No. 7, pp. 1020 - 1027, July 1992.

[93] Prince, B., "Semiconductor Memories: A handbook of Design, Manufacture & Application," 2nd Ed., John Wiley & Sons 1991

[94] Dungan, T. E., et. al., "One-Transistor GaAs MESFET- and JFET-Accessed Dynamic RAM Cells for High-Speed Medium Density Applications," *IEEE Trans. Electron Devices*, Vol. 37, No. 7, pp. 1599 - 1607, July 1990.

Appendix A: Comparison of Driving Logic Families

This appendix details the process of comparison of logic families suitable for driving the pre-decoded address lines, as discussed in Section 6.2.5. The logic families consist of common driving families, plus a new variation on the superbuffer family, here called Double Super Buffer FET logic. The classes evaluated were:

- Source-follower Direct Coupled FET Logic (SDCFL)
- Super Buffer FET Logic (SBFL)
- Ultra Buffer FET Logic (UBFL)
- Double Super Buffer FET Logic (DSBFL)

The circuits were all designed to drive an identical load, that of 128 enable signal inputs to address decoders. This represents an equivalent load of:

- A maximum of 1010 femtofarads capacitance, and
- A fan-out of 128 enhancement mode MESFETs with length 1.2 μm and width 6.4 μm.

This is a very large load to drive, and the biggest encountered in the DRAM design. To compare the circuits fairly, each one was designed such that it drove the load with approximately equal rise and fall times. They were then evaluated on the dissipated power, logic levels, noise margin and area used, and an optimal driver decided on. Unlike the read and write word drivers, these drivers will spend equal times (on average) in their respective high and low output states and so the average power dissipated was used to evaluate them.

The circuits were designed to function over a process variation of $\pm 2\sigma$ and at temperatures of up to 125 °C. As usual the 2σ slow process variation at 125 °C represented the worst case in all designs.

A.1 Design of the SDCFL driver

A schematic of the SDCFL driver designed is shown in Figure A.1. The

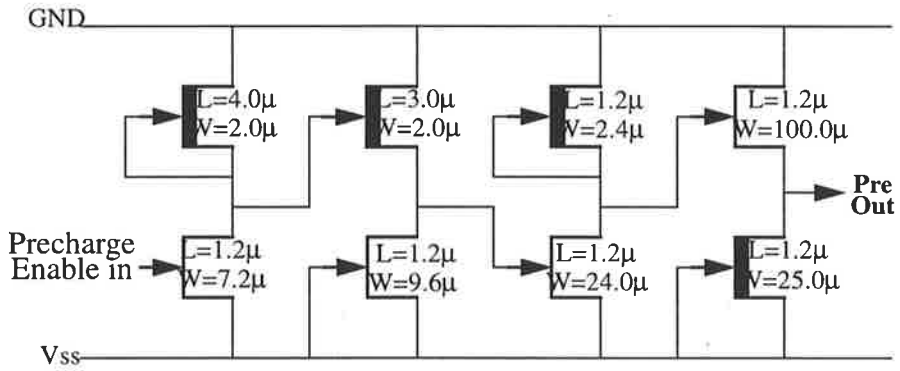


FIGURE A.1. Schematic of the SDCFL driver

SDCFL driver uses a large source follower output stage. The pull-down D-MESFET is permanently turned on, resulting in a low pull-down level, but high dissipation in the output high level. The pull-down transistor is smaller than the pull-down's in the other drivers evaluated, due to the better current sinking capability of the depletion mode MESFET relative to the enhancement-mode MESFET. The down side of this is that the pull-up MESFET must be very wide to obtain an acceptable pull-up to pull-down ratio and hence adequate logic high output value. This large MESFET

requires a large inverter stage to drive it, further increasing power dissipation. A simulation of the SDCFL driver at 75 °C using 2σ slow to 2σ fast transistor models is shown in Figure A.2.

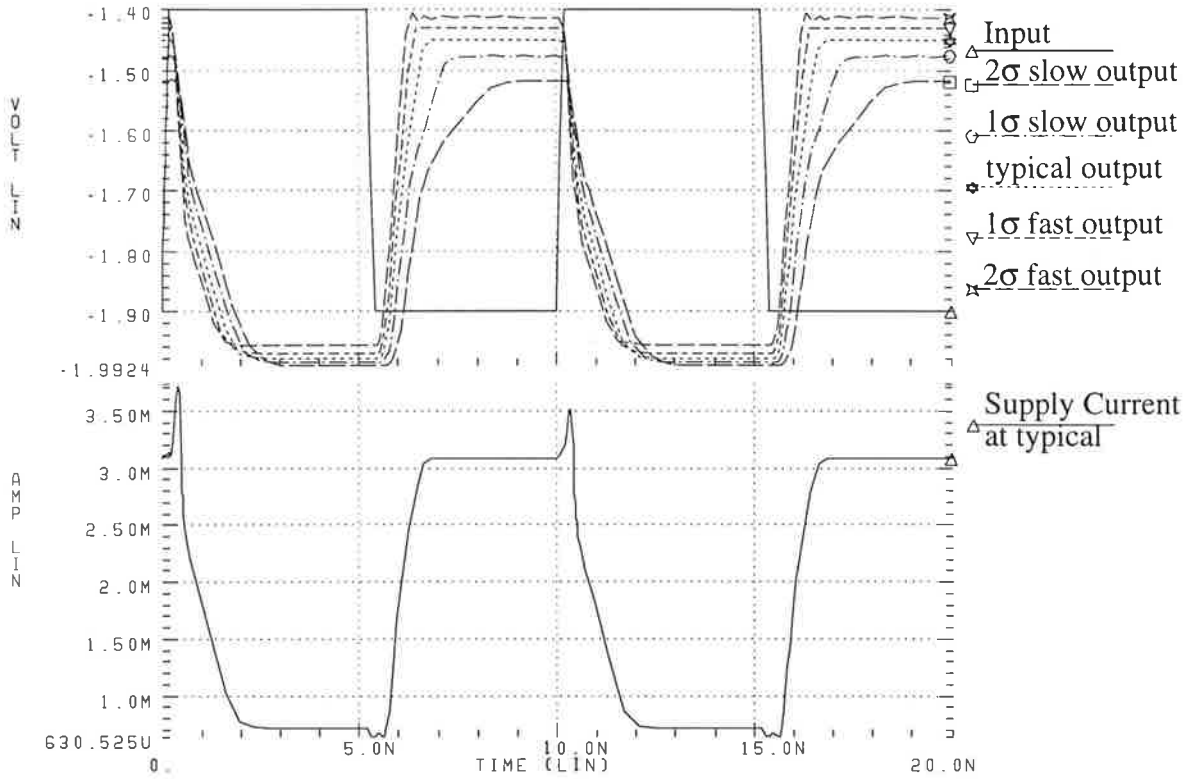


FIGURE A.2. Simulation of the SDCFL driver

A.2 Design of the SBFL driver

The SBFL driver schematic is shown in Figure A.3. It uses two large enhancement mode MESFETs, one each for pull-up and pull-down as its output stage. These MESFETs are very wide, due to their relatively small current sourcing/sinking capability, and hence a large inverter stage is required to drive them. However, due to the fact that only one output transistor is on at a time, the dissipation is significantly lower than that of the SDCFL driver.

Due to the single inverter delay driving the pull-up transistor relative to the pull-down transistor, when the driver is pulling down, both pull-up and pull-down transistor on momentarily, resulting in a current spike in the supply. For this reason, SBFL circuits are regarded as noisy. Efforts have

been made to avoid this effect by using UBFL, but this introduces other detrimental effects to the driver performance, as discussed in the next section. A simulation of the SBFL driver is shown in Figure A.4.

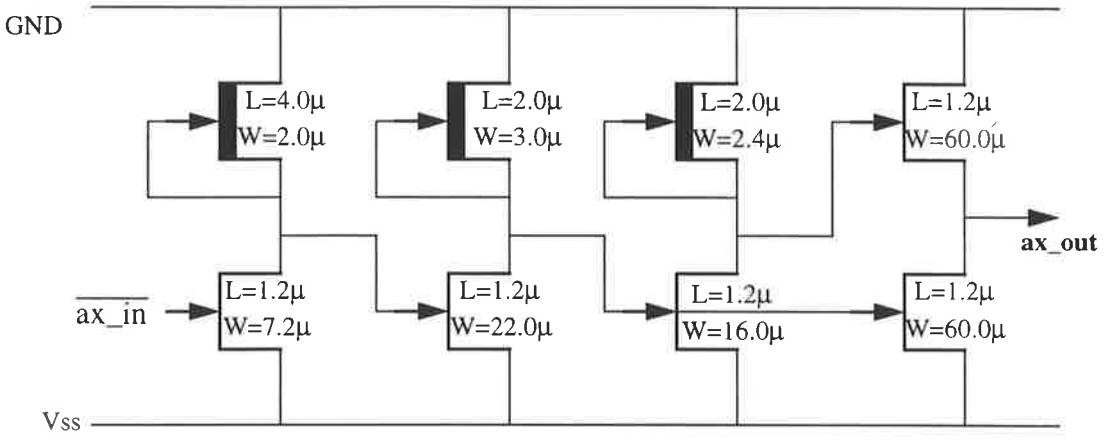


FIGURE A.3. Schematic of the SBFL driver

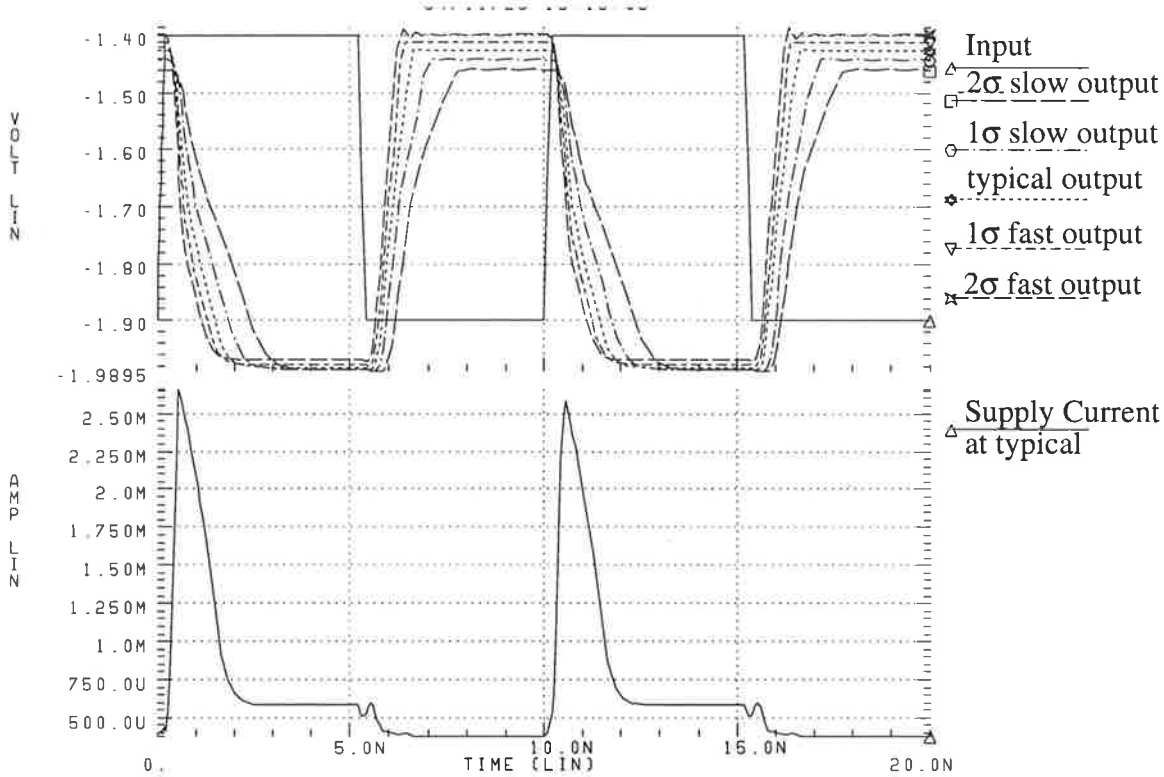


FIGURE A.4. Simulation of the SBFL driver

A.3 Design of the UBFL driver

A schematic of the UBFL driver designed is shown in Figure A.1. Due to the structure of the UBFL circuit, some serious performance flaws were found affecting both the pull-up and pull-down operation:

- During pull-up, the large pull-up enhancement mode MESFET is turned off slowly, resulting in a slower pull-up time. The high level is maintained by a small, permanently on, depletion mode transistor and as a result the high level is lower than both SBFL and DSBFL.
- During pull-down, the depletion mode pull-up is still on, resulting in a DCFL inverter effect at the output, and hence an increased output low voltage relative to SBFL, SDCFL and DSBFL. By reducing the size of the pull-up DFET, the pull down voltage can be reduced, but the pull-up voltage is also reduced and the rise-time is increased.

It should be noted that the performance of UBFL improves significantly when the fan-out (and hence current drawn by the load) is reduced, and a predominantly capacitive load is present. This is because of the limited current supply capability of the small pull-up DFET. The circuit was simulated at 75 °C, using 2σ slow to 2σ fast models. The results of the simulations are shown in Figure A.6. Extra static power dissipation can be observed due to the NOR gate. There is also a large current spike as the capacitance is discharged.

A.4 Design of the DSBFL driver

The DSBFL driver was designed with similar sizing to the UBFL and SBFL drivers previously discussed. The input stage was a double DCFL buffer, both stages with a β ratio of 12, because of its superior performance (particularly at slow process variation) and lower dissipated power when compared to an SFFL and SDCFL buffers. The input stage also allows the

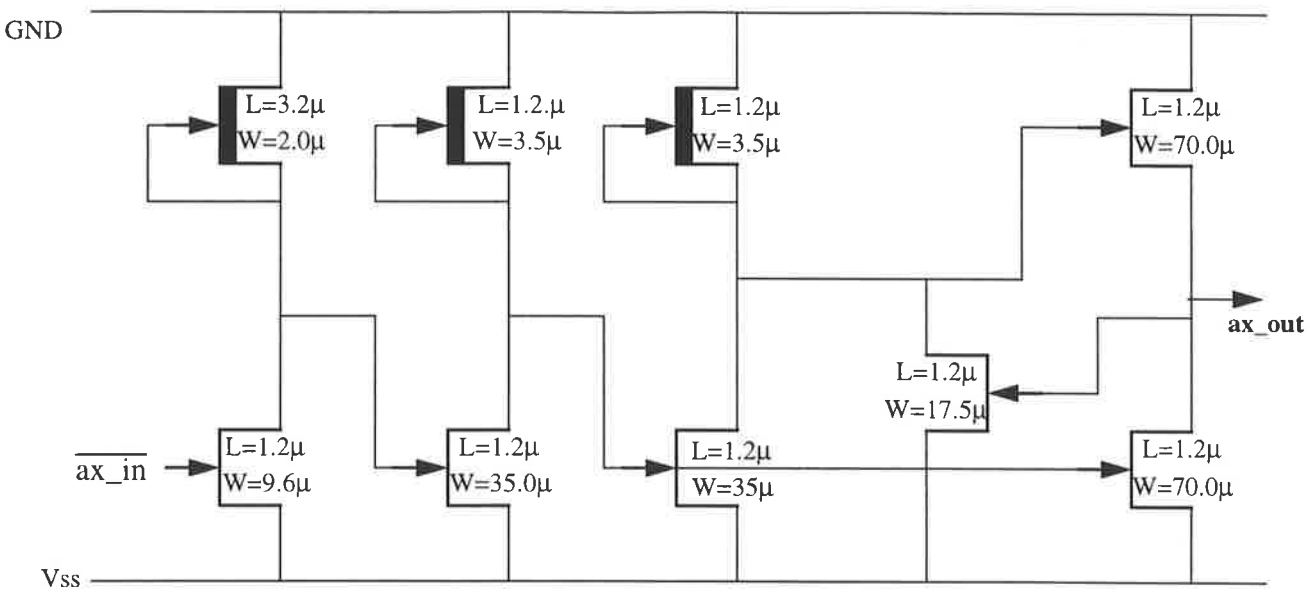


FIGURE A.5. Schematic of the UBFL driver

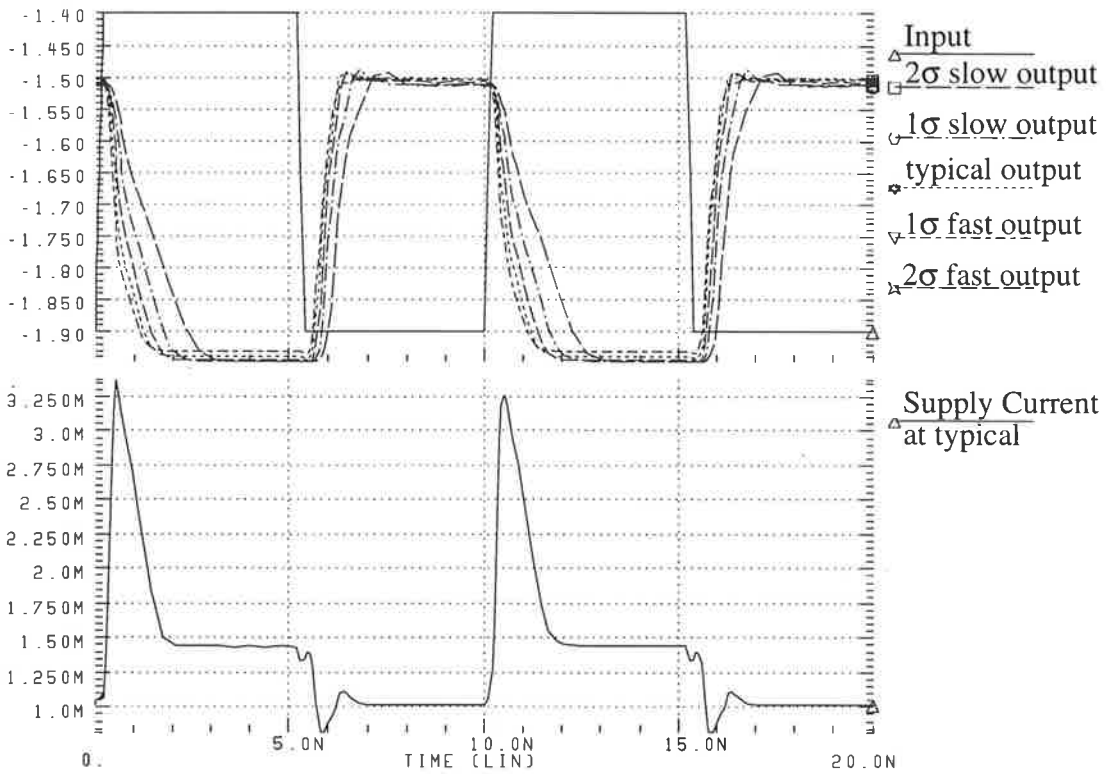


FIGURE A.6. Simulation of UBFL driver

driver to maintain an approximately constant static power dissipation, independent of logic level output.

Because of the double super buffer output stage, the size of the driving inverter can be substantially reduced, thus offering reduced power, while at the same time improving performance over the SBFL and UBFL drivers. Because of the fact that the inverter is smaller, the added first super buffer stage *does not* result in an increase in fan-in compared to the SBFL and UBFL drivers.

A schematic of the driver is shown below in Figure A.1.

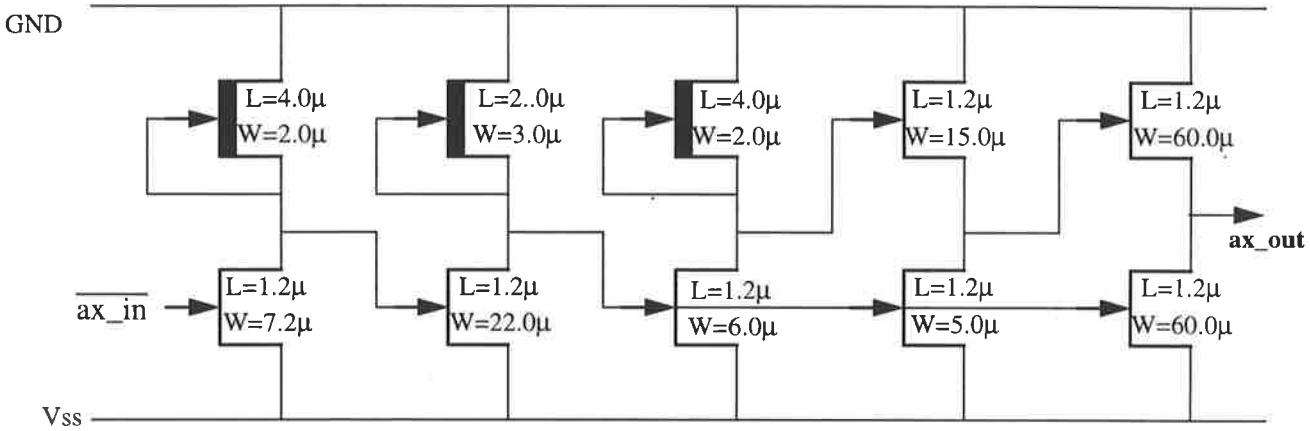


FIGURE A.7. Schematic of the DSBFL driver

The circuit was simulated at 75 °C and 125 °C, using 2σ slow to 2σ fast models. The results of the simulations are shown in Figure A.8. Figure A.8 also shows the current drawn from the supply at 125 °C using typical parameters. It can be seen that the average current drawn is about 500 µA, giving an average dissipation of 1 mW. There is a large current spike however as the highly capacitive output node is discharged. It can be seen that the worst case rise and fall times of the driver are approximately 2 ns and occur at 2σ slow process variation, 125 °C.

A.5 Evaluation and Comparison

A comparison table of results from driver simulations is shown in Table A.1. It can be seen that all drivers were designed to provide approxi-

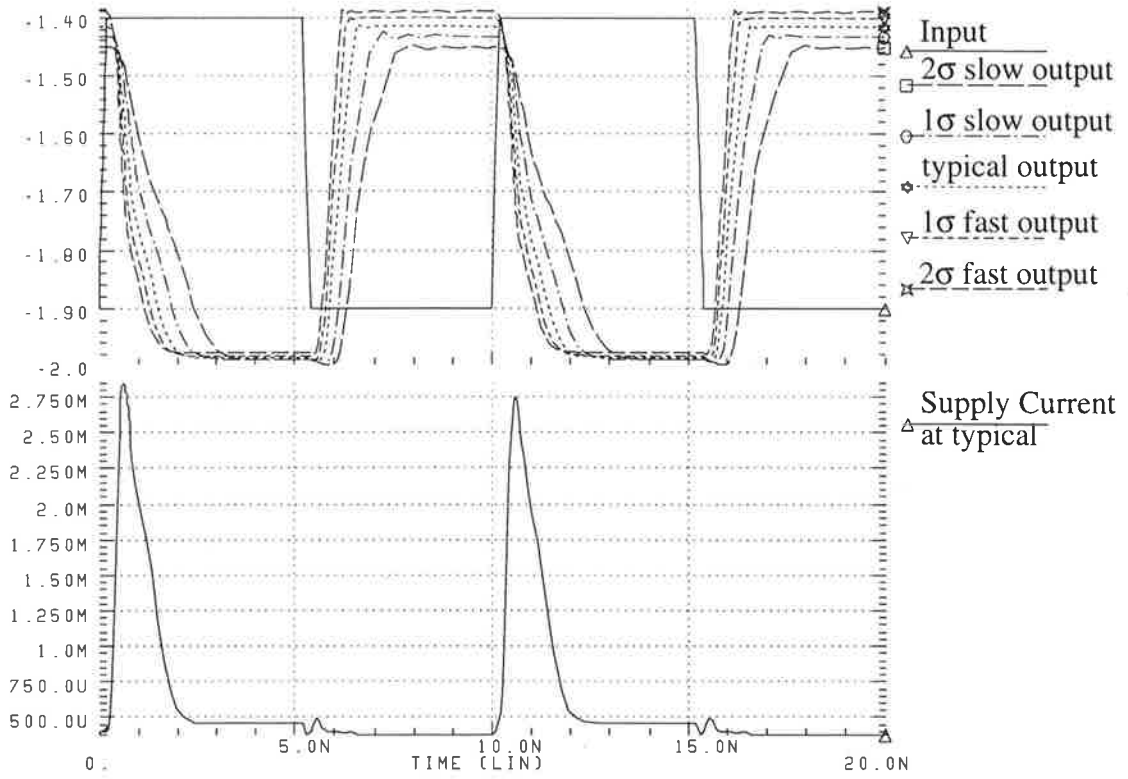


FIGURE A.8. Simulation of DSBFL driver at 75 °C using 2σ slow to 2σ fast parameters

mately the same timing performance. It can be seen that the DSBFL driver exhibits superior performance in the following areas: lower power dissipation, higher logic high output and lower logic low output. Its rise and fall times are also lower than all except UBFL. The reason for UBFL's fast rise and fall times are that its logic swing is the lowest of all (DSBFL's total swing is 0.57 volts or 32% higher than UBFL's). However, due to the extra stage, an additional delay is observed when using the DSBFL circuit. This delay is 8% and 4% more than SDCFL and SBFL respectively, and 20% more than UBFL. UBFL has the smallest rise delay due to the permanently turned on pull-up transistor.

The data demonstrates that DSBFL is a superior driver, provided a small increase in delay can be tolerated, offering a lower power dissipation, better logic swing and fast rise and fall times. Although these benefits are rela-

TABLE A.1 Comparison of driver performance at 75 C and typical parameters

Parameter	SDCFL	SBFL	UBFL	DSBFL
Rise Time (ps)	820	740	560	620
Fall Time (ps)	1180	1140	1020	1080
Rise Delay (ps)	750	780	660	810
Fall Delay (ps)	560	585	530	620
V_{High} *Out (V)*	-1.45	-1.43	-1.51	-1.42
V_{Low} *Out (V)*	-1.98	-1.98	-1.94	-1.99
Average Dissipation (μW)	3800	960	2420	820

tively small, they are significant in this application, particularly as a minimum power dissipation is critical.

The main benefit of UBFL is the lack of current spike as the driver is pulling down when both driving transistors of a SBFL or UBFL driver are on. However when driving a large capacitive load, a large spike is still present due to the rapid discharge of the capacitance. This can be seen in the simulation results. Therefore the UBFL driver has very little noise benefit over the other drivers.

The SDCFL driver offers a low logic low level, but this is bettered by DSBFL and matched by SBFL and so its high power dissipation cannot be justified.